Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT
FÜR MATHEMATIK, INFORMATIK
UND NATURWISSENSCHAFTEN

Characterization of interactions between Lassa virus nucleoprotein, matrix protein and RNA essential for RNP assembly and recruitment

Dissertation

with the aim of achieving a doctoral degree at the Faculty of Mathematics, Informatics and Natural Sciences

Department of Chemistry

University of Hamburg

submitted by

Lennart Sänger

born in Elmshorn

Hamburg, 05.04.2024

Thesis reviewers

Prof. Dr. Charlotte Uetrecht

Leibniz Institute of Virology, Hamburg, Germany; Centre for Structural Systems Biology, Hamburg, Germany; University of Lübeck, Lübeck, Germany; Deutsches Elektronen-Synchrotron, Hamburg, Germany.

Prof. Dr. Hartmut Schlüter

Institute of Clinical Chemistry and Laboratory Medicine, Section Mass Spectrometry and Proteomics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.

Examination board of the disputation

Prof. Dr. Kay Grünewald

University of Hamburg, Hamburg, Germany, Centre for Structural Systems Biology, Hamburg, Germany

Prof. Dr. Hartmut Schlüter

Institute of Clinical Chemistry and Laboratory Medicine, Section Mass Spectrometry and Proteomics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.

Prof. Dr. Maya Topf

Leibniz Institute of Virology, Hamburg, Germany, Centre for Structural Systems Biology, Hamburg, Germany, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.

Thesis submission

05.04.2024

Disputation

21.06.24

This thesis was prepared from 2020 to 2024 under the supervision of Dr. Maria Rosenthal at the Bernhard Nocht Institute, in the working group *Structural Virology* and Prof. Dr. Charlotte Uetrecht at the Leibniz Institute of Virology and the Centre for Structural Systems Biology, in the working group *Dynamics of Viral Structures*.

Contents

Publications

Parts of this thesis have been published or are in preparation for publication.

Abbreviations

°	degree (angle)
AC	alternating current
Ac-	acetylated
AUC	area under the curve
CAS	chemical abstract service registry number
CEM	chain ejection model
CID	collision-induced dissociation
cRNA	complementary RNA
CRM	charged residue model
CCHFV	Crimean-Congo hemorrhagic fever virus
Da	dalton
DC	direct current
DNA	deoxyribonucleic acid
DTT	dithiothreitol
e	elementary charge
eIF4E	eukaryotic translation initiation factor 4E
E. coli	*Escherichia coli*
e.g.	*exempli gratia*, Latin for “for example”
EDTA	ethylenediaminetetraacetic acid
EM	electron microscopy
ESCRT	endosomal sorting complex required for transport
ER	endoplasmic reticulum
ESI	electrospray ionization
et al.	*et alii*, Latin for “and others”
EtOH	ethanol
eV	electronvolt
FWHM	full width at half maximum
g	gravitational force equivalent
GST	glutathione-S-transferase
GHS	Globally Harmonized System of Classification and Labelling of Chemicals
GPC	glycoprotein precursor
H	hazard (statement)
H^+	hydrogen proton
HDX	hydrogen deuterium exchange
HCl	hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IEM	ion evaporation model
IFN	interferon
IPTG	isopropyl-β-D-1-thiogalactopyranosid
ISD	in-source dissociation
K	kelvin
K_a	association constant
K_d	dissociation constant
kDa	kilodalton
L	Large viral protein
LASV	Lassa virus
LC-MS	liquid chromatography mass spectrometry
m	mass
M	molecular weight
M_{exp}	experimental molecular weight
M_{th}	theoretical molecular weight
m/z	mass-to-charge ratio
MBP	maltose binding protein
MALDI	matrix-assisted laser desorption-ionization
mRNA	messenger ribonucleic acid
MS	mass spectrometry
MS^n	tandem mass spectrometry
MS^2	two-stage mass spectrometry
MS^3	three-stage mass spectrometry
MWCO	molecular weight cut-off
NP	nucleoprotein
NaCl	sodium chloride
$-NH_2$	amidated
NH_4^+	ammonium cation
nM	nanomolar
NMR	nuclear magnetic resonance (spectroscopy)
PMSF	phenylmethylsulfonyl fluoride
PTMs	post-translational modifications
Q	quadrupole
QE-UHMR	Q Exactive™ UHMR Hybrid Quadrupole Orbitrap™
RVFV	Rift Valley fever virus
RNP	ribonucleoprotein complex
RING	Really Interesting New Gene
RdRp	RNA-dependent RNA polymerase
SD	standard deviation
SDS-PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis
SEC	size-exclusion chromatography
ssRNA	single stranded RNA
SSP	stable signal peptide
t	time
term.	terminal
TOF	time-of-flight
Tris	tris(hydroxymethyl)aminomethane
UTR	conserved untranslated region
U	potential
v	velocity
vRNA	viral RNA
V	Volt
WHO	World Health Organization

In this thesis, amino acids are abbreviated according to the single-letter amino acid code.

alanine	Ala	A
arginine	Arg	R
asparagine	Asn	N
aspartic acid	Asp	D
cysteine	Cys	C
glutamic acid	Glu	E
glutamine	Gln	Q

Abbreviations

glycine	Gly	G
histidine	His	H
isoleucine	Ile	I
leucine	Leu	L
lysine	Lys	K
methionine	Met	M
phenylalanine	Phe	F
proline	Pro	P
serine	Ser	S
threonine	Thr	T
tryptophan	Trp	W
tyrosine	Tyr	Y
valine	Val	V

List of figures

List of tables

Abstract

Lassa virus (LASV) is a negative-sense single-strand RNA virus that causes periodic outbreaks in West Africa. LASV consists of only four structural proteins. The nucleoprotein (NP) encapsidates the viral genome, forming ribonucleoprotein complexes (RNPs) together with the viral RNA and the L protein. These RNPs are the minimal structural and functional units for viral genome replication and transcription in cells. The Z protein is important for membrane recruitment of RNPs, viral particle assembly, and budding and has also been shown to interact with the L protein. The glycoprotein GP is the receptor protein important for cell attachment and entry.

Our understanding of the interplay among the viral components NP, Z, and viral RNA, as well as the interaction between NP and L, remains limited. In this work structural mass spectrometry and other complementary methods were used to characterize the interactions between NP, Z, and RNA. The presence of RNA was identified as the driver for the disassembly of ring-like NP trimers, a storage form, into monomers to subsequently form higher order RNA-bound NP assemblies. An intermediate state of the trimeric NP bound to one RNA substrate was discovered, indicating a potential secondary binding site. The interaction site of Z and NP was located, while NP binds Z independently of the presence of RNA. The interaction between NP and Z appears to be pH-dependent that could hint to a pH dependent dissociation mechanism of the RNP from the viral matrix during endocytosis.

In this work nMS was used to screen for potential interaction partners of the viral RNP. NP does not directly bind to L indicating that interaction may be mediated by RNA. Different RNA substrates were used to reconstruct the RNP. Additionally, the known influenza virus host factor ANP32A was tested for potential direct interaction with the viral proteins. Contrary to influenza virus, no direct interaction between ANP32A and L was detected. However, ANP32A appears to bind to NP via its LRR domain.

The results of this work improved our understanding of RNP assembly, recruitment, and release in LASV and identified a potential host interaction partner for NP. The results can support the development of antiviral strategies.

Zusammenfassung

Das Lassa-Virus (LASV) ist ein Negativstrang-RNA-Virus, das in Westafrika periodische Ausbrüche verursacht. LASV besteht aus nur vier Strukturproteinen. Dem Nukleoprotein (NP), welches das virale Genom einkapselt und zusammen mit der viralen RNA und dem L-protein Ribonukleoproteinkomplexe (RNPs) bildet. Diese RNPs sind die minimalen strukturellen und funktionellen Einheiten für die virale Genomreplikation und Transkription in Zellen. Das Z-Protein ist wichtig für die Membranrekrutierung von RNPs, den Zusammenbau viraler Partikel und die Knospung und interagiert nachweislich mit dem L-Protein. Das Glykoprotein GP ist das für den Zelleintritt wichtige Rezeptorprotein.

Unser Verständnis des Zusammenspiels zwischen den viralen Komponenten NP, Z und viraler RNA sowie der Interaktion zwischen NP und L ist nach wie vor begrenzt. In dieser Arbeit wurden strukturelle Massenspektrometrie und andere ergänzende Methoden eingesetzt, um die Interaktionen zwischen NP, Z und RNA zu charakterisieren. Die Anwesenheit von RNA wurde als treibende Kraft identifiziert für die Zerlegung von ringförmigen NP Trimeren (einer Speicherform) in Monomere, um anschließend RNA-gebundene NP-Assemblierungen höherer Ordnung zu bilden. Ein Zwischenzustand des trimeren NP wurde identifiziert, welches an ein RNA-Substrat gebunden ist. Dies deutet auf eine mögliche sekundäre Bindungsstelle hin. Die Interaktionsstelle von Z und NP wurde lokalisiert. NP bindet unabhängig von der Anwesenheit von RNA an Z, jedoch ist die Interaktion zwischen NP und Z pH-abhängig, was auf einen pH-abhängigen Dissoziationsmechanismus des RNP von der viralen Matrix während der Endozytose hindeuten könnte.

In dieser Arbeit wurde strukturelle Massenspektrometrie verwendet, um nach potenziellen Interaktionspartnern des viralen RNP zu suchen. NP bindet nicht direkt an L, was darauf hindeutet, dass die Interaktion durch RNA vermittelt wird. Verschiedene RNA-Substrate wurden verwendet, um Faktoren für die Rekonstruktion des RNP zu finden. Zusätzlich wurde der Influenza Virus Wirtsfaktor ANP32A auf eine mögliche direkte Interaktion mit den viralen Proteinen untersucht. Im Gegensatz zum Influenza Virus wurde keine direkte Interaktion zwischen ANP32A und dem L Protein nachgewiesen. ANP32A scheint jedoch über seine LRR-Domäne an das NP zu binden.

Die Ergebnisse dieser Arbeit haben unser Verständnis des RNP-Aufbaus, der Rekrutierung und der Freisetzung in LASV verbessert und einen potenziellen Wirtsinteraktionspartner für NP identifiziert. Die Ergebnisse können dazu beitragen, neue antivirale Strategien zu entwickeln.

1 Introduction

1.1 Arenavirus family

Arenaviruses are a family of viruses that are part of the order of *Bunyavirales* which currently includes 14 different families of mostly insect or rodent borne viruses [1].
The family of *Arenaviridae* was first classified in 1976 as viruses with bi-segmented single strand (ss) RNA genome employing two genes in opposite direction on two genome segments (ambisense coding strategy) which is different to most other bunyaviruses. The name derived from its sandy (latin *arenosus*) appearance in electron microscopy images [2]. Arenaviruses have a broad range of hosts and can be divided into 5 different genera based on their host specificity: Fish (*Antennavirus*), snakes (*Hartmanivirus* and *Reptarenavirus*) and rodents (*Mammarenavirus*). Additionally, the genus *Innmovirus* was recently established with the discovery of Hailar virus from river sediment samples. [3-5].
Mammarenaviruses constitute the largest genus in the family of arenaviruses with currently 43 different species [3]. The natural hosts of mammarenaviruses are almost exclusively rodents apart from Tacaribe virus which is bat-borne [3, 6, 7]. Species from the genus *Mammarenavirus* are divided into Old World and New World viruses based on geographic, serologic, and phylogenetic criteria. Old World arenaviruses are endemic to Africa, Europe and Asia, whereas New World arenaviruses are found in the Americas. Notable exception is the Old World arenavirus lymphocytic choriomeningitis virus, which is distributed worldwide [8-10]. Some mammarenaviruses can spread from animals to humans and cause severe diseases: Lassa virus (LASV), Lujo virus, Machupo virus, Guanarito virus, Junin virus, and Sabia virus can cause life-threatening viral hemorrhagic fever. To date, there are no FDA-approved vaccines or antivirals available, and treatment is limited to off-label use of ribavirin, which has severe side effects. [11, 12].

1.2 Lassa virus

Multiple bunyaviruses can cause human diseases and are of public health concern emphasized by the WHO R&D Blueprint list of priority diseases, including the bunyaviruses Rift Valley fever virus (RVFV), Crimean-Congo hemorrhagic fever virus (CCHFV) and Lassa virus (LASV) [13].

LASV was first discovered in 1969 in the town of Lassa in Nigeria where it was first isolated [14]. Besides Nigeria, the virus is endemic to the West African countries Guinea, Sierra Leone and Liberia. LASV is the causative agent of Lassa fever which is a life-threatening hemorrhagic fever [15]. The main natural host species is the rodent *Mastomys nathalensis,* but *Mastomys erytholeucus* and *Hylomyscus pamfi* have been additionally linked as natural hosts [15, 16].

Human infections occur primarily via direct or indirect contact with infected animals, such as contaminated food, aerosols from excrements or hunting of infected rodents [6, 17, 18]. However, human-to-human transmission has been observed via direct contact with virus-containing body fluids [19].

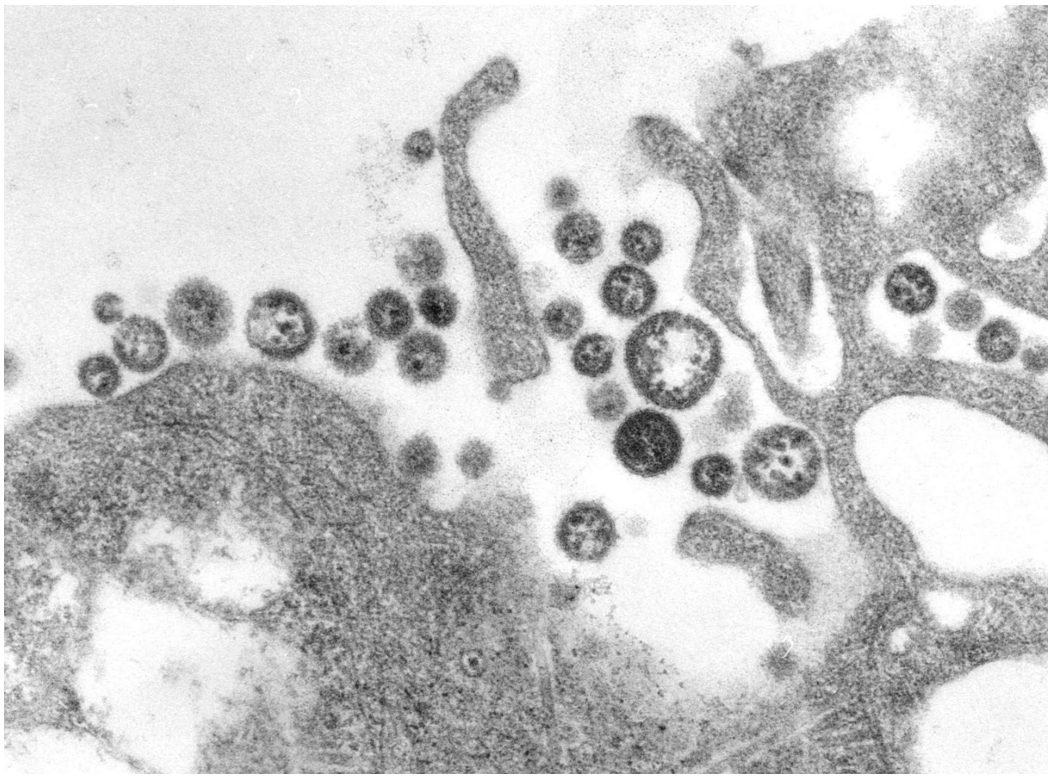

Figure 1: LASV virions

Transmission electron micrograph of LASV virions with cell debris. Source: CDC's Public Health Image Library Image #8700, C. S. Goldsmith

LASV infects between 100,000 and 300,000 humans in endemic countries each year leading to approximately 5,000 deaths [20]. Recent predictions estimate up to 186 million people at risk of LASV infections in West Africa by 2030 [21]. Especially pregnant women are at high risk when infected with LASV. Infections with LASV have been linked to a threefold increase in the risk of death of pregnant women and abortion in many cases [22, 23]. Initially, infections manifest with unspecific symptoms like fever, gastrointestinal symptoms, and headache. In approximately 20 % of cases, infected individuals develop severe disease. This can include thrombocytopenia, leucopenia, pleural effusion, pulmonary oedema, mucosal hemorrhages, kidney failure, lung failure, generalized circulatory failure, and neurological complications including encephalopathy, seizures, hearing loss and coma [24-26]. Recovered patients may suffer from long-term consequences of Lassa fever such as hearing loss, vision distortion and polyserositis [27]. However, numerous cases of LASV infections are mild or asymptomatic, leading to a considerable number of undiagnosed cases. [28].

1.2.1 Structure of the LASV particle

The LASV particle is enveloped and has a typical diameter of 40 to 200 nm. The appearance is spherical or pleomorphic, which is typical of Arenaviruses [29] (Figure 1). LASV contain a bi-segmented negative-sense single-stranded RNA genome divided into a large (L) and small (S) segment with a size of 7.2 kb and 3.4 kb, respectively [30, 31]. Each genome encodes for two structural proteins which are separated by a non-coding, highly structured intergenic region. The L genome segment encodes the ~250 kDa large (L) protein, which contains the RNA-dependent RNA polymerase, and the small ~ 11 kDa Really Interesting New Gene (RING) finger matrix protein Z. The S genome segment encodes the viral glycoprotein precursor (GPC, ~75 kDa) and the nucleoprotein (NP, ~63 kDa) [32]. Both genes on each segment are in opposite orientation (ambisense coding strategy). This implies that one gene on each genome segment can only be transcribed from a complimentary copy of the genome, called antigenome [32]. The conserved untranslated regions (UTR) at the 3’ and 5’ ends of the genome segment are complementary to each other, form a panhandle structure to some extend and are bound by the L protein sequence specifically [33].

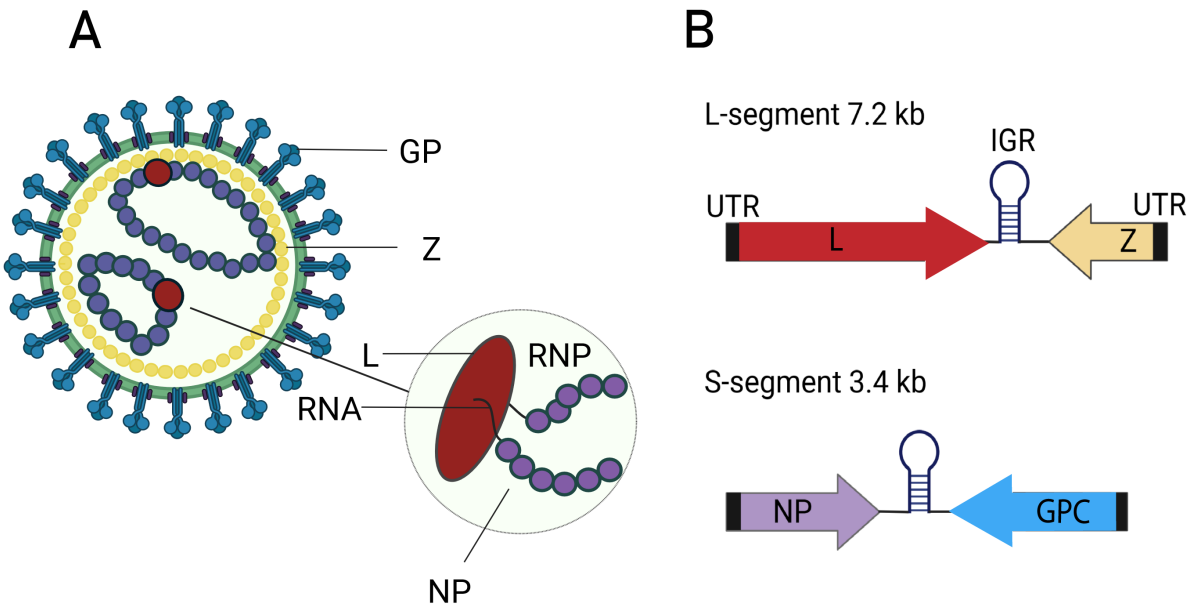

Figure 2: LASV particle and genome structure

(A) The LASV particle consists of 4 viral proteins GP, L, NP and Z. NP and L form the RNP complex together with the RNA genome. The RNP complex is the structural and functional unit of genome replication and transcription. L is the viral polymerase and performs both replication and transcription of the genome, whereas NP encapsidates the viral genome. Z lines the inside of the viral envelope and forms the matrix of the virus. GP is a heterotrimeric receptor complex on the surface of the virus. (B) Schematic overview of the LASV genome structure. The LASV genome is divided into two segments with a size of 7.2 kb (L-segment) and 3.4 kb (S-segment). Untranslated regions (UTRs) are located on both ends of the genome and have regulatory functions. Each segment is composed of two open reading frames in opposite orientation (ambisense coding strategy). Both open reading frames are separated by an intergenic region forming RNA secondary structures. The coding sequence of the L and Z protein are located on the L-segment, NP and GPC are located on the S-segment. The figure was created with the help of biorender.

NP encapsidates the viral RNA genome forming – together with L – the ribonucleoprotein complex (RNP). Within the RNP the viral RNA is protected from degradation and recognition by cellular pattern recognition receptors. The RNP is the structural and functional unit of viral genome replication and transcription [34]. The matrix protein Z lines the inner side of the viral envelope [35]. It is anchored into membranes by N-terminal myristoylation [36]. A homo-trimeric glycoprotein (GP) complex formed by heterotrimers of GP1, GP2 and the signal peptide (SSP) forms the viral receptor and is anchored in the envelope of the viral particle [37].

1.2.2 Replication cycle

Viral particles enter the cells via receptor mediated endocytosis. Entry is initiated primarily by interaction between GP-1 and the host cell receptor α-dystroglycan [37]. However, TAM family, C-type lectins and Axl have been implicated to play additionally a role as entry receptors [38]. Cellular entry is mediated by a clathrin-independent endocytic mechanism involving the endosomal sorting complex required for transport (ESCRT) proteins. Recently the micropinocytosis pathway has also been identified as entry mechanism for LASV [39, 40]. The endosome acidifies during endocytosis, causing a conformational change of the GP-1 and GP-2, which facilitates membrane fusion. This process is supported by a receptor switch of GP-1 to the lysosomal-associated membrane protein 1. The membrane fusion leads to the release of RNP complexes into the cytosol and genome transcription and replication can start [41].

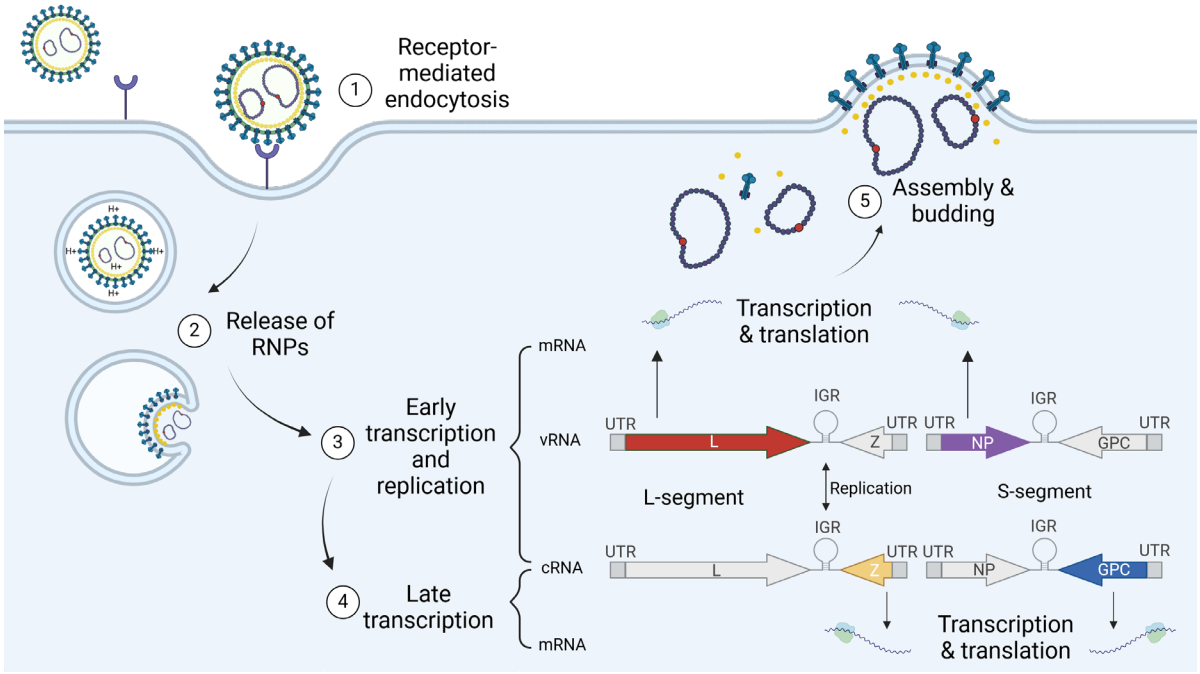

Figure 3: LASV replication cycle

The viral particle binds via GP-1 to the host cell receptor α-dystroglycan and enters the cell via receptor mediated endocytosis. During endocytosis, a lower pH induces a conformational change of GP-1 and GP-2 that mediates membrane fusion between endosomal and viral membrane. The RNP is subsequently released into the cytosol, where early viral transcription and genome replication occur. The L protein produces mRNA coding for the L and NP protein and replicates the viral RNA forming complementary RNA (cRNA). The antigenomic cRNA acts as a template for the late transcription of Z and GPC genes and the production of vRNA. All transcripts are translated into proteins by the host cell machinery. GPC is glycosylated and processed in the Golgi apparatus that results in the formation of a homo-trimeric receptor formed by heterotrimers of GP1, GP2 and SSP. Newly synthesized RNP complexes together with the Z proteins are transported to the cell membrane. Z mediates budding of newly formed virions that exit the cell. The figure was created with the help of biorender.

After the release of the RNPs into the cytosol the genomic/viral RNA (vRNA) can either be transcribed to form mRNA of the L and NP genes or replicated into an antigenomic/complementary RNA (cRNA) [42].

The viral mRNA is produced by the L protein involving a cap snatching mechanism. Host mRNAs are cleaved by the L protein yielding short RNA fragments with an m7GTP cap structure at the 5' end. These capped fragments directly prime the elongation of new viral mRNAs by the viral polymerase. By this process, the viral mRNA is composed of a 5'cap structure stolen from host mRNAs but misses a poly A tail. [43, 44].

The L protein initiates genome replication *de novo* by a prime-and-realign mechanism. The L protein binds a nucleotide at position 2 of the template, followed by addition of a second nucleotide at position 3. The dinucleotide is realigned to position -1 and 1 and the full genome is elongated. The L protein facilitates the replication process by synthesizing cRNA using vRNA as a template. Conversely, it utilizes cRNA as a template for the generation of new vRNA during the replication cycle. This process is fundamental to the viral life cycle, allowing for the production of both genomic and antigenomic RNA strands. [45, 46].

In the late transcription stage Z and GPC are transcribed from cRNA initiating the final phase of the replication cycle. When Z is produced to a certain amount in infected cells, it binds the L protein, thereby inhibiting its polymerase activity, potentially by blocking the formation of an elongation conformation of L [33, 45, 47-49]. Additionally, Z has been hypothesized to interact with NP directly or indirectly, thereby mediating recruitment of RNP complexes to the plasma membrane for virion assembly [35, 50, 51]. It remains unclear whether RNP recruitment is driven by the interaction between NP and Z, L and Z and/or via a host protein. The accumulation of Z at the plasma membrane enables the assembly of virions, with the involvement of ESCRT proteins that interact with Z and mediate budding of the LASV particle [52].

1.2.3 Lassa virus proteins

In the following chapters, details about the structure and function of the individual LASV proteins will be explained.

The most abundant viral protein is the NP (63 kDa) which is responsible for the encapsidation of the viral genome. It protects viral RNA from degradation and recognition by cellular pattern recognition receptors. NP forms together with L and the viral RNA the RNP complex [34]. Co-immunoprecipitation data has previously shown for arenaviruses, that NP and Z, either directly or mediated by other factors, interact with each other [50, 53, 54]. It has been hypothesized that the NP-Z interaction mediates virion assembly by recruiting viral RNPs to the plasma membrane.

NP is composed of two functional domains (N-terminal and C-terminal domain). The N-terminal domain contains an RNA binding pocket covering 6-8 nucleotides [55, 56]. Additionally, the N-terminal domain can be co-crystallized with a cap analogon [57]. However, whether NP binds m7GTP specifically remains unclear as interaction was not detectable by others. The C-terminal domain contains a 3'-5' exoribonuclease activity that cleaves double stranded RNA to protect it from recognition of the host cell. Double stranded RNA is generally detected during infection by the pattern recognition receptors of the host cell inducing an interferon (IFN) response [58, 59]. NP can additionally actively inhibit IFN response via several mechanisms: It interacts directly with I-kappa-B kinase epsilon and the Retinoic Acid Inducible Gene 1, which leads to the inhibition of the interferon regulatory factor 3 [59, 60]. Furthermore, the activity of NF-κB is inhibited by NP [61].

When not bound to RNA, NP forms homotrimeric rings with a head to tail arrangement of NP molecules. In this conformation, RNA binding is hindered as the respective pocket is blocked by two helices. [56]. To allow for RNA binding, an RNA gating mechanism was proposed where the C-terminal domain rotates slightly away allowing one helix to become partly unstructured (residues 112–122) whereas the other helix (residues 128-145) shifts away from the RNA binding pocket. As the trigger and mechanism of this conformational change are unknown, it was speculated that a (specific) host factor was needed. Notably, residues responsible for NP-NP interaction were shown to be critical for the virus [55, 62].

The glycoprotein GP is the receptor protein of the virus and is anchored in the envelope of the viral particle. Overall, the 3 different domains of GP have the following functions: GP-1 contains the receptor binding domain responsible for host cell binding for the viral entry. The N-terminal signal peptide (SSP) includes hydrophobic patches and is together with an N-terminal myristoilation site integrated into the membrane. The SSP interacts with the Z protein which is important for cell exit. GP-2 is partly membrane anchored and contains a fusion peptide for host cell fusion. [63, 64].

GP is synthesized as a precursor protein (GPC) with a size of 76 kDa including both GP-1 and GP-2 subunits and the SSP [65]. The SSP is cleaved from GP-1/GP-2 by signal peptidases in the endoplasmic reticulum (ER) and remains non-covalently bound [66, 67]. The non-covalent interaction of the SSP masks the ER localization of GP-2 that results in the transport of the complex to the Golgi compartment [68]. GP1-GP2 are cleaved by the host site 1 protease in the ER or Golgi [63]. Both subunits remain non covalently bound with each other forming the mature heterotrimeric GP complex [66]. Critical for maturation, function, transport and oligomerization of the GP complex are several N-glycosylation sites of the GP proteins [69]. The role of GP in cell entry is explained in chapter 1.2.2.

The LASV L protein contains an RNA-dependent RNA polymerase (RdRp) which catalyzes both genome replication (vRNA and cRNA) and transcription as explained in chapter 1.2.2. It consists of one peptide chain with a size of approximately 250 kDa, which is contrary to other negative strand RNA virus polymerases like the influenza virus polymerase complex, which is a multi-protein complex.

Due to the challenging size of the protein, detailed structural information of the L protein has been missing for a long time. The first low-resolution envelope of the whole L protein was provided by electron microscopy (EM) in 2010 and small-angle X-ray scattering in 2019 [45, 70]. But only recent achievements in the field of cryogenic electron microscopy (Cryo-EM) made it possible to obtain high-resolution structural information. In 2020, the structure of the LASV L protein was solved with a resolution of 3.9 Å and led to a classification of regions of the L protein [71]. In general, and similarly to the influenza virus polymerases, the L protein can be divided into three subunits: N-terminal PA-like region, the central RdRp (PB1-like) region and

the C-terminal PB2-like region. The RdRp region forms the core and has a fingers-palm-thumb right-hand motif (715-1612) [71]. The RdRp core is surrounded by the PA-like region (1 bis 715) and the PB2-like (1612 -2218) region that binds to the thumb of the RdRp core. The endonuclease is located on the opposite of the thumb and the PB-2 region. The structure is missing large parts of the PB2 like region including the cap binding domain (CBD) [71, 72]. Certain residues in the region for the potential CBD were shown to be essential for viral transcription but not genome replication [73]. In 2021, 9 additional cryo-EM structures were published showing the LASV L protein in apo-, promoter-bound pre-initiation, and active RNA synthesis states. The protein undergoes substantial conformational changes when binding promoter sequences at distinct promoter binding sites. Two different modes of function have been structurally resolved: A pre-initiation state and an elongation state. In the pre-initiation state, the terminal 12 nt of the 5’ promoter RNA forms a hook-like structure, which is bound to a specific site on the surface of the L protein. The distal region (nt 12-19) of the 5’ end forms a duplex with the distal region (nt 12-19) of the 3’ promoter RNA. Consequently, the 3’ terminus can enter the RdRp active site as a single stranded template reaching the active site motif ready for initiation of RNA synthesis. A template-product duplex is formed in the active site during RNA synthesis in the elongation state. The promoter duplex is disrupted by translocation of the template while the 5’ hook remains bound in its dedicated pocket. Additionally, a structure of the L protein only together with the 3’ promoter sequence identified a 3’ secondary binding site involving the pyramid domain. Presumably the secondary 3’ binding site rebinds the 3’end after it has left the RdRp active as elongation proceeds, which has been observed for other L proteins of other bunyaviruses [33, 74, 75].

The matrix protein Z has a size of 11 kDa and lines the inner side of the viral envelope [35]. It is anchored into membranes by N-terminal myristoilation [36]. It is composed of a central RING domain coordinating a zinc ion and has two terminal unstructured tails. Both monomeric and oligomeric forms of Z have been resolved [76, 77]. In the oligomeric form, 6 symmetric dimers are assembled as a ring with highly basic interactions. Beside the dimer interface, the oligomeric ring is stabilized by two interaction sites involving hydrophilic and hydrophobic patches [77]. A specific function of the Z dodecamer is not yet known.

The C-terminal tail of Z contains two so-called late domain motifs, which are proline rich motifs (PTAP and PPPY) interacting with the ESCRT proteins inducing virion budding and membrane fission [52, 78]. Repression of eukaryotic translation initiation factor 4E (eIF4E) dependent translation was shown by direct interaction between Z and eIF4E [76, 79].

Another important role of the Z protein is the regulation of viral genome replication and transcription. Biochemical assays demonstrated that Z inhibits RNA synthesis of L in a dose-dependent manner [45, 47]. Cryo-EM structures of the L-Z complex demonstrated that Z directly interacts with the L protein [48, 49]. Comparison to the elongation-state structure of L suggests that inhibition of the polymerase activity is potentially caused by blocking the formation of an elongation conformation of L [33].

1.3 Structural mass spectrometry

A key to develop antiviral strategies is the in depth understanding of the cell biology of viral infections. Due to the limited coding capacity in their genomes, viruses are dependent on protein-protein and protein-nucleic acid interactions with host factors and between the viral components [80]. However, cellular complexity often precludes precise investigations making biochemical investigations necessary. One fundamental way of biochemical investigations is structural biology. It provides detailed information on biological molecules, sometimes even with near atomic resolution. Recent developments in the field of Cryo-EM allowed the investigation of larger protein complexes, which was previously not possible with conventional structural biology methods such as X-ray crystallography or nuclear magnetic resonance spectroscopy. However, these high-resolution techniques have the disadvantage that they only capture snapshots of dynamic proteins. Integrating dynamic structural data of proteins into complex structural models can be achieved by mass spectrometry (MS).

The development of the first MS instrument was pioneered by J.J. Thomson and E. Everett around 1900 in Cambridge. The instrument was first used to measure the charge-to-mass ratio (e/m) and in parallel the charge (e) of electrons indirectly gives the mass of electrons [81]. The basic principle of this instrument has until now not changed to this day: Analytes are ionized, transferred into the gas phase, but now separated based on their mass-to-charge (m/z) ratio. Since then, MS was continuously improved and especially the work from Fenn, Gall, Hillenkamp and Karas in the 1980s led to the development of the soft ionization techniques electrospray ionization (ESI) and matrix-assisted laser desorption-ionization (MALDI) [82, 83]. Both techniques allowed to measure large molecules (over several thousands of Daltons) such as polymers, glycans or proteins without fragmenting them [84]. The ability to measure large molecules with these ionization methods led to the idea of not only investigating the mass of these large analytes, but also keeping their “native” state intact. Proteins and protein complexes are folded into diverse structural elements depending greatly on non-covalent interactions. Analyzing such non-covalent interactions of proteins with MS - or in other words:

structural MS - was mainly driven by the groups of Robbinson, Loo and Heck in the last decades. They developed methods based on ESI to analyze proteins in their native state, mainly overcoming the issue of preserving non-covalent interactions [85-87]. Other structural MS methods are rather based on the traditional MS set up for proteins, identifying and quantify them (bottom up proteomics) [88]. Such structural MS methods are grounded on labelling of proteins to detect their dynamics and conformations in their native environment. Commonly used labeling techniques are chemical crosslinking or isotopic exchange like hydrogen-deuterium exchange mass spectrometry (HDX-MS) [89, 90].

1.3.1 Native mass spectrometry

Native mass spectrometry (nMS) is an approach to analyze biological analytes based on their mass from non-denaturing solvent. Analysis of samples such as proteins in the gas-phase while noncovalent interactions are maintained, provides detailed information about composition dynamics and stoichiometry. The ionization event therefore needs to be as soft as possible to keep non-covalent interactions intact [91].

Whether the protein 3D structure is conserved during measurements has long been controversially discussed. Recent developments in the field of MS based soft matrix landing approaches, showed for the first time that protein structures from the gas phase of a mass spectrometer are very similar to structures from traditional cryo-EM workflows but are more condensed as water molecules are not present in the gas phase [92, 93]. Other nMS experiments with complete viruses shown that viruses were still infective after the measurement which demonstrated the ability of the method to preserve the functional 3D structure of proteins [94].

1.3.1.1 Sample application

Proteins are the most frequently analyzed samples when using nMS. Protein or protein complexes are typically recombinantly overexpressed and purified from the expression host prior nMS measurements. However, it is possible to analyze proteins directly from plant and organ tissues or cell extracts [95, 96].

To determine the exact masses of protein ions, they need to be free from adducts, which are commonly formed from components of the sample buffer and cause

signal suppression. A challenge here is to keep the conditions for proteins as native as possible without using buffer with non-volatile substances. Ammonium Acetate (AmAc) solutions as a buffer surrogate are frequently used in combination with nMS. AmAc can undergo a proton transfer from NH_4^+ to Ac^-, which results in NH_3 and HAc. Both products are volatile and ensure that proteins are adduct-free after ionization. AmAc solutions are buffering the pH at around 4.75 and 9.25 and not at physiological pH. To address a possible acidification during sample application, AmAc is usually used with concentrations higher than 150 mM [97]. Solutions with high concentrations of AmAc can tolerate a few milli molar of non-volatile substances such as co-factors or reducing agents. The use of a smaller diameter of the ESI emitter can further suppress the formation of adducts by non-volatile salts present in the sample buffer [98].

Protein buffers are normally exchanged with molecular weight cut-off (MWCO) filters, size-exclusion columns or dialysis devices. An alternative is the use of AmAc solution the final purification step. However, an exchange method needs to be carefully selected for every sample, as some proteins tend to aggregate during buffer exchange [99]. After buffer exchange, proteins are generally adjusted to a concentration of 0.5 – 20 µM. Higher concentrations may cause nonspecific artefacts which can lead to false positive complex formations [99]. A workable concentration needs to be evaluated for the analyzed protein and can vary depending on the instrument and protein itself.

1.3.1.2 Ionization of biological samples for the transition into the gas phase

The development of ESI as a soft ionization method enabled the ability to measure samples with large molecular weights, for example biological samples. Ions are usually generated from charged droplets formed with ESI in a positive ion mode. The analyzed sample is placed in a gold coated spray emitter with a flow rate in the range of a few nl/min with a nano ESI set up [100]. A potential of several kilovolts is applied between spray emitter and counter electrode, which results in the formation of a Taylor cone. At the tip of the cone, a spray of charged droplets is formed in the direction of the counter electrode [99, 101]. Each droplet has a positive charge due to ions from the sample solution (e.g., H^+, NH_4^+). These initial ESI droplets have a

diameter in the micrometer range [102]. Notably, recent investigations of the behavior of AmAc solutions during ESI reveal that protons get neutralized by Ac^- forming HAc at the tip of the Taylor cone. Because of the greater volatility of NH_3, HAc concentration increases and stabilizes the droplets by buffering the solution at pH 5.4. The pH drop of 1.6 units from initially pH 7.5 (commonly adjusted pH) is normally tolerated by biological analytes because of the short time scale of ESI [103]. The charge density of the droplets accumulates when solvent evaporates until the droplets fission into smaller droplets. The fission of droplets is balanced between surface tension and Coulomb repulsion. The process continues until the Rayleigh limit is reached, which depends on the molecular mass of the analyte. At the Rayleigh limit even smaller highly charged droplets are formed via jet fission. At this stage, the final ESI droplets are in the range of a few nanometers in diameter [102]. There are mainly three different ion release mechanisms which are discussed below.

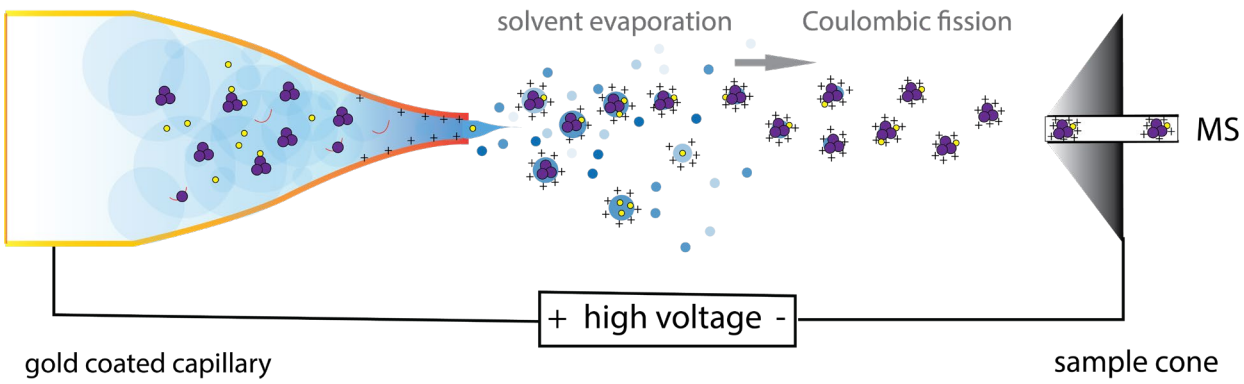

Figure 4: Schematic representation of ESI

In electrospray ionization (ESI), ions are generated by applying a potential of several kilovolts between a gold-coated spray emitter and a counter electrode, leading to the formation of a Taylor cone and the subsequent generation of charged droplets containing ions from the sample solution. The charge density of these droplets accumulates during solvent evaporation, causing the droplets to undergo fission into smaller droplets. This fission process is governed by a balance between surface tension and Coulomb repulsion, resulting in the formation of smaller, highly charged droplets. The final ESI droplets are typically in the range of approximately a few nanometers in diameter.

The theory for transferring peptides, sugars, lipids or other such low molecular mass molecules into the gas phase is the ion evaporation model (IEM) [104]. When the droplet charge reaches a certain level, the analyte is removed from the droplet, together with a small solvent shell. When the analyte enters the mass spectrometer, the remainder of the solvent shell evaporates. Larger and globular molecules (folded

proteins) are released into the gas phase via the charged residue model (CRM) as the IEM is not kinetically viable. Nanodroplets at the size of the Rayleigh limit contain a single analyte and evaporate to dryness, where the charge of the evaporating droplet is transferred mostly to the carboxyl groups of the analyte. The resulting charge state of the protein is independent of the intrinsic charge based on the amino acid composition but only dependent on Rayleigh charge. Unfolded proteins have different properties compared to folded proteins. Formerly non-polar residues at the hydrophobic core are now solvent accessible, which makes it unfavorable for unfolded proteins to be inside a droplet. Thus, unfolded chains drift to the surface of the droplet and the chain gets stepwise injected into the gas phase starting from one terminus of the chain. This process is known as chain injection model (CEM) [102].

1.3.1.3 Mass analysis in the gas phase of a mass spectrometer

Proteins enter the mass spectrometer in the gas phase as charged ions. Thus, the ions can be manipulated via electric fields. A mass spectrometer can have different designs, but all are based on the following components: sample introduction inlet, ion source, mass analyzer, and detector.

A commonly used instrument for nMS is a quadrupole time-of-flight (Q-ToF) mass spectrometer modified for high mass experiments in positive ion mode [105]. Here, a quadrupole mass filter is first used to select for a specific range of ions, a hexapole collision-gas cell for collision induced dissociation (CID) experiments, and a time of flight (TOF) mass analyzer. It is possible to perform tandem MS experiments with this set up. The function of different mass analyzer will be discussed below:

1.3.1.3.1 Mass analyzer

A quadrupole mass analyzer (or mass filter) consists of a set of 4 parallel rods creating a symmetrical electrical field. The voltage placed on one pair of rods has a positive direct current (dc) combined with a radio frequency (rf) voltage. The other pair of rods has a negative dc with a rf that is 180° out of phase. The rods oscillate constantly in a positive and negative polarity, which forces the ions to oscillate as well. Depending how voltages are applied, the corresponding electrical field only forces ions with a certain *m/z* on a stable trajectory to pass through [106].

A TOF mass analyzer determines the time of ions to travel through field free drift tube with a known length. The acceleration voltage (kinetic energy) applied is the

same for all ions that travel through the drift tube. Therefore, the TOF is the function of its momentum (*m/z)*. Ions with the lowest *m/z* arrive at the detector first and ions with the highest *m/z* arrive latest. As ESI produces a constant stream of ions, the flight tube is orthogonal to the ion stream. Pulsed voltages from a pusher can force ions into the analyzer tube. All excited ions need to travel through the analyzer tube before another pulse can be applied (duty cycle). Thus, part of the ions from the continuous stream are not sampled, which limits the sensitivity. The resolving power of a TOF analyzer is dependent on the arrival time of ions with different *m/z* values at the detector. TOF analyzers are often composed of reflectrons to increase the length of the drift tube, which ultimately increases separation of ions [106, 107].

The newest development of a mass analyzer is an orbitrap developed by A. Markarov [108]. The orbitrap is composed of a spindle like inner electrode and two outer electrodes facing each other. A radial electric field between outer and inner electrodes bends ions towards the inner electrode. At the time, ions are collected in an ion trap before being pushed into the orbitrap by high-speed pulses. The pulses provide the ions with enough tangential velocity to create a centrifugal force around the inner electrode. Ions remain in a circular orbit around the inner electrode. In parallel, the special conical form of the inner electrode induces an axial electric field that pushes ions to an axial oscillation. The outer electrodes detect a current based on the axial oscillation of the ions. The axial oscillation is proportional to the *m/z* of the ions and can be interpreted by Fourier transformation [109].

1.3.1.3.2 Analyzing nMS spectra

Usually, proteins have a couple of charge states depicted by a Gaussian distribution based on the most frequently detected charge states. Multiple detected charge states of the proteins allow the determination of the molecular weight of two neighboring peaks considering the *m/z* values and the number of charge states of both peaks. The number of charge states between neighboring peaks always differs by one.

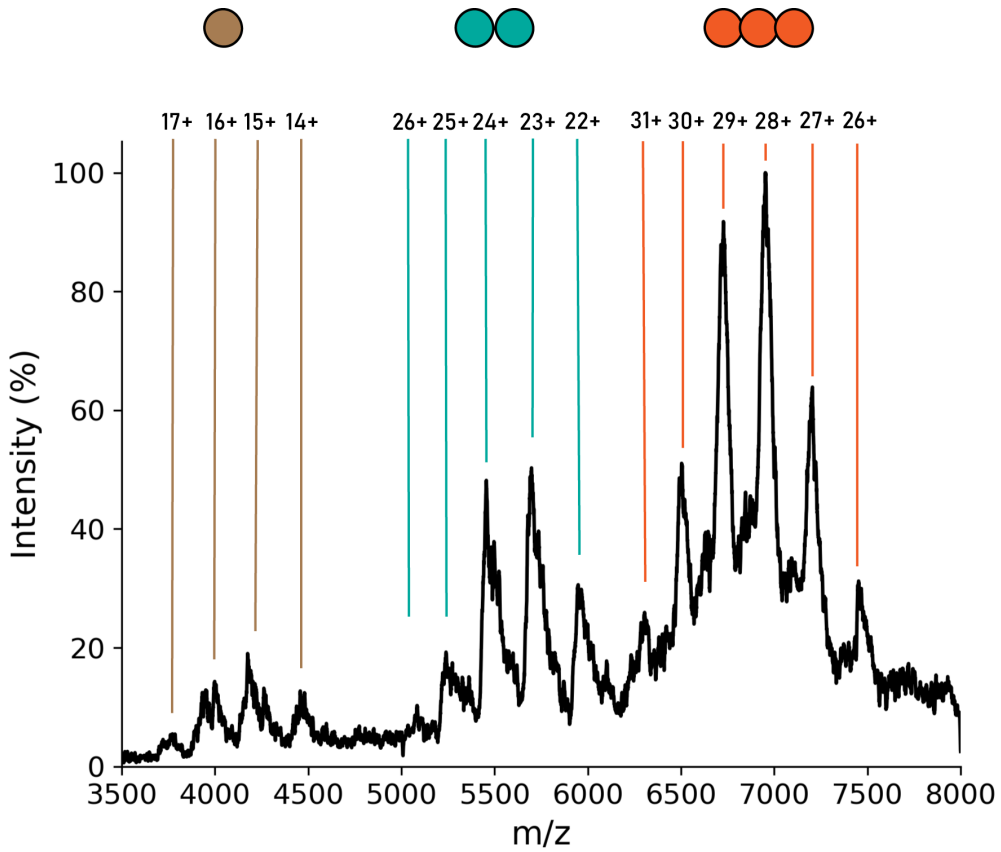

Figure 5: Native mass spectrum of multi complex protein

Shown is a native mass spectrum of a protein with three typical Gaussian charge state distributions (peak series). The protein forms different oligomeric states (indicated by brown, green and orange) which can be measured simultaneously. The mass of all different forms can be calculated from two neighboring peaks.

Experimentally determined masses are in most cases higher than the theoretical mass. Salt adducts or incomplete evaporation of the buffer can shift the observed mass to a higher value. Several automated deconvolution software can be used to help determine the mass of an analyte. However, manual evaluation of the mass and the corresponding annotation of charge states is needed in most cases. A native mass spectrum can further be used to evaluate its quality. Usually the full width half maximum (FWHM) is used to determine the resolution of a signal [110].

1.3.1.3.3 Tandem MS experiments and gas-phase dissociation of protein complexes

A tandem mass spectrometer consists of more than one mass analyzer (MS^n). By changing parameters of the mass spectrometer, it is possible to analyze full spectra or selected ions. A selected ion is called precursor ion and can be fragmented into product ions, which are then subjected to a mass analyzer for mass determination

[106]. The fragmentation of precursor ions is commonly achieved by collision induced dissociation (CID) inside a reaction chamber of the mass spectrometer. The precursor ions collide with heavy inert gas atoms (e.g., argon, xenon, nitrogen) converting a part of the kinetic energy into vibrational energy of chemical bonds. Increasing amounts of internal energy results in unfolding and dissociation (stripping) of subunits from the complex. The charge is asymmetrically distributed between the dissociated subunits and the complex [111, 112].

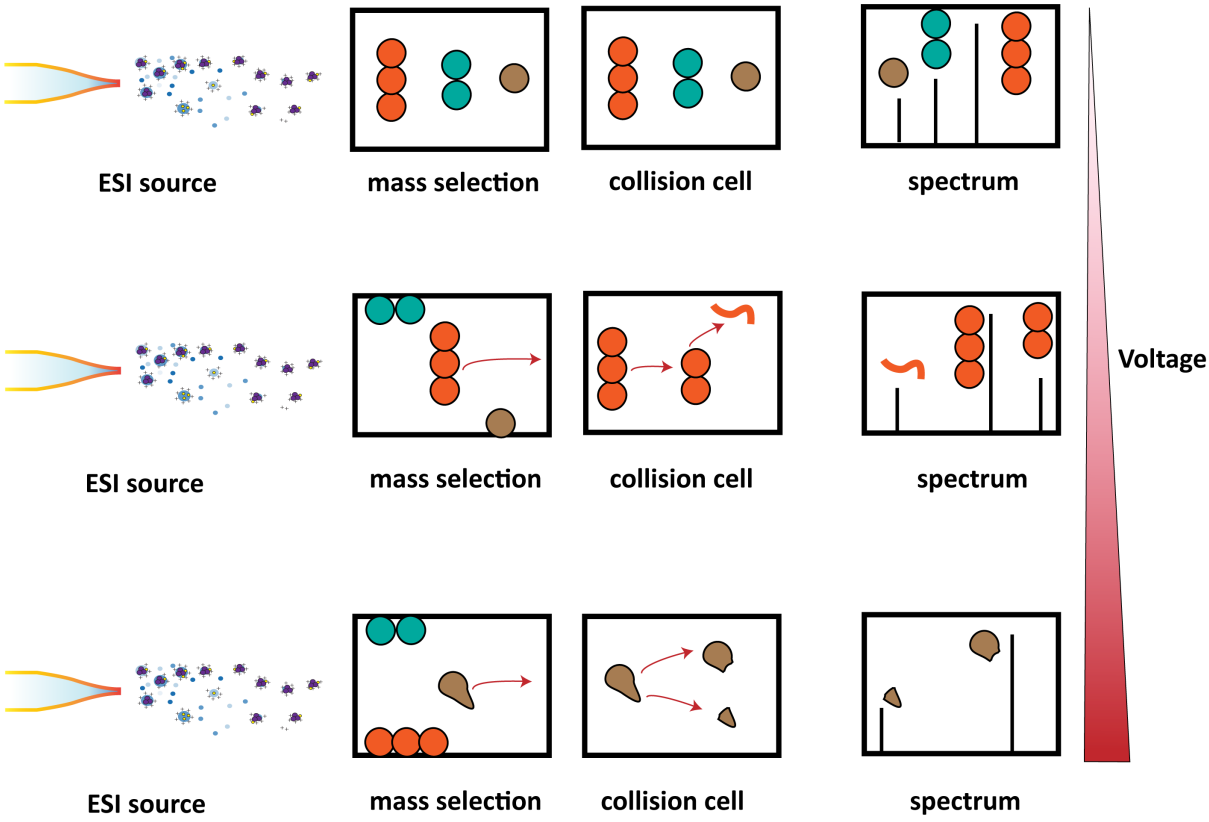

Figure 6: Tandem native MS experiments with different increasing amounts of CID
Ions generated by ESI enter the mass spectrometer passing through the mass filter without being selected. If a low acceleration voltage is applied in the collision cell, no CID is induced on the protein complexes. Masses are measured with a mass analyzer (top). Ions that enter the mass spectrometer are selected at the mass filter representing the precursor ion. The precursor ion enters the collision cell with higher acceleration voltage that leads to subunit unfolding and dissociation. Masses of the dissociated subunit and the remaining complex are measured in the mass analyzer (middle). Precursor ion is selected in the mass filter and passed through the collision cell. High acceleration voltage is applied on the protein causing fragmentation. The fragments and the remaining protein are selected in the mass analyzer (bottom).

The amount of energy transferred to the precursor ion can be controlled by changing gas pressure and acceleration voltage. In a standard MS/MS experiment of protein complexes, the precursor ion (protein complex) is initially measured at low

acceleration voltage. By gradually increasing the acceleration voltage, the collision energy on the (protein) complex is expanded and stripping occurs as described above. Subunits appear in a low *m/z* range, whereas the stripped complex appears in a higher *m/z* range.

Further increasing the energy settings in CID experiments can be used to induce fragmentation of covalent bonds. Proteins first fragment at the peptide bond, producing b-and y-ions. These top-down proteomics methods can be used to obtain information about the stability of a protein or protein complex. Usually, knowledge about the primary structure is collected including the detection of post-translational modifications (PTMs) [111].

1.3.2 HDX-MS

HDX-MS is a fundamental part of the structural MS repertoire. It provides structural data on the structural dynamics of proteins upon binding of ligands, co-factors or PTMs. It can also be used to identify interfaces of protein-protein interactions or ligand binding for drug discovery [113, 114]. HDX-MS tracks the exchange of hydrogen from the protein backbone with deuterium from the solvent. Analysis of proteins dynamics is carried out in standard buffers using deuterated water as solvent. The exchange rate of hydrogen from the protein backbone to deuterium is monitored by LC-MS workflows based on the mass difference between deuterium and hydrogen [115]. The dynamic exchange of hydrogen atoms in biomolecules, influenced by solvent accessibility, structural features, and environmental factors, provides an understanding of protein dynamics.

1.3.2.1 The HDX reaction of proteins

Hydrogen atoms of biomolecules are in constant exchange with hydrogen atoms of the solvent. The accessibility of hydrogen atoms of biomolecules by the solvent determines the exchange rate. Due to the structure of a protein, different parts have different access to the solvent and therefore have different exchange rates. The exchange of hydrogen in proteins to deuterium atoms from the solvent is possible with all labile hydrogen atoms (e.g. -NH, -OH groups) but usually only the peptide backbone (evenly distributed hydrogen atoms) is included in an HDX-MS experiment

[115]. Beside the solvent accessibility to hydrogen atoms, which is influenced by the structure (secondary structure, core or surface), the exchange rate also depends on the temperature and pH [116]. The chemical exchange rate (k_{ch}) of an amino acid is a process catalyzed by a base (OH^-) or acid (H_3O^+). Local parts of proteins are in constant motion and transition between an open and closed state [117]. The actual exchange reaction can only occur in the open state when the hydrogen-bonding network is not built up. The exchange kinetic can be therefore described with the rate for "opening" (k_{op}) and "closing" (k_{cl}). Furthermore, it is assumed that the exchange step is unidirectional as there is an excess of deuterium in the experimental set up [115].

Equitation 1:

$$k_{ex} = \frac{k_{op} * k_{ch}}{k_{op} + k_{ch}}$$

However, it should be considered that two different kinetic events are generally observed in HDX experiments (EX1 and EX2). EX1 kinetic is observed as a bimodal distribution of the isotopic spectrum. The main reason for this is that the opening event of a region or peptide of a protein is much slower than the actual exchange reaction. This effect is usually observed for non-structured regions of a protein. The more predominant EX2 kinetic normally occurs for structured parts of the proteins. These regions are inaccessible most of the time for HDX and may require hours before being exchanged. This exchange appears gradually with a single-peak distribution [118, 119].

1.3.2.2 Experimental design of bottom-up HDX-MS experiments

Most commonly, continuous labeling is used for bottom-up HDX-MS experiments. The protein under investigation is first equilibrated at room temperature (RT) in a physiological buffer system based on normal H_2O. At this stage any ligands, co-factors or other proteins may be included. Usually, two states of the protein are measured in parallel to observe differences in the uptake of deuterium based on a reference (e.g., protein alone vs. protein + ligand). The labeling reaction is started by diluting the sample in a deuterated buffer and stopped at different time points (e.g., 15 s, 10 min, 6 h). The labeling reaction is quenched by lowering the pH to 2.5 and the temperature to 0°C, which ensures minimal DHX. The different, quenched

reactions can be analyzed directly using liquid chromatography mass spectrometry (LC-MS), where the proteins are cleaved into peptides via a flow-through column with immobilized acid proteases (e.g., pepsin). This is conducted under quench conditions to ensure no D/H back exchange [114, 120].

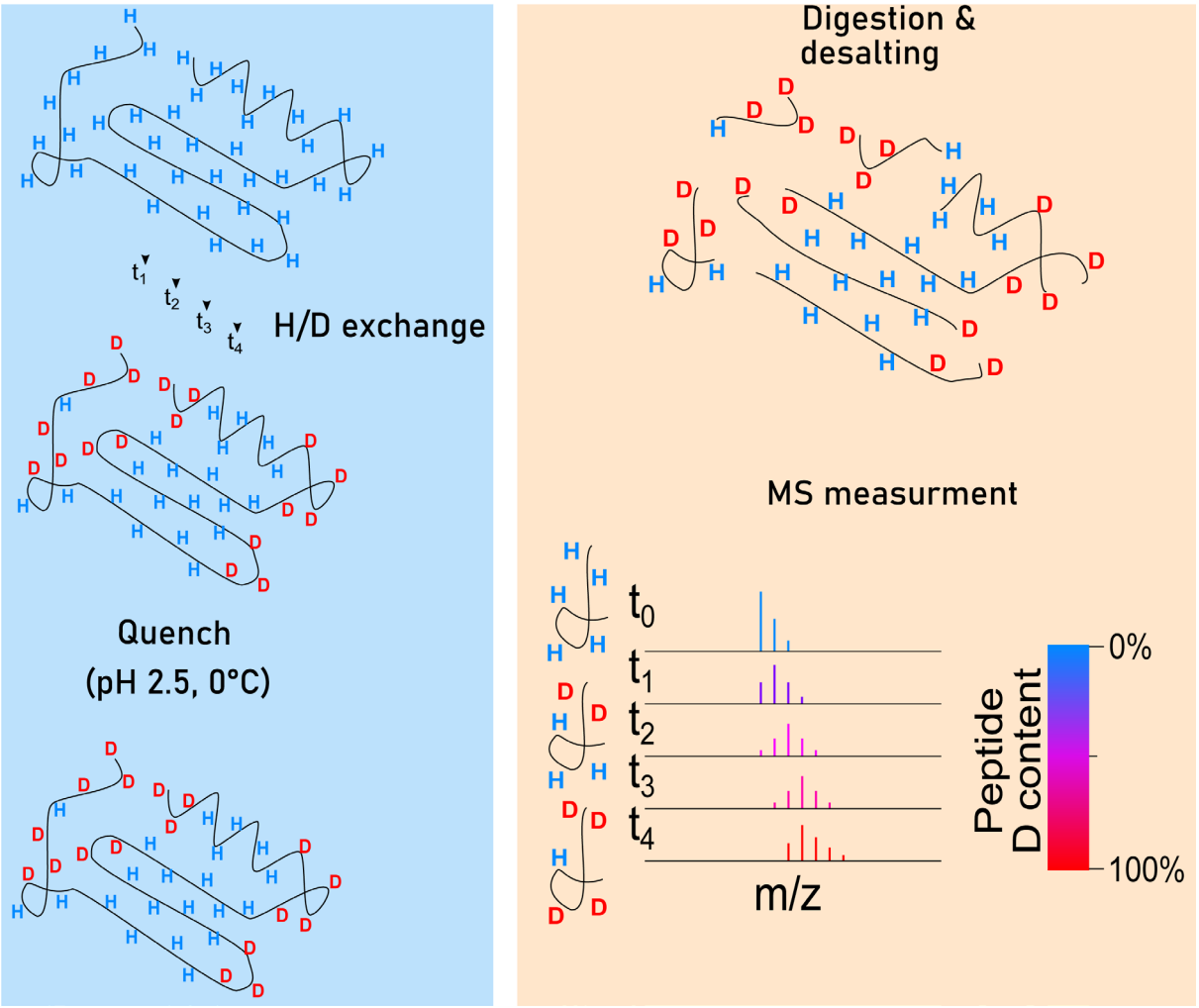

Figure 7: HDX-MS workflow

Proteins are equilibrated under physiological buffer conditions. The HDX reaction is started by diluting the protein sample in deuterated buffer. The pH and temperature need to be controlled thoroughly. After different timepoints (in the order of 4 magnitudes), the exchange reaction is quenched by adding a quench buffer. The final solution has a pH of 2.5 and a temperature of 0°C. The samples can either be frozen or directly subjected to LC-MS. Here, the proteins are digested into peptides by an immobilized protease and desalted afterwards. The peptides are separated by a reverse-phase column and subjected to mass spectrometry. Adapted from D. Wollenberg (CC-BY-SA 4.0)

The digested peptides are further processed by the LC-MS system. First, any buffer components are removed using a desalting step, then a final separation is conducted via a reverse-phase column. The separated peptides are then sprayed with ESI into the mass spectrometer. For later identification of the peptides, non-deuterated samples are screened first by MS/MS approaches for fragmentation [114].

Recent approaches also aim to conduct HDX-MS experiments with intact proteins. Here, quenched reactions are directly subjected to a mass spectrometer and fragmented later. These approaches mainly aim to characterize heterogenous samples or PTMs by comparing the global uptake of deuterium [121].

2 Objective

Due to the limited number of 4 structural proteins, LASV is greatly dependent on protein-protein and protein-nucleic acid interactions between the viral components to facilitate all functions in the viral replication cycle. However, potential interactions between NP and Z, NP and L or NP and RNA have been only indirectly implicated and knowledge of the exact mechanism how the viral components interact are only partially understood. These interaction sites are often potential drug targets, as for example the correct assembly of the full RNP complex is essential for viral genome replication and transcription. Therefore, this work investigates the quaternary structure of the LASV protein complexes to identify direct interactions between the viral components and gain insights about the dynamic behavior of the proteins. The viral components will be recombinantly expressed, purified separately and investigated by structural MS approaches. First, possible multimeric complexes by direct interaction of the viral proteins are investigated by nMS. Here, stoichiometry and topology of complexes are realized in a time-resolved manner. Second, this work will address potential interaction sites between viral components by HDX-MS to complement further structural information. The structural MS methods will be supported by and combined with mutational analysis and functional assays.

3 Results

The first step of this work was the recombinant production of viral proteins. We focused on the production of NP, Z and L which are all involved in viral replication. The expression and purification of the viral proteins were in part previously established by the Rosenthal group [45].

3.1 Recombinant expression and purification of viral proteins

3.1.1 Expression and purification of NP

NP was expressed in *E. coli* as a maltose binding protein (MBP) fusion protein to increase solubility. Additionally, the MBP fusion protein was used to purify the protein via amylose affinity chromatography. At first, *E. coli* transformed with an NP-MBP expression plasmid were cultured at 37°C to an OD_{600} of 0.8, at which point protein expression was induced by adding 500 µM isopropyl-β-D-1-thiogalactopyranosid (IPTG) with a subsequent reduction in temperature to 17°C. Under these conditions, NP-MBP was expressed at a relatively low rate. Therefore, an improved expression protocol was established to increase the yield of NP. The amount of IPTG was lowered from 500 to 250 µM and the temperature prior induction was reduced, resulting in a stronger expression of NP (compare Figure 8 A and B). However, after cleavage of the MBP-NP fusion protein, a significant amount of NP remained attached to the amylose beads indicating aggregation of the protein (Figure 8B). Therefore, the NP-MBP fusion protein was eluted first with maltose containing buffer and later incubated with 3C protease to separate the MBP-tag from NP (Figure 8C). This resulted in a higher amount of soluble NP. Low amounts of protein aggregates still visible after cleavage of the MBP were removed by centrifugation.

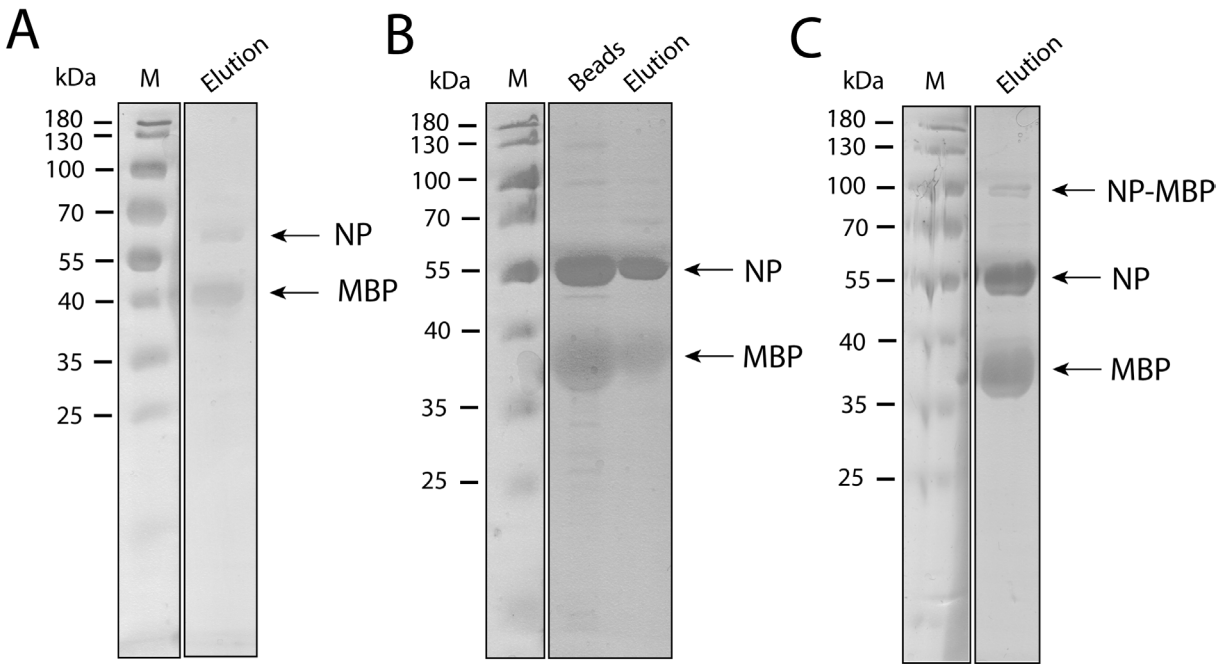

Figure 8: Increasing the expression of LASV NP

LASV NP was expressed in *E. coli* strain BL21 (DE3). 1 L of liquid culture were grown until an OD_{600} of 0.8 at 37°C. The culture was cooled down in an ice-cold water bath and 500 µM (A) or 250 µM (B) IPTG were added to induce protein expression. The cultures were further incubated at 17°C for 18 h and harvested by centrifugation. The cell pellet was resuspended in lysis buffer NP and sonicated. After clearing the lysate by centrifugation, amylose beads were added to bind NP-MBP. The MBP was removed by addition of the 3C protease. The supernatant after 3C protease cleavage is visible in (A) and (B) (Elution). A sample from the amylose resin after cleavage shows NP still bound to the amylose resin (Beads). (C) The fusion protein NP-MBP (NP-MBP) was eluted from the amylose beads by maltose containing buffer and cleaved by the 3C protease into NP and MBP. The size of the marker is indicated in kDa.

NP was further purified by heparin chromatography to remove unspecific RNA. For that, NP was diluted from 500 mM to reach 150 mM NaCl concentration to enable binding of NP to the column. NP was loaded on a heparin column and eluted in one fraction by a linear NaCl gradient (Figure 9B). NP was concentrated by centrifugal filter columns for further purification by size-exclusion chromatography (SEC). 500 µl of concentrated protein was loaded on an SD200 10/300 column and NP eluted in one monodisperse peak (Figure 9C). The purity of the protein as analyzed by SDS-PAGE; NP migrated at around 70 kDa with a high grade of purity (Figure 9D).

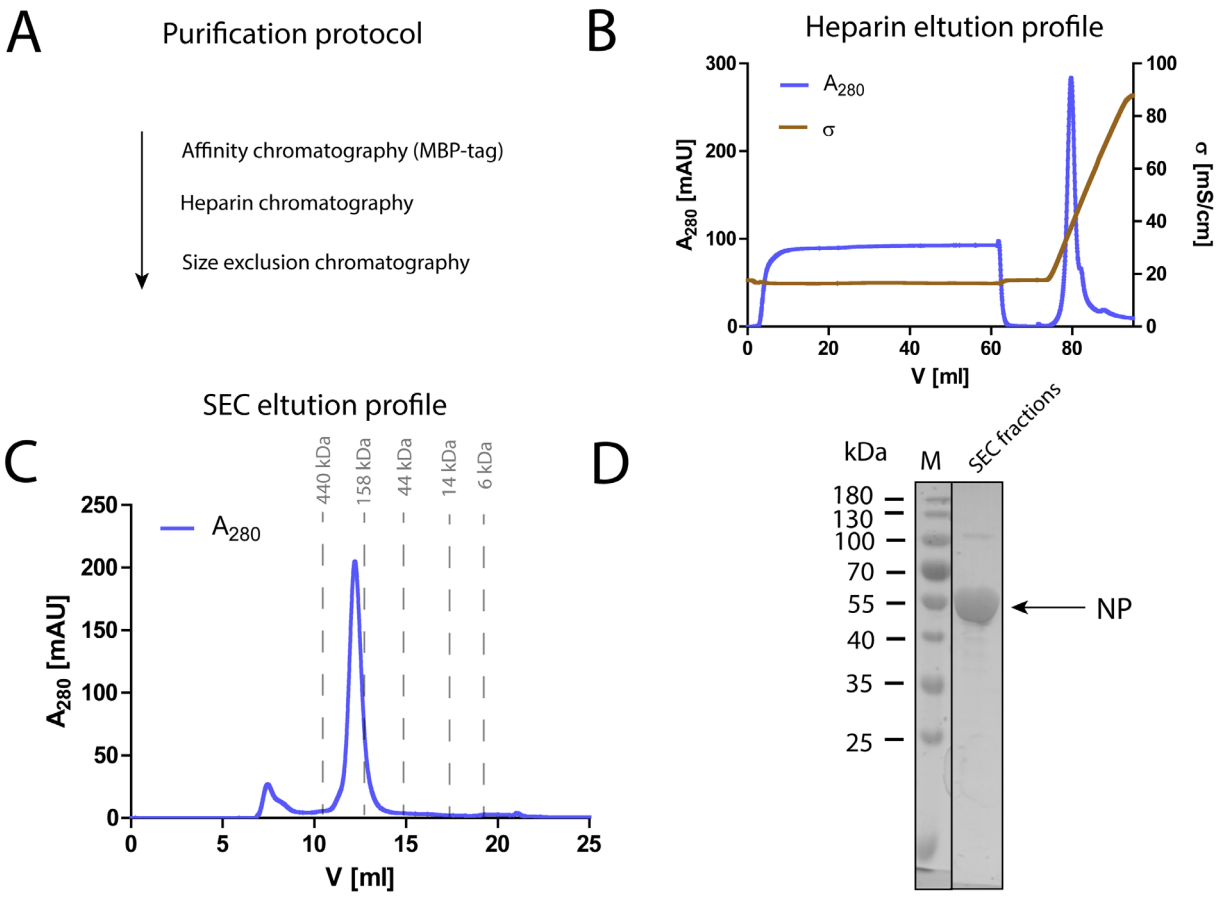

Figure 9: Purification of LASV NP

(A) Overview of the purification protocol. LASV NP was expressed as MBP fusion protein and purified via amylose affinity chromatography. After removing the MBP tag, NP was loaded on a Heparin column to remove unspecific bound RNA from the expression host. Next, NP was further purified via SEC in 50 mM Tris pH 7.5, 150 mM NaCl, 2 mM 1,4-dithiothreitol (DTT) with a SD200 10/300 SEC column. (B) Elution profile from the heparin chromatography. The absorption at 280 nm (A_{280}) is shown at different volumes (V). NP was loaded at low NaCl concentration and eluted via a linear gradient with increasing NaCl concentration. The brown curve is the conductivity measured in mS/cm (right y axis). (C) Elution profile of the SEC, NP eluted at a retention volume of 12 ml in one peak. The grey dotted lines indicate the elution volume of the calibration proteins: Ferritin (440 kDa), Aldolase (158 kDa), Ovalbumin (44 kDa), Ribonuclease A (13.7 kDa) and Aprotinin (6.5 kDa). (D) 12 % SDS-PAGE of the NP peak after SEC. Sizes of the marker bands are indicated. NP migrated at the size of around 55 kDa.

3.1.2 Expression and purification of Z

LASV Z was expressed as a glutathione-S-transferase (GST) fusion protein in *E. coli* BL21 (DE3) and purified via affinity chromatography, ion-exchange chromatography (IEX) and SEC (Figure 10A).

E. coli cells carrying the Z expression plasmid were grown in 1 L cultures at 37°C until and OD_{600} of 0.8-1.2. Protein expression was induced with 500 µM IPTG after which cells were incubated at 17°C for 18-20 h. Cells were harvested by centrifugation and resuspended in lysis buffer Z. Cells were lysed by sonication and the lysate was cleared by ultra centrifugation. The soluble supernatant was incubated with glutathione sepharose beads GST beads to bind the Z protein via the GST tag to the beads. The glutathione sepharose beads with bound GST-Z were then incubated with the 3C protease to cleave the GST-Z fusion protein so that the untagged Z eluted from the column. Next, the solution with tag free Z protein was diluted to 100 mM NaCl at pH 8.5 and passed through an anion exchange chromatography column to further remove impurities which bound to the IEX column whereas Z remained in the flow-through (Figure 10B). The flow-through of the IEX column was then concentrated and loaded on a SD200 10/300 SEC column. Z eluted in two peaks indicating two different oligomeric states (Figure 10C). Fractions were analyzed by an SDS-PAGE and showed Z protein migrating with a size between 10 and 15 kDa (Figure 10D).

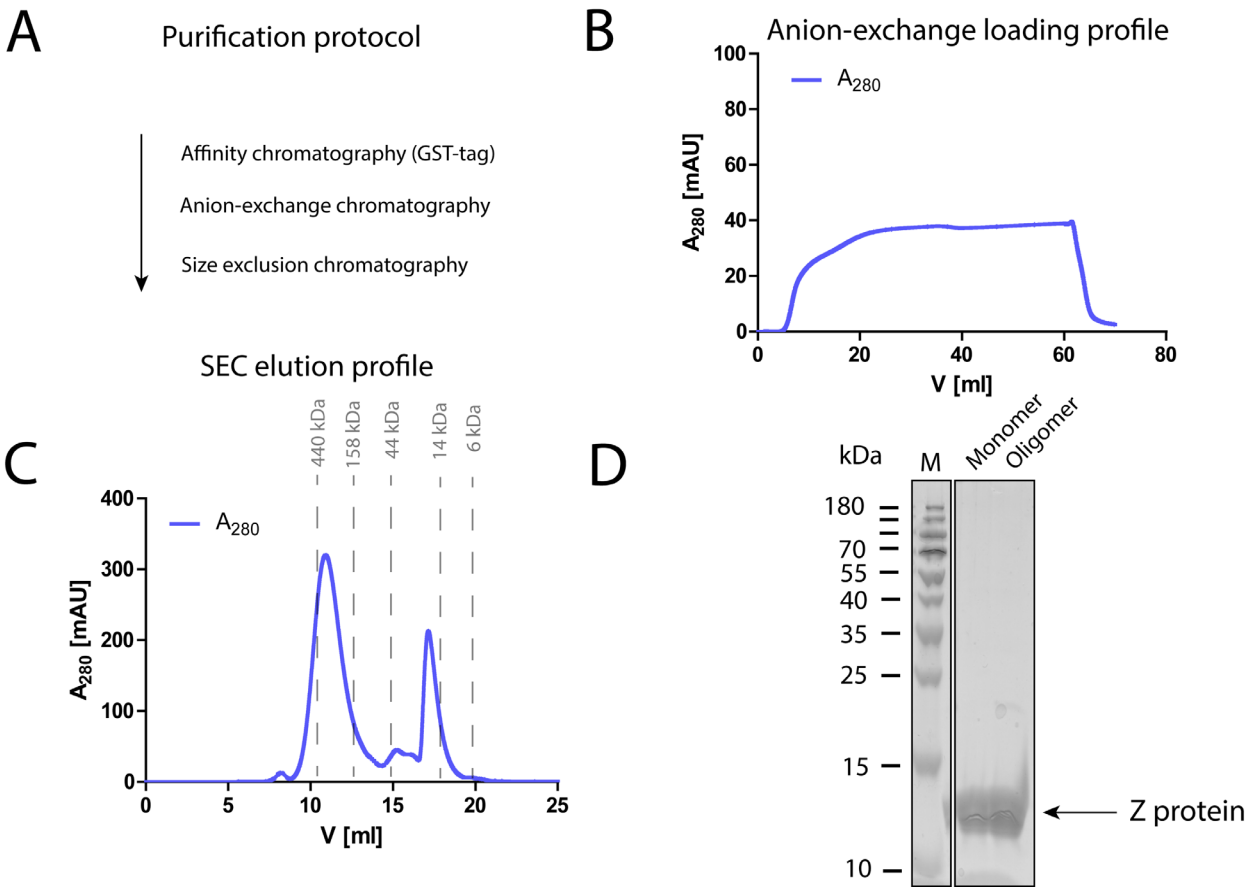

Figure 10: Purification of LASV Z protein

(A) Overview of the purification protocol. LASV Z protein was expressed as a GST fusion protein and purified via GST affinity chromatography. After removing the GST tag, Z was passed through an anion exchange chromatography column to further purify the Z protein. As a final purification step, LASV Z was purified via SEC in 50 mM Tris pH 7.5, with a SD200 10/300 SEC column. (B) Loading profile from the anion exchange chromatography. The absorption at 280 nm (A_{280}) is shown at different loading volumes (V). Z did not bind to the column and the Z containing flow through was collected. (C) Elution profile of the SEC. The absorption at 280 nm (A_{280}) is shown at different retention volumes (V). Z eluted with a retention volume of 12 ml and 17 ml. The grey dotted lines indicate the elution volume of the calibration proteins: Ferritin (440 kDa), Aldolase (158 kDa), Ovalbumin (44 kDa), Ribonuclease A (13.7 kDa) and Aprotinin (6.5 kDa). (D) Both fractions were loaded on a 15 % SDS PAGE indicated by monomer and oligomer. Z migrated between 10 and 15 kDa. Sizes of the marker (M) are indicated.

3.1.3 Expression and purification of L

The size of the LASV L protein is around 250 kDa and therefore not suitable for an expression in *E. coli* cells. Based on [45], a baculovirus expression system in insect cells was chosen with two different L protein constructs. Both L protein constructs contain a StrepII-tag but at different positions (C-terminal (L-C) and at position 407 (L-404)). Both L protein constructs retain polymerase activity.

A pFastBac^{TM}HTB vector containing the sequence of the L protein and the tag was used as a donor vector to transfer the expression construct into the *E. coli* strain DH10EMBacY [122]. The DH10EMBacY strain includes the baculoviral genome and a transposase enzyme which helps to integrate the expression construct into the baculoviral genome. The construct is inserted into a β-galactosidase gene which allows a blue-white selection for successful integration. The expression construct is under control of a late promoter together with a yellow fluorescent protein (YFP). The corresponding YFP fluorescence signal can be used to monitor L protein expression. The bacmids containing successfully integrated L protein expression constructs were extracted and directly used for transfection in Sf21 cells. Three colonies per expression construct were used to produce recombinant baculoviruses and compared regarding their expression yield of the L protein. For screening purposes, 50 ml Hi5 cells were infected with 0.5 % (v/v) baculoviruses expressing L-C, derived from 3 different colonies (named V1-3). Cell density and YFP fluorescence was monitored every 24 h.p.i., and samples were analyzed by SDS-PAGE. At 72 h.p.i., all cells were harvested.

L protein expression was visible after 48 h.p.i. and was strongest after 72 h.p.i. V3 had the highest expression level of L after 72 h.p.i. and was chosen for subsequent expression in larger volumes (Figure 11A). 400 ml of Hi5 cells with a concentration of 0.5 x 10^5 cells/ml were infected with 0.5 % (v/v) and incubated for 72 h. Cells were harvested and lysed in lysis buffer L. The lysate was incubated with Strep-Tactin XT beads, washed, and eluted with biotin-containing elution buffer. The elution fractions showed high amounts of purified L protein (Figure 11B).

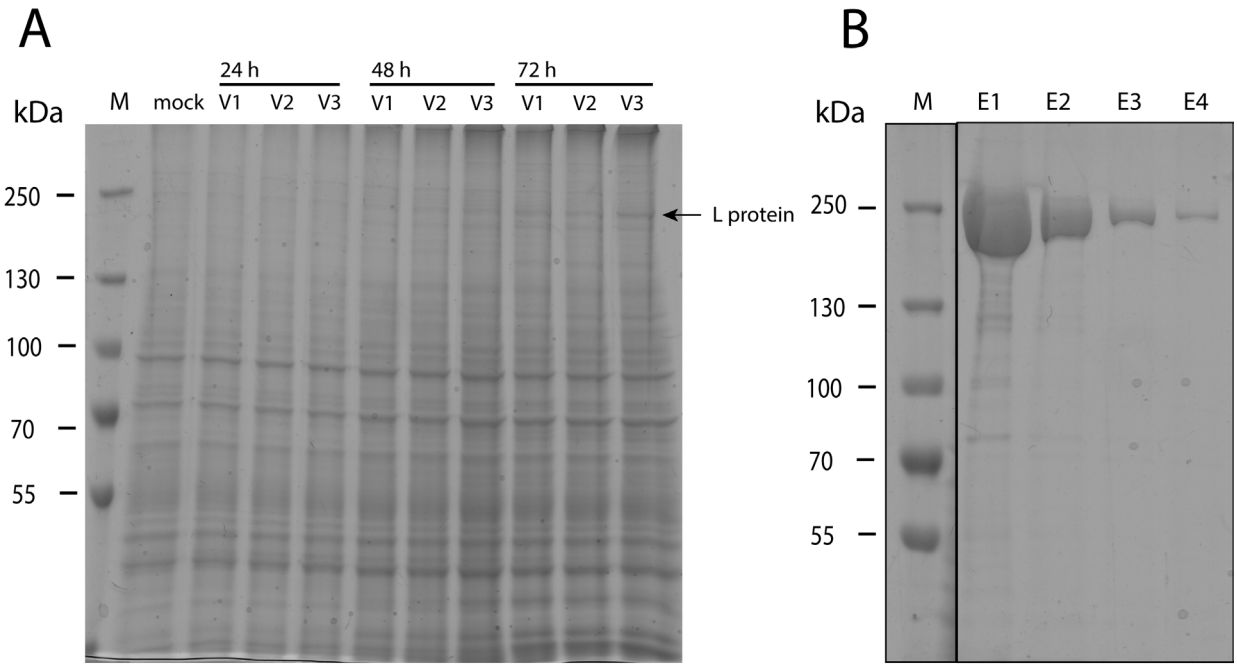

Figure 11: Testexpression LASV L

(A) 50 ml Hi5 cell were infected with 0.5 % (v/v) of 3 different baculoviruses containing the L-C expression construct (V1-V3). 1 x 10^6 cells were harvested every 24 h and analyzed via an 8% SDS-PAGE. Sizes of the marker (M) are indicated one sample corresponds to uninfected cells (mock). The mock control has a band at the same high as the L protein. Over the time, the band appeared to get thicker indicating L protein expression (B) 400 ml Hi5 cells were infected with 0.5 % of V3 and incubated for 72 h. Cells were harvested by centrifugation and resuspended in 50 ml lysis buffer L. After sonication, the cell lysis was incubated with Strep-Tactin XT beads and eluted with biotin containing buffer. The lanes show different elution fractions (E1-E4, 1.5 ml each).

Larger scale expressions were carried out in a total volume of 1-2 L with the same conditions as established in Figure 11. After cell lysis and affinity chromatography with Strep-Tactin XT beads, L-C was eluted with biotin, diluted to reach 250 mM NaCl concentration, and loaded onto a heparin column to remove any unspecific bound nucleic acids and other impurities (Figure 12A). After washing, L was eluted with a buffer containing 1 M NaCl. Fractions containing high amounts of pure L protein were pooled and stored at -80 °C until further use (Figure 12B).

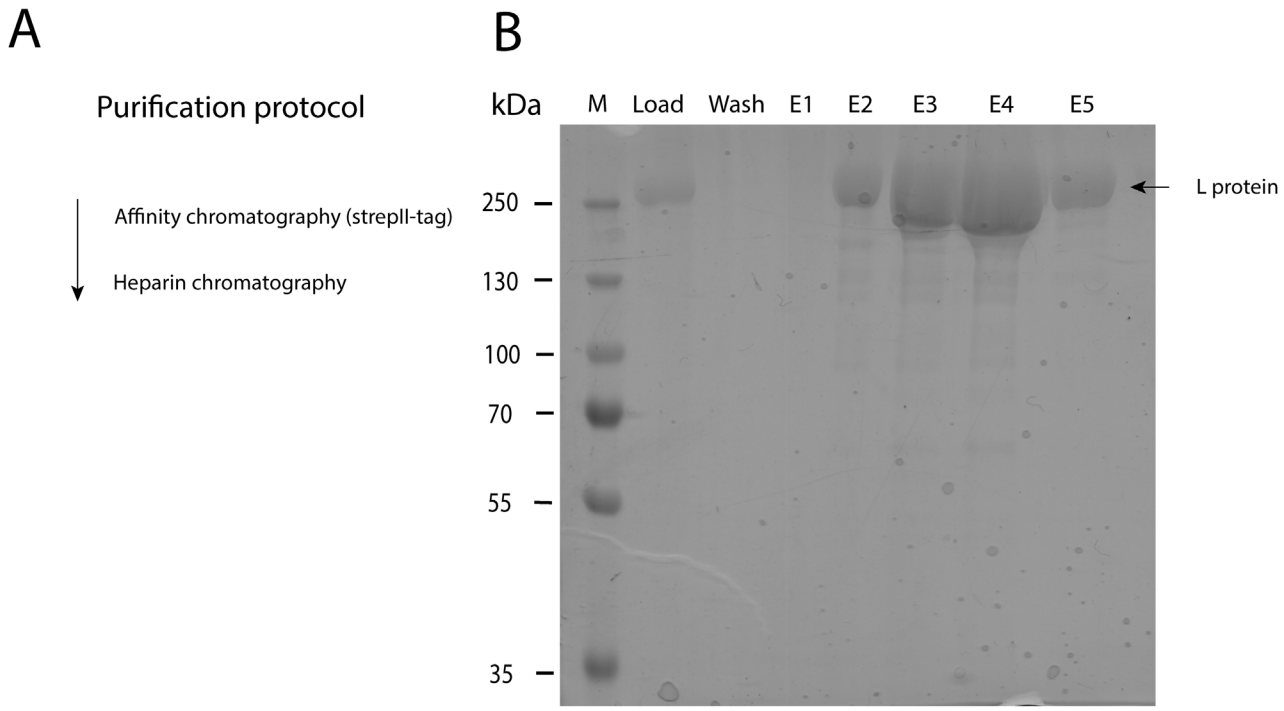

Figure 12: Purification of the LASV L protein

(A) Overview of the purification protocol. LASV L protein was expressed as a strep fusion protein and purified via Strep-Tactin XT affinity chromatography. L was eluted with biotin, and as a final purification step, loaded on a heparin column and eluted in 20 mM HEPES (NaOH) pH 7, 1000 mM NaCl, 10 % glycerol, 2 mM DTT. (B) Final SDS-PAGE of the completed L purification. After affinity chromatography, L protein was diluted to reach 250 mM NaCl and loaded on a heparin chromatography column (Load). After washing with 10 CV of washing buffer (Wash), the protein was eluted by using a buffer containing 1 M NaCl (E1-E5, 1 ml each). Fractions 2-5 were pooled, and flash frozen in liquid N_2.

Overall, it was possible to express and purify all three viral proteins in adequate amounts suitable for biochemical and MS-based analysis.

Establishing nMS with LASV proteins

3.1.4 Behavior of NP in AmAc solution

nMS was used for analyzing the recombinantly expressed and purified viral proteins in terms of their overall stoichiometry and topology. To perform nMS, proteins need to be in a volatile and sodium-free buffer. AmAc solutions are commonly used for nMS set ups and allow the measurement of proteins in their native state.

First, two different workflows were tested for suitable preparation of NP for nMS analysis. A workflow encompassed the purification and storage of proteins in a conventional buffer, followed by a subsequent buffer exchange to AmAc and nMS measurement. The second workflow encompassed a change to AmAc solution during the final purification step. This would allow to use the protein for nMS measurements directly after thawing without performing a time-consuming buffer exchange step (Figure 13A).

To test the suitability of both workflows for NP, the general behavior of NP in AmAc solution was first investigated by performing SEC. One sample contained NP in AmAc solution, and a second sample contained NP in a conventional Tris-HCl buffer. Additionally, one NP sample was flash frozen in AmAc solution to check the behavior of NP upon storage in AmAc solution. All samples showed a similar elution profile after SEC, as the main peak from all samples appeared at 1.7 ml retention volume (Figure 13B). Therefore, AmAc was used as solution for the final purification step and as storage buffer. However, after a couple of weeks of NP stored at -80°C, signs of aggregation appeared after thawing. A nMS spectrum showed a heterogeneous sample compared to a measurement where NP was stored in conventional buffer and exchanged prior nMS measurement (compare Figure 13 C and D). Therefore, the second workflow to store NP in conventional buffer and perform buffer exchange always prior the nMS measurements was used from here on. This workflow resulted in reproducible spectra of NP and a homogenous sample similar to the SEC elution profile of NP purified in conventional buffer (Figure 9C). This workflow was used for all purified proteins.

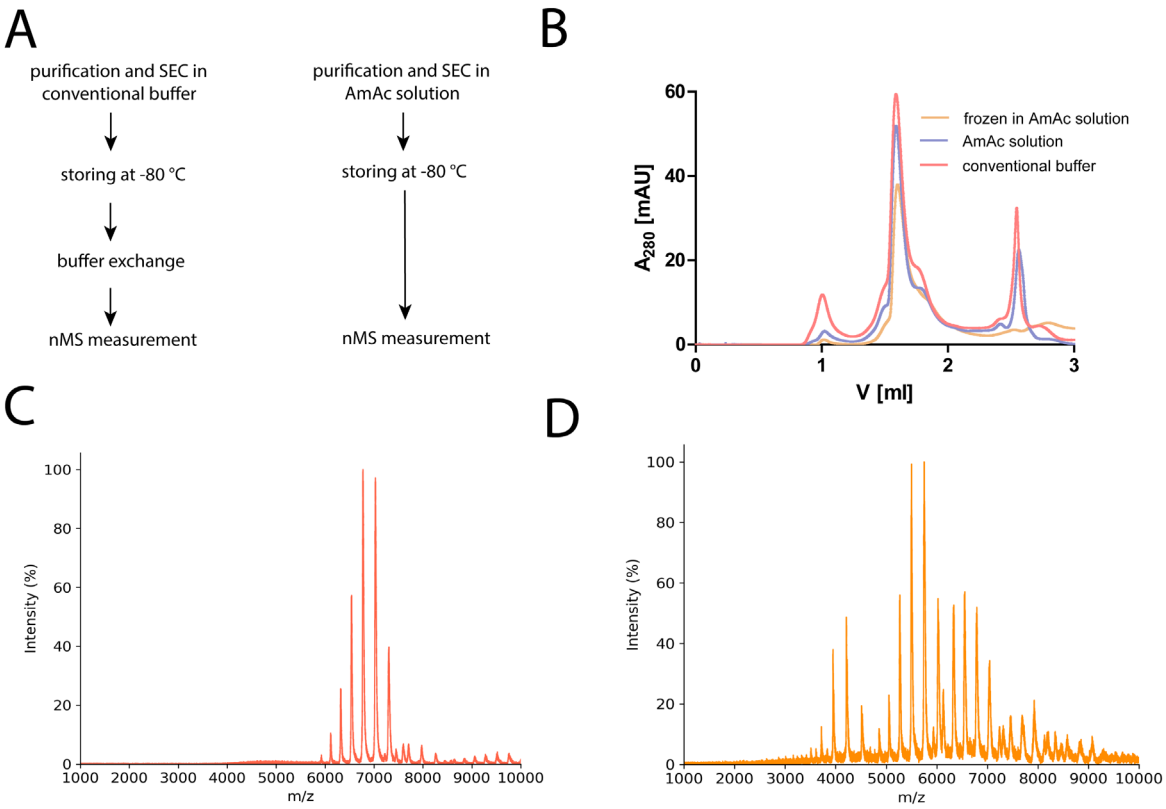

Figure 13: Behavior of NP in AmAc solution

(A) Two different workflows of sample preparation for nMS measurements. Proteins are stored in conventional buffer after purification. Before performing nMS, proteins are buffer exchanged to nMS compatible solutions e.g., AmAc (left). Final SEC of the protein purification are directly performed in AmAc solution and stored at -80 °C. After thawing they can be directly used for nMS measurements (right). (B) SEC of three different NP samples: NP purified in AmAc solution (blue), NP purified and frozen in AmAc solution (yellow), and NP purified in conventional buffer (red). (C) nMS spectrum of 7.5 µM NP which was purified and stored in conventional buffer and buffer exchanged to AmAc before the measurement. (D) nMS spectrum of 8.3 µM NP which was purified and stored in AmAc solution.

3.1.5 Screening for bufferexchange conditions for the L protein

Establishing the measurement of the L protein with nMS appeared to be more challenging than for the other viral proteins. The L protein has a mass of around 250 kDa as one chain and is most stable at high NaCl concentrations. Exposing the L protein to MS-compatible solutions without NaCl, negatively influenced the stability of the protein. The lowest possible concentration of soluble L protein for nMS measurement was 400 nM. Lower concentrations resulted in unusable spectra as the resulting signals were too low to assign any charge states.

First experiments using 150 mM AmAc solutions at neutral pH were unsuccessful as the protein aggregated during buffer exchange. The concentration of the remaining

soluble protein was too low to record usable nMS spectra. Therefore, a screening was performed for suitable buffer exchange devices and buffer compositions to increase the amount of soluble protein. The L-C protein was used for screening purposes as it appeared to be more stable than the L-407. For the buffer exchange device, Amicon Ultra Filter Units with 30 kDa MWCO were best suited for the recovery of the L protein. Filter units with an MWCO of 100 kDa were not able to retain the L protein. Other tested buffer exchange devices were not suitable for the L protein as there was no protein left after the buffer exchange (Table 1).

Table 1: Buffer exchange testing for L protein

Buffer exchange devices were screened according to their ability to retain soluble L-C protein in 500 mM AmAc solution. The amount of recovered soluble proteins is compared between the devices and indicated by +. One + represents the lowest possible concentration of the L protein to record usable nMS spectra.

Buffer exchange device	Recovery
Amicon Ultra Filter Units 100 kDa MWCO	+
Amicon Ultra Filter Units 30 kDa MWCO	+++
VivaSpin 500 Filter Units 30 kDa MWCO	-
Micro Bio-Spin 6 Columns	-
Dialysis	-

Next, different buffer compositions were screened to recover the highest amount of soluble protein. The AmAc concentration used for the NP and Z protein were not suitable for the L protein as no soluble protein was recovered. However, increasing the amount of AmAc had a positive effect on the recovery (Table 2). Furthermore, low amounts of detergent was added to the AmAc solution as this has been reported to increase protein recovery and not interfere with nMS [123]. Combining 1 M AmAc with 0.01 % Tween-20 resulted in the highest recovery rate of L protein. However, the high AmAc concentration could have an influence on polar interactions with other molecules. Therefore, 500 mM AmAc in combination with 0.01 % Tween-20 was used for all following experiments (Table 2). Of note, the internally tagged L protein (L-407) could only be recovered in sufficient amount using 1 M AmAc. Overall, even with the improved conditions a lot of protein aggregation remained for both L protein variants

after buffer exchange, which needed to be removed by centrifugation before the measurement.

Table 2: nMS compatible solutions for the L protein

Different buffer compositions were screened according to their ability to retain soluble L protein in AmAc solution. All buffers were used in combination with Amicon Filter Units with 30 kDa MWCO. * The amount of the recovery of soluble proteins is compared between the different buffer and indicated by +. One + represents the lowest possible concentration of the L protein to record usable nMS spectra.

Buffer ingredients	Recovery
150 mM AmAc pH 7	-
500 mM AmAc pH 7	+
300 mM AmAc pH 7, 0.05 % Tween-20	+
500 mM AmAc pH 7, 0.05 % Tween-20	++
1 M AmAc pH 7, 0.05 % Tween20	+++
500 mM AmAc pH 7, 0.05 % Tween-20, 0.5 mM MnAc	++

Characterization of the LASV proteins by nMS

3.1.6 NP forms a homo trimer and shows high gas phase stability

The first step of the characterization of the viral proteins was to measure the proteins individually to gain first insights into their stoichiometry and behavior. NP was subjected to nMS with concentrations ranging from 5 to 15 µM. Here, NP appeared as a homo trimer with a mass of 189.56 +/- 0.03 kDa (referred to as NP_3) which is in line the current literature reporting NP as a trimer in absence of nucleic acids [56, 62]. Additionally small amounts of a hexameric state were present. It is yet unclear whether the hexamer is an artifact from *in vitro* expression or rather biologically relevant.

Applying varying amounts of energy to the proteins in nMS, i.e., different acceleration voltage in the collision cell, allows to investigate the complex stability in the gas phase of the mass spectrometer. These experiments revealed a particularly high noncovalent complex stability of NP_3, as no collision-induced dissociation (CID) was observed at up to 200 V acceleration. At acceleration voltage higher than 200 V, NP_3 started to dissociate into peptides which appeared in the low *m/z* range, but no masses corresponding to NP monomers were present. The absence of the monomers indicates (i) strong non-covalent interaction between the NP protomers in the trimer and (ii) a potential instability of the monomeric protein. Of note, reducing agents were used during purification to eliminate any covalent disulfide links.

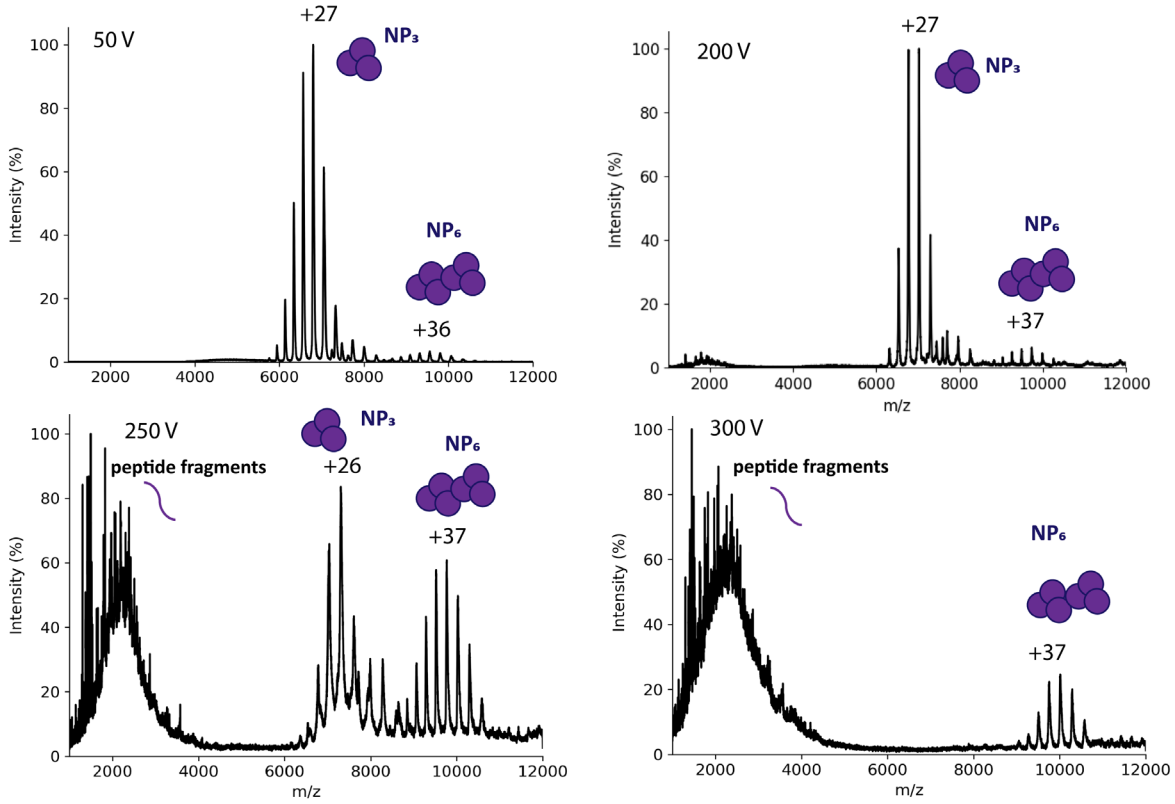

Figure 14: NP forms stable trimers

nMS measurement of 6 µM of LASV NP in 150 mM ammonium acetate solution at pH 7.5 on a QToF2 mass spectrometer. The different graphs show spectra at different acceleration voltages in the collision cell (50-300 V, indicated in the graphs). The shown range is between 1000 and 12,000 *m/z*. Signals between *m/z* 6,000-8,000 relating to a mass species of 189.56 +/- 0.03 kDa which corresponds to an NP_3. At around 10000 *m/z* a second mass species is visible, corresponding to an NP hexamer. The charge state for the strongest signal of all mass species is indicated. Above 200 V peptide fragments appear in the low mass range.

Next, the Z protein was investigated by nMS. After purification Z appeared in two peaks, indicating two oligomeric states (Figure 10C). Both peaks were analyzed separately by nMS, one peak belonged to a mass of 11.35 +/- 0.04 kDa which corresponds to the theoretical mass of a Z monomer (Figure 15A). The other peak appeared as heterogenous oligomers with masses at around 120 kDa (Figure 15B). The exact masses of all oligomeric species could not be determined, but the oligomeric complex completely dissociated into Z monomers above 100 acceleration voltage in the collision cell (Figure 15C). The Z dodecamer has previously been crystallized and its structure determined [77]. Our nMS data suggest that Z does not exclusively form dodecamers but also smaller and bigger oligomeric forms and there is still a lack of knowledge about the biological role of the oligomeric forms. The monomeric Z was used for all following experiments.

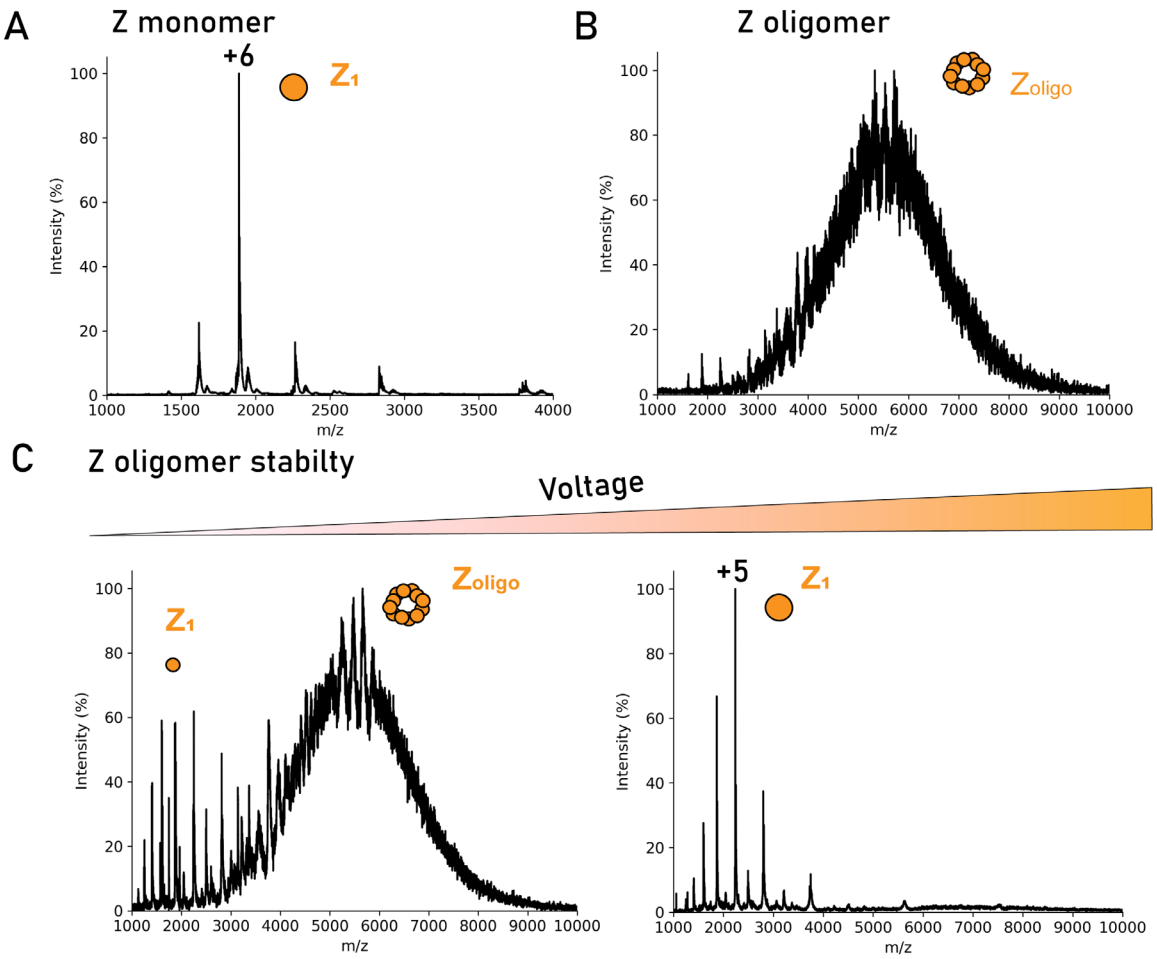

Figure 15: Z proteins monomeric and oligomeric states

(A) The monomeric fraction of Z was measured with native MS in a 150 mM AmAc solution at pH 7.5 on QToF2 mass spectrometer. Capillary and cone voltage were constantly at 1450 V and 150 V, respectively. Acceleration voltage in the collision cell was at 50 V. Z appeared primarily in the monomeric state (11.35 +/- 0.04 kDa) between 1000 and 4000 *m/z*. The charge state for the strongest signal is indicated. Small fractions of dimers were traceable. (B) The oligomeric fraction of Z was measured with the same setting as described in A but with an acceleration voltage of 20 V. Heterogenous peaks were measured between 3000 and 8000 *m/z*. (C) The oligomeric fraction was measured at higher acceleration voltage (50 V and 150 V). At higher acceleration voltage (150 V, right panel), peaks in the low *m/z* range were measurable, which corresponded to Z monomers.

Next, the overall stoichiometry of the L protein was analyzed. Both versions of the L protein (L-407 and L-C) appeared as monomers. The masses were determined for L-C with 255.97 +/- 0.68 kDa and L-407 with 254.74 +/- 0.31 kDa. This matches the theoretical masses expected from these proteins with 255.75 kDa for L-C and 254.50 kDa for L-407, indicating no major post-translational modification e.g., glycosylation.

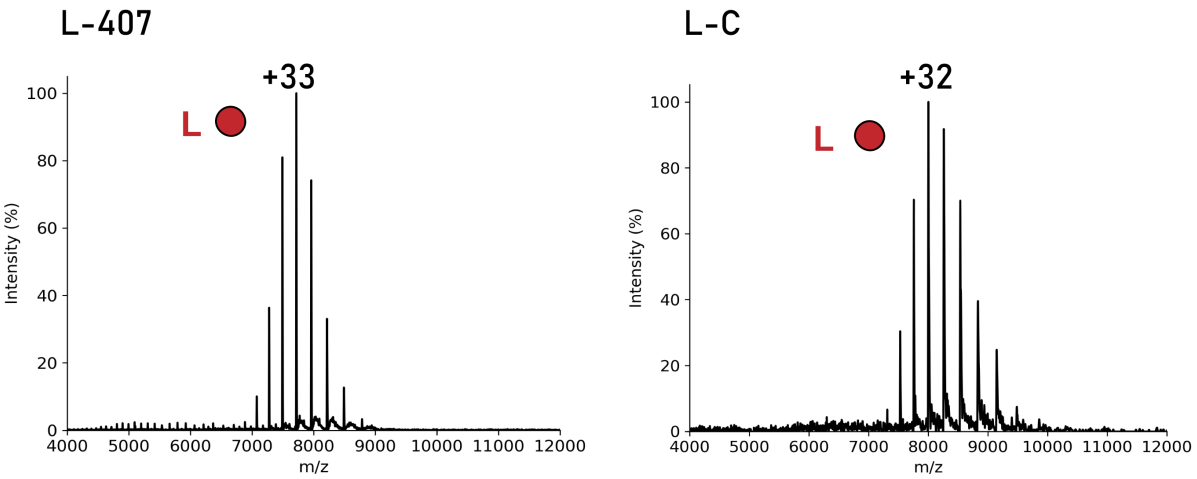

Figure 16: L protein is monomeric in nMS set up

2 – 3 µM of L protein were subjected to a QE-UHMR mass spectromer. L-407 was measured in 1 M AmAc, 0.05 % Tween-20, pH 7 and L-C in 500 mM AmAc, 0.05 % Tween-20, pH 7. The shown range is between 4000 and 12,000 *m/z*. L appeared appeared between 7000 and 9500 *m/z* with charge states around +33. The charge state for the strongest signal is indicated.

3.2 NP directly interacts with the Z protein

Directing the RNP complex to the cell membrane prior to assembly and budding is a process that likely involves an interaction between NP and Z. When simultaneously overexpressed in cells, both proteins colocalized in the cytoplasm [35]. Furthermore, both proteins were also found to co-immunoprecipitated, indicating either a direct interaction or mediated by other factors [50, 53, 54].

3.2.1 NP_3 binds up to three Z proteins

nMS was used to search for a potential direct interaction using recombinantly expressed and purified proteins. Both proteins were buffer exchanged to an AmAc solution and mixed. The proteins were directly subjected to a QToF2 mass spectrometer, and the spectra were searched for potential complex formation. Both proteins were analyzed by measuring Z in a 10-fold excess over NP (Figure 17A). Several new mass species were detected between 6,000-8,000 *m/z* which is in the range of NP_3 (compare Figure 17B and C). To identify the exact composition of the newly formed complexes, tandem MS (MS/MS) with CID was used. Here, the mass spectrometer isolates a single peak (Figure 17D), which corresponds to a newly formed complex and applies high acceleration voltage inside the collision cell to the mass spectrometer. The resulting spectrum includes CID products of the complex which allows to identify the exact composition of the originally isolated peak. Dissociation products of NP_3 (189.56 +/- 0.03 kDa), NP_3 + Z (201.00 +/- 0.09 kDa) and NP_3 + 2xZ (212.37 +/- 0.13 kDa) were detected when isolating three different peaks and applying CID-MS/MS (Figure 17E). This corresponds to 1, 2 or 3 Z proteins attached to an NP_3.

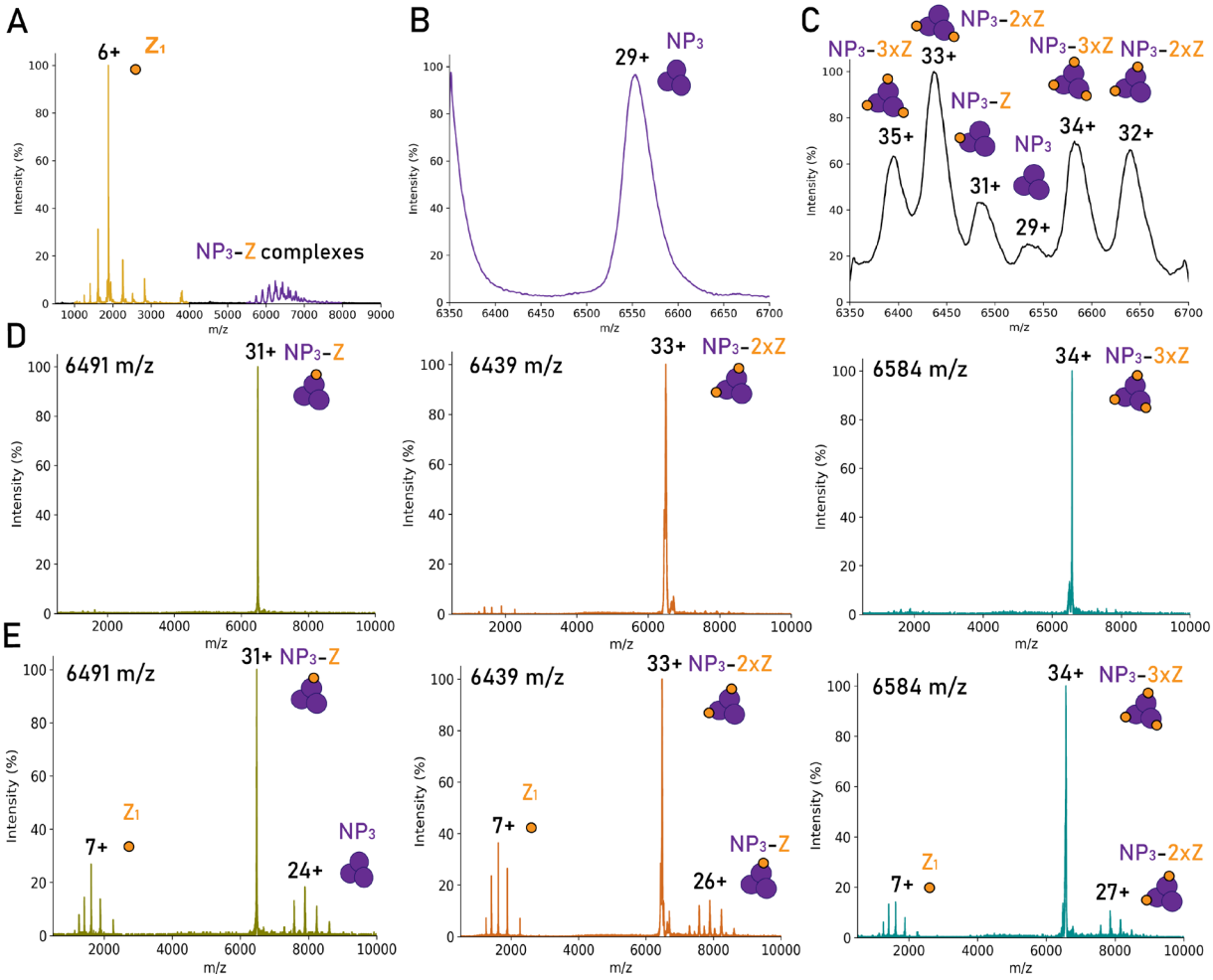

Figure 17: NP_3 binds up to 3 Z monomers

NP and Z were incubated in a 1:10 molar ratio and subjected to nMS in 150 mM ammonium acetate solution at pH 7.5. (A) Overview spectrum recorded with 50 V acceleration. Two peak series are observable between 1,000 and 4,000 *m/z* (corresponding to the Z protein) and between 5,500 and 7,000 *m/z* (NP_3 complexes). Zoom to NP_3 29+ charge state without (B) and with Z protein (C). Several new species corresponding to NP-Z complexes are recorded. CID-MS/MS at 50V (D) and 100 V (E) acceleration for NP_3-Z (31+ charge state), NP_3-2xZ (33+ charge state) and NP_3-3xZ (34+ charge state). CID products correspond to masses for Z (11.35 +/- 0.04 kDa), NP_3 (189.56 +/- 0.03 kDa), NP_3-Z (201.00 +/- 0.09 kDa) and NP_3-2xZ (212.37 +/- 0.13 kDa), respectively. The charge states of the most abundant peaks are indicated.

3.2.2 Monomeric NP is able to bind Z

NP binds as a homo trimer up to three Z proteins, the question remained whether the oligomeric state of NP influences the interaction to Z. Therefore, a trimerization mutant was designed to test exclusively a monomeric version of NP in their ability to bind Z. Arginine at position 52 had been identified previously to be crucial NP trimer formation [62].

3.2.2.1 Purification of a trimerization mutant

The residue arginine 52 was changed to alanine by site-directed mutagenesis PCR. The protein was expressed and purified analogously to the wild-type (wt)-NP (Figure 18A). NP-R52A behaved similar to the wt-NP during purification but eluted with a retention volume of 11.7 ml and 13.4 ml (Figure 18BC). Both fractions were analyzed by SDS-PAGE and showed no differences as bands were migrating at around 60 kDa, which corresponds to the theoretical mass of NP (Figure 18D). As SEC indicates different oligomeric states which are not detectable by SDS-PAGE, both fractions were subjected to nMS. Here, one monomeric fraction and one fraction with monomers, dimers and trimers was present (Figure 18EF). Notably, the spectrum revealed nucleotide monophosphates (NMPs) bound to NP. This suggests that the RNA binding site of monomeric NP is in a conformation that enables RNA binding which is not possible in the trimeric conformation (Figure 18E).

Overall, it was possible to recombinantly express and purify the NP-R52A trimerization mutant. A fraction could be recovered that is exclusively monomeric and can be used for interaction studies.

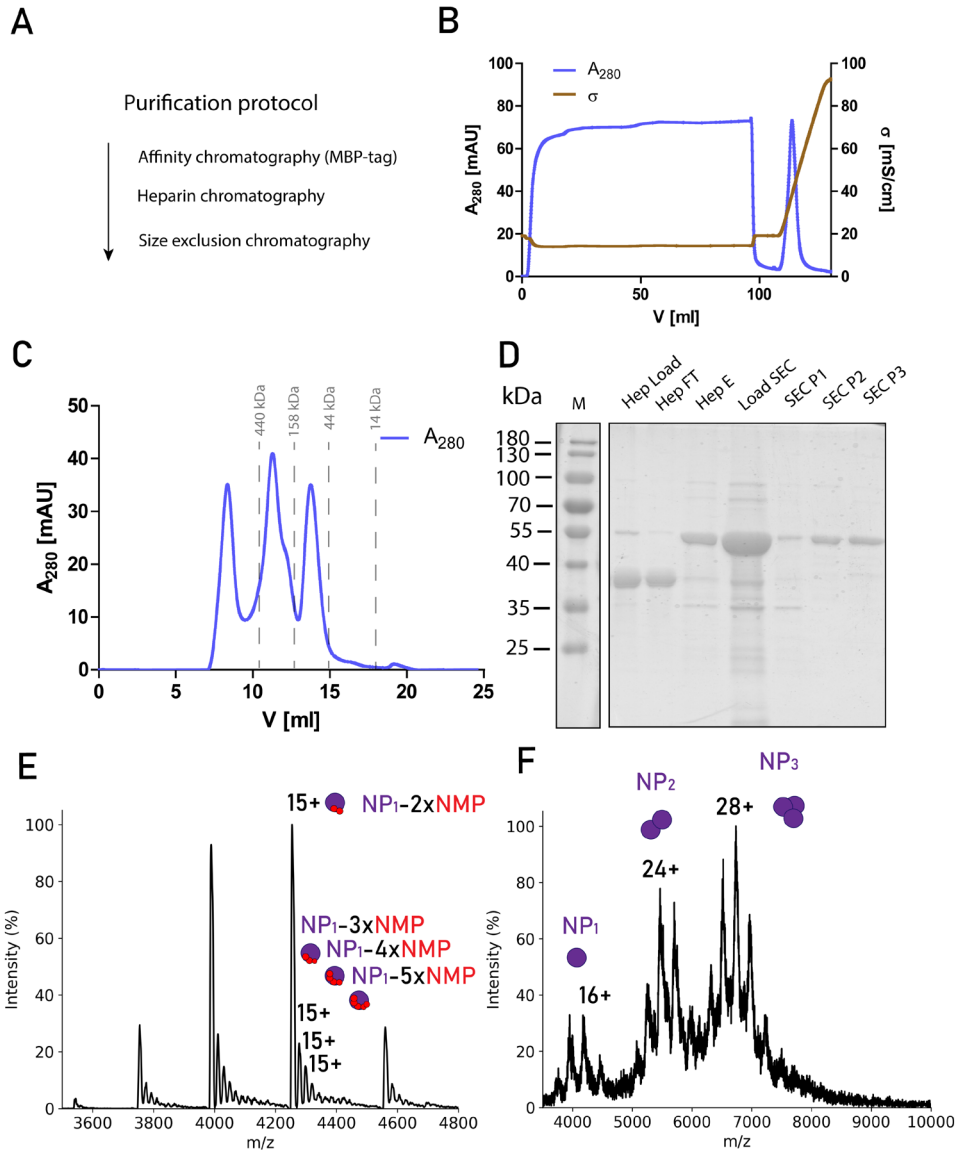

Figure 18: Characterization of the NP-R52A trimerization mutant

(A) Overview of the purification protocol. LASV NP_R52A was expressed as MBP fusion protein and purified via amylose affinity chromatography. After removing the MBP tag, NP was loaded on a Heparin column to remove unspecific bound RNA from the expression host. NP was further purified via SEC. (B) Elution profile from the heparin chromatography, NP was loaded at low NaCl concentration and eluted via a linear gradient with increasing NaCl concentration. The brown curve is the conductivity measured in mS/cm (right y axis). The blue curve is the absorption at 280 nm (A_{280}, right y-axis) (C) Elution profile of the SEC, the protein eluted in two fractions at 11.4 ml and 13.4 ml elution volume. (D) 12% SDS-PAGE of NP-R52A purification. Solution loaded to the heparin column (Hep Load), flow through from the heparin column (FT Hep), elution form the Heparin column (Hep E), solution loaded on the SEC column (Load SEC), peak fractions from the SEC (SEC P1-P3). Sizes of the marker bands are indicated. NP migrated at the size of around 55 kDa. (E) Measurement of 4.1 µM NP-R52A monomer fraction on a QE UHMR instrument in 150 mM AmAc solution at pH 7.5. (D) nMS measurement of the peak at 11.36 ml elution volume on a QToF2 instrument is shown in 150 mM ammonium acetate buffer surrogate at pH 7.5. Capillary and cone voltages were constantly at 1450 and 150 V, respectively. Acceleration voltage in the collision cell was at 50 V. The most abundant mass species was trimeric NP with the depicted +28 charge states. Between 4000 and 6000 *m/z* two peak series were visible with masses corresponding to NP dimers (+24 charge state) and NP monomers (+16 charge state), respectively.

3.2.2.2 Monomeric NP binds the Z protein

Next, the monomeric NP was tested in their ability to bind the Z protein. At a 1:3 ratio between NP-R52A and Z, native mass spectra clearly showed the 75.5 +/- 0.3 kDa monomeric NP-R52A-Z complex (compare Figure 19 A and B). CID-MS/MS of the NP-R52A-Z complex further confirmed this interaction. These data show that the monomeric NP-Z interaction is stronger compared to trimeric NP, either influenced by the monomeric conformation or by the R52A mutation.

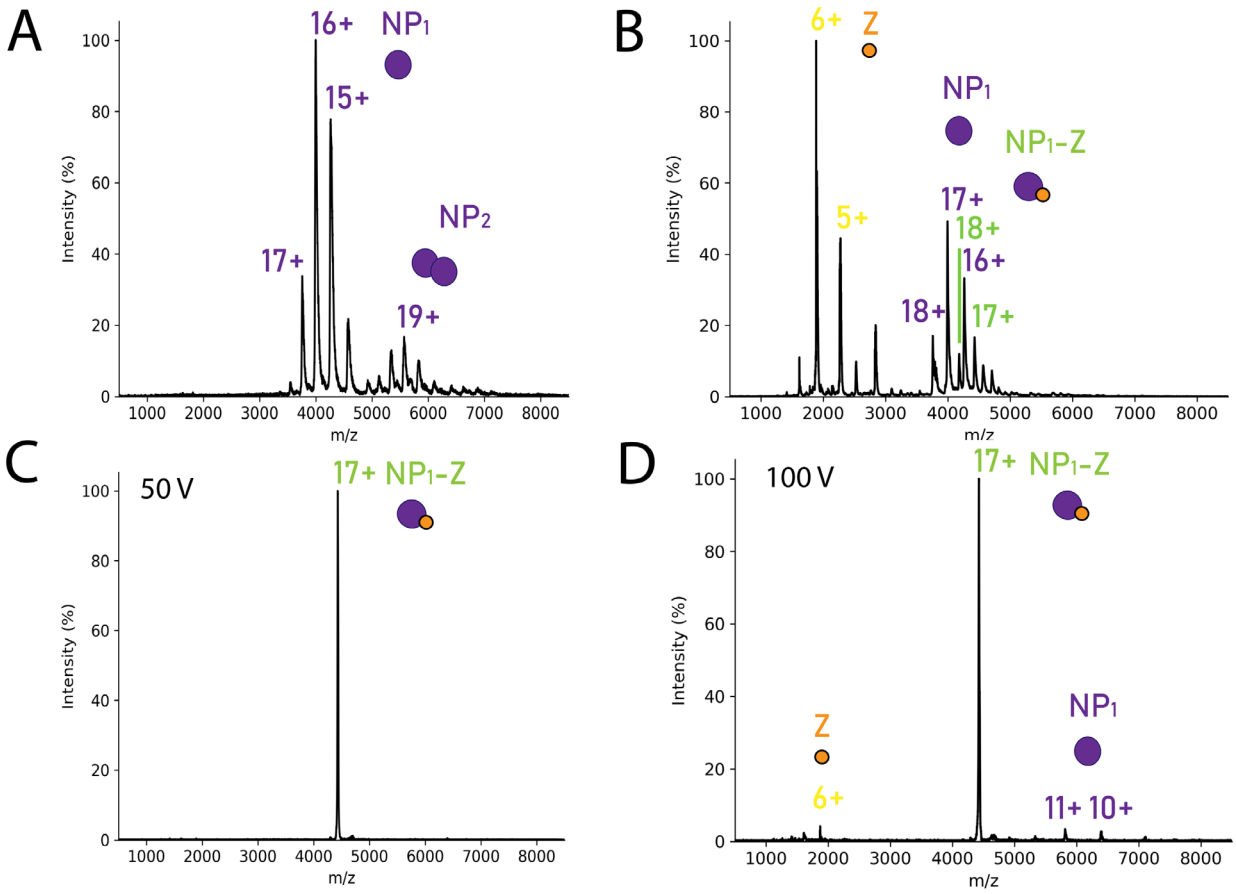

Figure 19: NP-Z interaction is independent of the oligomerization state of NP in nMS

(A) NP trimerization mutant R52A was measured with nMS in a 150 mM AmAc solution at pH 7.5 and a concentration of 5.2 µM. Acceleration voltage in the collision cell was at 50 V for the depicted spectrum. NP-R52A appears primarily in the monomeric state between 3,500 and 4,500 with charge states between +17 and +14 and a corresponding mass of 63.90 +/- 0.08 kDa. A small fraction of dimeric NP is visible in the *m/z* range between 5,000 and 6,000 with the corresponding charge state of +20, +19 and +18. The most abundant species was normalized to 1. (B) NP-R52A and Z were incubated in a 1:3 ratio and measured with the same instrument settings as described in (A). Spectra show new peak series appearing for unbound Z protein in the *m/z* range between 1,500 and 3,000. New peaks also appear in the range between 4,000 and 5,000 *m/z* with +18, +17 and +16 charge states corresponding to a complex of NP-R52A and Z (75.5 +/- 0.3 kDa). CID-MS/MS of one peak corresponding to a NP-R52A-Z complex at 4431 *m/z* with 50 (C) and 150 (D) acceleration voltage. CID products at high acceleration voltage correspond to masses of a Z monomer and NP-R52A.

The affinity of the NP-Z interaction was further evaluated by calculating the dissociation constant (K_D) based on the law of mass action. For that, nMS spectra were recorded in triplicates at an NP to Z molar ratio of 1:3. the deconvoluted spectra were used to determine the area under the curve (AUC). A K_D of 110 µM (± 10) was determined for binding of a Z protein to an NP_3. For the monomeric NP-R52A interaction with the Z protein, a K_D of 33 (± 2) µM was determined (Figure 20). Overall, this indicates a higher affinity of the monomeric NP to Z compared to the trimeric NP.

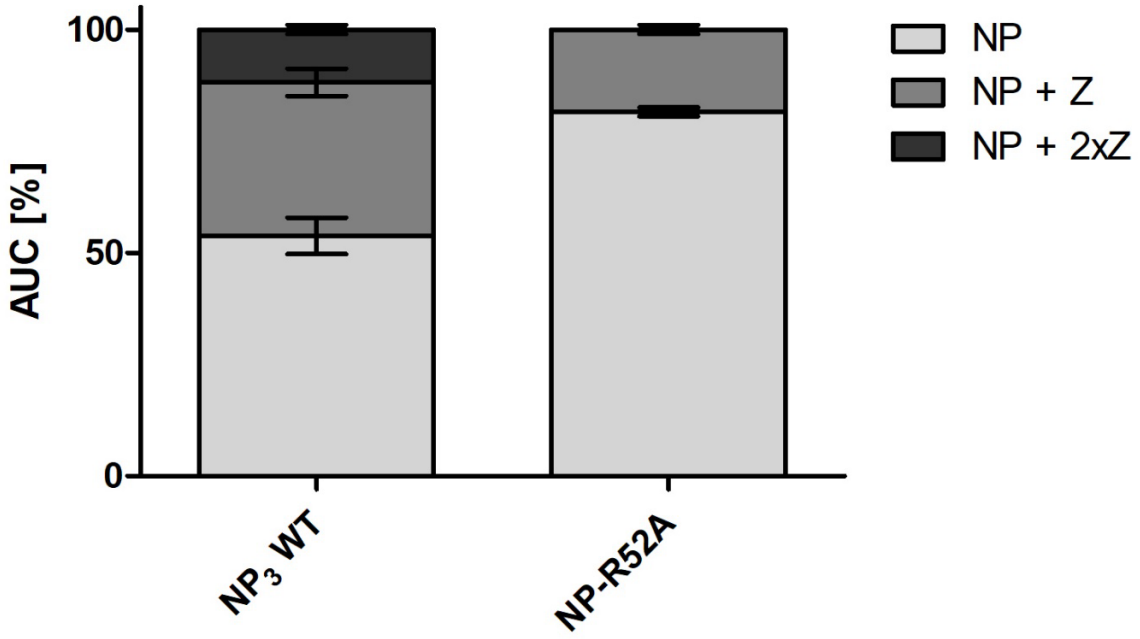

Figure 20: K_D determination of the NP-Z interaction

NP_3 WT and NP-R52A were incubated in a 1:3 (NP:Z) with a NP monomer concentration of 9 and 3 µM respectively. Samples were subjected to nMS. Capillary voltages were held at 1.2 kV, source temperature at 50°C and HCD voltage at 100 V. Resulting spectra from at least 3 independent measurements were deconvoluted to a zero-charge mass spectrum with Unidec and AUC was determined for the respective mass species. The K_D was determined after [124] taking into account that wt-NP contains as a trimer three Z binding pockets and the mutant as a monomer only one Z binding pocket.

3.2.3 Mapping the NP-Z interaction site

An integrative structural approach was then used to map the interaction site between NP and Z using 3 different methods: (i) Artificial intelligence (AI)-powered structure complex prediction (in collaboration with the Kosinski group at EMBL Hamburg), (ii) HDX-MS to map changes in the protein dynamics caused by protein-protein

interactions [125], and (iii) mutational *in vitro* studies for validation of potential interfaces.

3.2.3.1 Establishing HDX-MS to map the NP-Z interaction site

Hydrogen atoms of biomolecules are undergoing a constant exchange with hydrogens from their solvent. The accessibility of hydrogens of biomolecules by the solvent determines the exchange rate. Different parts of a protein have varying access to the solvent and have therefore varying exchange rates. HDX-MS is based on the exchange reaction in which hydrogen atoms from proteins are exchanged by deuterium of deuterated solvent. MS can determine the mass difference of deuterium and hydrogen of peptides from proteins after H/D exchange in deuterated buffer. This method allows to identify interaction sites as in the event of an interaction, the exchange rate of peptides near the interaction site is altered.

Proteins are labeled in deuterated buffer in presence of a ligand for different amounts of time. The reaction is stopped with a denaturing buffer that quenches the labeling reaction. Choosing a suitable quenching buffer is crucial for the amounts of detected peptides. A screening with different quenching buffers was performed to detect as many peptides as possible (Table 3). 100 pmol of NP was mixed in 50 µl labeling buffer with 50 µl quenching buffer. The denatured samples were analyzed by LC-MS/MS. The buffer 1 M glycine pH 2.5, 2 mM DTT resulted in the highest number of detected peptides with a coverage of 100 % (Table 3) and was chosen for following HDX experiments.

Table 3: Screening for HDX quenchbuffer

Prior to the HDX-MS experiments, different quench buffers were tested and compared regarding the number of detected peptides and protein sequence coverage.

Buffer	Peptides detected	Coverage (%)
1M Glycine pH 2.5, 2 mM DTT	577	100
1M Glycine pH 2.5, 6M Urea, 2 mM DTT	468	100
1M Glycine pH 2.5, 6M GndHCl, 2 mM DTT	421	100

3.2.3.2 Integrating HDX-MS in the predicted NP-Z complex

HDX-MS was then used to experimentally identify the interaction site between NP and Z. Labeling experiments in deuterated buffer were performed with NP both in the presence and absence of Z protein. Labeling timepoints were chosen between 15 s and 6 h. A summary of the experimental conditions is listed in Table 4. Differences in deuterium uptake were only considered to be significant if passing a t-test within 95 % confidence interval and were higher than 0.5616 D which is the variance across all species. The sequence coverage of the HDX-MS experiment was 97.9 %.

Table 4: HDX summary table

Conditions for the HDX-MS experiments were chosen according to the HDX-MS community recommendations [126].

	unbound	**50 µM Z**
HDX reaction details	40 mM Tris, pH 7.5, 150 mM NaCl, 25°C	
HDX time course (min)	0.25, 1, 10, 60, 360	
HDX control samples	Fully deuterated control, labeled for 24 h, buffer 40 mM Tris, 6 M urea, pH 7.5	
Back-exchange (mean)	28.442 % +/- 8.537 %	
# of Peptides	285	282
Sequence coverage	97.90%	97.90%
Average peptide length / Redundancy	13.80 / 6.89	13.71 / 6.77
Replicates	3 (technical)	3 (technical)
Repeatability (average standard deviation)	0.0716	0.0659
Significant differences in HDX (delta HDX > X D)	t-test 95 % confidence interval and delta D < 0.5616 (variance across all peptides)	

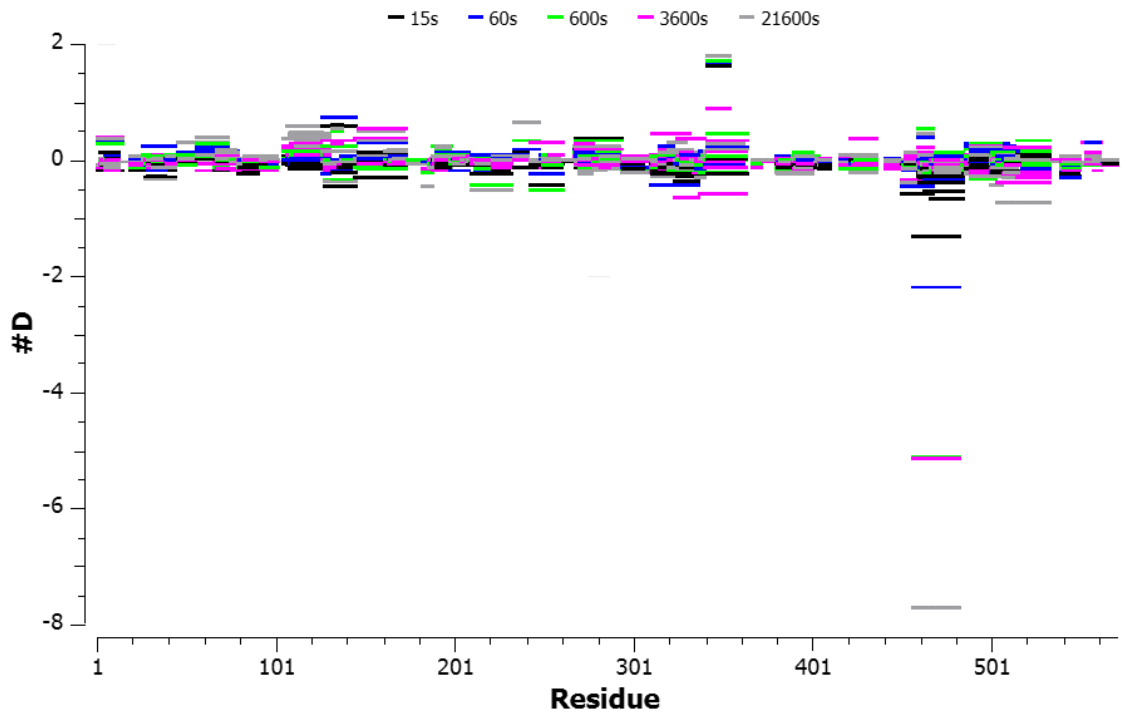

Figure 21 Woodplot of HDX-MS data of NP together with Z:

HDX-MS dataset was analyzed with HDExaminer Version 3.3. Differences of deuterium uptake in deuterons (#D) of NP peptides (residue) in presence of Z compared to an NP only control. For most peptides, there is no difference of deuterium uptake between both states. Differences in deuterium uptake are visible in the area between residue 450 and 500 indicating a potential interaction site.

Next, the structure of NP-Z complex was predicted by AlphaPulldown [127], a python package built upon AlphaFold [128] and AlphaFold Multimer [129]. The structural model obtained high local and global quality scores as returned by AlphaFold and AlphaPulldown and predicted that Z interacts with the C-terminal domain of NP (Figure 22AB).

The significant differences in deuterium uptake detected for the peptides 450-483 and 456-483, align with the predicted NP-Z interface (Figure 22C). This was well visible when the differences based on the deuterium uptake at the 6h timepoint were plotted to the predicted NP-Z complex structure (Figure 22D).

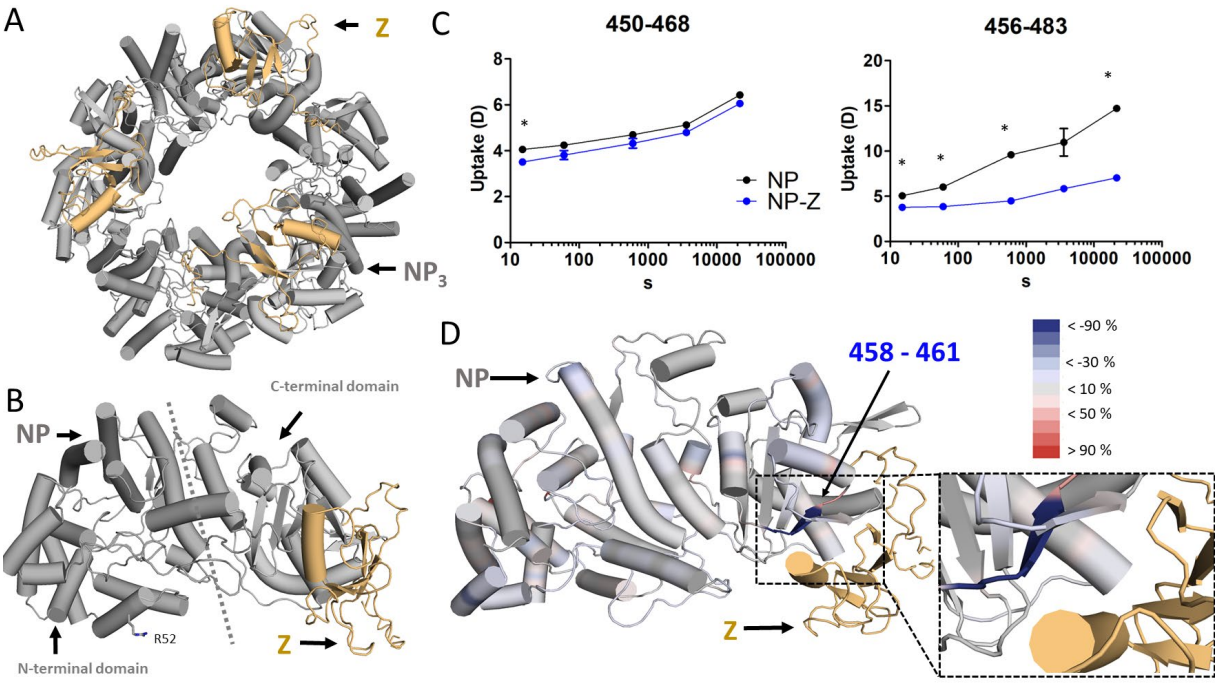

Figure 22: Identification of the NP-Z interface

The trimeric model of NP-Z (A) was generated by superposing the model of the NP-Z dimer predicted by AlphaFold (Version 2.1.0) (B) onto the crystal structure of the NP trimer (PDB ID 3R3L). NP is colored grey, and the Z protein is shown in gold. (C) HDX differences in NP peptides in presence of Z. Deuterium uptake plots showing both conditions NP alone and NP-Z mixture with significant differences (*delta HDX > X D – 95 % CI, see methods, Table 4) near to the predicted NP-Z interaction site (D). Differences in deuterium uptake between both states were mapped to the NP sequence and can be seen in Figure 21. The differences are based on the atomic range which considers overlapping peptides with no differences. The differences at the 6 h timepoint were plotted onto the predicted NP-Z complex. Regions that are protected and exposed in HDX in the presence of Z are highlighted in blue and red, respectively. Regions in grey show no change in deuterium uptake in the presence of Z. The Z protein is highlighted in gold. The AlphaFold NP-Z model was produced in collaboration with the Kosinski Lab at EMBL Hamburg.

3.2.3.3 Arginine 16 and 74 of Z are important for interaction with NP

To further validate the proposed interface, point mutations were introduced into the Z protein based on the NP-Z structural model. Polar residues were exchanged to alanine or glycine at or near the predicted interface. These were R4, R16, S59, S61, N62, R74, and T82. The Z mutants were recombinantly expressed in *E. coli* and purified analogous to wt-Z.

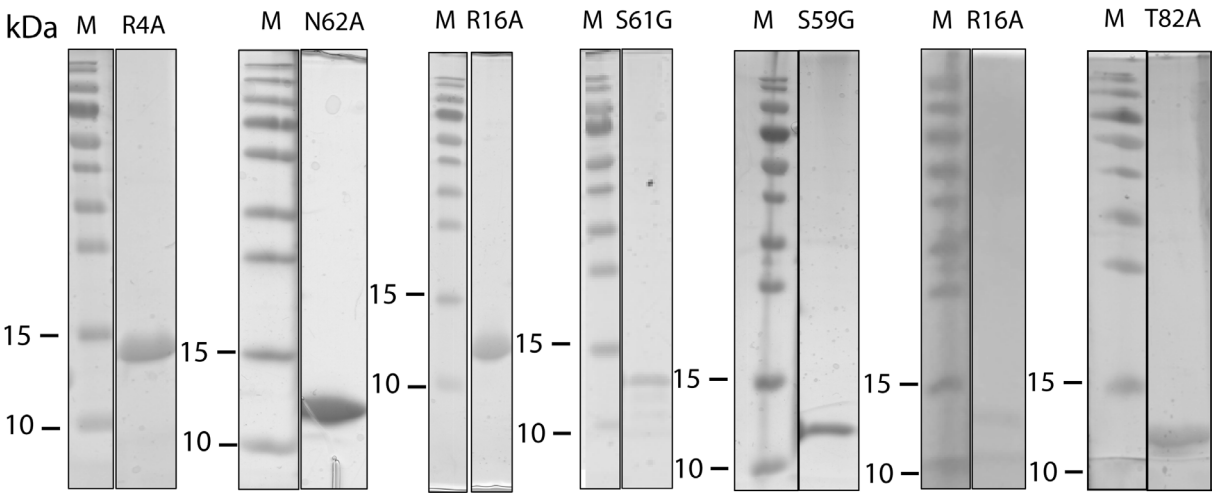

Figure 23: Purification of Z mutants

LASV Z protein mutants were expressed as a GST fusion protein and purified via GST affinity chromatography. After removing the GST tag, Z was passed through an anion exchange chromatography column to further purify the Z protein. As a final purification step, LASV Z was purified via SEC. Shown is a 15 % SDS PAGE of all purified Z mutants.

All Z mutants were measured together with NP in a 1:3 molar ratio by nMS and compared in the ability to form NP-Z complex by using the deconvoluted spectra of at least 3 independent measurements (Figure 24AB). Z mutants R16A and R74A showed strong reduction of NP-Z complex formation compared to the wild-type Z, whereas the other mutations had no effect on NP-Z interaction (Figure 24B). This leads to the conclusion that NP-Z interaction is strongly dependent on arginine 16 and 74 of Z, which supports the predicted model (Figure 24C). The other point mutations of polar residues seem to be less critical and were not able to influence NP-Z interaction.

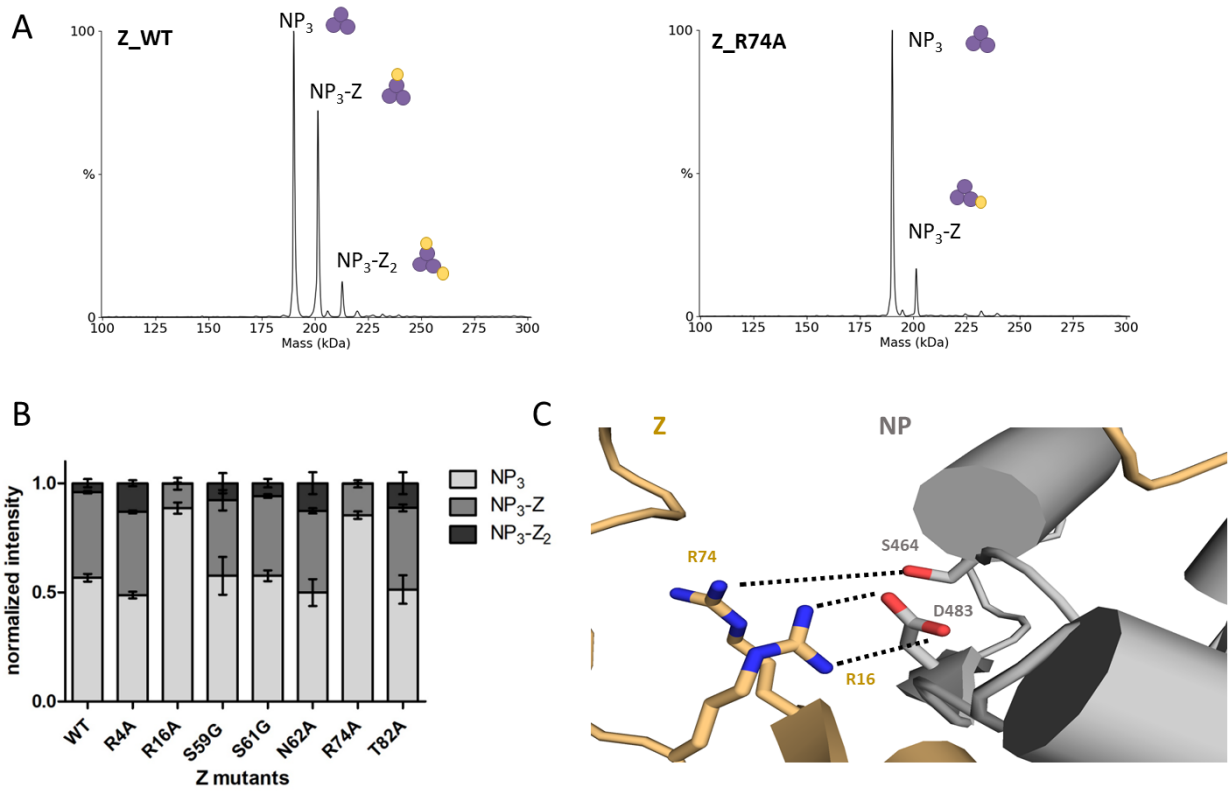

Figure 24: Z mutant screening identifies important residues for NP-Z interaction

The different Z mutants were recombinantly expressed and purified. (A) nMS screening of the Z mutants together with NP (6 µM) in a 1:3 (NP:Z) molar ratio in a 150 mM ammonium acetate solution at pH 7.5. (B) The sum of the intensities of every mass species of one measurement was normalized to 1. The normalized intensities for every measurement were plotted. Error bars represent standard deviation of at least 3 independent measurements. Measurements with the mutants R16A and R74A show reduced abundance of NP-Z complex species compared to wildtype (WT). (C) Detailed closeup of the Z residues R74 and R16 in the AlphaFold Multimer model. Z residues R74 and R16 are in close proximity to NP-S464 and NP-D483, respectively.

Taken together, our structural integrative approach including AI-based protein complex prediction, HDX-MS and mutational analysis in combination with nMS lead to detailed information about the direct NP-Z interaction. Z seems to bind at the C terminus of NP, where R16 and R74 of Z are likely important for the NP-Z interaction.

3.2.4 NP-Z interaction is pH dependent

If the NP-Z interaction is responsible for RNP recruitment into budding virions, it is unclear what triggers the release of the RNPs upon virus entry into a host cell and membrane fusion. A potentially important factor could be the pH as LASV particles are taken up via the endosomal pathway. It was recently shown that acidification of LASV virions by membrane permeability occurs before membrane fusion due to secondary receptor binding and GP rearrangement [130, 131]. A low pH could lead to

dissociation of RNPs from the Z protein lining the inner side of the viral membrane. Therefore, the pH-dependence of the NP-Z interaction was evaluated by incubating the proteins at varying pH between 5.0 and 7.5 and NP-Z complex formation was subsequently measured by nMS (Figure 26A). The first step was to evaluate the protein homogeneity after changing the buffer to lower pH. NP showed a similar oligomerization behavior at pH 5.0 and pH 7.5, with the masses still corresponding to a trimer (Figure 26B). In contrast, Z started to homo-oligomerize after changing the pH to 5 forming up to homohexameric complexes (Figure 25, Figure 26B). This is in line with an observation on influenza virus matrix protein that also undergoes oligomerization at low pH [132]. However, similar protein concentrations were observed at both pH 5 and pH 7.5 after buffer exchange and high-speed centrifugation indicating that no unspecific protein aggregation happens during buffer exchange.

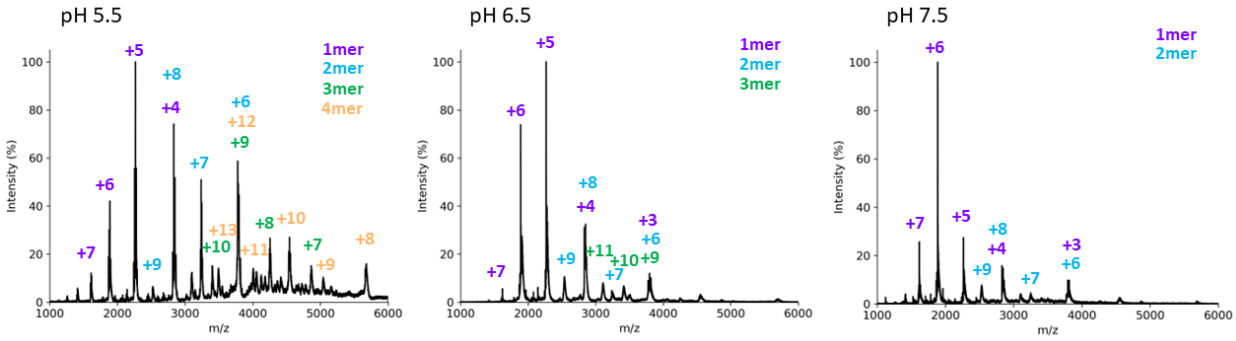

Figure 25: nMS spectra of Z at different pH

Z protein buffer was exchanged to 150 mM ammonium acetate solution with adjusted pH at 7.5, 6.5 and 5.5. The proteins were subjected to nMS with capillary and cone voltages were constantly at 1450 and 150 V, respectively. Acceleration voltage in the collision cell was at 50 V. Z appeared in the mass range between 1000 and 6000 *m/z*. Charge states are depicted for the Z protein and corresponded to a mass of an 1mer, 2mer, 3mer and 4mer, respectively.

To test the pH-dependence of the NP-Z interaction, 3 independent spectra of a mixture of NP and Z were recorded at a 1:3 (NP:Z) molar ratio at pH 5.0, 5.5, 6.5 and 7.5. We found that almost no NP-Z interaction was observed at pH 5.0 while complex formation at pH 7.5 was as observed in previous experiments (Figure 26C). At pH values between 5.5 and 6., NP-Z interaction was also less abundant compared to pH 7.5 (Figure 26D). These experiments show that the interaction between NP and Z is highly pH-dependent. This supports our hypothesis that the endosomal acidic pH reduces NP-Z interaction and, as a result, RNPs are released from the viral matrix.

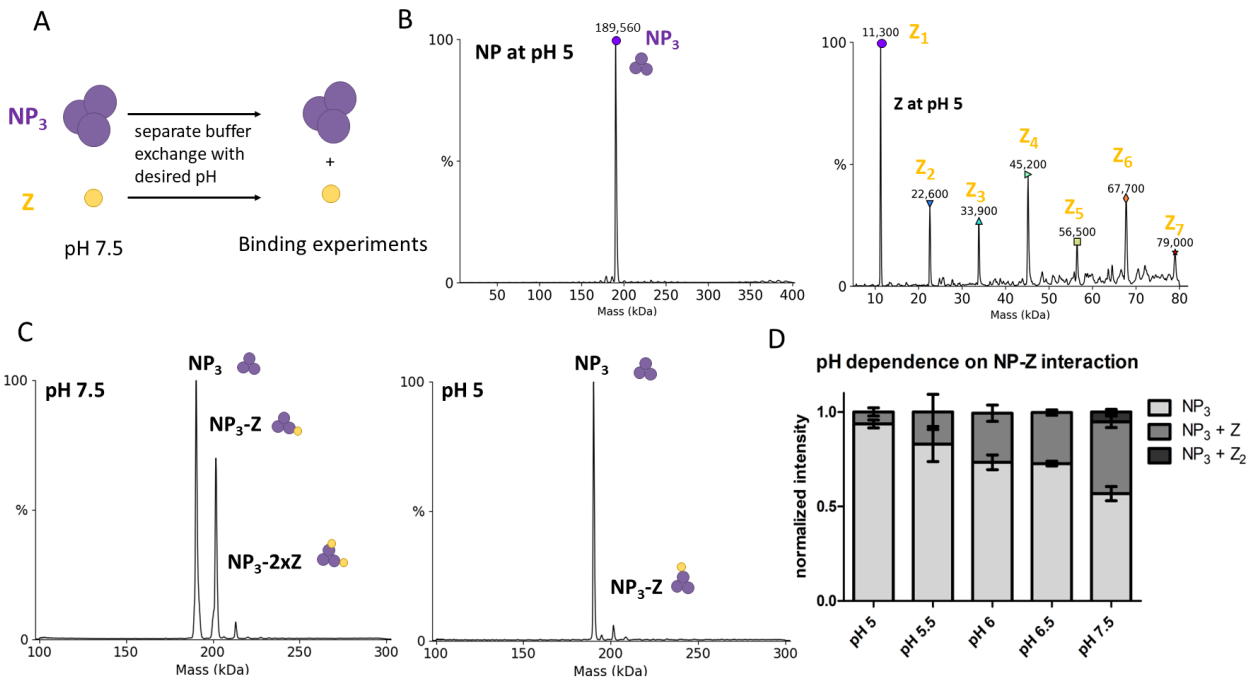

Figure 26: NP-Z interaction is less efficient at lower pH

(A) Scheme of the pH-dependent interaction assay. Proteins were separately exchanged to 150 mM ammonium acetate surrogate with adjusted pH. NP (6 µM) and Z with the same pH were incubated in a 1:3 (NP:Z) molar ratio and measured by nMS. (B) Deconvoluted spectra of NP and Z after changing the buffer to pH 5. NP oligomerization status was unaltered whereas Z starts to oligomerize at pH 5 and 5.5. (C) Spectra of NP and Z in a 1:3 molar ratio at different pH. (D) The normalized intensity for every mass species from the deconvoluted spectra was plotted according to the different pH conditions. The sum of all mass species of one measurement was set to 1. Error bars represent standard deviation of at least 3 independent measurements.

3.3 NP-RNA interaction

NP is incapable of binding RNA when in a trimeric conformation. The respective RNA binding pocket is blocked by two helices [56]. A proposed RNA gating mechanism suggests RNA binding when the C-terminal domain slightly rotates, affecting one helix (α5) to become partially unstructured, while another helix (α6) shifts away from the RNA binding pocket. However, the mechanism behind the conformational change remains unknown. A specific (host-) factor may trigger the RNA binding pocket to be open [55, 62]. The aim was to search for possible factors that could allow NP-RNA interaction.

3.3.1 NP_3 dissociates in the presence of short RNAs

To investigate the effect of RNA binding on the NP quaternary structure, NP_3 was incubated with short RNAs. As no sequence specificity of NP for RNA has been described so far, RNAs of different sequences and lengths were used. First, a 9 nt RNA was mixed in an equimolar ratio with NP and the changes in quaternary structure were monitored over at least 500 s in the ESI capillary (Figure 27A). The acquisition started approximately 30 s after mixing the components and mounting the capillary. During the measurement, the fractions belonging to the NP_3 decreased whereas peaks assigned to a mass of 66.08 +/- 0.02 kDa, corresponding to NP monomer bound to one RNA molecule (NP_1-RNA_1), increased. Importantly, NP_3 interacting with one RNA molecule (193.00 +/- 0.09 kDa) was observed at intermediate timepoints (Figure 27B). As the RNA binding groove of NP is inaccessible in the trimeric ring conformation the observed complex of one NP trimer and one RNA molecule (NP_3-RNA_1) presumably constitutes an open conformation of NP_3, allowing the RNA to bind. After 500 s the NP_3 was mostly disassembled.

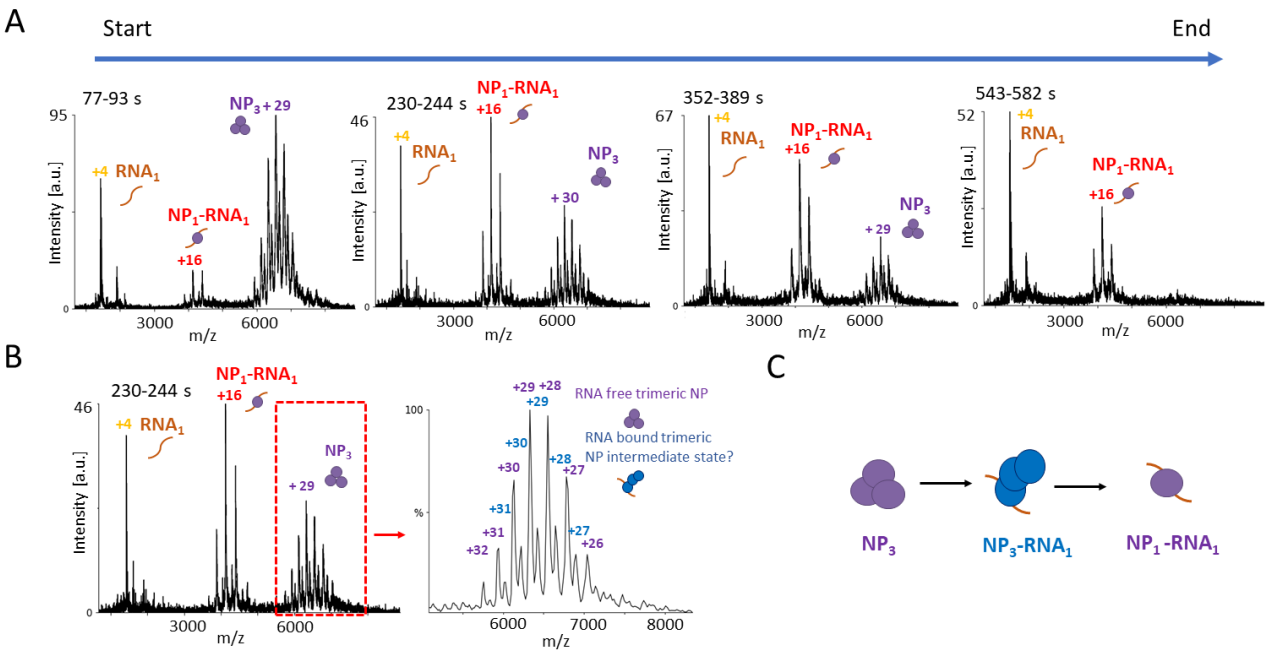

Figure 27: RNA triggers the dissociation of NP3 into monomers

NP_3 and a 9 nt RNA were mixed together in a 150 mM ammonium acetate solution at pH 7.5 on ice in a 1:1 (NP_1:RNA) molar ratio. The mixture was placed into an ESI capillary and the measurement was started approximately 30 s after mixing RNA and NP together. Spectra were recorded over 550 s. (A) Each diagram represents combined spectra recorded at 77-93 s, 230-244 s, 355-389 s, and 543-582 s with the intensity in arbitrary units (a.u.). Main charge states are labeled. Shown is one representative measurement of at least 3 independent measurements. (B) Recorded spectrum at timepoint 230-244 s. Zoom into the *m/z* range between 5000 and 8000 shows an RNA bound NP3 (193.00 +/- 0.08 kDa) species.

To further support our results, negative-stain electron microscopy was used to confirm the disassembly of NP_3 in the presence of RNA. As in our nMS data, NP was mainly observed as a ring-shaped trimer in the absence of RNA (Figure 28). In the presence of RNA, however, the morphology changes to smaller rounded particles and more heterogenous species as NP_3 is not detectable anymore (Figure 28).

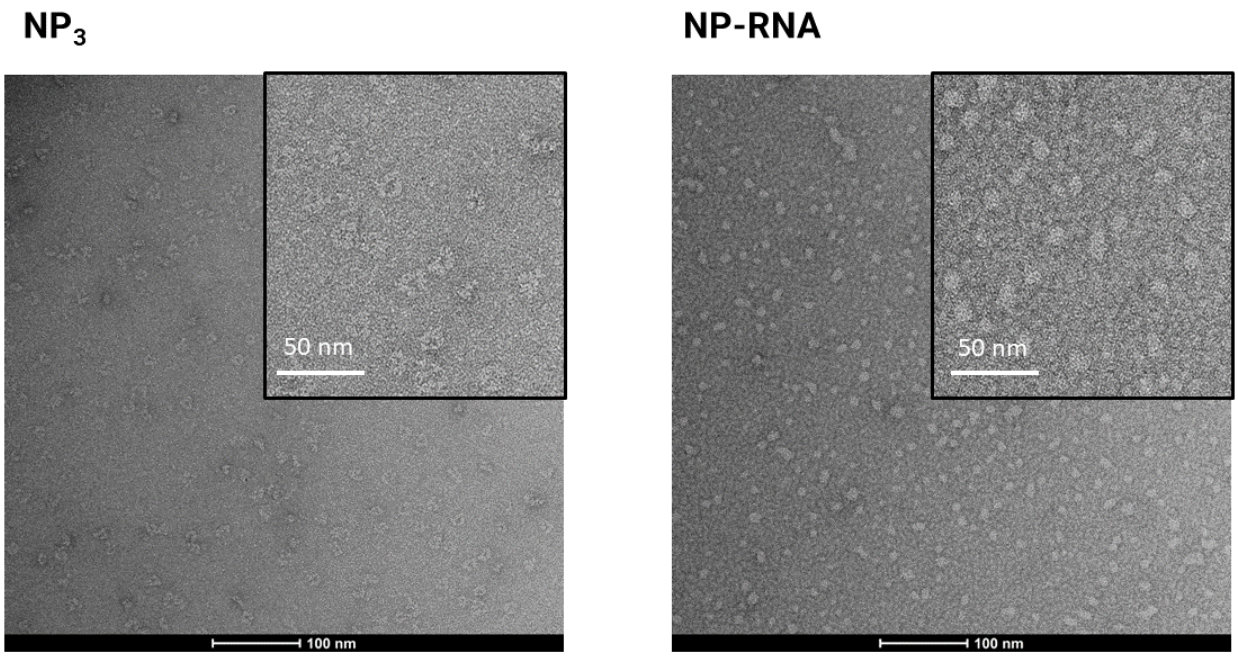

Figure 28: Negative staining electron microscopy of NP in presence or absence of RNA

NP at 160 nM was mixed with a single-stranded 25 nt RNA at a 1:2 molar ratio (NP:RNA) and incubated at room temperature for approximately 15 min. NP was applied to glow-discharged carbon-coated copper grids and stained with uranyl acetate immediately before imaging. Images were collected with a transmission electron microscope. The EM pictures were recorded in collaboration with Harry Williams (BNITM).

Furthermore, the results were validated by mass photometry. Analogously to our previous experiments, NP was observed in absence of RNA with a mass of 184 kDa which corresponds to a trimer and is comparable to the nMS experiments (189.56 +/- 0.03 kDa) (Figure 29A). For mass photometry measurements, NP at 500 nM was mixed with a single stranded 9-mer in a 4 molar-excess to the NP monomer. The mixture was diluted to an NP concentration of 25 nM and subsequently measured. After the addition of RNA, NP monomer formation was observed indicated by the mass of 69 kDa. Notably, the slightly higher mass of 192 kDa could indicate the NP_3-RNA_1 intermediate state which was observed in the nMS experiments (Figure 29B).

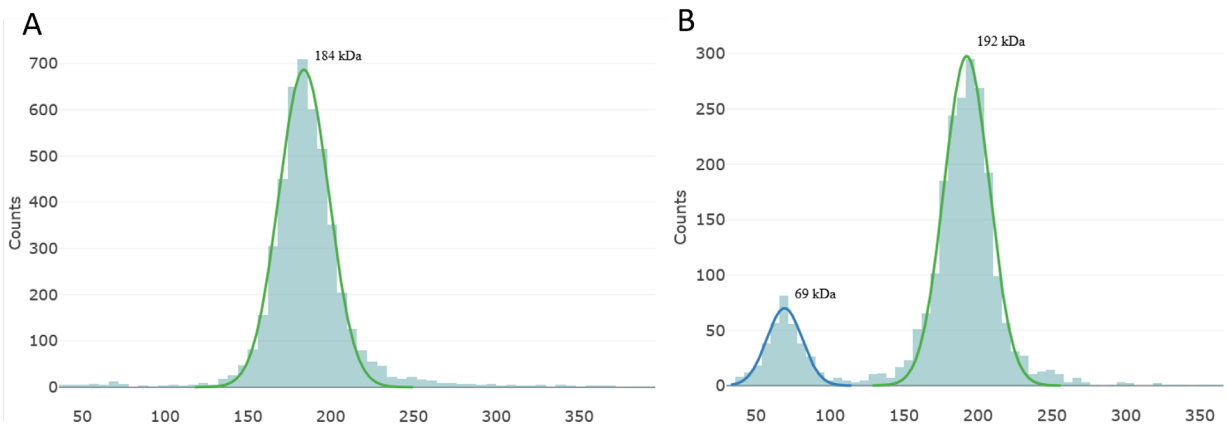

Figure 29: Mass photometry of NP in presence or absence of RNA

(A) Mass photometry was performed on a Refeyn OneMP mass photometer. NP was measured in a concentration of 25 nM in a conventional buffer condition. 4291 counts for NP were detected which represents 91 % of the total counts. The mass of NP was determined with 184 kDa and a sigma of 15 kDa. (B) NP at 500 nM was mixed with a single stranded 9mer in a 4 molar-excess to the NP monomer. The NP + RNA sample was diluted to an NP concentration of 25 nM and subsequently measured. 1880 counts for a mass of 192 kDa (sigma 15 kDa) were detected, which represents 78 % of the total counts. Additionally, 344 counts for a mass of 69 kDa (sigma 12 kDa) were detected representing 14 % of the total counts.

From these experiments it was concluded that short RNA is sufficient to trigger conformational changes leading to the disassembly of the trimeric ring structure of NP into RNA-bound NP monomers.

3.3.2 Monitoring higher order NP-RNA assembly

The question remained why only NP-RNA as monomers were observed and not higher order NP-RNA assembly as NP would make NP-NP interaction and assemble with viral genomic RNA. Therefore, a slightly longer RNA of 12 nts was used to investigate a potential assembly of RNA bound NP monomers, corresponding to the beginning of higher order NP-RNA assemblies. The components were mixed, and the complex species were monitored over a duration of 45 min in an ESI capillary. The measurement was started approximately 30 sec after mixing RNA and NP together. An equimolar ratio of NP to the RNA of 12 nts was tested first (Figure 30). After the start of the measurement 3 species were visible: (i) NP_3, (ii) a trimeric NP bound to one RNA molecule (NP_3-RNA_1,193.63 +/- 0.09 kDa) and (iii) unbound RNA (3.9 kDa) (Figure 30B). One minute later, a signal corresponding to an NP monomer bound to one RNA molecule (67.03 +/- 0.05 kDa) became visible. Between 5 and 30 min after the start of the measurement NP-RNA complexes assembled corresponding to

NP_2-RNA_2 (134.03 +/- 0.05 kDa), and NP_3-RNA_3 (201.17 +/- 0.13 kDa) species. As in the previous experiment, the signal for the ring-structured NP_3 decreased over time. The NP_3-RNA_3 complexes were most abundant after around 30 min of measurement. Notably, NP:RNA ratios of the complexes were always stoichiometric after initial disassembly of the ring-structured NP_3 via NP_3-RNA_1. The RNA concentration was increased to a molar ratio of 1:2 (NP:RNA) to further evaluate the formation of higher-order RNA-NP complexes, which was significantly faster compared to using an equimolar ratio (Figure 30). Although the NP_3 and the NP_3-RNA_1 complex (193.3 +/- 0.09 kDa), likely an intermediate, were observed at early timepoints of the measurement, the decay rate of the trimeric ring structure was higher than observed in the previous experiment with an equimolar NP:RNA ratio and NP_3 had completely disappeared after 40 min. In this experiment, formation of NP_1RNA_1 (67.03 +/- 0.05 kDa), NP_2-RNA_2 (134.03 +/- 0.05 kDa), NP_3-RNA_3 (201.17 +/- 0.13 kDa) and NP_4-RNA_4 (268.25 +/- 0.05 kDa) complexes were observed.

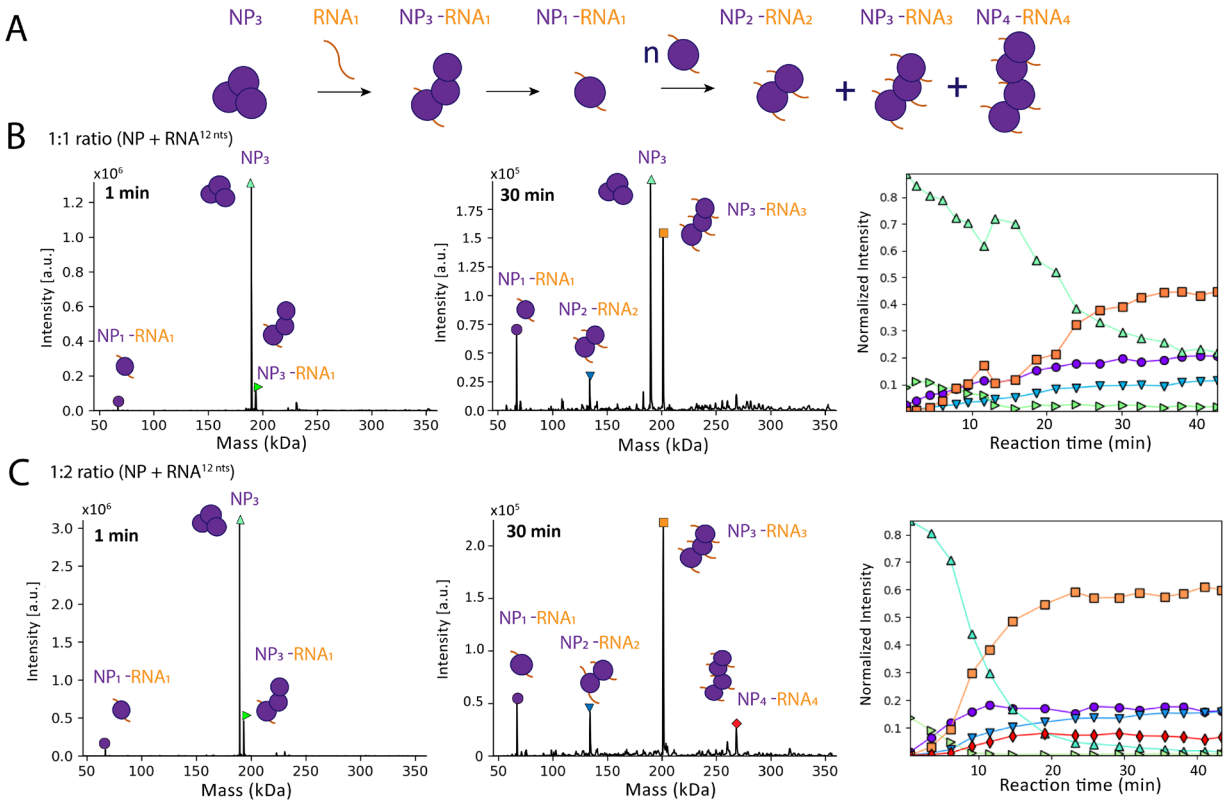

Figure 30: NP_3 dissociation and step-wise assembly of nucleocapsids

(A) Schematic summary of NP_3 dissociation and NP-RNA complex assembly over time as detected in the experiment. (B) Representative measurement of NP (9 µM) together with a 12 nt RNA in a 1:1 (NP:RNA) molar ratio recorded for 45 min in a 150 mM AmAc solution at pH 7.5. Spectra of all timepoints were deconvoluted with Unidec 30 shown are spectra at 1 min after starting the measurement (left) and at 30 min after starting the measurement (middle) with the intensity in arbitrary units (a.u.). The cartoons represent the complexes of NP and NP-RNA. Masses corresponding to NP_3 (189.56 +/- 0.02 kDa), NP_3-RNA_1 (193.63 +/- 0.09 kDa), NP_1RNA_1 (67.03 +/- 0.05 kDa), NP_2-RNA_2 (134.03 +/- 0.05 kDa), and NP_3-RNA_3 (201.17 +/- 0.13 kDa). The normalized intensity for every mass species from the deconvoluted spectra were plotted according to the different timepoints. The sum of all mass species of one timepoint was normalized to 1 (right). (C) Representative measurement of NP together with a 12 nt RNA in a 1:2 (NP:RNA) molar ratio. Spectra were analyzed as in (B). Masses correspond to NP_3 (189.56 +/- 0.02 kDa), NP_3-RNA_1 (193.63 +/- 0.09 kDa), NP_1-RNA_1 (67.03 +/- 0.05 kDa), NP_2-RNA_2 (134.03 +/- 0.05 kDa), NP_3-RNA_3 (201.17 +/- 0.13 kDa) and NP_4-RNA_4 (268.25 +/- 0.05 kDa).

In our measurements small fractions of NP_5-RNA_5 (335.78 kDa, FWHM: 0.90 kDa), and NP_6-RNA_6 (393.18 kDa, FWHM: 1.12 kDa) were detected, indicating the formation of higher oligomers (Figure 31A). In addition to the 12 nt RNA-NP complexes, NP monomers bound to truncated RNAs (RNA_{tr}) were detected. As no RNA_{tr} as free species were detected, the NP_1-RNA_{tr} complexes were most likely a result of an onset of CID, clipping off nucleotides not covered by the RNA-binding groove of NP (Figure 31B). These NP_1-RNA_{tr} complexes corresponding to 8-9 bases bound to NP were observed with different RNA substrates of varying lengths. This is in line with

previously reported crystal structures of NP bound to RNA, where 6 nts were fit into the RNA-binding groove [55]. Furthermore, no assembly between NP molecules happened via base pairing of overhanging RNAs, as with our acceleration voltage settings only single RNAs were observed (compare Figure 31 C and D).

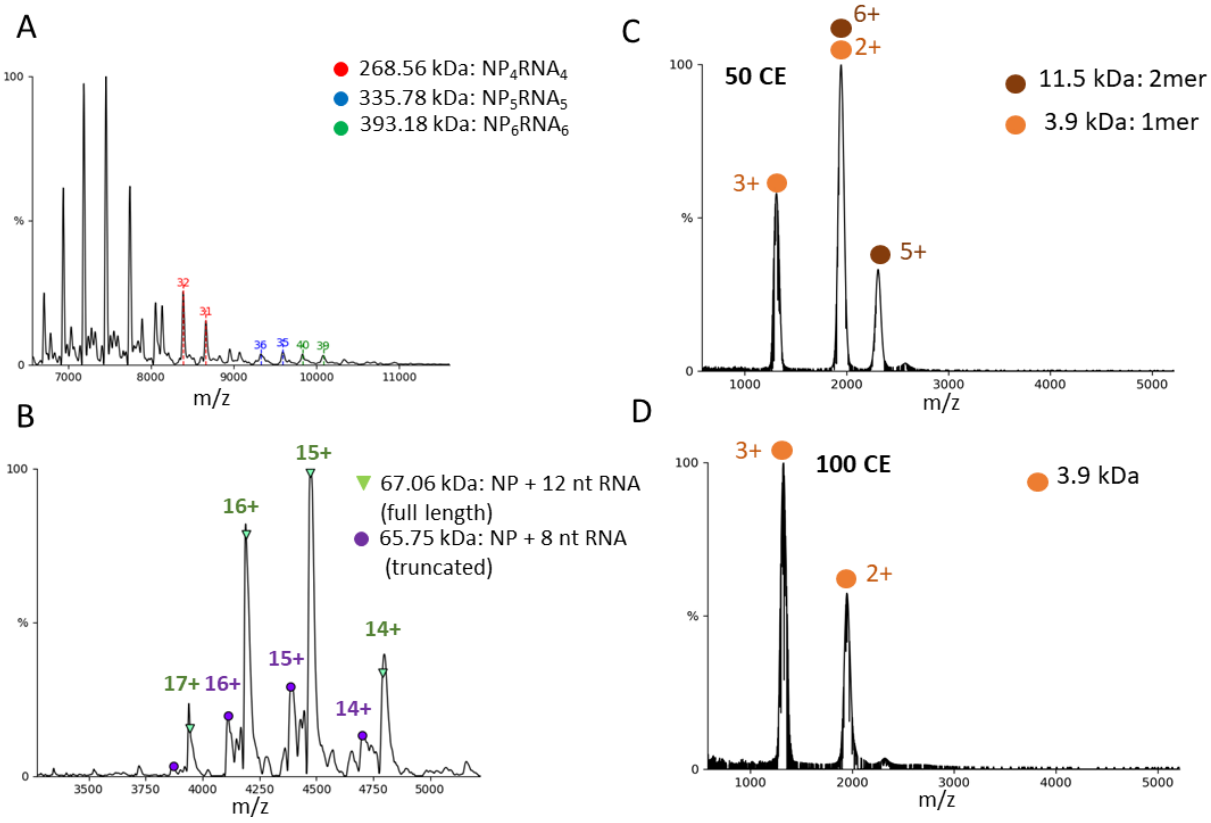

Figure 31: nMS spectra of NP-RNA species

(A) NP was measured together with a 12 nt RNA in a 1:2 ratio (NP:RNA). Shown are spectra at a late timepoint of the reaction. Small fractions of higher order NP-RNA assemblies appear between 9000 and 11000 *m/z*. Masses corresponding to NP_5RNA_5 335.78 kDa (FWHM: 0.90 kDa) and NP_6RNA_6 393.18 kDa (FWHM: 1.12 kDa). Charge states of 36+, 35+ and 40+, 39+ are indicated. (B) Close up of the mass range of NP monomers. The main monomer-RNA species is indicated in green with a corresponding mass of 67.06 +/- 0.05 kDa. Smaller amounts of NP monomers with truncated bound RNAs are indicated in purple with a corresponding mass of 65.75 kDa (FWHM: 0.21 kDa). The mass difference is about 4 nucleotides. (C) 25 µM of the 12 nt RNA was measured at different collision energies. The theoretical mass of this RNA is 3.78 kDa. At 50 collisional energy two mass species were identified in the range between 1000 and 3000 *m/z*. Charge states of 3+, 2+ and 6+, 5+ are indicated. Masses correspond to 3.9 kDa and 11.5 kDa indicating single and a trimeric RNA structure. (D) Only one peak series was observable at 100 CE corresponding to a mass of a single RNA (3.9 kDa).

It appeared that the NP-RNA complexes consistently had a 1:1 ratio. Therefore, longer RNAs were included in our analysis to see if those would be able to bind to two NPs. For this, three different RNAs with lengths of 18, 25 and 27 nucleotides

were used (Figure 32). As observed with the previously used 9 and 12 nt RNAs at early timepoints, the most abundant species was NP_3. Masses corresponding to NP_3-RNA_1 and NP_1-RNA_1 were detected at the 1 min timepoint. Also here, NP assembled with RNA in higher order NP-RNA assemblies. However, NP-RNA complexes were observed, where RNA was bound substoichiometrically. At the 20 min timepoint, signals corresponding to NP_2-RNA_1 and NP_3-RNA_2 assemblies were detectable with all tested RNAs. NP_4-RNA_2 assemblies were detectable with 25 and 27 nt RNAs. The signal for NP_3 and NP_3RNA_1 decreased. It can be concluded that a certain RNA length is needed to allow multiple NPs to bind to the same RNA molecule. The critical RNA length lies between 13 and 18 nucleotides. Notably, NP-RNA assembly kinetic seemed to be dependent on the RNA length. More assembly products were detected with longer RNA (27 nt) compared to the shorter RNA (18 nt), suggesting a more efficient process.

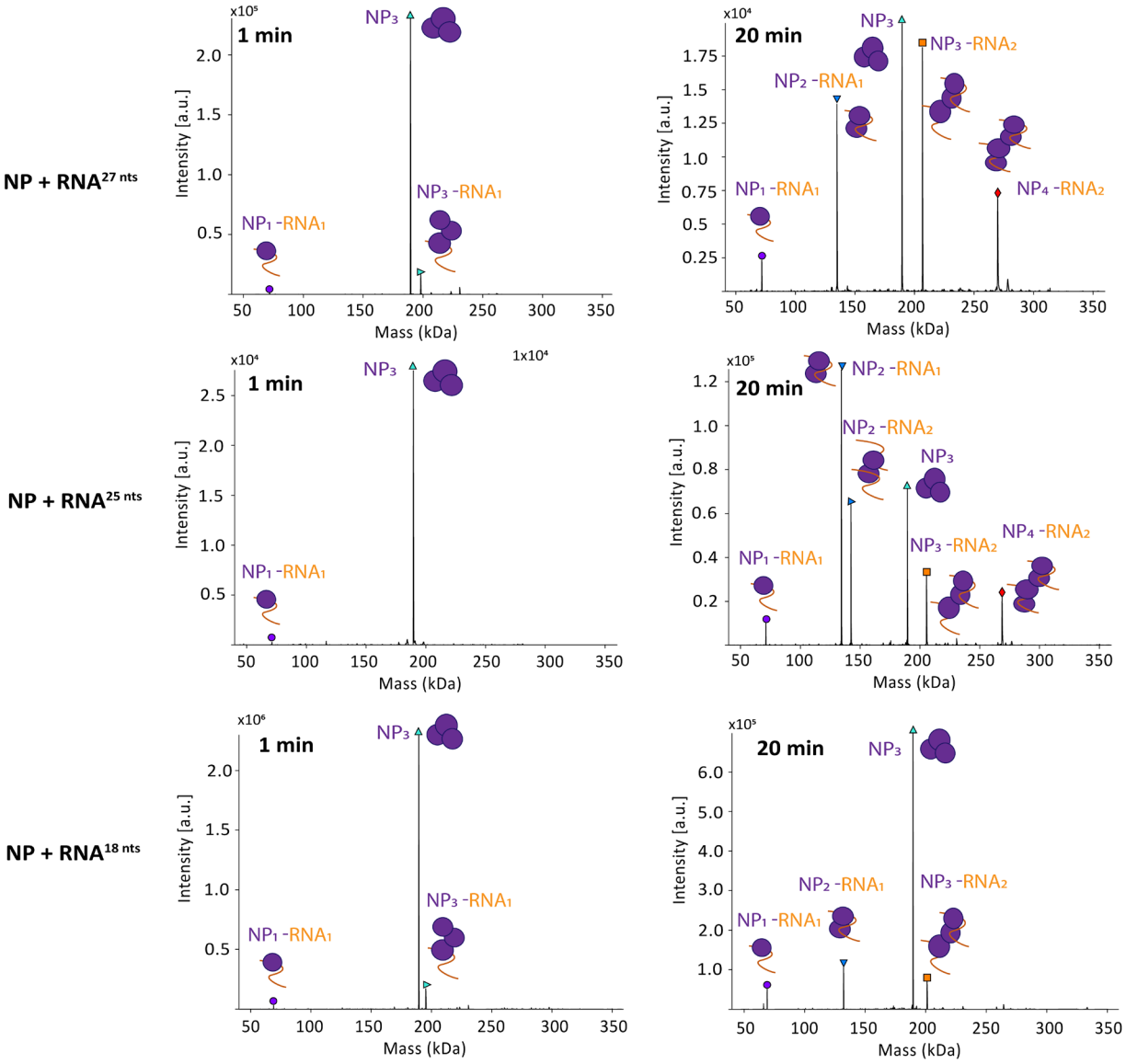

Figure 32: Representative measurement of NP together with a 27, 25 and 18 nt RNA in a 1:1 (NP:RNA) molar ratio

Spectra were analyzed as in Figure 30. Timepoints at 1 min and 20 mins are depicted. For the 27 nt RNA, masses corresponding to NP_3 (189.56 +/- 0.02 kDa), NP_3-RNA_1 (198.19 +/- 0.05 kDa) NP_1-RNA_1 (71.79 +/- 0.03 kDa) were detectable at the 1 min timepoint. At the 20 min timepoint, additional masses corresponding to NP_2-RNA_1 (134.95 +/- 0.04 kDa), NP_3-RNA_2 (206.69 +/- 0.04 kDa) NP_4-RNA_2 (269.92 +/- 0.05 kDa) were visible. For the 25 nt RNA, masses corresponding to NP_3 (189.56 +/- 0.02 kDa) and NP_1-RNA_1 (71.16+/- 0.01 kDa) were detectable at the 1 min timepoint. At the 20 min timepoint, additional masses corresponding to NP_2-RNA_1 (134.29 +/- 0.01 kDa), NP_3-RNA_2 (205.44 +/- 0.02 kDa) NP_4-RNA_2 (268.62 +/- 0.02 kDa) were visible. For the 18 nt RNA, masses corresponding to NP_3 (189.56 +/- 0.02 kDa), NP_3-RNA_1 (195.27 +/- 0.01 kDa) and NP_1-RNA_1 (68.97 +/- 0.03 kDa) were detectable at the 1 min timepoint. At the 20 min timepoint, additional masses corresponding to NP_2-RNA_1 (132.12 +/- 0.03 kDa) and NP_3-RNA_2 (201.09 +/- 0.03 kDa) were visible.

Altogether, these results show the higher-order assembly of NP and RNA, possibly resembling RNP formation during viral replication. Based on our data, we hypothesize that, as a first step, one molecule of NP within the RNA-free NP trimer binds to RNA leading to an opening of the ring-like conformation, which is less stable than the closed ring-like trimer. After this trigger, the open trimer further disassembles into RNA-bound NP monomers. Starting from that, higher-order NP-RNA complexes are then formed (Figure 30B). We could demonstrate that this process is dependent on the RNA concentration as an excess of RNA relative to NP led to a faster trimer disassembly and higher-order RNP-like NP-RNA complex formation (Figure 30C). For multiple NP protomers binding to the same RNA, a certain RNA length is needed suggesting a critical distance between two NPs inside RNPs. With an extended RNA length, the assembly process appears to be more efficient (Figure 32).

3.3.3 NP can interact with RNA and Z simultaneously

As described in chapter 3.2.3, Z binds to the C-terminal domain of NP while the RNA binding pocket of NP is located in the N-terminal domain. Co-immunoprecipitation data has previously shown for other arenaviruses that NP and Z, either directly or mediated by other factors, interact with each other [50, 53, 54]. It has been hypothesized that the NP-Z interaction mediates virion assembly by recruiting viral RNPs to the plasma membrane. We therefore investigated if LASV NP could indeed bind to Z and RNA simultaneously. Mixing a 9 nt RNA, NP and Z in a 1:1:2 ratio for subsequent nMS measurements, the NP_3 ring conformation again dissociated into RNA-bound monomers (Figure 33A) (66.08 +/- 0.03 kDa), some of which were associated with Z (77.38 +/- 0.06 kDa) (Figure 33B). CID-MS/MS confirmed these results (Figure 33CD). In both cases, the RNA remained bound to NP at high collision voltage whereas Z dissociated. This indicates a strong ionic interaction between RNA and NP compared to a weaker interaction between NP and Z. This observation fits with our proposed model of NP-Z interaction, which shows a rather small protein-protein interface between the C-terminal domain of NP and Z (Figure 22B). The affinity between wild-type NP and Z for the NP_1-Z_1-RNA_1 complex appeared to be in the same range when using the NP-R52A trimerization mutant with a K_D around 30 µM. Taken together, these results show that the NP_1-RNA_1 complex can indeed

interact with Z, supporting the idea that Z mediates the recruitment of RNPs to the plasma membrane by direct interaction of Z with RNA-bound NP.

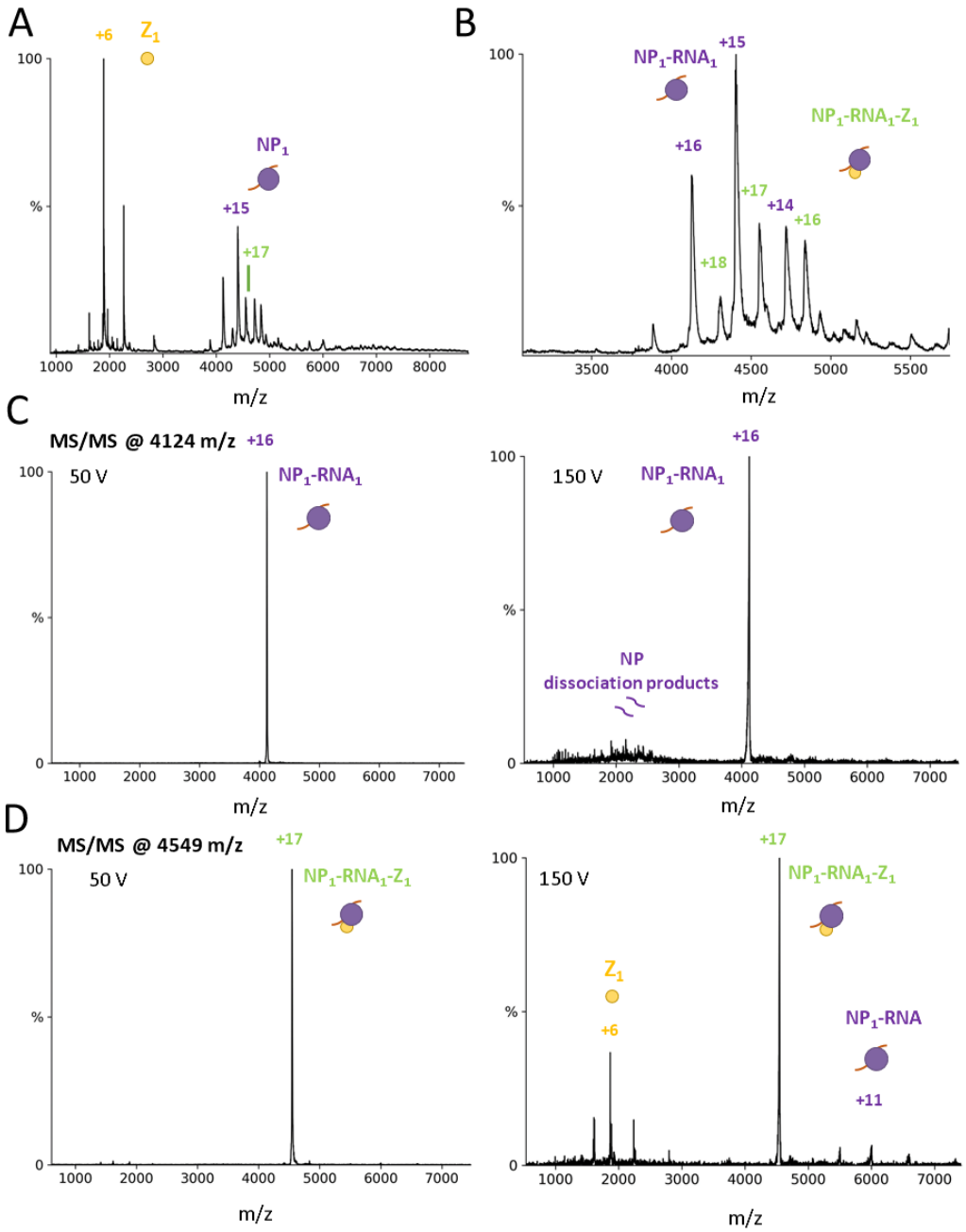

Figure 33: NP can bind Z and RNA simultaneously

(A) Full nMS spectrum of NP with Z and 9 nt RNA in a 150 mM AmAc solution at pH 7.5. NP3 completely dissociates into RNA-bound monomers. (B) Closeup of the *m/z* range 3,000-6,000 for determination of masses corresponding to NP-Z and RNA-NP-Z complexes. MS/MS of peaks corresponding to NP-Z isolated at 4,124 *m/z* (C) and NP_1-Z_1-RNA_1 isolated at 4,549 *m/z* (D) at low and high acceleration voltage in the collision cell. CID products for the NP_1-Z_1-RNA_1 complexes led to masses corresponding to Z monomers and NP_1-RNA_1 complexes, whereas no CID products for the NP-RNA complex were observed at higher acceleration voltage.

3.3.4 pH has no observable effect on NP-RNA binding affinity

As shown in chapter 3.2.4, NP-Z interaction was highly pH-dependent, which is likely relevant for the release of the RNPs from the matrix during the endosomal entry pathway. On the other hand, RNA needs to be still bound by NP at lower pH. The effect of a lower pH on NP-RNA binding was further investigated. NP-RNA interaction with 9 nt and 12 nt RNA was observed down to pH 5.0. In both cases, trimer dissociation and formation of NP-RNA complexes was observed analogous to pH 7.5. There is no indication of altered affinity. Notably, dissociation kinetics of the trimer in presence of RNA were accelerated (Figure 34).

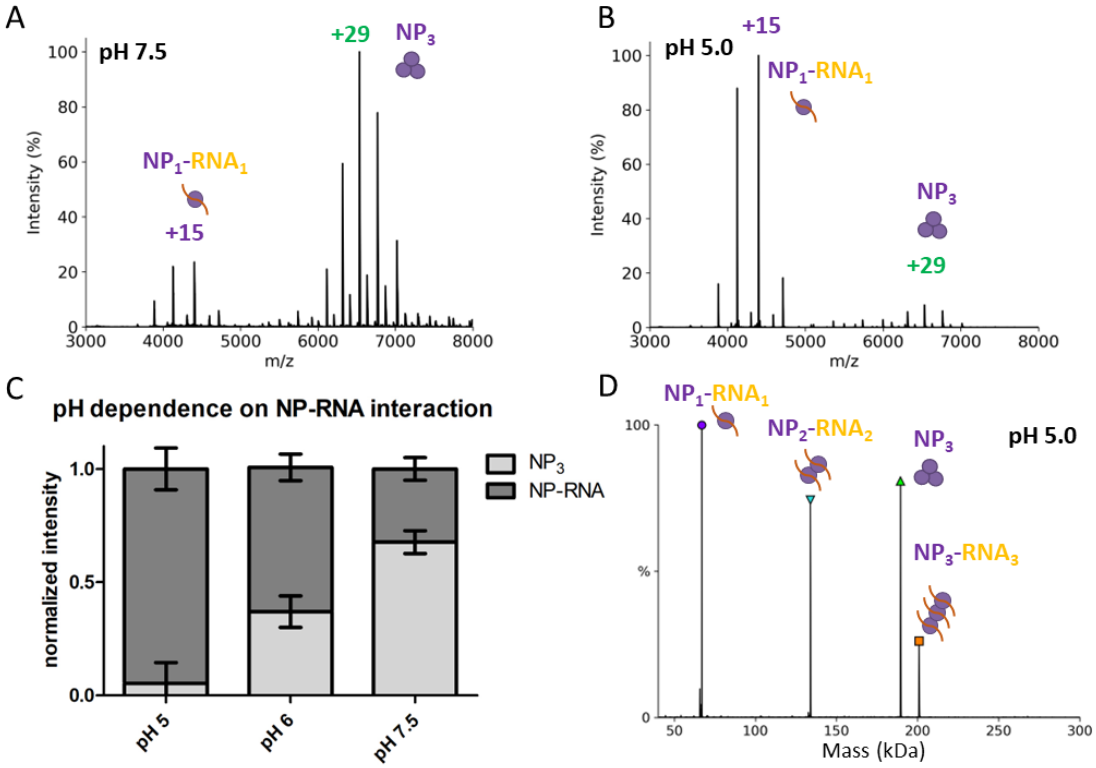

Figure 34: Influence of pH on the NP-RNA interaction

NP_3 and a 9 nt RNA were mixed together in a 150 mM AmAc solution at pH 7.5, 6.0 and 5.0 on ice in a 1:2 (NP_1:RNA) molar ratio. The measurement started approximately 30 s after mixing RNA and NP together. Shown is one representative measurement at 1 min after starting the measurement at pH (A) 7.5 and (B) pH 5.0. Main charge states are labeled with corresponding masses of NP_1-RNA_1 (66.08 +/- 0.02 kDa) and NP_3 (189.56 +/- 0.03 kDa). (C) The normalized intensity for NP_3 or NP-RNA (9 nt RNA) mass species from the deconvoluted spectra were plotted according to the different pH conditions. The sum of mass species corresponding to NP_3 or NP-RNA of one measurement was set to 1. Error bars representing standard deviation of at least 3 independent measurements. (D) NP_3 and a 12 nt RNA were mixed together in a 150 mM AmAc solution at pH 5.0 on ice in a 1:2 (NP_1:RNA) molar ratio. The measurement started approximately 30 s after mixing RNA and NP together. Shown is one deconvoluted spectra of a representative measurement at 6 min after starting the measurement. NP-RNA assembly products of masses corresponding to NP_1RNA_1 (67.03 +/- 0.05 kDa), NP_2-RNA_2 (134.03 +/- 0.05 kDa), and NP_3-RNA_3 (201.17 +/- 0.13 kDa) are visible.

3.4 Characterizing L-Z interaction by nMS

Z has been shown to interact with the L protein mostly via hydrophobic patches inhibiting its polymerase activity potentially by blocking the formation of an elongation conformation [45, 47, 133]. The aim was to further characterize the L-Z interaction via nMS together with RNA. As the L protein is present in high AmAc concentration and detergent in our nMS workflow, the L-Z interaction can be used as a positive control for later interaction studies with the L protein.

3.4.1 L binds Z with high affinity

First, L together was measured with Z in a 1:2 molar ratio. Under these conditions L was completely bound by one Z protein represented by a mass of 267.65 +/- 0.18 kDa (Figure 35A). The Z mutant W35A was additionally tested, which has an essential residue for L-Z interaction exchanged. No interaction was detected with this Z mutant (Figure 35B). By lowering the molar ratio to 1:1 with the WT Z protein, it was possible to detect both free L protein and L bound by Z. A K_D for the L-Z complex of 0.2 +/- 0.1 µM was determined (Figure 35CD), which is in line with already published affinity data [133].

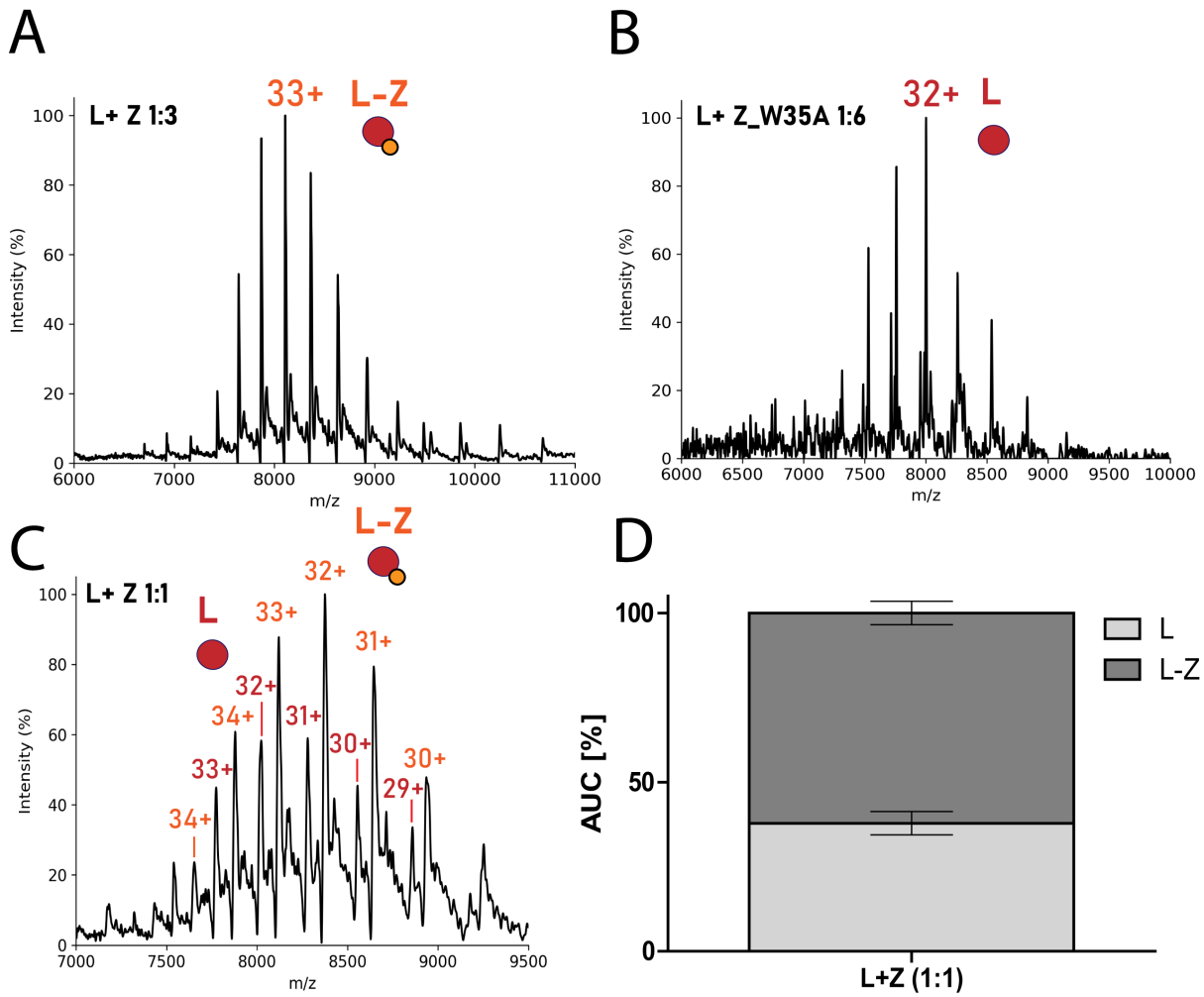

Figure 35: **Characterizing L-Z interaction by nMS**

(A) 0.8 µM L-C in 500 mM AmAc, 0.01 % Tween-20 at pH 7.0 was mixed with 2.4 µM Z in 150 mM AmAc pH 7.5 and subjected to nMS. Shown is a nMS spectra in the range between 6000 and 11000 *m/z*. The charge state for the strongest signal of all peaks is indicated. (B) 1.5 µM L-C were measured together with 10 µM Z_W35A. The charge state for the strongest signal of all peaks is indicated (C) 0.7 µM L-C and 0.7 µM Z were mixed together and measured with nMS. Shown is a nMS spectra in the range between 7000 and 9500 *m/z*. The charge states for L-Z (267.65 +/- 0.18 kDa) and L-C (255.97 +/- 0.68 kDa) are indicated. (D) L-C and Z were incubated in a 1:1 ratio (as in B) and resulting spectra from at least 3 independent measurements were deconvoluted to a zero-charge mass spectrum with Unidec and the AUC was determined for the respective mass species.

Next, higher molar ratios between Z and L were tested. When Z is present in high excess over L, an additional Z protein bound to the L-Z complex (278.78 +/- 0.06 kDa) was detected (Figure 36). A second Z bound to L could either indicate a secondary binding site of the Z at the L protein, or could result from a Z dimer bound to L. As free Z dimers are also observed at this concentration, a L-Z_2 complex via Z-Z interaction is most likely. This would imply that Z-Z interaction, i.e., Z oligomerization is not hindered when one Z protein is bound to the L protein.

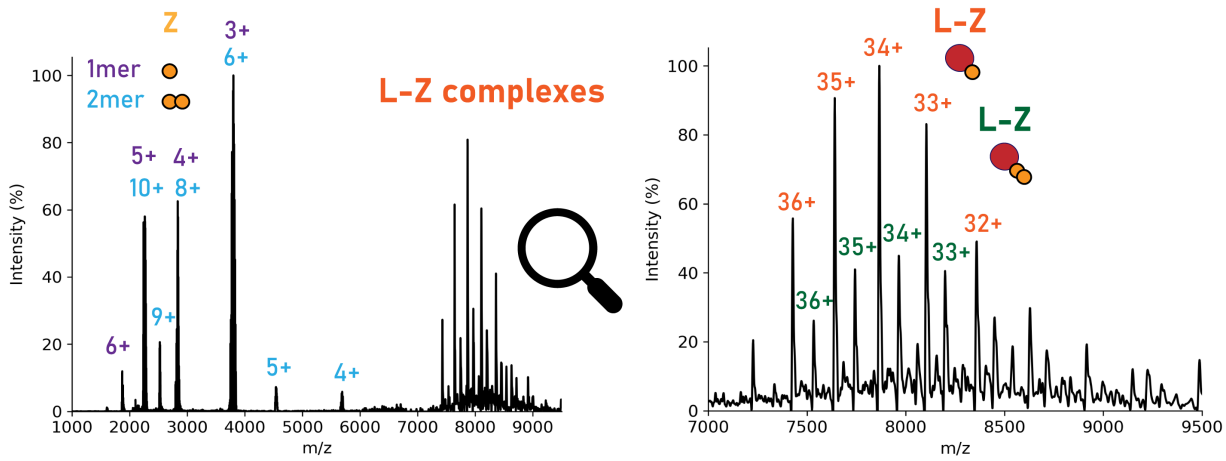

Figure 36: Observation of a L-Z_2 complex

L was measured together with Z and Z being present in an 80-molar excess over L with nMS. Overview spectrum between 1,000 and 9,500 *m/z* (left). The charge states for Z and Z dimers are indicated. L-Z complexes appeared around 8000 *m/z*. Detailed spectrum of the L-Z complexes between 7000 and 9500 *m/z* (right). Two peak series with masses corresponding to L-Z (267.65 +/- 0.18 kDa) and L-Z_2 (278.78 +/- 0.06 kDa) were present. Charge states are indicated.

3.4.2 L-Z complex can bind promoter RNA

Next experiments aimed to characterize L-RNA interaction. In general, the principal detection of RNA bound to L serves additionally as a control that polar interaction is still stable under the nMS conditions used in this work. It was possible to detect L-RNA interaction as mass for the L protein increased corresponding to the mass of the used RNAs (267.93 +/- 0.06 kDa) (Figure 37A).

The next experiment should further characterize the relationship between the L-Z and RNA binding. As described above, Z binds to the L protein inhibiting its polymerase activity by blocking the formation of the elongation conformation. However, it is an open question if the L-Z complex is still able to bind viral promoter RNA as well as nascent RNA. This was tested by incubating the L protein first with the Z protein for 5 min at RT and a later addition the primary 18 nt of the 3’ and 5’ LASV promoter

sequence. The mass of the L-Z complex (267.65 +/- 0.18 kDa) shifted according to the mass of the two RNAs (279.58 +/- 0.08 kDa) (Figure 37B).

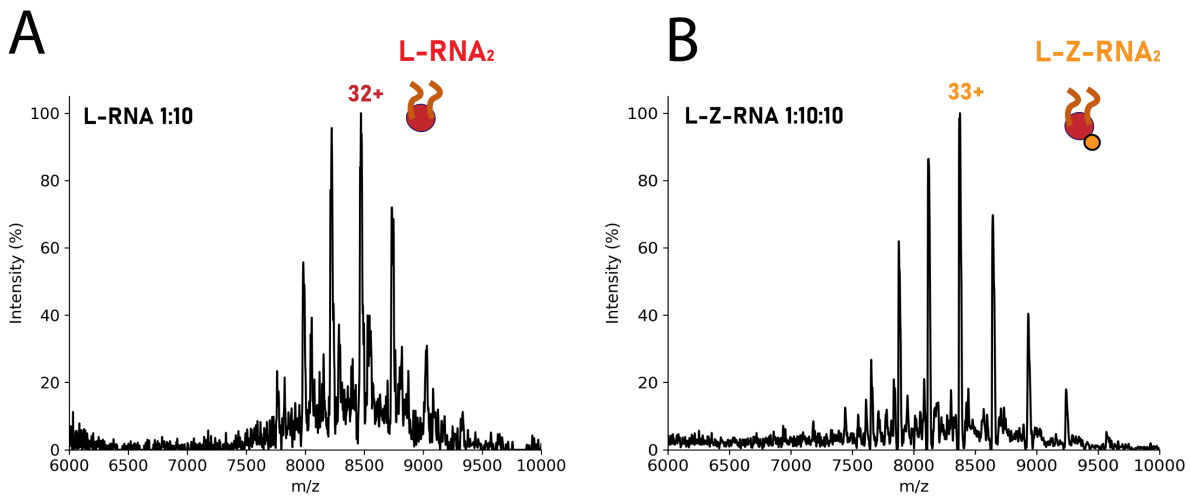

Figure 37: L-Z complex is able to bind RNA

(A) 1 µM L-C in 500 mM AmAc, 0.01 % Tween-20 at pH 7.0 was mixed with 10 µM 5' and 3' 1-18 Lassa promoter sequence. Shown is a nMS spectrum in the range between 6000 and 10000 *m/z*. The charge state for the strongest signal of all peaks is indicated. Mass for the L-RNA_2 complex corresponds to 267.65 +/- 0.18 kDa. (B) 1 µM of L-C protein was mixed with 10 µM Z and incubated for 5 min at RT. 10 µM of 3' and 5' 1-18 RNA was added and directly subjected to nMS. Shown is a nMS spectrum in the range between 6000 and 10000 *m/z*. The charge state for the strongest signal of all peaks is indicated. Mass for the L-RNA_2 complex corresponds to 279.58 +/- 0.08 kDa.

Overall, our established workflow investigating the LASV L protein by nMS is suitable to characterize binding partner of the L protein. Interaction of the L protein with the Z protein was demonstrated which is mainly based on hydrophobic interaction. Interaction with RNA which is polar based was also demonstrated. Complexes of L together with two Z protomers were observed indicating that Z-Z interaction might be possible on the L. Furthermore, promoter RNA binding seems to be possible when Z is present but further experiments are needed to evaluate if nascent RNA stays bound to L or is released upon inhibition of RNA synthesis.

3.5 NP does not directly interact with the L protein

Next experiments aimed to reconstitute the RNP complex involving NP, L and RNA. It is still unknown whether L is interacting directly with the NP during viral replication or interaction is mediated by RNA or other factors. The exact composition of the RNP has been previously controversially discussed. A study suggested that NP and L can be co-immunoprecipitated in presence of viral RNA [134], while another study reported co-immunoprecipitation also in absence of viral RNA [135]. Additionally, a study suggested a potential direct interaction [136].

Our first experiments aimed to identify a possible direct interaction between NP and L. NP_3 were tested together with both L variants in different molar ratios. However, no direct interaction between both proteins were detected (Figure 38A and B). It is possible that the trimeric conformation has an influence on an interaction between NP and L. Therefore, the NP trimerization mutant R52A (see chapter 3.2.2.1) was used to test this. L-C were mixed with NP-R52A in a 1:3 molar ratio and subjected to nMS. There was no direct interaction observable with the L protein when NP was present in a monomeric conformation, as only masses for NP-R52A (63.90 +/- 0.08 kDa) and L-C (255.97 +/- 0.68 kDa) were detected (Figure 38D).

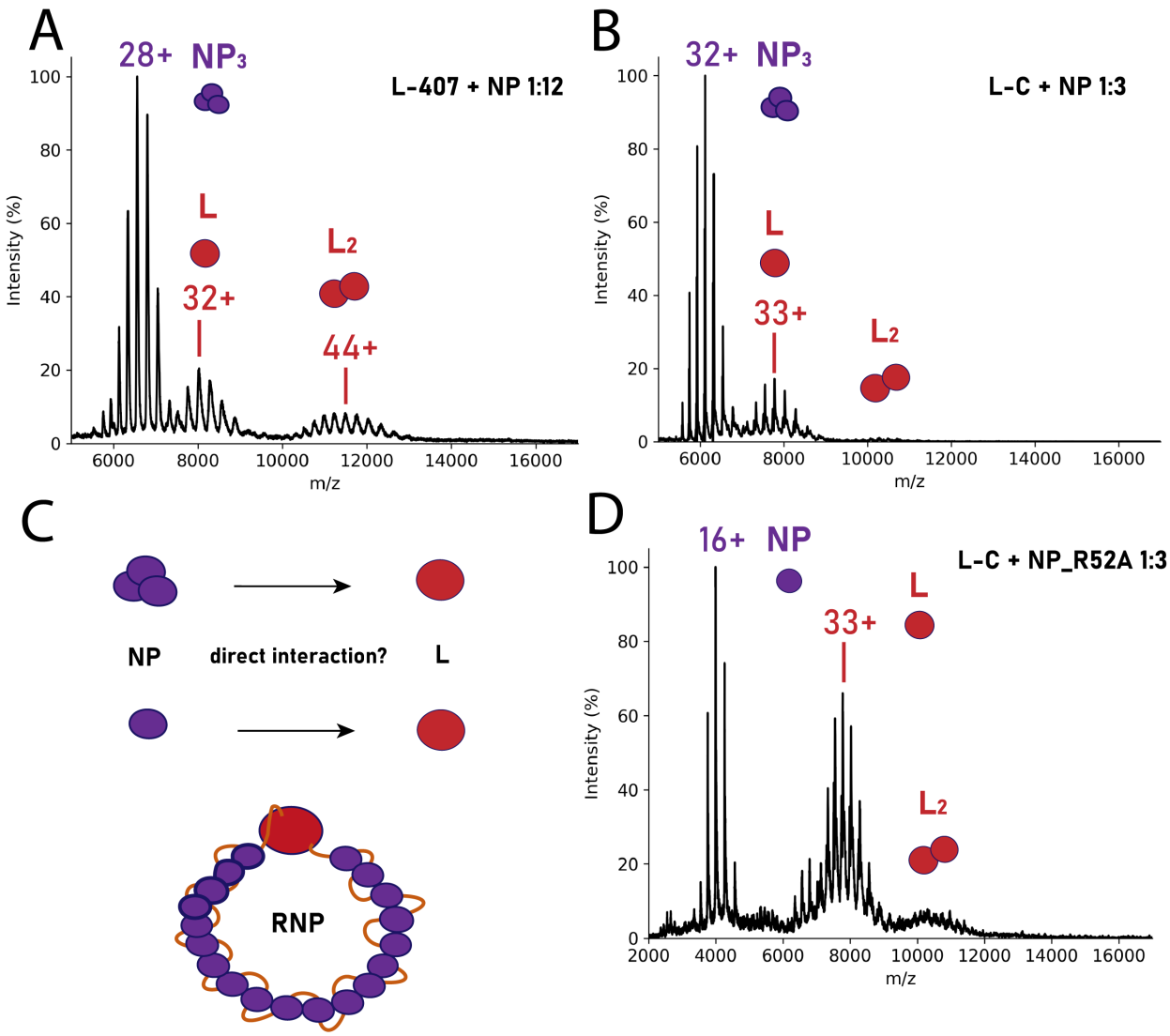

Figure 38: NP does not directly interact with the L protein

(A) 0.5 µM L-407 in 500 mM AmAc, 0.01 % Tween-20, 500 µM manganese acetate at pH 7.0 were mixed with 6 µM NP in 150 mM AmAc pH 7.5 and subjected to a QToF2 mass spectrometer. Shown is a nMS spectrum in the range between 5000 and 17000 *m/z*. The charge states for the strongest signal of every peak series are indicated. Masses for the L protein (254.74 +/- 0.31 kDa) and NP_3 (189.56 +/- 0.03 kDa). (B) 1 µM L-C in 407 in 500 mM AmAc, 0.01 % Tween-20, at pH 7.0 were mixed with 3 µM NP in 150 mM AmAc pH 7.5 and subjected to a QE-UHMR mass spectrometer. As in (A), masses for the L and NP_3 were traceable. (C) Model of the RNP and aim of this experiment to analyze a potential direct interaction between NP and L. (D) 1 µM L-C in 407 in 500 mM AmAc, 0.01 % Tween-20, at pH 7.0 were mixed with 3 µM NP-R52A in 150 mM AmAc pH 7.5 and subjected to a QE-UHMR mass spectrometer. Masses of NP-R52A (63.90 +/- 0.08 kDa) and L-C (255.97 +/- 0.68 kDa) were detected.

As the Z protein is interacting with both L protein and NP, next experiments aimed to reconstitute a complex of all three proteins. Z would act as a bridge between NP and L. However, despite the biological relevance, it would be interesting to see if a complex composed of three proteins is sterically possible. L, NP, and Z were mixed in

a 1:2:6.6 molar ratio and measured by nMS. Under these conditions a large fraction of L was present in complex with Z (267.65 +/- 0.18 kDa) whereas NP_3 (189.56 +/ 0.03 kDa) was only partially occupied by Z (201.00 +/- 0.09 kDa). This can be explained by the different affinities of the two proteins to the Z protein. One NP_3-Z complex and traces of NP-Z_2 complexes were detected. However, no masses corresponding to NP-Z-L complex were found (Figure 39A). Next, it was speculated if RNA could influence an interaction as promoter RNA was shown to induce substantial conformational changes of the L protein. A 3' LASV promoter sequence (1-18) was included into the mixture and again measured with nMS. As observed in our previous experiments, L showed a high affinity to the RNA as L-Z was entirely complexed by RNA (279.58 +/- 0.76 kDa). However, also in complex with RNA no interaction with NP or NP-Z was detected (Figure 39B). Notably, no NP-RNA interaction was observed, either RNA was completely bound by the L protein with a higher affinity, or the NP-RNA affinity was reduced by the high amount of AmAc present (because of the L protein).

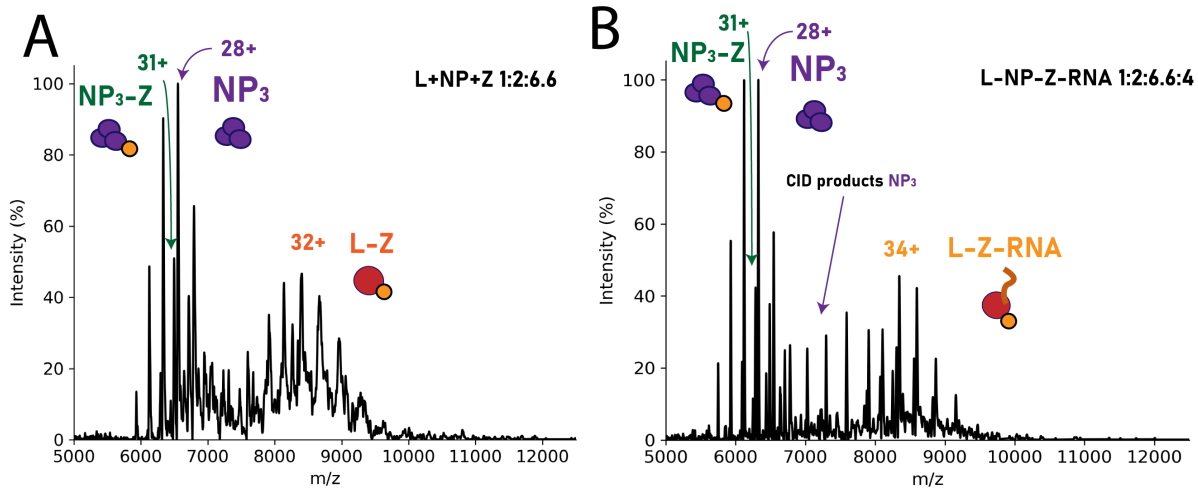

Figure 39: Indirect interaction between NP and L via Z not observed

(A) 1.2 µM L-C, 2.4 µM NP and 8 µM Z where mixed and subjected to a QE UHMR instrument. Shown is a nMS spectrum in the range between 5000 and 13000 *m/z*. The charge states for the strongest signal of every peak series are indicated. Masses for the L-Z complex (267.65 +/- 0.18 kDa) NP_3 (189.56 +/- 0.03 kDa) and NP_3-Z are visible. (B) 1.2 µM L-C, 2.4 µM NP, 8 µM Z and 5 µM 3' 1-18 Lassa promoter RNA was mixed and subjected to a QE UHMR. nMS spectrum is presented as in (A). Masses for the L-Z-RNA complex (279.58 +/- 0.08 kDa), NP_3 (189.56 +/- 0.03 kDa) and NP_3-Z (201.00 +/- 0.09 kDa).

Another hypothesis was addressed in the following experiment. NP might need to have RNA bound before interacting with the L protein. RNA binding induces substantial conformational changes in both proteins which could make the interaction site accessible. However, as explained above, there were difficulties inducing NP-RNA binding when L was present. Therefore, NP was incubated with RNA first to induce NP-RNA interaction and L protein was added later for potential interaction with NP-RNA assemblies. After subjecting the mixture to nMS, NP-RNA assemblies of monomers (69.34 +/- 0.17 kDa) and dimers (138.01 +/- 0.01 kDa) were found. Additionally, the mass for L bound by two RNAs (267.68 +/- 0.06 kDa) was observed. However, also with this experimental set-up, the detection of an NP-L complex was not possible (Figure 40A).

Additionally, another RNA was used which has the first 7 nts of the LASV 3’ promoter sequence following 20 random nts (LASV 3'1-7-r20). A specific binding of the 3’ promoter sequence to the 3’ RNA secondary binding site of the L protein was anticipated, which would ideally result in the random part of the RNA sticking out of the L protein. This freely accessible part of the RNA could act as a binding platform for NP. The LASV 3'1-7-r20 RNA was incubated together with NP for 5 min at RT and the L protein was added afterwards. Again, NP-RNA assemblies were observed with masses corresponding for NP_1-RNA_1* (67.76 +/- 0.04 kDa) NP_2-RNA_1 (134.95 +/- 0.04 kDa) NP_3-RNA_2 (206.69 +/- 0.04 kDa). L bound to one or two* RNAs (264.79 +/- 0.02 kDa, 270.22 +/- 0.02 kDa) was detected. However, there was no binding of NP to RNA which is attached to L (Figure 40B). Notably, some truncated RNA species bound to L and NP were detected (indicated by an asterisk). This could hint to a possible active endonuclease of the L protein or exonuclease activity of the NP.

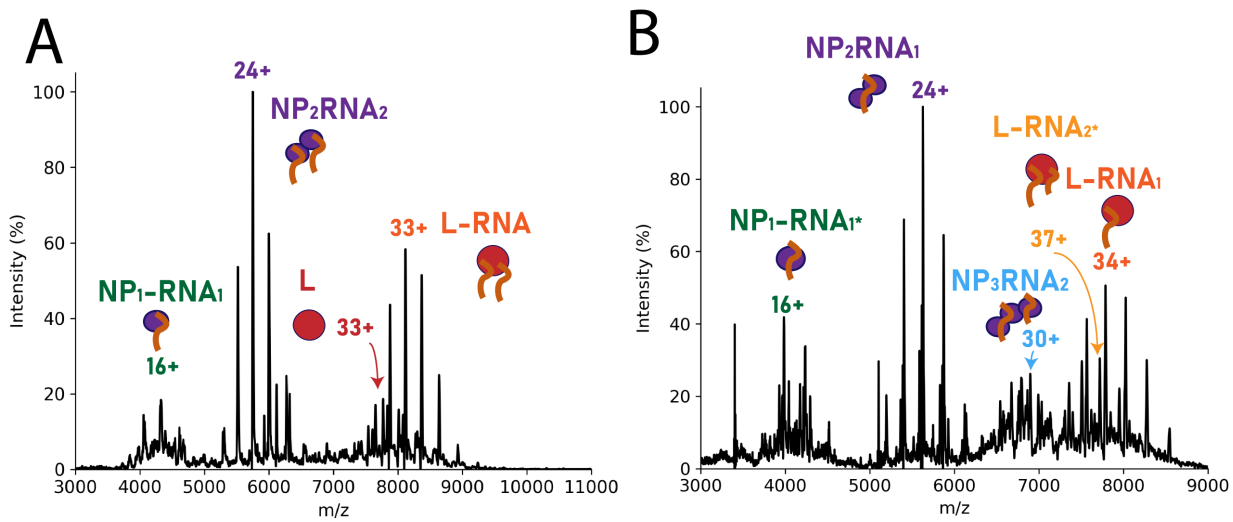

Figure 40: RNA does not induce NP-L interaction

(A) 0.8 µM NP were incubated with 1.4 µM 3' 1-18 Lassa promoter RNA for 5 mins at RT. 1.4 µM L-C was added on ice and directly subjected to nMS. Shown is a nMS spectrum in the range between 3000 and 11000 *m/z*. The charge states for the strongest signal of every peak series are indicated. Masses for the L protein (255.97 +/- 0.68 kDa), L-RNA_2 complex (267.65 +/- 0.18 kDa) NP_1-RNA (69.34 +/- 0.17 kDa) and NP_2RNA_2 (138.01 +/- 0.01 kDa) are observed (B) 0.3 uM L + 5 µM LASV 3'1-7-r20 RNA were mixed with 1.5 µM NP and incubated for 5 min at RT. 0.3 µM L-C was added on ice and directly subjected to nMS. Shown is a nMS spectrum in the range between 3000 and 9000 *m/z*. The charge states for the strongest signal of every peak series are indicated. Masses for the L protein (255.97 +/- 0.68 kDa), L-RNA_2* complex (270.22 +/- 0.02 kDa) L-RNA_1 complex (264.79 +/- 0.02 kDa), NP_1-RNA_1* (67.76 +/- 0.04 kDa) and NP_2RNA_1 (134.95 +/- 0.04 kDa), NP_3RNA_2 (206.69 +/- 0.04 kDa) are observed. An asterisk (*) truncated RNA species

Altogether, these results show that it was not possible to reconstitute the RNP complex *in vitro* via nMS. Evidence was provided that NP either as trimer nor monomer is directly interacting with the L protein under our experimental conditions. Different RNA substrates were used to induce a possible indirect interaction between NP and L. However, it was not possible to observe interaction with different RNA substrates. As described above, the Z protein binds both L and NP. It was tested if Z could act as bridge between NP and L, however no complexes were detected composed of all three proteins which indicates that Z is not able to bind NP and L simultaneously.

3.6 nMS as a platform to detect interaction between LASV and host proteins

Many steps of the viral replication cycle depend largely on host–pathogen interactions [137]. However, little is known about how host factors are directly involved in the replication of the LASV. The identification of direct host interaction partners can increase our understanding of how the virus is functioning in cells and finding targets for potential anti-viral drug development.
The establishment of nMS with the LASV proteins can be used as a platform to analyze possible direct interaction between the LASV proteins and host factors.

3.6.1 ANP32A does not directly bind to the L protein

The acidic leucine-rich nuclear phosphoproteins (ANP32) are a family of proteins that play a role in multiple cellular processes including chromatin remodeling and intracellular transport [138]. ANP32A have been identified as an essential host factor for influenza virus genome replication, directly binding to the polymerase mediating dimerization [139]. Here, it was investigated whether ANP32A has a similar function for the LASV polymerase. It was investigated whether ANP32A can directly interact with the LASV L protein.
ANP32A was successfully recombinantly expressed and purified in *E. coli* (see supplement). ANP32A was buffer exchanged to an AmAc solution and subjected to nMS to analyze a possible complex formation with the LASV L protein. Both L-C and L-407 protein variants were tested, as the C-terminal StrepII-tag might influence a potential interaction. ANP32A and both L protein variants were subjected to nMS in different molar ratios. However, no complex formation was observed as only masses for monomeric ANP32A (28.65 +/- 0.02 kDa) and the L proteins (255.97 +/- 0.68 kDa and 254.74 +/- 0.31 kDa) were detected. Overall direct interaction between L and ANP32A seems to be unlikely. However, other factors or conditions might influence an interaction.

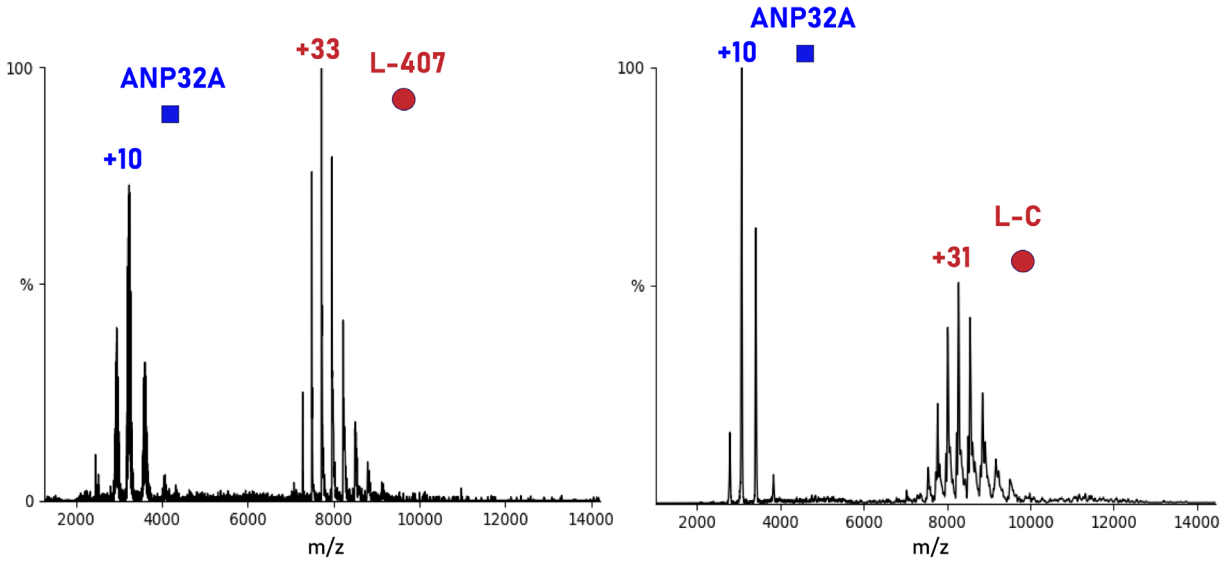

Figure 41: Both L protein variants are not interacting with ANP32A

(A) ANP32A were buffer exchanged to 150 mM AmAc pH 7.5. 12 µM ANP32A were incubated with 1.5 µM L-407 and subjected to the mass spectrometer after short incubation. Shown is a nMS spectrum in the range between 1500 and 14000 *m/z*. The charge states for the strongest signal of every peak series are indicated. Masses for the L-407 protein (254.74 +/- 0.31 kDa) and ANP32A (28.65 +/- 0.02 kDa). (B) 4.5 µM ANP32A were incubated with 1.5 µM L-C and measured as in (A). Masses for L-C protein (255.97 +/- 0.31 kDa) and ANP32A (28.65 +/- 0.02 kDa) are observed.

3.6.2 ANP32 interacts with NP

Parts of the results and figures of this chapter were produced by Elisabeth Eckelmann who did her master thesis under my supervision.

Recently, a direct interaction between the influenza virus NP and ANP32 has been observed. The C-terminal low complexity acidic region (LCAR) domain was shown to bind NP to put it into proximity to RNA synthesized by the viral polymerase [140]. ANP32 was previously also identified to play a critical role during influenza virus replication by mediating dimerization of the polymerase [139]. However, we detected no direct interaction between LASV L and ANP32A. We tested for a possible direct interaction between LASV NP and ANP32A instead.

We incubated ANP32A and NP in a 1:0.5 molar ratio for 5 mins at RT and subjected the mixture to nMS. As observed in previous experiments NP forms a homo-trimer in the absence of RNA (189.56 +/- 0.03 kDa). ANP32A appeared as monomer (28.65 +/ 0.02 kDa). Additionally, we could detect a third mass species corresponding to a NP_3-ANP32A complex (218.26 +/- 0.12), (Figure 42A). This

indicates a direct interaction between NP and ANP32A. We further characterized the potential interaction between NP and ANP32A and applied increasing amounts of CE on the complex. The complex remained relatively stable, whereas unbound ANP32A was not visible anymore (Figure 42B). Furthermore, we tested different molar ratios and observed a concentration dependent increase of complex formation between NP_3 and ANP32A. At a 1:1 molar ratio, approximately 10 % of NP_3 was bound by ANP32A (Figure 42C).

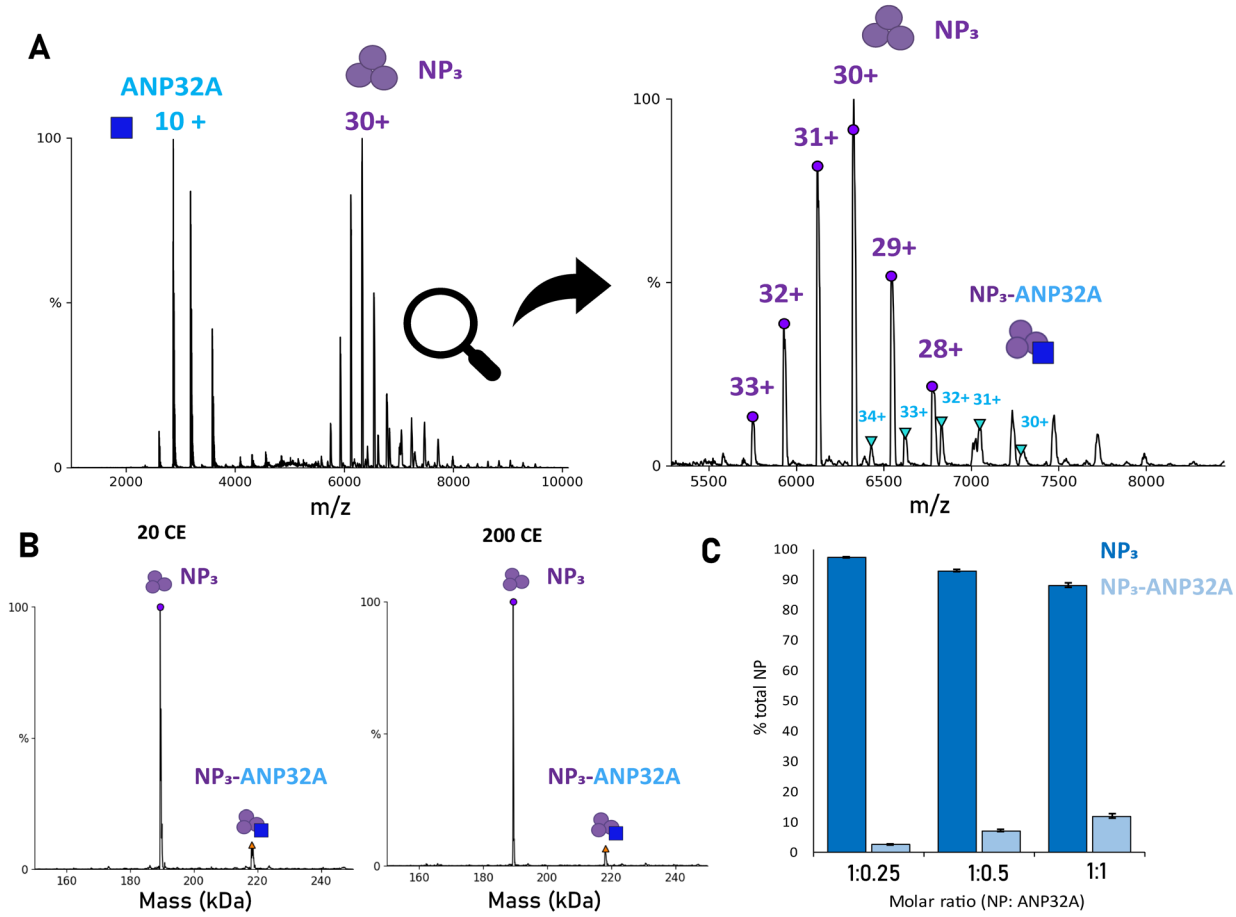

Figure 42: Characterization of the NP ANP32A interaction

(A) 10 µM ANP32A were incubated with 10 µM of NP and subjected to nMS. Shown is a nMS spectrum in the range between 1000 and 10000 *m/z*. The charge states for the strongest signal of every peak series are indicated with masses corresponding to ANP32A (28.65 +/- 0.02 kDa) and NP_3 (189.56 +/- 0.03 kDa) (left). Closeup to the NP peak series where another peak series is visible, corresponding to the mass of a NP_3-ANP32A complex (218.26 +/- 0.12) (right). (B) NP and ANP32 were measured in a 1:1 molar ratio. Different CE were applied on the NP_3-ANP32A complex (indicated in the graph). The complex remains stable at 200 CE. (C) ANP32A and NP were measured with different molar ratios. The nMS measurements with a molar ratio of 1:0.25 and 1:0.5 were measured with low CE, whereas the molar ratio 1:1 was measured with 200 CE. The spectra were deconvoluted and the total NP signal including NP_3 and NP_3-ANP32A was normalized to 100 %. Shown is the percentage of NP_3 and NP_3-ANP32A. The data represents two replicates. The data from this figure was produced by Elisabeth Eckelmann

3.6.2.1 RNA disrupts the interaction between NP and ANP32A

We continued to characterize the NP- ANP32A interaction. In influenza virus, ANP32A works similar to a flexible connector between NP and elongating polymerase. ANP32A binds NP via the flexible LCAR domain and the polymerase via its N-terminal leucin rich repeat (LRR) domain. NP dissociates from ANP32A and binds newly synthesized RNA produced by the polymerase [140]. We wondered if a LASV NP-RNA complex is still able to bind ANP32A. An interaction would be not possible anymore based on the mechanism proposed in influenza virus.

ANP32 was incubated for 20 mins with NP at RT. After that RNA was added to induce formation of NP-RNA complexes, The mixture was subjected to the mass spectrometer. We determined masses for ANP32A (28.65 +/- 0.02 kDa) and NP_1-RNA_1 (66.08 +/- 0.02 kDa) but no interaction between ANP32A and NP was observed. The same conditions were tested without RNA, where we observed ANP32A (28.65 +/- 0.02 kDa), NP_3 (189.56 +/- 0.03 kDa) and NP_3-ANP32A (218.26 +/- 0.12), (compare Figure 43A and B).

The data indicates that NP-ANP32A interaction is not observed when RNA is present. Apparently, NP has a higher affinity to RNA compared to ANP32A. The data might suggest that the binding site between RNA and NP could potentially overlap. This is in line with the role of ANP32A in influenza virus as described above.

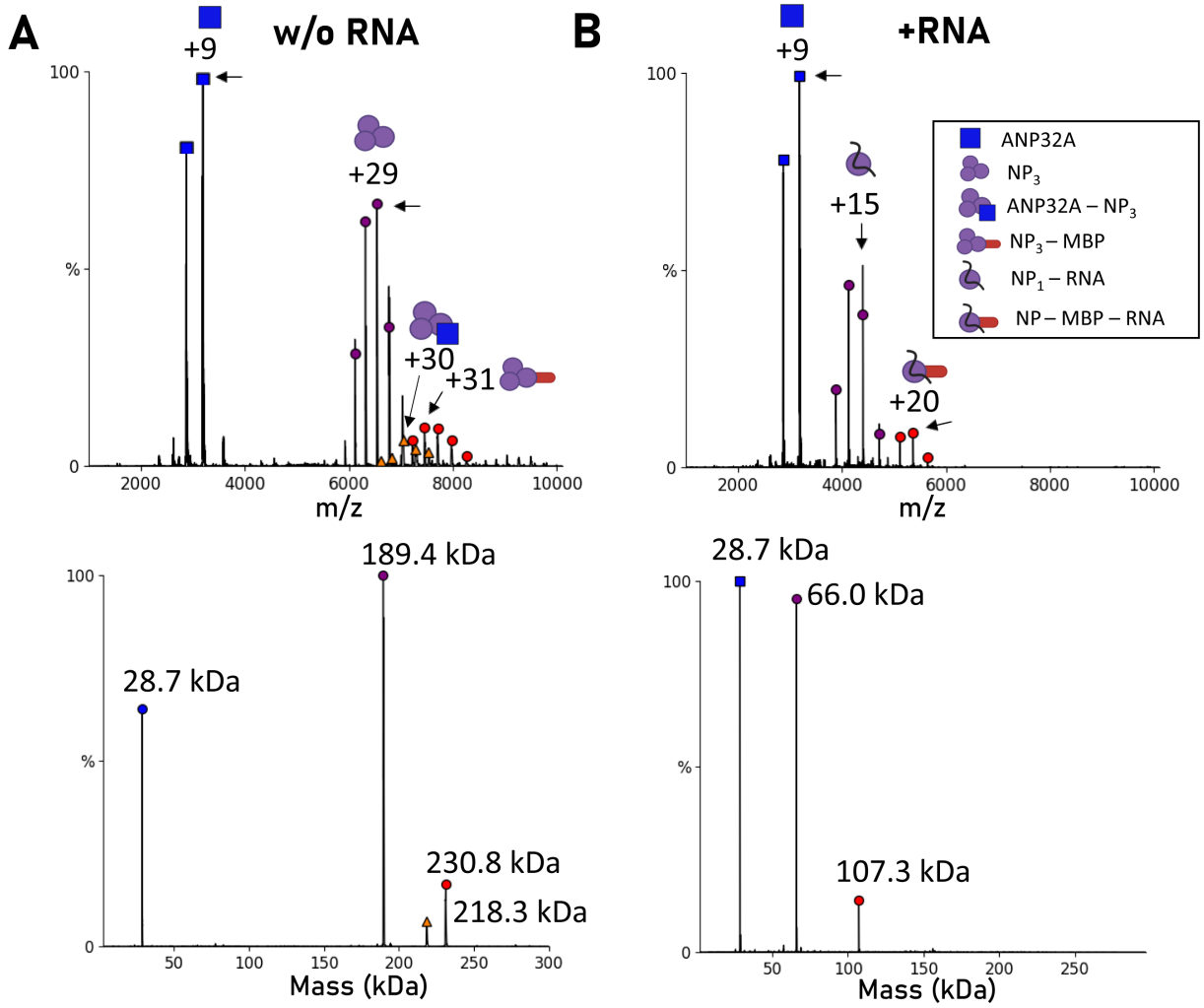

Figure 43: NP-RNA is not interacting with ANP32A

nMS spectra of 2.5 µM ANP32A and 10 µM NP with the corresponding deconvoluted mass. (A) ANP32A and NP were mixed on ice and incubated at RT for 20 min before being introduced into the mass spectrometer. (B) A 9 nt RNA was added to NP in a 2-fold molar excess over NP and incubated for 5 mins at RT. ANP32A was added and measured with nMS. The charge states of highest peaks are indicated. The data from this figure was produced by Elisabeth Eckelmann under my supervision.

3.6.2.2 The LRR domain of ANP32 mediates interaction with NP

In Influenza virus, the interaction between NP and ANP32A is mediated by the C-terminal LCAR domain of ANP32A. We wondered if the binding site is similar to the binding site for LASV NP. For that we recombinantly expressed and purified both the C-terminal LCAR domain and the N-terminal LRR domain of ANP32A. To ensure the solubility of the unstructured LCAR domain, we purified the LCAR domain together with a N-terminal GST tag.

Either ANP32A$_{LRR}$ and ANP32A$_{LCAR}$ were mixed with LASV NP in a 1:2 ratio and incubated for 10 min before subjecting to nMS. For the ANP32A$_{LRR}$ measurement, we observed masses for the single proteins ANP32A$_{LRR}$ (16.83 +/- 0.01 kDa), NP$_3$ (189.56 +/- 0.03 kDa) and a complex formed between ANP32A$_{LRR}$ and NP$_3$ (206.47 +/- 0.18). For the ANP32A$_{LCAR}$ measurement, we observed masses for ANP32A$_{LCAR}$ (39.47 +/- 0.03) and NP$_3$ (189.56 +/- 0.03 kDa), but no other complexes were formed.

Overall, this data suggests that ANP32A binds NP via its LRR domain. This would be a contrary mechanism compared to the binding of ANP32A via its LCAR domain to the influenza virus NP.

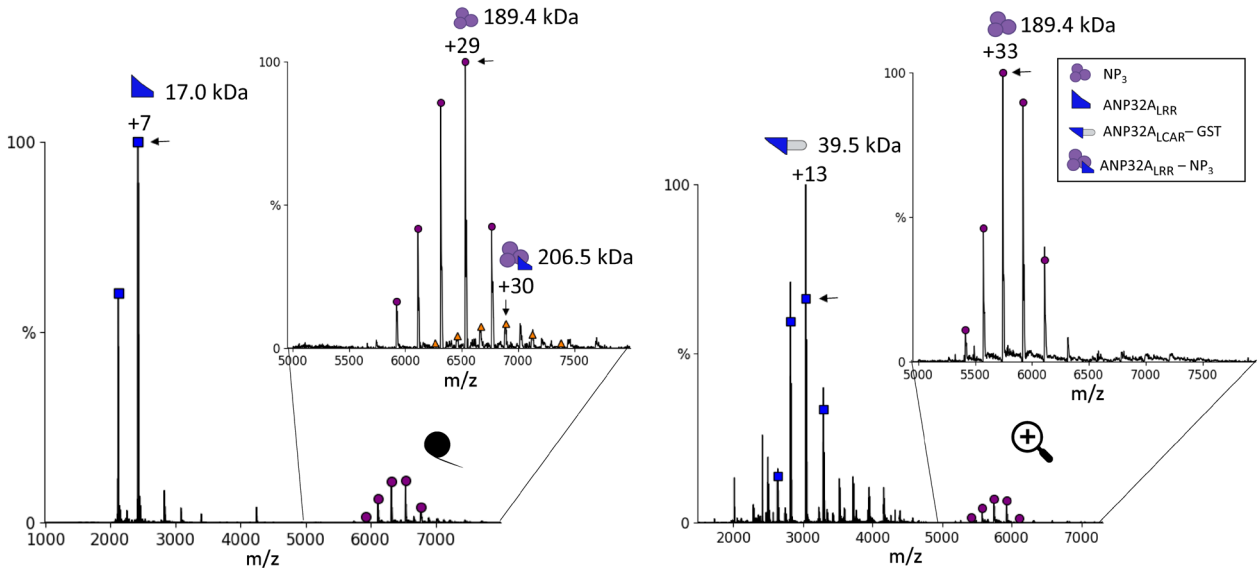

Figure 44: ANP32 binds via the LRR domain to NP

ANP32A$_{LRR}$, ANP32A$_{LCAR}$ and NP were buffer exchanged into 150 µM AmAc, pH 7.5. All proteins were mixed on ice and incubated at RT for 10 min before being introduced into the mass spectrometer. (A) 20 µM ANP32A$_{LRR}$ and 10 µM NP and (B) 10 µM ANP32A$_{LCAR}$ and 5 µM NP. The charge states for the strongest signal of every peak series are indicated. The spectra from this figure are shown as representative of two replicates. The data from this figure was produced by Elisabeth Eckelmann

3.7 Python-based graphical user interface for nMS spectra plotting (MSänger)

Several software tools for analyzing mass spectrometry data are available. Conventional datasets for MS workflows are usually analyzed and processed and not visualized as raw spectra. Contrary to that, nMS spectra are often used directly in publication for visualization. The rationale behind this project was to develop a simple user interface for plotting nMS spectra in high resolution and in different formats ready to use for publications. Several parameters should be easily adjustable including color, size of the axis description and smoothing of the mass spectrum. To have such a simple tool at hand would make nMS easier to access for beginners and a broader range of scientists. The code is packaged into an executive data format and can be distributed and used without having a python environment installed on the operating system.

Python as programming language was used for this software tool. It is most suitable for handling MS based data sets. The overall architecture of the tool is based on the python package *PySimpleGUI* which allows to build an adjustable graphical user interface (GUI). Further packages are *numpy* for data management, *matplotlib* for graph generation and *scipy* for integrating a smoothing function.
Overall, the window of the tool is divided into three parts. On the left side, the user can load files into the tool. In general, a txt file format was chosen as the accepted format which can be exported from most MS instrument software. File outputs from two different instruments can be loaded into the tool. The txt files are handled with the following code which picks out the data from the txt file, normalizes the intensity and plots the spectrum.

```
        if os.path.isfile(file_path):
                        data = np.loadtxt(file_path, skiprows=8)
                        x = data[:, 0]
                        y = data[:, 1]

                        ymax = np.amax(y)
                        y = y / ymax * 100
                        changed_color = values["-LINE-COLOR-"].strip()
                        ax.clear()
                        ax.plot(x, y, linewidth=1, color=changed_color)
                        ax.set_xlabel("m/z")
                        ax.set_ylabel("Intensity (%)")
                        ax.set_ylim(bottom=0)
```

On the right side, the user can specify different graph settings including different colors for different parts of the spectrum and size of the graph label. The following code takes user input to specify which part of the spectrum should be plotted in which color.

```
elif event == "Add X-Range":
    x_range_start = values["-X-RANGE-START-"]
    x_range_end = values["-X-RANGE-END-"]
    x_range_color = values["-X-RANGE-COLOR-"].strip()

    if x_range_start and x_range_end and x_range_color:
      x_ranges_and_colors.append((float(x_range_start), float(x_range_end), x_range_color))
      x_range_list_text = "\n".join(
      [f"Range: {start}-{end}, Color: {color}" for start, end, color in x_ranges_and_colors])
      window["-X-RANGE-LIST-"].update(x_range_list_text)

elif event == "Apply Changes":
  ax.clear()

  for x_start_range, x_end_range, color in x_ranges_and_colors:
       if x_start is not None and x_end is not None:
       mask1 = (data[:, 0] >= max(x_start_range, x_start)) & (data[:, 0] <= min(x_end_range,
       x_end))
       y_segment = y[mask1]
       ax.plot(data[mask1, 0], y_segment, linewidth=1, color=color)
```

The graph is plotted in the middle of the tool which can be adjusted on the x-axis scale and the spectrum can be smoothed by interface elements at the bottom of the graph. All input elements have an interface button which updates the plot and makes the tool interactive. The following code is responsible for specifying the x axis range and the normalization in the chosen x axis range. The second part includes the smoothing function.

```
elif event == "Update X-Axis":
     x_start = float(values["-X-START-"]) if values["-X-START-"] else None
     x_end = float(values["-X-END-"]) if values["-X-END-"] else None

     if data is not None and x_start is not None and x_end is not None:
          x_indices = np.where((data[:, 0] >= x_start) & (data[:, 0] <= x_end))
          if len(x_indices[0]) > 0:
               y_range = data[x_indices, 1]
               y_max_in_range = np.max(y_range)
               y_normalized = y_range / y_max_in_range * 100
               y[x_indices[0]] = y_normalized

          ax.clear()
          if normalized_smoothed_y is not None:
          ax.plot(data[:, 0], normalized_smoothed_y, linewidth=1, color=changed_color)

elif event == "Smooth spectrum!":
       changed_window_length = int(values["-WINDOW_LENGTH-"])
        smoothed_y    =    signal.savgol_filter(y,    window_length=changed_window_length,
        polyorder=3)
       ymax_after_smooth = np.amax(smoothed_y)
       normalized_smoothed_y = smoothed_y / ymax_after_smooth * 100
       ax.clear()
       ax.plot(data[:, 0], normalized_smoothed_y, linewidth=1, color=changed_color)

       ax.set_xlabel("m/z")
       ax.set_ylabel("Intensity (%)")
       ax.set_ylim(bottom=0)
       ax.set_xlim(left=x_start, right=x_end)
       ax.spines['right'].set_visible(False)
       ax.spines['top'].set visible(False)
```

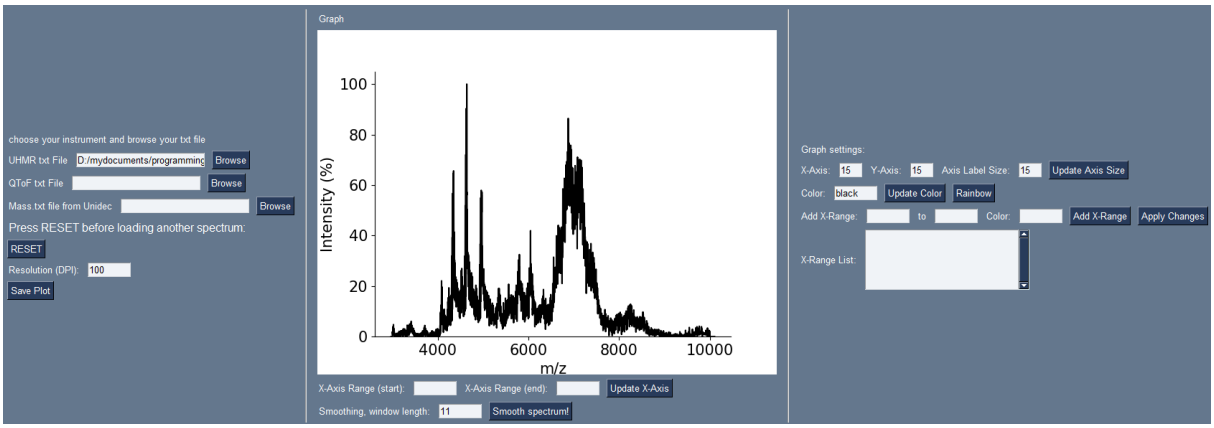

Figure 45: GUI for plotting nMS spectra

The structure of the tool is divided into three parts. On the left side, the user can choose a txt file to be loaded into the tool. It can be chosen between UHMR and QToF txt file outputs. There is the option to save the plot in different data formats and resolution. In the middle, the graph is plotted, and the x axis range can be chosen. Additionally, there is the option to smooth the spectrum. On the right side, several settings of the graph can be adjusted, including color (at different x ranges) and label.

The tool provides a platform for plotting nMS spectra with adjustable settings. It can be easily extended to have additional functions included (e.g., plotting multiple datasets, editable line width or font type).

4 Discussion

The development of antiviral strategies requires knowledge about the cell biology of viruses. In order to inhibit the viral replication in host cells, researchers aim to identify specific points or mechanisms during the viral life cycle to be targeted. Typical targets are specific enzymatic functions and ligand-binding sites, but also protein-protein interaction sites can be promising targets [141]. The goal of this work was the identification and characterization of LASV complexes to strengthen our knowledge of how the virus functions and to potentially identify targets for antiviral strategies.

4.1 Production of LASV proteins and establishment for nMS

The strategy for the investigation of viral complexes was to reconstruct the complexes from the individual components. The advantage of generating those components individually makes production much easier as complete complexes are often difficult to isolate in large quantities. Another advantage is to be able to analyze all critical factors needed for the formation of complexes. Viral proteins and protein complexes have been recombinantly produced for decades. The combination of producing viral proteins recombinantly and using structural MS to characterize them is a relatively new approach [142]. As many viral proteins are often difficult to produce in high amounts, MS based methods are most suitable as they consume only 10-40 pmol of sample per measurement.

However, a reasonable amount of protein must be available for an in-depth analysis of the samples. Part of this work focused on increasing the amount of LASV proteins and establishing the analysis by nMS methods. Our focus was to improve the workflow at two points: (i) Increasing the production of the recombinant proteins and (ii) finding suitable MS-compatible solutions as high amounts of sample loss were noticed when proteins were transferred to the conventional nMS solutions.

4.1.1 Establishing a nMS workflow for NP by increasing the protein production

Concerning the first point, the expression and purification protocol of NP was improved. At the beginning of this work, only low yields of NP protein were achieved with the previously established protocol. A key step was to decrease the amount of IPTG to induce expression in *E. coli*. A lower amount of IPTG increased the expression of NP significantly. A possible explanation is that a lower amount of IPTG could reduce the stress and metabolic burden of the *E. coli* cells. This is in line with a study investigating different IPTG concentrations and different temperatures on the production of an example protein. IPTG concentrations between 0.05 and 0.1 mM resulted in the highest protein yield at all investigated temperatures [143].
Another improvement was an adjustment in the affinity chromatography protocol. High amounts of aggregated protein were encountered on the affinity resin after cleavage of the MBP-tag. In general, MBP-tags increase solubility and expression of proteins that tend to aggregate. Protein aggregation is influenced by several factors including the properties of the individual proteins [144]. Parameters that can be easily adjusted to decrease protein aggregation *in vitro* are a lower protein concentration and higher NaCl concentration. Both parameters were adjusted at the step when the MBP-tag was cleaved from the protein. The NP-MBP fusion protein was eluted from the resin in a higher volume to lower the protein concentration. Additionally, the elution contained a higher amount of NaCl to stabilize the protein after cleavage. Adjustment of the buffer conditions during MBP cleavage resulted in a higher yield of soluble NP.

4.1.2 Establishing nMS of the L protein

The measurement of the L protein by nMS turned out to be more challenging compared to the other viral proteins. The first attempts to exchange the purification buffer to AmAc solution using desalting columns resulted in a total loss of protein. Additionally, fast protein liquid chromatography (FPLC) coupled SEC was tested. However, it resulted in aggregation of the L protein as well. The newest state of the art buffer exchange method for nMS is a rapid online buffer exchange system where

samples are injected onto a short size-exclusion chromatography column directly coupled to the mass spectrometer [145]. However, considering the limited success using traditional SEC devices it is questionable whether this system works in combination with instable samples like the L protein.

The aggregation of the L protein could be a result of a buffer exchange that is too fast. The buffer exchange with centrifugal filter units is a slower process including several rounds of adding new buffer. The longer process could be beneficial for the L protein to slowly equilibrate to the new buffer. However, protein aggregation with centrifugal filter units was observable as well. Yet, significantly better results were observed using Amicon filter units. The centrifugal filter units are composed either of regenerated cellulose (Amicon) or polyethersulfone (Vivaspin) which could explain the different results. Additionally, Amicon filter units with a MWCO of 100 kDa were found unsuitable for the L protein, as the L protein was not retained by the filter, despite the size of ca. 250 kDa. The L proteins highly flexible nature may help the protein to run through the filter. On the other hand, the 30 kDa MWCO filter units were able to retain the protein and resulted in the highest yield of L protein after buffer exchange.

Another important factor was a screening with nMS-compatible solutions to retain high amounts of soluble L protein after buffer exchange. The AmAc solution (150 mM AmAc, pH 7.5) that was used for the other LASV proteins resulted in high amounts of protein aggregation and no measurable protein concentration. The LASV L protein is most stable in high salt concentrations but nonvolatile salts (e.g. NaCl, KCl) are not suitable for nMS [45, 97]. Therefore, the concentrations of AmAc were increased to work in a milieu of similar ionic strength. Additionally, it was shown that low concentrations of nonionic detergent Tween-20 are nMS compatible and can increase the recovery of soluble protein after buffer exchange [146]. The recovery was significantly increased by using 500 mM AmAc and 0.01 % Tween-20 for the C-terminally tagged L protein. However, it was only possible to recover high amounts of the internally tagged L protein by using 1 M AmAc and 0.01 % Tween-20. The lower stability of the internally tagged L protein matches our observation of reduced expression rates compared to the C-terminally tagged version.

It was possible to record high quality spectra for the L protein with this reproducible workflow. However, the use of high AmAc concentrations and detergent in the L protein buffer could influence complex formation or binding affinities. Similar to NaCl or KCl, high concentrations of AmAc have been implicated in affecting binding affinities in context of nMS [147]. The same applies to Tween-20, which could negatively influence hydrophobic interactions. Therefore, known interactions were tested with the L protein and used as positive controls.

LASV L and Z are known to interact, Z inhibits the polymerase function supposedly by blocking the formation of an elongation conformation of L [33]. The interaction is mainly based on a hydrophobic interaction site [33, 45, 47-49]. This known interaction is useful to test the influence of Tween-20 on hydrophobic interactions.

It was possible to reconstruct the L-Z complex and binding affinities were determined similar to published data which were produced by bio-layer interferometry. The determined K_D of 0.2 µM +/- 0.1 µM by nMS is around 100 fold higher compared to the published data produced by bio-layer interferometry (K_D of 1 nM) [133]. However, in this area of high affinity-binding, the determination of the K_D by nMS is error prone, observable by the high standard deviation. Similar to the conditions in this work, the K_D of the L-Z interaction determined by Xu et al. was measured in 500 mM NaCl and 0.05 % Tween-20 which could also have an influence on the binding affinity compared to cellular conditions [133]. Notably, a K_D was determined additionally for Machupo virus (12 nM) [133] and Junin virus (393 nM) [48], indicating a broader range of affinities between L and Z within different arenaviruses.

Additionally, L-RNA interaction was characterized to test if high concentrations of AmAc influence polar interaction. In general, L showed a high binding affinity to different RNA substrates. It was possible to identify L-RNA complexes with multiple RNA substrates bound to L protein. However, LASV promoter sequences were exclusively used to reconstruct L-RNA complexes. Whether L also binds RNA nonspecifically remains to be seen. Notably, when NP was present together with L protein in some of our assays, no trimer dissociation and RNA binding of NP was observed. L could bind RNA with a higher affinity compared to NP and not enough free RNA is present anymore to be bound by the NP. Additionally, it is possible that high AmAc concentrations reduce the binding affinity to NP and prevent RNA binding. However, NP-RNA complexes did not dissociate when L (present in high

AmAc concentration) was added to the NP-RNA complexes. To circumvent this, capillaries with smaller diameters can be used to analyze the proteins in conventional buffers including NaCl. Smaller droplets resulted from these submicron capillaries are reducing the chance of salt ions present in the same droplet as the protein analyte [148]. However, submicron capillaries are prone to clog during measurement because of the smaller diameter and are difficult to reproduce. Another method to measure proteins in non-volatile buffers is to add high concentration of AmAc to proteins present in conventional buffer (e.g., Tris-HCl) which resulted in usable spectra for the investigated alcohol dehydrogenase (1.6 M AmAc final concentration) or nucleosome core particle (400 mM AmAc final concentration) [149]. However, no NaCl was present in the buffer, which also normally causes suppression of analyte ionization. Notably, both protein complexes were stable at high AmAc concentration, suggesting that high AmAc concentration are not preventing non-covalent protein-protein or protein-DNA interaction [149]. High AmAc concentrations (usually around 1 M) are used to characterize overexpressed proteins directly from cell lysate. Cells containing overexpressed proteins are lysed in AmAc solutions and directly subjected to nMS [150]. However, this is limited to proteins that are very abundant and expressed with high yields. Protein-ligand or protein-protein interaction are probably difficult to characterize when measuring the components together with the lysate. Also with this technique, non-covalent interactions between proteins were shown to be stable at high AmAc concentration [150]. There are furthermore alternative solutions that are compatible with nMS e.g., ammonium bicarbonate or ammonium formate and can be tested for the L protein. However, AmAc is commonly used in the field with the best properties [99].

Overall, it was possible to establish L-RNA measurements with nMS. A high AmAc concentration does not prevent polar interactions of L with RNA. However, it cannot be ruled out that high AmAc concentrations can have an influence on polar interactions to a certain extent.

4.2 NP trimer dissociates in presence of RNA and forms higher order NP-RNA assemblies.

NP is known to form trimers in the absence of RNA and during recombinant purification. The atomic structure has been previously solved [56]. The trimeric NP can be isolated from RNP containing cells alongside full RNP complexes demonstrating existence of this complex in infected cells [62, 136]. It was speculated that the trimer serves as a storage form for NP in cells. The ability of NP to trimerize was found to be essential for viral genome replication and transcription but was dispensable for interferon antagonism and NP recruitment into budding virions [62]. Many viruses use specific complexes (e.g., NP-phosphoprotein complex in non-segmented RNA viruses) to prevent unspecific RNA binding to the nucleoprotein. It was therefore suggested that the trimeric storage complex of LASV NP acts in a similar way [55, 151]. Structurally, the trimeric ring conformation of arenavirus NP is unable to bind RNA. To enable this, the C-terminal domain must first rotate away from the RNA-binding pocket allowing one helix to become partly unstructured (α5) and a second helix (α6) to shift away from the RNA binding pocket [55, 152]. To induce this conformational change, it was speculated that specific (host-) factors are needed [55]. In our experiments, it was demonstrated that short RNAs are sufficient to induce this conformational change and trigger NP trimer dissociation. This was demonstrated by nMS and validated by negative stain electron microscopy and mass photometry which were performed in physiologically more relevant buffers. Yet, how does NP primarily associate with viral RNA instead of host RNA? No sequence specificity was observed as NP bound to all random RNA sequences used in our experiments, which is in line with other groups reporting randomly bound RNAs to LASV NP [55, 57]. However, it cannot be ruled out that the LASV RNA genome contains some specific structural and sequence motifs (packaging signals) or additional modifications that could lead to a lower threshold for RNA binding. These signals would have to be different in the genomic RNA compared to viral mRNA, which is usually not encapsidated. It also remains unclear whether and how specific recruitment of all genome segments is controlled in bunyaviruses as specific RNA secondary structures, as present in influenza viruses [153], seem to be missing.

Overall, it is likely that another (host-) factor is needed to block unspecific RNA binding to NP, potentially by further stabilizing the trimeric NP. For Rift Valley fever virus, a non-selective recruitment of viral genome segments into budding virions was found to be the most likely scenario [154-156]. Another factor favoring viral RNA binding could be the local RNA concentration. In our experiments, relatively high RNA concentrations (1:1 or 1:2 molar ratio) were used and the dissociation and assembly process was observed in real time, pointing to rather slow RNA binding kinetics. Several scenarios are conceivable in support of the idea of local viral RNA concentration as a key determinant for preferred and efficient packaging of viral genomic RNA:

(i) Direct interaction of NP with the L protein for close proximity to nascent viral RNA. NP is a critical component of the arenavirus replication machinery. A direct interaction of NP and L was suggested based on co-immunoprecipitation studies, although an interaction mediated by RNA or host factors could not be ruled out in this case [135, 157].
(ii) Indirect interaction of NP with L during viral genome replication via a host protein. This has been observed for influenza virus polymerase complex, where the host acidic nuclear protein (ANP32) enables L-L interaction and also binds to NP to support encapsidation of nascent viral RNA and re-encapsidation of the template [139, 140].
(iii) Formation of local microenvironments for viral genome replication and transcription within the cell. Bunyavirus NPs have been detected in non-membranous compartments such as processing bodies and stress granules in cells [158-161]. Other studies reported NP containing cytosolic puncta forming upon arenavirus infection associated with cellular membranes. These puncta were concluded to contain full-length genomic and antigenomic RNAs along with proteins involved in cellular mRNA metabolism reminiscent of classical replication-transcription complexes [162]. Interestingly, phosphorylation of lymphocytic choriomeningitis virus (LCMV) NP at position T206 was found to be critical for the formation of these cytosolic puncta [163]. However, there is no evidence that the L protein is indeed present in these puncta and the presence of these local puncta would not provide an explanation on why mRNAs are not encapsidated. It therefore remains unclear what

exactly determines viral RNA packaging by NP in addition to the presence of the RNA molecule itself.

In our experiments, it was possible to observe NP trimer dissociation into RNA-bound NP monomers, which then assembled to higher-order NP-RNA structures. NP-RNA higher-order complex formation seemed to require a certain RNA length (12 nt), indicating an additional stabilizing effect of the RNA or even a secondary RNA binding site in NP as observed for other segmented negative-strand RNA viruses [164]. Moreover, only 2 NPs per RNA with RNA oligos of 18, 25 and 27 nt length could be identified. This demonstrates that a specific RNA length is needed and possibly a secondary binding site for multiple NP binding to one RNA. It was not possible to observe NP_3-RNA_1 complexes with the 27 nt RNA after initial disassembly, which would be in theory enough to cover 3 binding sites as NP_2RNA_1 complexes with the 18 nt RNA were observed. This further supports a possible secondary RNA binding site required for assembly.

Interestingly, during the NP trimer disassembly process a complex of one RNA molecule bound to the NP trimer (NP_3-RNA_1) was observed although according to the interpretation of the available structural data, RNA binding is not possible in the trimeric ring conformation [56]. It is likely that the observed NP_3-RNA_1 state is therefore a disassembly intermediate, or RNA is bound via a potential secondary binding site. Mechanistically, the process can be explained by two hypotheses: The trimer needs to disassemble, allowing the canonical RNA binding sites to be open. This can either be achieved by (i) RNA binding to a secondary binding site destabilizing the trimer or (ii) an “open” trimer conformation, which allows one RNA molecule to be bound by the canonical binding site. The open conformation is less stable and therefore the trimer quickly disassembles into monomers which immediately bind RNA. This indicates overall a cooperative dissociation mechanism. Notably, the influenza virus NP is composed of two RNA-binding sites by two patches of multiple arginine residues. The RNA is at one binding site exposed and could favor the interaction of the viral polymerase to bind on the template strand during replication [165]. Further experiments are needed to characterize the initial LASV NP trimer dissociation and identify a potential secondary binding site.

4.3 Characterization of the NP-Z interaction

4.3.1 NP binds directly to Z and is independent of its oligomeric state

It has long been known that arenavirus NP and Z are acting together, based on cellular experiments, such as VLP assays, co-immunoprecipitation and colocalization studies [35, 50, 166, 167]. Here, NP and Z interact directly and independent of any host factor with each other. Importantly, monomeric NP in presence or absence of RNA has a substantially higher affinity for Z compared to the NP_3 storage form, which is not recruited to the membrane.

4.3.2 Mapping the interaction site between NP and Z

The NP-Z interaction site was mapped with an integrative structural approach of nMS and Alphafold Multimer-powered complex prediction. The model was further validated with HDX-MS which was performed in a physiologically more relevant buffer. Our data point towards Z binding to the C-terminal domain of NP likely involving the side chains of R16 and R74 of Z. Previous co-immunoprecipitation studies already suggested the C-terminal domain of NP to be important for the interaction with Z protein, either directly or mediated by cellular factors [53]. Several conserved Z residues, including R16 and T73, were also demonstrated to be important for inhibition of viral RNA synthesis and virus-like particle (VLP) infectivity although R16 and T73 are not involved in the binding of Z to the L protein [133, 168, 169]. This points towards the importance of R16 and T73 conservation for Z localization and/or proper folding. However, a difference in the protein folding was not detectable comparing Z R16A to the wild-type Z *in vitro*.
It was hypothesized that Z residues R16, T73 and others were involved in RNP recruitment via binding to NP and/or viral RNA [168]. Our experimental data and the proposed model support this hypothesis as R16 and T73 located in the interface between NP and Z, while an interaction between Z and RNA was not detectable *in vitro* as observed for the matrix protein of other negative strand RNA viruses [169]. Upon RNA packaging, RNPs are recruited into nascent virions and Z has been shown

to be critical for budding of most arenaviruses [78]. While Z alone is sufficient to drive the formation of VLPs, the process is significantly enhanced when NP is present [50]. For Tacaribe virus, the so-called late-domain motifs ASAP and YxxL of Z have an influence on the NP incorporation, with the Y56 in the YxxL motif being critical. These data suggest a role of the late-domain motif YxxL in the NP-Z interaction [170]. Interestingly, although the YxxL motif is also present in the LASV Z, the Y48 of the YxxL motif is not in the proposed NP-Z interface in our model but rather on the opposite side of the Z molecule. However, the proposed location of this motif in our model is in agreement with another study reporting the binding of Z to a host protein required for budding via the YxxL motif [171].
The data shown here support the concept, that the NP-Z interaction is important for the recruitment of RNPs to the plasma membrane and incorporation into nascent virions. RNA-bound NP is able to bind Z as would be expected for RNP recruitment.

4.3.3 A possible role of NP and Z during assembly and recruitment

Arenavirus cell entry is mediated by the glycoprotein complex. After receptor binding, virions enter the host cell via the endosomal pathway [37, 41]. Our data show that the interaction between NP and Z is highly pH-dependent with a decreased NP-Z interaction at pH values lower than pH 7.5. Therefore, it is conceivable that RNP-Z dissociation triggered by a lower pH in the endosome leads to the release of RNPs into the cytosol after viral membrane fusion (see model Figure 46).

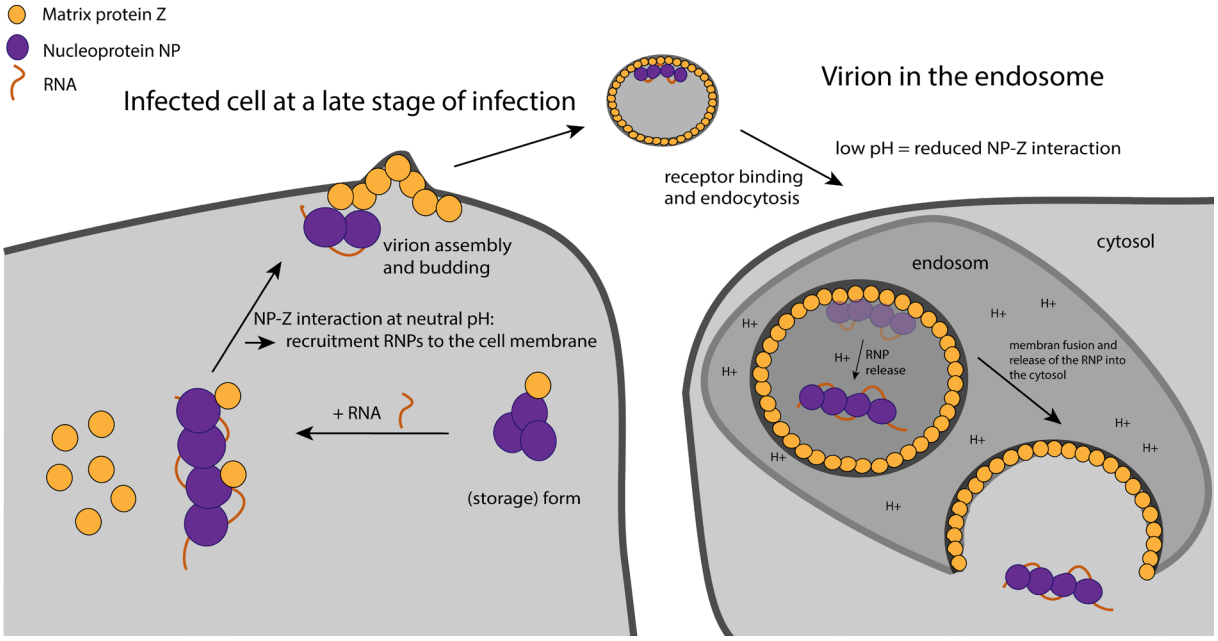

Figure 46: Proposed model of the NP-Z interaction during the viral life cycle.

At a late stage of infection, the cellular concentration of Z is high favoring NP-Z interaction. This leads to recruitment of the RNPs to the cell membrane by myristoylated Z for virion assembly (left). During virus entry into the host cell via the endosomal pathway, the lower pH leads to dissociation of NP from Z, which coats the inner virion membrane. This leads to the release of the RNPs into the cytosol.

Unlike in influenza viruses, arenaviruses do not have proton channels in their lipid envelope that actively acidify the virion upon endocytosis. However, it has recently been shown for LASV that protons passively cross the viral membrane which leads to acidification of virus interior prior to membrane fusion. The membrane permeability is induced by a conformational change in the GP triggered by endosomal acidification and potential interaction of GP with the endosomal membrane [130, 131]. Similar observations have been reported for Ebola virus, were acidification of virions leads to viral matrix reorganization [172]. Besides the pH-dependent NP-Z interaction, it was observed that Z starts to oligomerize between pH 5 and 5.5. It is possible that stronger Z-Z interaction at low pH, as indicated by oligomerization, further reduces availability of Z for the low affinity Z-NP interaction. This is in line with several studies that report pH as a critical factor for matrix protein oligomerization and protein conformation [132, 173-176]. Matching our observations, a lower pH at around 5.4 also induced weakening of influenza virus matrix protein M1-vRNP interactions [177].

Our findings support a model, in which viral RNA is necessary and sufficient to transform trimeric NP into monomeric RNA-bound, RNP-like NP assemblies. Z then binds to the RNP at neutral pH in the cytosol or a cellular microenvironment via direct interaction with NP. This interaction facilitates targeted transport of the RNP to the plasma membrane at high cellular concentration of Z protein during late stages of the viral “life” cycle and subsequent virion budding. Upon virus entry into a host cell, the RNPs are released into the cytosol from the virion matrix triggered by the low pH in the endosome and following membrane fusion. Subsequently viral genome replication and transcription take place.

4.4 Reconstruction of the RNP complex

A fundamental question about the composition of the Arenavirus RNP complex is how NP interacts with L protein. Interaction could be either mediated indirectly by RNA or directly by physical interaction between NP and L. The aim was to reconstruct the RNP complex by nMS by combining NP, L and different RNAs.

Initial experiments aimed to identify a potential direct interaction between NP and L. Experiments were performed incubating NP and L and screened for a complex formation between both proteins. However, it was not possible to observe direct interaction under the used experimental conditions. NP did not directly interact with L either in a trimeric or monomeric conformation. It cannot be ruled out that a direct interaction is too transient to detect complex formation under the used measurement conditions, (see 4.1.2 about buffer composition of the L protein). However, an interaction between NP and L needs to be quite robust and should be clearly detectable considering the importance of the RNP for genome replication and transcription [34]. On the other hand, a specific conformation of the L protein could be essential for a direct interaction with NP. L undergoes several conformational changes when conducting different functions [33]. Additionally, binding of Z to the L protein specifically blocks the L protein in a replication-inactive conformation making it convincible that a direct interaction between NP and L could be conformation-specific [133].

Promoter sequences have been identified as a driver for conformational change in the L protein [33]. These promoter sequences were included in the measurements to

test whether conformational changes could induce direct interaction between NP and L. Similar to L, RNA induces conformational change of NP which could be needed for a direct interaction between NP and L. However, no signs were observed that RNA induces a direct interaction between NP and L in our preliminary experiments.

Furthermore, RNA could not only induce a direct interaction between NP and L but also mediate an indirect interaction by connecting NP and L. To connect L and NP by RNA several things need to be considered: (i) A certain RNA length is needed that is long enough to be bound both by L and NP. (ii) RNA can bind to L at several sites unspecifically, which could make it difficult for NP to access RNA bound by L (iii) NP trimer dissociation needs to be induced first before RNA binding is possible.

To tackle these points, an RNA was designed that has the first 7 nts of the LASV 3' promoter sequence as which binds specifically in the 3' binding site of L. Additionally, the RNA contains 19 random nts that could stick out of the L protein and can be bound by NP. In our first experiments, NP was incubated first with the RNA to induce trimer dissociation and the L protein was added afterwards. However, no interaction was observable with L, the RNA was bound to both proteins NP and L. Incubating the RNA first with L and adding NP afterwards resulted in no trimer dissociation and RNA binding of NP. It is possible that more RNA needs to be accessible for NP to induce trimer dissociation, a higher RNA concentration or a longer RNA substrate could be used for this.

Overall, these initial experiments laid the foundation for further analysis to characterize protein interaction within the RNP. The experiments demonstrated that NP is not directly interacting with L. It is possible that RNA bridges interaction between NP and L, which has been previously reported by co-immunoprecipitation experiments, where the presence of viral RNA promotes efficiently NP-L interaction [134]. It is therefore likely that viral RNA bridges the interaction between NP and L. However, an interaction mediated via host factors cannot be ruled out from this experiment. Another study reported efficient NP-L interaction for LASV, LCMV and Mopeia virus via co-immunoprecipitation also in absence of viral RNA [135]. The contrary results might be explained by the different use of expression promoters for overexpression in both studies which can result in a non-physiological intracellular concentration of NP and L and unspecific binding [134]. For Machupo virus,

interaction between NP and L has been shown by isolating L from RNP expressing cells. RNPs were observed by negative staining microscopy in a large variety of L-NP formations including L in complex with a single NP suggesting direct interaction [136]. However, as viral RNA is also present in this case, interaction between NP and L via RNA cannot be ruled out.

Overall, all these experiments have in common, that RNP formation happened in a cellular context prior to the isolation and mostly involving viral RNA. This is contrary to the nMS study of this work, involving the investigation of the isolated components NP, L and short (viral) RNAs without the cellular context. This could suggest that RNP formation is only formed in cells by the help of an unknown (host) factor that mediates interaction between NP and L.

4.5 nMS as a platform to detect direct interactions between LASV proteins and host factors

The viral replication cycle depends largely on host–pathogen interactions. The identification of interaction partners is critical for our understanding of how the virus is functioning in cells and for the identification of targets for potential anti-viral drug development. Many methods for the identification of protein interactions fail to distinguish between direct or indirect protein-protein interaction (e.g. Co-immunoprecipitation, colocalization). Analyzing recombinant proteins by nMS is ideal to detect direct protein interactions or screen for factors needed for interaction.

The establishment of nMS with the LASV proteins can be used as a platform to analyze possible direct interaction between the LASV proteins and host factors.

We tested ANP32A as a potential host factor for LASV as it was previously identified as part of the influenza virus replicase complex. The influenza virus replicase complex is composed of a heterotrimer including an asymmetric dimer of the polymerase bound to the LRR domain of ANP32A bridging both polymerases [139].

We hypothesized whether ANP32A has a similar function in LASV by binding directly to L. We tested both L protein variants at different molar ratios with ANP32A but were not able to detect any direct interaction. As observed in influenza virus, ANP32

appears to not have the ability to form a complex with two polymerases. However, further experiments with other techniques are needed to validate this result.

4.5.1 ANP32A directly interacts with NP via the LRR domain

We continued to characterize a potential direct interaction between LASV NP and ANP32A. While bound to the influenza virus polymerase, the C-terminal LCAR of ANP32 is responsible for NP recruitment to nascent RNA by putting NP into proximity to newly produced RNA from the polymerase [140]. Although we did not detect an interaction between the LASV L protein and ANP32A, an interaction between NP and ANP32A could be possible and important for an alternative function. Notably ANP32B was identified to bind the matrix protein of different paramyxoviruses indicating also other roles besides influenza virus [178, 179]. Furthermore, ANP32A and B have been shown to be critical for the transport of mRNA transcripts from the nucleus to the cytoplasm of the human immunodeficiency virus 1 (HIV-1) by interacting with the HIV-1 Rev protein via the LCAR domain of ANP32A [180]. It is not an uncommon feature for many viruses to target the same host factor but with different viral proteins. An example is the interferon pathway which is downregulated by many viruses (interferon antagonism). Direct binding-dependent inhibition of proteins like the signal transducer and activator of transcription (STAT) 1 or 2 has been demonstrated for numerous viruses including p protein of rabies virus or non-structural protein 5 of yellow fever virus [181].

In our experiments we were able to demonstrate that NP_3 directly interacts with ANP32A via its LRR domain. This is contrary to the binding mechanism in influenza virus, as the LRR domain seems to be responsible for LASV NP ANP32A interaction and not the LCAR domain [140]. It is conceivable that the unstructured LCAR domain might recruit other cellular factors to NP which may be needed for replication. The RNA binding pocket of NP is potentially the area where ANP32A binds as it was shown for influenza virus [140]. Similar to LASV, the influenza NP forms trimers in absence of RNA and assembles in higher-order structures in presence of RNA but consist of two binding pockets and lacks a 3’-5’ exoribonuclease activity [164]. Both identified RNA binding pockets of the influenza NP were shown to be interacting with

ANP32A [140]. A second binding pocket has not yet been identified for the LASV NP. Similar to influenza virus, RNA replaces the interaction between NP and ANP32. It was additionally suggested that multiple influenza virus NPs are binding to a single ANP32A as observed in our experiments [140]. However, the exact function of how ANP32A acts as host factor for LASV is completely unclear and needs to be further validated.

5 Outlook

Our NP-RNA assembly experiments gave indications about a possible RNA secondary binding site. An intermediate of the NP_3 bound to one RNA was detectable in all NP-RNA experiments. However, there is no clear experimental proof whether a secondary binding site exists. Further experiments could aim to identify a potential secondary binding site. A possible procedure would be the production of NP mutant where RNA binding at the primary binding site is impaired. Measurements together with RNA and the NP mutant would clarify if RNA binding is still possible via the RNA secondary binding site. Additionally, it would be interesting whether trimer dissociation is possible or if RNA binding at the primary binding site is needed for this process. If RNA binding is observed via the secondary binding site, HDX-MS could be used to identify the location of the secondary binding site. Based on the location, a mutational analysis could be conducted to identify the importance of the secondary binding site. Secondary binding site mutants can be screened in cellular assays to observe how critical the mutations potentially influence e.g., viral replication. Furthermore, it would be interesting to test if the mutant still forms the trimeric ring structure and dissociates in the presence of RNA. As a complementary method, Cryo-EM could be used to receive high resolution structural data of NP-RNA assemblies. However, the assemblies need to have a specific size that is suitable in combination with Cryo-EM.

An open question remains, how components of the RNP exactly interact with each other. This study demonstrated that NP does not directly interact with L, further experiments with RNA are needed to prove whether RNA mediates interaction between NP and L. Our experiments laid the foundation for further analysis investigating different RNA substrates. Our procedure would include the use of a partially LASV specific promoter RNA in combination with a random sequence that does not form any RNA secondary structures. The LASV specific promoter sequence would be bound by L and the random sequence would be the part where NP can interact. Our approach allows us to screen different RNA substrates that can be screened on their ability to mediate interaction between NP and L.

6 Materials and methods

6.1 Materials

6.1.1 *E. coli* strains

Table 5: List of *E. coli* strains

DH5α	[Genotype: fhuA2 Δ(argF- lacZ)U169 phoA glnV44 Φ80 gyrA96 recA1 relA1 endA1 thi-1 hsdR17 Δ(lacZ)M15], New England Biolabs
BL21-Gold(DE3)	[Genotyp: F- ompT hsdS(r_B–m_B–) dcm+ Tetr galλ(DE3) endA Hte], Agilent Technologies
Rosetta™2(DE3)	[Genotyp: F- ompT hsdSB(r_B- m_B-) gal dcm pRARE2 (CamR)], Merck
DH10EMBacY	[F-mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ-rpsL nupG/pMON14272($Kana^R$)/pMON7124(Tet^R)], EMBL Grenoble

6.1.2 Insect cells

- Sf21 (Sf900-II), Thermo Scientific
- High Five (Express Five), Thermo Scientific

6.1.3 Plasmids

Table 6: List of plasmids

Plasmid	Origin
popinM NP (LASV AV)	Dominik Vogel
popinJ Z (LASV AV)	Dominik Vogel
pFastBac L-C-strep (LASV Ba289)	Dominik Vogel

pFastBac L-407-strep (LASV Ba289)	Dominik Vogel
popinJ Z_F30A (LASV AV)	Dominik Vogel
popinJ Z_W35A (LASV AV)	Dominik Vogel
popinM NP_R52A (LASV AV)	This work
popinM NP_D483A (LASV AV)	This work
popinJ Z_F36A (LASV AV)	This work
popinJ Z_W35A_F36A (LASV AV)	This work
popinJ Z_S33G (LASV AV)	This work
popinJ Z_K32A (LASV AV)	This work
popinJ Z_R4A (LASV AV)	This work
popinJ Z_R16A (LASV AV)	This work
popinJ Z_N62A (LASV AV)	This work
popinJ Z_S61G (LASV AV)	This work
popinJ Z_S59G (LASV AV)	This work
popinJ Z_R74A (LASV AV)	This work
popinJ Z_T82G (LASV AV)	This work
popinJ, popinM	Oxford Expression Facility

6.1.4 DNA oligonucleotides

DNA oligonucleotides to perform DNA amplification were ordered from IDT. A list can be found in the appendix.

6.1.5 RNA oligonucleotide

DNA oligonucleotides to perform DNA amplification were ordered from biomers. A list can be found in the appendix.

6.1.6 Media

6.1.6.1 Media for *E. coli* cultivation

Table 7: Media for *E. coli*

Description	Ingredients
LB	Peptone (10 g/l), yeast extract (5 g/l), NaCl (10 g/l)

LB-Agar	Peptone (10 g/l), yeast extract (5 g/l), NaCl (10 g/l), agarose (20 g/L)
SOC	Peptone (20 g/l), yeast extract (5 g/l), 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$
TB	Peptone (12 g/l), yeast extract (24 g/l), glycerol (4 ml/l), 0.17 M KH_2PO_4, 0.72 M K_2HPO_4,

6.1.6.2 Media for insect cell cultivation

Table 8: Media for insect cells

Description	Additives
Express Five	penicilline, streptomycin
Sf900-II	

6.1.7 Buffer and solutions

6.1.7.1 Buffer for electrophoresis

Table 9: Buffer for electrophoresis

Description	Ingredients
TAE (10x)	2 M Tris, 50 mM Na_2EDTA pH 8
Gel red loading buffer (6x)	From NEB
SDS sample buffer (10x)	
SDS running buffer (10x)	1 % (w/v) SDS, 250 mM TRIS, 2 M glycine
Coomassie safe staining solution	0.008 % (w/v) Brilliant Blue G250, 0.126 % (v/v) HCl
Stacking gel (4 %)	125 mM Tris, 10 % Rotiphorese® Gel 40, 0.1 % SDS, 0.1 % APS, 0.1 % TEMED, pH 8.8
Resolving gel (8 % -15 %)	125 mM Tris, 20 % -37 % Rotiphorese® Gel 40, 0.1 % SDS

	0.1 % APS, 0.3 % TEMED, pH 6.8

6.1.7.2 Buffer for protein purification

Table 10: Buffer for protein purification

Description	Ingredients
Lysis buffer NP	50 mM NaH_2PO_4 pH8, 500 mM NaCl, 5 % (v/v) glycerol, 1 mM PMSF, 0.05 % Triton x 100, 2 mM DTT, 0,025% (w/v) lysozyme
Wash buffer NP	50 mM NaH_2PO_4 pH 8, 500 mM NaCl
Amylose elution buffer NP	50 mM NaH_2PO_4 pH 8, 500 mM NaCl, 20 mM maltose
Heparin binding buffer NP	20 mM NaH_2PO_4 pH 8, 150 mM NaCl
Heparin elution buffer NP	20 mM NaH_2PO_4 pH 8, 1M NaCl
SEC buffer NP and Z	50 mM Tris (HCl) pH 7.5, 150 mM NaCl, 5% (v/v) glycerol, 2 mM DTT
Lysis buffer Z	50 mM Tris (HCl) pH 8, 300 mM NaCl, 5 % (v/v) glycerol, 1 mM PMSF, 0.05 % Triton x 100, 0,025% (w/v) lysozyme, 2 mM DTT
Wash buffer Z	50 mM Tris (HCl) pH 8, 300 mM NaCl
QFF buffer	20 mM Tris (HCl) pH 8.5, 25 mM NaCl
Lysis buffer L	50 mM HEPES (NaOH) pH 7, 500 mM NaCl, 5 % (v/v) glycerol, 0.5 % Tween-20, cOmplete™, 2 mM DTT
Strep washing buffer L	50 mM HEPES (NaOH) pH 7, 500 mM NaCl
Strep elution buffer L	50 mM HEPES (NaOH) pH 7, 500 mM NaCl, 50 mM biotin, 2 mM DTT
Heparin binding buffer L	20 mM HEPES (NaOH) pH 7, 100 mM NaCl, 10 % glycerol
Heparin elution buffer L	20 mM HEPES (NaOH) pH 7, 1000 mM NaCl, 10 % glycerol, 2 mM DTT

6.1.7.3 Buffer for mass spectrometry

Table 11: Buffer for mass spectrometry

Description	Ingredients
nMS standard buffer	150 mM AmAc pH 7.5
nMS buffer L	500 mM AmAc pH 7, 0.01 % (v/v) Tween-20
HDX buffer	40 mM Tris (HCl) pH 7.5 150 mM NaCl
HDX quench buffer	1 M glycine pH 2.3
HDX FD buffer	40 mM Tris(HCl), 6 M urea, pH 7.5

6.1.8 Transfection reagents

FuGENE® HD, Roche

6.1.9 Enzymes, kits and markers

Table 12: Enzymes kits, and markers

Name	Supplier
Nucleobond Xtra Midi Plus Kit	Machery Nagel
NucleoSpin Plasmid, Mini Kit	Machery Nagel
Restrictionenzymes	New England Biolabs
NEBuilder HiFi DNA Assembly	New England Biolabs
Z-competent™E. coli Buffer Set	Zymo Research
Mix & Go! E.coli Transformation Kit	Zymo Research
Quick Load 1kb Ladder	New England Biolabs
Quick Load 100bp Ladder	New England Biolabs
Page Ruler Prestained Protein Ladder	Thermo Fisher Scientific
Page Ruler Plus Prestained Protein Ladder	Thermo Fisher Scientific
dNTP Set, 100mM	Thermo Fisher Scientific
Lysozyme	Roth
NucleoSpin Gel and PCR clean up	Machery Nagel
Q5® High-Fidelity DNA Polymerase	New England Biolabs

6.1.10 Chemicals

Chemicals from Roth: Acetonitrile, 1,4-dithiothreitol (DTT), formic acid, 2-Propanol, 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid (HEPES), acetic acid, ammonium persulfate (APS), ammonium sulfate, Brilliant Blue G250, carbenicillin, chloramphenicol, di-potassium hydrogen phosphate (K2HPO4), ethanol, ethidium bromide, ethylenediaminetetraacetic acid (EDTA), glycerol, H_2O (MS grade), hydrochloric acid (HCl) (37 %), imidazole, isopropyl-β-D-thiogalacto-pyranoside (IPTG), peptone, phenylmethylsulfonyl fluoride (PMSF), potassium chloride (KCl), potassium dihydrogen phosphat (KH2PO4), Roti®Block, Rotiphorese® Gel 40, sodium chloride (NaCl), sodium dodecyl sulphate (SDS) ultra-pure, sodium hydroxide (NaOH),
tetramethylethylenediamine (TEMED), tris-(hydroxymethyl)-amino methane (TRIS), Triton® X 100, Tween-20, yeast extract, β-mercaptoethanol. Chemicals from Merck: Magnesium chloride (MgCl2), magnesium sulfate (MgSO4).
Chemicals from Serva: Bromphenol blue, Xylencyanol FF.
Chemicals from Life Technologies: SYBR® Green I, SYPRO® Orange.
Chemicals from Sigma: D-Biotin, Fast Green, ammonium acetate (99.99% purity)
Chemicals from Biozym: Agarose LE.
Chemicals from Fresenius Kabi: Ampuwa® water.
Chemicals from Roche: cOmplete™ EDTA-free Protease Inhibitor Cocktail.

6.1.11 Consumables

All consumables not listed were purchased from Sarstedt.

Table 13: Consumables

Name	Supplier
Amicon Ultra Filter Units	Millipore
VivaSpin 500 Filter Units	Sartorius
Micro Bio-Spin 6 Columns	Bio-Rad
6-/24-well cell culture plates	TPP
Pierce™ LTQ Velos ESI Positive Ion Calibration Solution	Thermo Scientific

6.1.12 Column material for chromatography

Table 14: Column material for chromatography

Name	Supplier
Strep-Tactin-XT Superflow Matrix	Iba
Superdex™200 Increase 3.2/300	Cytiva
Superdex™200 Increase 10/300 GL	Cytiva
HiTrap™Heparin HP	Cytvia
Glutathione Sepharose™ 4 Fast Flow	Cytvia
Amylose Resin	BioLabs
HiTrap™ Q FF	Cytvia
Porozyme 629 immobilized pepsin beads	Thermo Scientific
Trap column OPTI-TRAP for peptides	Optimize Technologies
Analytical column PLRP-S for Biomolecules	Agilent 632 Technologies

6.1.13 Devices

Table 15: Devices

Name	Company
ÄKTAprime	Cytiva
ÄKTApure	Cytiva
Fluorescence microscope LEICA DMIL	Leica
Electrophorese chamber Mini-PROTEAN®	Biorad
Energy source EV231	Peqlab
Ultraviolet (UV)/Vis-Spektralphotometer Nanodrop2000c	Thermo Scientific
(UV)/Vis-Spektralphotometer Colibri+	Berthold
refrigerated centrifuge 5415R	Eppendorf
Centrifuge 5424	Eppendorf
Centrifuge Rotina 420R	Hettich
scale	Sartorius
Micro scale	Kern&Sohn GmbH
Membrane vacuum pump	Vacuubrand
pH meter FiveEasy	Mettler Toledo

Pipettes	Gilson
Thermomixer 5436	Eppendorf
Shaking incubator Innova 42	Eppendorf
Bio photometer	Eppendorf
Mass photometer OneMP	Refeyn Ltd.
Talos L120c 671 microscope	Thermo Fisher Scientific
Centrifuge Evolution RC	Sorvall
Sonifier	Branson
Shaking flask 250-200 ml	Corning, pyrex, schott
Q-Tof 2 hybrid quadrupole time of flight mass spectrometer	Waters/Micromass
Q Exactive™ UHMR Hybrid Quadrupole Orbitrap™ mass spectrometer	Thermo Fisher Scientific
Infinity 1260 HPLC system	Agilent
Orbitrap Fusion mass spectrometer	Thermo Scientific
Sputter coater CCU-010	safematic
Micropipette puller P-2000	Sutter Instruments

6.1.14 Software

Table 16: Software

Name	Supplier
Excel, Powerpoint, Word	Microsoft Office
GraphPad Prism 5	Graphpad Software, Inc.
Unicorn 7.0	GE Healthcare
Snap Gene Viewer	GSL Biotech
ProtParam	https://web.expansy.org/protparam
MassLynx 4.1	Waters
Xcalibur	Themo Scientific
Massign	Nina Morgener
mMass	Martin Strohalm
UniDec	Michael T. Marty
MSänger	This work

Acquire and Discover MP v2.5.0	Refeyn Ltd
PhotoMol	EMBL
HDExaminer Version 3.3	Sierra Analytics
MaxQuant 2.1.2.0	Max-Planck-Institute of Biochemistry
PyMOL	Schrödinger

6.2 Methods

6.2.1 Microbiology and cell biology

6.2.1.1 Production of chemically *E. coli* cells

Chemically transformation potent *E. coli* DH5α and Bl21 cells were produced after company recommendations (Z-competent™ *E. coli* Buffer Set, Zymo Research). Cells were frozen in a volume of 50 µl in liquid nitrogen and stored in a -80 freezer.

6.2.1.2 Transformation of *E. coli*

100 ng plasmid DNA or 4 µl of NEB-Assembly reaction mix was added to a 50 µl of chemically competent *E. coli* cells. After 30 min of incubation on ice, the cells were incubated for 45 s at 42°C to perform a heat shock. 100 µl of SOC media was added after the cells were cooled down for 2 min on ice. The *E. coli* cells were cultivated for 1 h at 37°C and 300 rpm and spread either on LB agar or inoculated in LB media.
After the transformation of plasmids in *E. coli* DH10EMBacY, the cells were incubated in 1 ml SOC media for 6-8 h and 30°C without antibiotics. The *E. coli* cultures were centrifuged for 2 mins at 2000 x g and resuspended in 90 µl SOC media. The cells were spread on LB-agar with suitable antibiotics.

6.2.1.3 *E. coli* cultivation

E. coli cells spread out on LB agar were incubated with selective antibiotics (100 µg/ml carbenicillin overnight at 37°C or for 3 days at RT. LB media with selective antibiotics (100 µg/ml carbenicillin and 34 µg/ml chloramphenicol for Rosetta 2 (DE3)) was used for cultivation of *E. coli* in liquid culture. The media was inoculated and incubated at 37°C and 180 rpm.

6.2.1.4 Protein expression in *E. coli*

Chemically competent Bl21 (DE3) or Rosetta 2 (DE3) cells were transformed as described in 6.2.1.2. and cultivated over night as described in 6.2.1.3. TB media was used as an expression media in a volume ranged from 50 ml to 1 l including selective antibiotics (100 µg/ml carbenicillin). The overnight culture was used to inoculate the expression culture with a ratio of 1:100. The culture was incubated at 37°C and 180 rpm while monitoring the growth by measuring the optical density at 600 nm (OD_{600}). The cultures were cooled down after reaching an OD_{600} between 0.6 and 1. To induced protein expression, 0.25 – 0.5 mM IPTG was added to the expression cultures. The final incubation for protein expression was performed at 17 °C for approximately 18 h. The cells were harvested by centrifugation at 6000 x g and 4°C for 15 min and supernatant discarded. The cell pellet was stored at -20°C until further use.

6.2.1.5 Cultivation of insect cells

Sf21 cells were used for the generation of baculoviruses. The cells were cultivated in Sf900II SFM medium without additives. The cells were kept as suspension cultures in Erlenmeyer tubes at 27°C and 120-130 rpm. To ensure constant exponential growth. The cells were diluted two times a week to a concentration of 3-4 x 10^5 cells/ml
Protein expression by baculoviruses was performed in High Five™ cells in Express Five™ medium with 100 mM L-glutamine and 100 µg/ml penicillin and 100 µg/ml streptomycin. The cells were cultivated analogous to the Sf21 cells. Both cell types were always diluted to 5 x 10^5 cells/ml the day before usage to ensure constant exponential growth.

6.2.1.6 Protein expression in insect cells

High Five™ cells as described in 6.2.1.5. were used for the expression of L protein constructs. Smal scale expressions were performed to test the protein yield of different viruses. Each virus was tested in an expression culture with a cell concentration of 5 x 10^5 cells/ml and a volume of 50 ml. Virus concentrations between 0.025 and 0.5 % (v/v) were tested. The cell density was monitored after 24 h.p.i.. The cells were diluted to a concentration of 1 x 10^6 cells/ml if the cells were

still growing. Viral growth and protein production were monitored by harvesting 1 x 10^6 cells every 24 h and the measurement of the yellow fluorescence protein (YFP) signal. The best conditions were adjusted to a bigger volume to express larger amounts of proteins. Usually, 400 ml culture per 2 L flask were used. Cells were harvested after approximately 72 h.p.i. when cells were still intact and the highest YFP signal was measurable. For harvesting, the cells were centrifuged at 500 x g at 4°C for 15 min and stored at -20°C until further usage.

6.2.2 Molecular biology

6.2.2.1 Preparation of plasmid DNA

5 ml or 200 ml LB medium were inoculated with transformed *E. coli* and cultivated (see 6.2.1.2 and 6.2.1.3). *E. coli* cells from overnight cultures were harvested by centrifugation. The NucleoSpin® Plasmid Mini Kit was used for 5 ml cultures and NucleoBond® Xtra Midi Plus Kit for 200 ml cultures. Both preparations were performed according to the manual instructions.

6.2.2.2 Polymerase chain reaction (PCR)

A preparative PCR was used to introduce single amino acid mutations in the NP and Z protein. Two DNA oligonucleotides are used to define the start and end of the sequence to be amplified. A DNA polymerase amplifies the DNA in a 3-step cycle with 30 repeats. First, double stranded DNA is made single stranded at 95°C. Second, the DNA primer anneals to the template sequence at 55°C. Finally, the polymerase amplifies the DNA fragment at 72°C. If not stated otherwise the PCR reaction mix was prepared based on Table 17 and executed according to Table 18.

Table 17 Components and concentration for a 50 µl PCR

Component	Concentration
Q5 reaction buffer	1x
Forward primer	0.2 µM
Reverse primer	0.2 µM
Q5 Polymerase	0.02 U/µl
DNA template	0.2 ng/µl

dNTP mix	200 µM
Ampuva water	

Table 18 Reaction program for PCR

Step	Temperature [°C]	Duration	
Initial denaturation	98	1 min	
Denaturation	95	20 s	30 cycles
Annealing	55	30 s	
Elongation	72	30 s/kb	
Final elongation	72	5 min	

6.2.2.3 Agarose gel electrophoresis

Agarose gel electrophoresis was used to separate DNA by its size. Agarose gels with a percentage of agarose ranging from 0.8 % to 1.5 % were used depending on the size of the DNA. 6x DNA loading dye were loaded together with the samples onto the agarose gels. DNA separation was performed with an electrophoresis cell horizontally with TAE buffer. The voltage ranged from 90 to 150 V. 0.01 % ethidium bromide in TAE buffer was used to stain analytic gels for 20 mins. The gels were analyzed under UV light at 254 nm. Preparative gels were stained in a solution of 0.05 % SBYR Green in TAE buffer.

6.2.2.4 Purification of DNA from PCR reactions and agarose gel preparation

The NucleoSpin® Gel and PCR Clean-up Mini Kit was used to purify DNA after PCR (6.2.2.2) or extracted form an agarose gel (6.2.2.3) and was performed according to the manual instructions.

6.2.2.5 Concentration determination of DNA

The determination of DNA concentrations was performed using a NanoDrop spectrophotometer by measuring the absorbance at 260 nm. The concentration was

calculated based on the circumstance that 50 ng/µl of double stranded DNA has an absorbance of 1 at 260 nm.

6.2.2.6 Restriction of DNA

Restriction endonucleases hydrolyses DNA at a defined sequence leading to cleavage at this position. They were used for cloning and analysis and according to manufacturer's recommendations. Separation of DNA products were performed by agarose gel electrophoresis (6.2.2.3)

6.2.2.7 Assembly cloning

Assembly cloning was used to construct DNA fragments together with linearized vectors. Amplified DNA fragments by PCR were designed to have 15 base pairs on both sides matching the ends of a linearized vector. NEBuilder® HiFi DNA Assembly tool was used according to the manual instructions. Vector and insert were mixed in 1:2 ratio using 100 ng of the vector. A 3'-5' exonuclease creates single stranded overhangs to ensure annealing of vector and insert at these positions. A ligase seals nicks in the assembled DNA and a polymerase fills the gaps. The assembled DNA was transformed into competent *E. coli* (6.2.1.2)

6.2.2.8 Sequencing of DNA

For Sequencing, 1 µg of DNA and 20 pmol of sequencing primer in 14 µl Ampuwa® water were prepared. The sequencing was carried out by LGC Genomics GmbH.

6.2.2.9 Production of recombinant bacmids

Bacmids were used to express recombinant proteins in insect cells. pFastBac^{TM}HTP with a gentamycin resistance was used as a donor vector to transfer the expression construct into the *E. coli* strain DH10EMBacY. Besides the baculoviral genome, the DH10EMBacY strain includes a helper plasmid that contains a tetracycline resistance and a transposase enzyme. The transposase integrates the expression construct into the baculoviral genome inside a *lacZ* gene. This allows to identify successfully integrated expression constructs into the baculoviral genome by blue/white screening, as integration results in the disruption of the β galactosidase gene. The

donor vector was transformed into *E. coli* as described in 6.2.1.2 and streaked out on LB-agar containing gentamycin (7 µg/ml), kanamycin (50 µg/ml) and tetracycline (10 µg/ml) as selection antibiotics and additionally, Blue-gal (50 µg/ml) and IPTG for blue/white screening. White colonies were selected after 48 h incubation at 37°C and streaked out on new LB-agar plates to prevent mixed colonies. After 48 h at 37°C they were again checked if they produced blue color. 3 ml of LB medium containing the selection antibiotics were inoculated with white colonies and incubated over night at 37°C. The bacmid DNA was extracted and directly used for transformation in Sf21 cells (see 6.2.3.1). 3 colonies per expression construct were used to produce recombinant baculoviruses.

6.2.3 Virology methods

6.2.3.1 Production of recombinant baculoviruses

For the generation of baculoviruses to express protein in insect cells, bacmid DNA from a 3 ml overnight culture of the *E. coli* strain DH10EMBacY containing an expression construct were used. DNA was extracted by alkalic lysis, neutralization (buffer P1-3 Qiagen) and isopropanol precipitation. The DNA was washed with 70 % (v/v) ethanol and dissolved in 30 µl H_2O and mixed with 200 µl of medium. The bacmid DNA was checked in parallel of the successful integration of the expression construct by PCR (6.2.2.2). The Sf21 cells were seeded in a 6 well plate with 0.7×10^6 cells per well in a volume of 3 ml medium per well. 10 µl of Fugene (Promega) transfection reagent were diluted with 100 µl of medium. The bacmid DNA in medium was mixed with the transfection mix and incubated for 5 min. The transfection reagent/DNA mix was added to 2 wells incubated until signs of virus production were visible (magnified and fluorescent cells) which was usually after 72 h. The supernatant was used as the first virus passage (P0) for further amplification (P1).

6.2.3.2 Amplification of recombinant baculoviruses.

Amplification of baculoviruses was carried out in Sf21 suspension cells. 3 ml of the first virus passage was added to 45 ml of 5×10^5 cells/ml. The cells were monitored every 24 hours and diluted to 5×10^5 cells/ml. After 1-2 days the cells stopped growing and were incubated for another 24 h. The cells were centrifuged at 500 x g

for 5 min and 4 °C and the supernatant divided in several batches which corresponded to the second virus passage (P1). The virus was flash frozen in liquid nitrogen and stored at -80°C until it was used for protein expression in High Five™ cells (6.2.1.6).

6.2.4 Protein biochemistry methods

6.2.4.1 Determination of protein concentration

Protein concentrations were determined photometrically with a NanoDrop spectrophotometer. The absorption at the wavelength of 280 nm is the maximum absorption of aromatic amino acids. By using the Lambert-Beer's law, the concentration in mg/ml was determined. The tool ProtParam was used to calculate the extinction coefficient of the proteins.

$$A = \varepsilon * c * l$$

A: absorbance; *ε*: extinction coefficient in ml * mg^{-1} * cm^{-1}; *c*: concentration in mg/ml, and *l*: path length in cm.

6.2.4.2 SDS polyacrylamide gel electrophoresis

SDS polyacrylamide gel electrophoresis (SDS-PAGE) was used to separate proteins based on their molecular weight. Proteins were denaturated by SDS sample buffer containing SDS and β-mercaptoethanol. Anionic SDS binds to the proteins and results in a negative net charge of the protein which is proportionally to its MW. β-mercaptoethanol disrupts any disulfide bonds. Additionally, the proteins samples were heat denaturated at 95°C for 3 mins. A polyacrylamide gel is composed of a stacking and resolving gel. Depending on the size of the proteins, an acrylamide concentration of the resolving gel was used between 8 and 15 %. The stacking gel had an acrylamide concentration of 4 %. The samples were loaded into pockets of the stacking gel and the gel was placed in an electrophoresis chamber containing SDS running buffer. A current was applied, and the samples were separated at 150 V for 90 mins. A MW marker was used to identify the size of the proteins.
The SDS gel was shortly boiled and washed in H_2O and stained by Safe Stain solution by shortly boiling and incubating for 10 mins. The gel was destained in H_2O.

6.2.4.3 Purification of LASV NP

LASV NP (AV strain) was expressed as a cleavable MBP-fusion protein from a popinM vector in Bl21(DE3) *E. coli* cells. For buffer ingredients see 6.1.7.2. Cell pellets were thawed at room temperature and cells were resuspended in lysis buffer NP with 3 times the volume of the cell pellet weight. The lysate was cooled in ice water while introduced to 3 cycles of sonication with an interval of 10 min sonication and 10 min break. The lysate was clarified by centrifugation at 30,000 g for 30 min at 4°C and incubated while stirring for 45 min with 2.5 ml column volume (CV) amylose beads. The beads were transferred to a column and washed with washing buffer NP with 15 CV. NP-MBP was eluted with maltose containing NP elution buffer. The fusion protein was cleaved by GST-tagged 3C protease at 4°C overnight. The NP containing buffer was diluted to reach 150 mM NaCl concentration by adding 20 mM NaH_2PO_4 pH 8 and loaded on a heparin column to remove bound bacterial RNA. NP was eluted by an NaCl gradient and concentrated for further purification by size-exclusion chromatography on an SD200 column in SEC buffer NP/Z. Proteins were concentrated using filter units (Amicon Ultra, 30,000 MWCO), flash frozen in liquid nitrogen and stored at -80°C.

6.2.4.4 Purification of LASV Z

LASV Z (AV strain) was expressed as a cleavable GST-fusion protein from a popinJ vector in Bl21(DE3) *E. coli* cells. The Z protein was purified base on [45]. For buffer ingredients see 6.1.7.2. Cell pellets were thawed at room temperature and cells were resuspended in lysis buffer Z with 3 times the volume of the cell pellet weight. The lysate was cooled in ice water while introduced to 3 cycles of sonication with a duration of 10 min sonication and 10 min break. The lysate was clarified by centrifugation at 30,000 g for 30 min at 4°C and incubated while stirring for 45 min with 1.5 ml CV GST beads. The beads were transferred to a column and washed with washing buffer Z with 10 CV and 10 CV SEC buffer NP/Z. The fusion protein was cleaved by GST-tagged 3C protease at 4°C overnight on column. The Z containing buffer was diluted to reach 100 mM NaCl concentration by adding 20 mM Tris pH 8.5 and passed on a HiTrap Q FF column. The Z containing flow through was concentrated for further purification by size-exclusion chromatography on an SD200

column in 50 mM in SEC buffer NP/Z. Proteins were concentrated using filter units (Amicon Ultra, 3,000 MWCO), flash frozen in liquid nitrogen and stored at -80°C.

6.2.4.5 Purification of LASV L

LASV L (Ba289 strain) was expressed with a baculovirus expression system in insect cells either with an internal StrepII-tag at position 407 or at the C-Terminus. The purification was performed based on [45]. For buffer ingredients see 6.1.7.2. Cell pellets were thawed at room temperature and cells were resuspended in 50 ml lysis buffer L. The lysate was cooled in ice water while introduced to 3 cycles of sonication with a duration of 10 min sonication and 10 min break. The lysate was clarified by centrifugation at 30,000 g for 30 min at 4°C and incubated while stirring for 45 min with 1.5 ml CV Strep-Tactin-XT beads. The beads were transferred to a column and washed with 15 CV washing buffer L and the L protein was eluted with biotin containing elution buffer L. The buffer containing the L protein was diluted to reach 250 mM NaCl concentration and loaded on a heparin column remove unspecifically bound RNA. The column was washed with 10 CV heparin buffer A and eluted with 1 M NaCl. Proteins were flash frozen in liquid nitrogen and stored at -80°C.

6.2.4.6 Mass photometry

Mass photometry was performed on a Refeyn OneMP mass photometer using the programs Acquire MP v2.5.0 and Discover MP v2.5.0 (Refeyn Ltd.) and PhotoMol. The workflow was published previously [182]. NP was measured in 50 mM Tris (HCl), 150 mM NaCl, 5 % glycerol and a concentration of 25 nM. For the NP + RNA condition, NP at 500 nM was mixed with a single stranded 9mer at a 4 molar-excess to the NP monomer. NP + RNA was incubated at room temperature for approximately 5 min. The sample was diluted with 50 mM Tris (HCl), 150 mM NaCl, 5 % Glycerol to a concentration of 25 nM and subsequently measured.

6.2.4.7 Negative staining electron microscopy

Purified NP at 0.01 mg/mL was applied to glow-discharged carbon-coated copper grids (Electron Microscopy Sciences) and stained with uranyl acetate immediately before imaging. Transmission electron microscopy images were collected using a

Talos L120c microscope (120 kV) with a LaB6 thermionic source at x92,000 magnification and a defocus of between -0.2 and 0.5 µm. For the NP + RNA condition, NP at 0.01 mg/mL was mixed with a single-stranded 25mer at a 2 molar-excess to the NP monomer. The NP + RNA was incubated at room temperature for approximately 15 minutes and then imaged as for the NP without RNA condition.

6.2.5 Mass spectrometry methods

6.2.5.1 Native mass spectrometry

6.2.5.1.1 Sample preparation

To perform nMS, proteins need to be in sodium free and volatile buffers. In this work, an AmAc solution was used as buffer surrogate to perform nMS. Ammonium hydroxide and acetic acid were used to adjust the pH of the ammonium acetate solution. The purified proteins were in common buffer systems and needed to be exchanged to the nMS compatible ammonium acetate solution. For that, different centrifugal devices were used.

Amicon® Ultra 0.5 mL centrifugal filter units (*Merck Millipore*)

The purified protein was added to the centrifugal device and concentrated if necessary to 50 µl volume. 450 µl AmAc solution was added to centrifugal device and concentrated to 50 µl at 13500 x g and 4°C. This took about 10 to 30 minutes depending on the sample and the filter size of the concentrators. This procedure was repeated 3-5 times by adding 450 µl of new AmAc solution. The sample was recovered by mixing the sample and washing the membrane with the sample solution to recover attached proteins from the membrane.

Micro Bio-Spin 6 Columns (*Bio-Rad*)

The gel matrix was resuspended in packing buffer and the tip and cap were separated from the device to remove the packing buffer by gravity drain. Additionally, the device was centrifuged for two minutes at 1000 x g and 4°C to remove the remaining packing buffer. The column was loaded with 500 µl AmAc solution and the column were spun down for one minute at 1000 x g and 4°C to equilibrate the gel matrix. This step was repeated 3 times. The column was then loaded with purified protein sample at the center of the gel matrix and spun down for four minutes at

1000 g and 4 °C. This procedure was repeated with a fresh column to ensure complete desalting.

Samples were put into new micro reaction tubes and centrifuged for 10 minutes at 20.000 x g and 4 °C to clean the solution from protein aggregates which might have developed during buffer exchange. The concentration of the proteins ranged from 0.5 µM to 150 µM. AmAc concentrations and devices used for the respective proteins after are listed in Table 19.

Table 19: nMS analysis solutions and exchange devices

Sample	Analysis solution	Buffer exchange device	MWCO
LASV L-C	1 M ammonium acetate pH 7, 0.01 % Tween-20	Amicon® Ultra 0.5 mL centrifugal	30 kDa
LASV L_407	500 mM ammonium acetate pH 7, 0.01 % Tween-20	Amicon® Ultra 0.5 mL centrifugal	30 kDa
LASV NP and mutants	150 mM ammonium acetate pH 7.5	Amicon® Ultra 0.5 mL centrifugal	30 kDa
LASV Z and mutants	150 mM ammonium acetate pH 7.5	Amicon® Ultra 0.5 mL centrifugal	3 kDa
ANP32 and mutants	150 mM ammonium acetate pH 7.5	Amicon® Ultra 0.5 mL centrifugal	10 kDa

6.2.5.1.2 Data acquisition

Data acquisition was either performed on a nanoESI Q-ToF II mass spectrometer (Waters/Micromass) modified for high mass experiments in positive ion mode [105] or on a Q Exactive™ UHMR Hybrid Quadrupole Orbitrap™ mass spectrometer (Thermo Fisher Scientific).

A spectrum of cesium iodide (25 g/l, 99.999% trace metals basis) was recorded every day for calibrating the data when using the Q-Tof 2 instrument. QE-UHMR was calibrated based on the manufacturer's recommendations every month.
Samples were introduced to the mass spectrometer by gold-coated capillaries made from borosilicate glass tubes (inner diameter: 0.68 mm, outer diameter: 1.2 mm; World Precision Instruments). Capillaries were pulled in a two-step program using a micropipette puller (P-1000, Sutter Instruments) with a squared box filament (2.5 mm by 2.5 mm, Sutter Instruments) Capillaries were afterwards gold-coated using a sputter coater (5.0 × 10−2 mbar, 30.0 mA, 100 s, three runs to vacuum limit 3.0 × 10−2 mbar argon, CCU-010, Safematic). 2 µl of analyzed sample were loaded into a capillary and subsequently introduced to a mass spectrometer.

6.2.5.1.3 Data analysis

Raw data was analyzed with MassLynx V4.1, (Waters) for QToF2 datasets or Xcalibur for UHMR datasets by combining spectra and exporting them into txt file format. The following evaluation regarding full width at half maximum (FWHM) was done with mMass (Martin Strohalm). Experimental mass and area under the curve (AUC) were evaluated with Unidec (Michael T. Marty [183]). Data was further analyzed in Excel. For Visualization, Unidec, MSänger and Illustrator 2023 (Adobe) were used. Diagrams were created with Excel (Microsoft) and Prism 5.03 (*GraphPad Software*). The values for all recorded masses and FWHM resulted from at least 3 independent measurements and are provided in (Table 20)

6.2.5.2 HDX-MS

HDX-MS is based on the chemical reaction in which hydrogen atoms are exchanged by deuterium. The reaction rate is dependent on how the exchangeable hydrogen atoms are exposed to the deuterium solvent. MS can determine the deuterium content of peptides from proteins after H/D exchange in deuterated buffer. We used HDX-MS to identify the NP-Z interaction site by comparing H/D exchange rates of NP peptides when Z is bound to NP or not. The workflow is based on [125].

6.2.5.2.1 Sample preparation

NP protein (50 pmol) was mixed with the Z protein in a 1:5 molar ratio (NP:Z). The deuterium exchange reaction was started by a 1:9 dilution into 99% deuterated buffer (40 mM Tris(HCl) pH 7.5, 150 mM NaCl) at 25°C. The exchange reaction was quenched after the following time points: 15 s, 1 min, 10 min, 1 h, 6 h. The quenching was initiated by a 1:1 addition of ice-cold quench buffer (1 M glycine pH 2.3). After that, samples were flash-frozen in liquid nitrogen. Additionally, a fully deuterated (FD) control was prepared where NP was diluted 1:9 in 99 % deuterated 40 mM Tris(HCl) pH 7.5, 150 mM NaCl, 6 M urea buffer. The reaction was quenched after 24 h incubation at room temperature as described above. All timepoints were analyzed in three technical replicates.

6.2.5.2.2 Data acquisition

The samples were shortly centrifuged after thawing and injected onto a cooled (2°C) HPLC System (Agilent Infinity 1260, Agilent Technologies, Santa Clara, CA, USA) which includes a home packed pepsin column (IDEX guard column with 60 µl Porozyme immobilized pepsin beads, Thermo Scientific, Waltham, MA, USA), in a column oven 5°C) a peptide trap column (OPTI-TRAP for peptides, Optimize Technologies, Oregon City, OR, USA) and a reversed-phase analytical column (PLRP-S for Biomolecules, Agilent Technologies, Santa Clara, CA, USA). Peptide digestion was carried out online at a flow rate of 75 µl/min (0.4 % formic acid in water) and washed after every run with 2 M urea, 2 % acetonitrile, 0.4 % formic acid. Peptide separation prior MS analysis was performed on an analytical column at a flow rate with 150 µl/min and a gradient of 8-40 % buffer B in 7 min (buffer A: 0.4% formic acid in water, buffer B: 0.4 % formic acid in in acetonitrile). The analytical column was washed with 100 % of buffer B after every sample. The HPLC was connected to an Orbitrap Fusion Tribrid in positive ESI MS only mode (Orbitrap resolution 120,000, 4 microscans, Thermo Scientific, Waltham, MA, USA).

Peptide identification was evaluated with 100 pmol of non-deuterated samples. A gradient from 8-40 % buffer B in 27 min was used for the analytical column with a flow rate of 150 µl/min. MS analysis was performed in data-dependent MS/MS acquisition mode (Orbitrap resolution 120,000, 1 microscan, HCD 30 with dynamic exclusion).

6.2.5.2.3 Data analysis and statistics

Peptides were searched against the protein sequence in MaxQuant (version 2.1.2.0) with the Andromeda search engine [184]. Following MaxQuant settings were used: digestion mode was set to unspecific and peptides between 5 and 30 amino acids length were accepted. The minimum score for identification of peptides was 20. The default mass tolerance for precursor (4.5 ppm) and fragment (20 ppm) ions were used according to the Thermo Orbitrap instrument. The retention time for peptides of the HDX runs were adjusted to the shorter gradient of the analytical column compared to the longer gradient of the peptide identification runs.

Datasets were analyzed according to deuterium uptake of peptides via the automated centroid analysis by HDExaminer Version 3.3 (Sierra Analytics). All peptides were checked manually based on the presence of overlapping peptide envelopes, correct retention time, *m/z* range, and charge state. All peptides at different states and timepoints are plotted as wood plot. A Summary of all experimental conditions and statistics can be found in (Table 4).

GraphPAD Prism (GarphPAD Software, Inc) and PyMOL (Schrödinger) software were used for visualization.

6.2.6 Bioinformatic methods

6.2.6.1 Structural modelling

Structural modeling was performed in collaboration by the Kosinski group at EMBL Hamburg.

To model the NP-Z complex, we used the ‘custom’ mode of AlphaPulldown [127], a python package built upon AlphaFold [185] and AlphaFold Multimer [129]. Both the initial release of AlphaFold Multimer (version 2.1.0) and the latest version (version 2.3.0) were used. All parameters were set to default except for the max_recycles, which was increased from 3 to 12 for better model quality and to avoid steric clashes. The local quality of the model was assessed by predicted local distance difference test (pLDDT) scores as returned by AlphaFold. The confidence in the relative arrangement between NP and Z proteins was evaluated by predicted aligned errors (PAE), also as returned by AlphaFold. Other evaluations of the model quality and properties, including Predicted DockQ score (pDockQ) [186], protein-interface

score (PI-score) [187], and biophysical properties of the interaction interface (using PI-score program), were calculated by AlphaPulldown.

6.2.6.2 Calculating affinities

The affinities of protein-protein interaction were calculated according to the law of mass action and based on [124]. 3 binding pockets were considered when calculating the K_D for the NP_3-Z interaction.

Equitation 2:

$$P1 + P2 \rightleftharpoons (P1 + P2)$$

The speed of the reaction is described by the dissociation constant:

Equitation 3:

$$K_D = \frac{[P1] \times [P2]}{[P1 + P2]}$$

[P1] is the concentration of free protein 1 without a ligand. This corresponds to the monomer concentration of $[P1]_0$ which was introduced to the mass spectrometer and multiplied by the value formed by the $P1_{unbound}$ signal, which corresponds to the AUC values of the native mass spectrometry spectrum and divided by value formed by the total P1 signal (P1-P2 + $P1_{unbound}$):

Equitation 4:

$$[P1] = [P1]_0 \times \frac{P1_{unbound}}{P1\text{-}P2 + P1_{unbound}}$$

[P2] is the concentration of the free protein 2 which is calculated by using the concentration which was introduced to the mass spectrometer [$P2_0$] subtracted by the (P1-P2) signal which, corresponds to the AUC values from the native mass spectrometry spectra:

Equitation 5:

$$[P2] = [P2]_0 - P1\text{-}P2$$

[P1-P2] is the concentration of the protein1-protein2 complex. This corresponds to the monomer concentration of P1 which was introduced to the mass spectrometer and multiplied by the value formed by the P1-P2 signal, which corresponds to the AUC values from the native mass spectrometry spectrum, divided by value formed by the total P1 signal (P1 + P1-P2):

Equitation 6:

$$[P1 - P2] = [P1]_0 \times \frac{P1\text{-}P2}{P_{unbound} + P1\text{-P2}}$$

6.2.7 Statistics and reproducibility

The determined masses, FWHM and AUC from nMS analysis were taken from at least 3 independent measurements and are shown together with the resulting standard deviation. Values for quantification are shown as mean from three independent measurements. The error bars represent the standard deviation. The shown standard deviation for K_D values are calculated according to the rules of Gaussian error propagation. HDX-MS experimental design and data analysis was evaluated according to HDX-MS community-recommendations [126]. The quench conditions were optimized for maximum sequence coverage. The back exchange was tested with a fully deuterated control. The labeling timepoints cover 3-4 orders of magnitude. The deuterium uptake differences between two states were statistically evaluated by using a two paired t-test with the p-value < 0.05. Peptides showing differences were only considered to be significant if passing the t-test and have additionally a difference higher than 0.561 deuteron, which is the variance across all replicates of all peptides.

References

1. Maes, P., et al., *Taxonomy of the family Arenaviridae and the order Bunyavirales: update 2018.* Archives of virology, 2018. **163**: p. 2295-2310.
2. Fenner, F., *Second Report of the International Committee on Taxonomy of Viruses.* Intervirology, 2008. **7**(1-2): p. 1-3.
3. Charrel, R., et al., *ICTV Virus Taxonomy Profile: Arenaviridae 2023.* Journal of General Virology, 2023. **104**(9).
4. Chen, Y.-M., et al., *RNA viromes from terrestrial sites across China expand environmental viral diversity.* Nature Microbiology, 2022. **7**(8): p. 1312-1323.
5. Ly, H., *The ever-expanding diversity and complexity of the Arenaviridae family*. Virulence, 2023. **14**(1): p. 2279353.
6. Salazar-Bravo, J., L.A. Ruedas, and T.L. Yates, *Mammalian Reservoirs of Arenaviruses*, in *Arenaviruses I: The Epidemiology, Molecular and Cell Biology of Arenaviruses*, M.B.A. Oldstone, Editor. 2002, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 25-63.
7. Price, J.L., *Serological evidence of infection of Tacaribe virus and arboviruses in Trinidadian bats.* Am J Trop Med Hyg, 1978. **27**(1 Pt 1): p. 162-7.
8. Bowen, M.D., C.J. Peters, and S.T. Nichol, *Phylogenetic Analysis of theArenaviridae:Patterns of Virus Evolution and Evidence for Cospeciation between Arenaviruses and Their Rodent Hosts.* Molecular Phylogenetics and Evolution, 1997. **8**(3): p. 301-316.
9. Salvato, M.S., *The arenaviridae*. 2012: Springer Science & Business Media.
10. Ruo, S.L., et al., *Antigenic relatedness between arenaviruses defined at the epitope level by monoclonal antibodies.* Journal of General Virology, 1991. **72**(3): p. 549-555.
11. Günther, S. and O. Lenz, *Lassa virus.* Crit Rev Clin Lab Sci, 2004. **41**(4): p. 339-90.
12. Salam, A.P., et al., *Ribavirin for treating Lassa fever: A systematic review of pre-clinical studies and implications for human dosing.* PLOS Neglected Tropical Diseases, 2022. **16**(3): p. e0010289.
13. Kieny, M.P. and P. Salama, *WHO R&D Blueprint: a global coordination mechanism for R&D preparedness.* The Lancet, 2017. **389**(10088): p. 2469-2470.
14. Buckley, S.M., J. Casals, and W.G. Downs, *Isolation and Antigenic Characterization of Lassa Virus.* Nature, 1970. **227**(5254): p. 174-174.
15. Olayemi, A., et al., *New Hosts of The Lassa Virus.* Scientific Reports, 2016. **6**(1): p. 25280.
16. Lecompte, E., et al., *Mastomys natalensis and Lassa Fever, West Africa.* Emerging Infectious Disease journal, 2006. **12**(12): p. 1971.
17. Meulen, J.T., et al., *Hunting of Peridomestic Rodents and Consumption of Their Meat as Possible Risk Factors for Rodent-to-Human Transmission of Lassa Virus in the Republic of Guinea.* The American Journal of Tropical Medicine and Hygiene, 1996. **55**(6): p. 661-666.
18. Stephenson, E.H., E.W. Larson, and J.W. Dominik, *Effect of environmental factors on aerosol-induced lassa virus infection.* Journal of Medical Virology, 1984. **14**(4): p. 295-303.

19. McCormick, J.B., et al., *A Prospective Study of the Epidemiology and Ecology of Lassa Fever.* The Journal of Infectious Diseases, 1987. **155**(3): p. 437-444.
20. Hallam, H.J., et al., *Baseline mapping of Lassa fever virology, epidemiology and vaccine research and development.* npj Vaccines, 2018. **3**(1): p. 11.
21. Klitting, R., et al., *Predicting the evolution of the Lassa virus endemic area and population at risk over the next decades.* Nature Communications, 2022. **13**(1): p. 5596.
22. Price, M.E., et al., *A prospective study of maternal and fetal outcome in acute Lassa fever infection during pregnancy.* British Medical Journal, 1988. **297**(6648): p. 584-587.
23. Kayem, N.D., et al., *Lassa fever in pregnancy: a systematic review and meta-analysis.* Transactions of The Royal Society of Tropical Medicine and Hygiene, 2020. **114**(5): p. 385-396.
24. Moraz, M.-L. and S. Kunz, *Pathogenesis of arenavirus hemorrhagic fevers.* Expert Review of Anti-infective Therapy, 2011. **9**(1): p. 49-59.
25. Cummins, D., et al., *Acute sensorineural deafness in Lassa fever.* Jama, 1990. **264**(16): p. 2093-6.
26. Yun, N.E. and D.H. Walker, *Pathogenesis of Lassa Fever.* Viruses, 2012. **4**(10): p. 2031-2048.
27. Garnett, L.E. and J.E. Strong, *Lassa fever: With 50 years of study, hundreds of thousands of patients and an extremely high disease burden, what have we learned?* Current Opinion in Virology, 2019. **37**: p. 123-131.
28. Asogun, D.A., et al., *Lassa Fever: Epidemiology, Clinical Features, Diagnosis, Management and Prevention.* Infectious Disease Clinics of North America, 2019. **33**(4): p. 933-951.
29. Radoshitzky, S.R., et al., *ICTV Virus Taxonomy Profile: Arenaviridae 2023.* Journal of General Virology, 2023. **104**(9).
30. Hallam, S.J., et al., *Review of Mammarenavirus Biology and Replication.* Frontiers in Microbiology, 2018. **9**.
31. Lukashevich, I., et al., *Ribonucleic acids of Machupo and Lassa viruses.* Archives of virology, 1984. **79**: p. 189-203.
32. Auperin, D.D., et al., *Sequencing studies of pichinde arenavirus S RNA indicate a novel coding strategy, an ambisense viral S RNA.* Journal of Virology, 1984. **52**(3): p. 897-904.
33. Kouba, T., et al., *Conformational changes in Lassa virus L protein associated with promoter binding and RNA synthesis activity.* Nature Communications, 2021. **12**(1): p. 7018.
34. Pinschewer, D.D., M. Perez, and J.C.d.l. Torre, *Role of the Virus Nucleoprotein in the Regulation of Lymphocytic Choriomeningitis Virus Transcription and RNA Replication.* Journal of Virology, 2003. **77**(6): p. 3882-3887.
35. Eichler, R., et al., *Characterization of the Lassa virus matrix protein Z: electron microscopic study of virus-like particles and interaction with the nucleoprotein (NP).* Virus research, 2004. **100**(2): p. 249-255.
36. Perez, M., D.L. Greenwald, and J.C.d.L. Torre, *Myristoylation of the RING Finger Z Protein Is Essential for Arenavirus Budding.* Journal of Virology, 2004. **78**(20): p. 11443-11448.
37. Cao, W., et al., *Identification of alpha-dystroglycan as a receptor for lymphocytic choriomeningitis virus and Lassa fever virus.* Science, 1998. **282**(5396): p. 2079-81.

38. Shimojima, M., et al., *Identification of cell surface molecules involved in dystroglycan-independent Lassa virus cell entry.* Journal of virology, 2012. **86**(4): p. 2067-2078.
39. Pasqual, G., et al., *Old world arenaviruses enter the host cell via the multivesicular body and depend on the endosomal sorting complex required for transport.* PLoS pathogens, 2011. **7**(9): p. e1002232.
40. Oppliger, J., et al., *Lassa virus cell entry via dystroglycan involves an unusual pathway of macropinocytosis.* Journal of virology, 2016. **90**(14): p. 6412-6429.
41. Jae, L.T., et al., *Lassa virus entry requires a trigger-induced receptor switch.* Science, 2014. **344**(6191): p. 1506-1510.
42. Rojek, J.M. and S. Kunz, *Cell entry by human pathogenic arenaviruses.* Cellular Microbiology, 2008. **10**(4): p. 828-835.
43. Williams, H.M., et al., *Structural snapshots of phenuivirus cap-snatching and transcription.* bioRxiv, 2023: p. 2023.11.29.569195.
44. Olschewski, S., S. Cusack, and M. Rosenthal, *The Cap-Snatching Mechanism of Bunyaviruses.* Trends in Microbiology, 2020. **28**(4): p. 293-303.
45. Vogel, D., et al., *Biochemical characterization of the Lassa virus L protein.* Journal of Biological Chemistry, 2019. **294**(20): p. 8088-8100.
46. Günther, S. and O. Lenz, *Lassa Virus.* Critical Reviews in Clinical Laboratory Sciences, 2004. **41**(4): p. 339-390.
47. Kranzusch, P.J. and S.P.J. Whelan, *Arenavirus Z protein controls viral RNA synthesis by locking a polymerase–promoter complex.* Proceedings of the National Academy of Sciences, 2011. **108**(49): p. 19743-19748.
48. Kang, H., et al., *Structural basis for recognition and regulation of arenavirus polymerase L by Z protein.* Nature Communications, 2021. **12**(1): p. 4134.
49. Ma, J., S. Zhang, and X. Zhang, *Structure of Machupo virus polymerase in complex with matrix protein Z.* Nature Communications, 2021. **12**(1): p. 6163.
50. Casabona, J.C., et al., *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.* Journal of Virology, 2009. **83**(14): p. 7029-7039.
51. Loureiro, M.E., et al., *Uncovering Viral Protein-Protein Interactions and their Role in Arenavirus Life Cycle.* Viruses, 2012. **4**(9): p. 1651-1667.
52. Wolff, S., H. Ebihara, and A. Groseth, *Arenavirus Budding: A Common Pathway with Mechanistic Differences.* Viruses, 2013. **5**(2): p. 528-549.
53. Levingston Macleod, J.M., et al., *Identification of two functional domains within the arenavirus nucleoprotein.* Journal of virology, 2011. **85**(5): p. 2012-2023.
54. Ortiz-Riaño, E., et al., *The C-Terminal Region of Lymphocytic Choriomeningitis Virus Nucleoprotein Contains Distinct and Segregable Functional Domains Involved in NP-Z Interaction and Counteraction of the Type I Interferon Response.* Journal of Virology, 2011. **85**(24): p. 13038-13048.
55. Hastie, K.M., et al., *Crystal structure of the Lassa virus nucleoprotein-RNA complex reveals a gating mechanism for RNA binding.* Proceedings of the National Academy of Sciences, 2011. **108**(48): p. 19365-19370.
56. Brunotte, L., et al., *Structure of the Lassa Virus Nucleoprotein Revealed by X-ray Crystallography, Small-angle X-ray Scattering, and Electron Microscopy *.* Journal of Biological Chemistry, 2011. **286**(44): p. 38748-38756.
57. Qi, X., et al., *Cap binding and immune evasion revealed by Lassa nucleoprotein structure.* Nature, 2010. **468**(7325): p. 779-783.

58. Hastie, K.M., et al., *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, 2011. **108**(6): p. 2396-401.
59. Zhou, S., et al., *Induction and Inhibition of Type I Interferon Responses by Distinct Components of Lymphocytic Choriomeningitis Virus.* Journal of Virology, 2010. **84**(18): p. 9452-9462.
60. Pythoud, C., et al., *Arenavirus Nucleoprotein Targets Interferon Regulatory Factor-Activating Kinase IKKε.* Journal of Virology, 2012. **86**(15): p. 7728-7738.
61. Rodrigo, W.W.S.I., et al., *Arenavirus Nucleoproteins Prevent Activation of Nuclear Factor Kappa B.* Journal of Virology, 2012. **86**(15): p. 8185-8197.
62. Lennartz, F., et al., *The role of oligomerization for the biological functions of the arenavirus nucleoprotein.* Archives of virology, 2013. **158**(9): p. 1895-1905.
63. Lenz, O., et al., *The Lassa virus glycoprotein precursor GP-C is proteolytically processed by subtilase SKI-1/S1P.* Proceedings of the National Academy of Sciences, 2001. **98**(22): p. 12701-12705.
64. Capul, A.A., et al., *Arenavirus Z-glycoprotein association requires Z myristoylation but not functional RING or late domains.* Journal of virology, 2007. **81**(17): p. 9451-9460.
65. York, J. and J.H. Nunberg, *Role of the Stable Signal Peptide of Junín Arenavirus Envelope Glycoprotein in pH-Dependent Membrane Fusion.* Journal of Virology, 2006. **80**(15): p. 7775-7780.
66. Eichler, R., et al., *Signal peptide of Lassa virus glycoprotein GP-C exhibits an unusual length.* FEBS Letters, 2003. **538**(1-3): p. 203-206.
67. York, J. and J.H. Nunberg, *Distinct requirements for signal peptidase processing and function in the stable signal peptide subunit of the Junín virus envelope glycoprotein.* Virology, 2007. **359**(1): p. 72-81.
68. Agnihothram, S.S., J. York, and J.H. Nunberg, *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, 2006. **80**(11): p. 5189-5198.
69. Eichler, R., et al., *The role of single N-glycans in proteolytic processing and cell surface transport of the Lassa virus glycoprotein GP-C.* Virology Journal, 2006. **3**(1): p. 41.
70. Pyle, J.D. and S.P.J. Whelan, *RNA ligands activate the Machupo virus polymerase and guide promoter usage.* Proceedings of the National Academy of Sciences, 2019. **116**(21): p. 10518-10524.
71. Peng, R., et al., *Structural insight into arenavirus replication machinery.* Nature, 2020. **579**(7800): p. 615-619.
72. Brunotte, L., et al., *Domain structure of Lassa virus L protein.* Journal of virology, 2011. **85**(1): p. 324-333.
73. Lehmann, M., et al., *Role of the C Terminus of Lassa Virus L Protein in Viral mRNA Synthesis.* Journal of Virology, 2014. **88**(15): p. 8713-8717.
74. Arragain, B., et al., *Structural snapshots of La Crosse virus polymerase reveal the mechanisms underlying Peribunyaviridae replication and transcription.* Nature Communications, 2022. **13**(1): p. 902.
75. Williams, H.M., et al., *Structural insights into viral genome replication by the severe fever with thrombocytopenia syndrome virus L protein.* Nucleic Acids Research, 2023. **51**(3): p. 1424-1442.

76. Volpon, L., et al., *Structural characterization of the Z RING-eIF4E complex reveals a distinct mode of control for eIF4E.* Proceedings of the National Academy of Sciences, 2010. **107**(12): p. 5441-5446.
77. Hastie, K.M., et al., *Crystal Structure of the Oligomeric Form of Lassa Virus Matrix Protein Z.* Journal of Virology, 2016. **90**(9): p. 4556-4562.
78. Strecker, T., et al., *Lassa virus Z protein is a matrix protein sufficient for the release of virus-like particles.* Journal of virology, 2003. **77**(19): p. 10700-10705.
79. Kentsis, A., et al., *The RING domains of the promyelocytic leukemia protein PML and the arenaviral protein Z repress translation by directly inhibiting translation initiation factor eIF4E11Edited by D. Draper.* Journal of Molecular Biology, 2001. **312**(4): p. 609-623.
80. Brito, A.F. and J.W. Pinney, *Protein–Protein Interactions in Virus–Host Systems.* Frontiers in Microbiology, 2017. **8**: p. 1557.
81. Griffiths, J., *A brief history of mass spectrometry.* Anal. Chem, 2008. **80**(15): p. 5678-5683.
82. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons.* Analytical Chemistry, 1988. **60**(20): p. 2299-2301.
83. Yamashita, M. and J.B. Fenn, *Electrospray ion source. Another variation on the free-jet theme.* The Journal of Physical Chemistry, 1984. **88**(20): p. 4451-4459.
84. De Hoffmann, E. and V. Stroobant, *Mass spectrometry: principles and applications*. 2007: John Wiley & Sons.
85. Loo, J.A., *Observation of large subunit protein complexes by electrospray ionization mass spectrometry.* Journal of Mass Spectrometry, 1995. **30**(1): p. 180-183.
86. Rostom, A.A. and C.V. Robinson, *Detection of the Intact GroEL Chaperonin Assembly by Mass Spectrometry.* Journal of the American Chemical Society, 1999. **121**(19): p. 4718-4719.
87. Van Berkel, W.J.H., et al., *Detection of intact megaDalton protein assemblies of vanillyl-alcohol oxidase by mass spectrometry.* Protein Science, 2000. **9**(3): p. 435-439.
88. Patterson, S.D. and R.H. Aebersold, *Proteomics: the first decade and beyond.* Nature Genetics, 2003. **33**(3): p. 311-323.
89. Katta, V., B.T. Chait, and S. Carr, *Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry.* Rapid Communications in Mass Spectrometry, 1991. **5**(4): p. 214-217.
90. Rossi, V., et al., *Structure of the catalytic region of human complement protease C.hivin.1s: Study by chemical crosslinking and three-dimensional homology modeling.* Biochemistry, 1995. **34**(22): p. 7311-7321.
91. Leney, A.C. and A.J.R. Heck, *Native Mass Spectrometry: What is in the Name?* Journal of the American Society for Mass Spectrometry, 2017. **28**(1): p. 5-13.
92. Esser, T.K., et al., *Cryo-EM of soft-landed β-galactosidase: Gas-phase and native structures are remarkably similar.* bioRxiv, 2023: p. 2023.08.17.553673.
93. Westphall, M.S., et al., *Three-dimensional structure determination of protein complexes using matrix-landing mass spectrometry.* Nature Communications, 2022. **13**(1): p. 2276.

94. Hogan, C.J., et al., *Charge Reduced Electrospray Size Spectrometry of Mega- and Gigadalton Complexes: Whole Viruses and Virus Fragments.* Analytical Chemistry, 2006. **78**(3): p. 844-852.
95. Gan, J., et al., *Native Mass Spectrometry of Recombinant Proteins from Crude Cell Lysates.* Anal Chem, 2017. **89**(8): p. 4398-4404.
96. Griffiths, R.L., et al., *Native mass spectrometry imaging of intact proteins and protein complexes in thin tissue sections.* International Journal of Mass Spectrometry, 2019. **437**: p. 23-29.
97. Konermann, L., *Addressing a Common Misconception: Ammonium Acetate as Neutral pH "Buffer" for Native Electrospray Mass Spectrometry.* Journal of the American Society for Mass Spectrometry, 2017. **28**(9): p. 1827-1835.
98. Susa, A.C., Z. Xia, and E.R. Williams, *Small Emitter Tips for Native Mass Spectrometry of Proteins and Protein Complexes from Nonvolatile Buffers That Mimic the Intracellular Environment.* Analytical Chemistry, 2017. **89**(5): p. 3116-3122.
99. Barth, M. and C. Schmidt, *Native mass spectrometry—A valuable tool in structural biology.* Journal of Mass Spectrometry, 2020. **55**(10): p. e4578.
100. Wilm, M. and M. Mann, *Analytical properties of the nanoelectrospray ion source.* Analytical chemistry, 1996. **68**(1): p. 1-8.
101. Taylor, G.I., *Disintegration of water drops in an electric field.* Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences, 1964. **280**(1382): p. 383-397.
102. Konermann, L., et al., *Unraveling the Mechanism of Electrospray Ionization.* Analytical Chemistry, 2013. **85**(1): p. 2-9.
103. Konermann, L., et al., *On the Chemistry of Aqueous Ammonium Acetate Droplets during Native Electrospray Ionization Mass Spectrometry.* Analytical Chemistry, 2023. **95**(37): p. 13957-13966.
104. Iribarne, J. and B. Thomson, *On the evaporation of small ions from charged droplets.* The Journal of chemical physics, 1976. **64**(6): p. 2287-2294.
105. Van Den Heuvel, R.H.H., et al., *Improving the Performance of a Quadrupole Time-of-Flight Instrument for Macromolecular Mass Spectrometry.* Analytical Chemistry, 2006. **78**(21): p. 7473-7483.
106. Greaves, J. and J. Roboz, *Mass spectrometry for the novice*. 2013: CRC Press.
107. Boesl, U., *Time-of-flight mass spectrometry: Introduction to the basics.* Mass Spectrometry Reviews, 2017. **36**(1): p. 86-109.
108. Makarov, A., *Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis.* Analytical Chemistry, 2000. **72**(6): p. 1156-1162.
109. Zubarev, R.A. and A. Makarov, *Orbitrap Mass Spectrometry.* Analytical Chemistry, 2013. **85**(11): p. 5288-5296.
110. Lössl, P., J. Snijder, and A.J. Heck, *Boundaries of mass resolution in native mass spectrometry.* Journal of the American Society for Mass Spectrometry, 2014. **25**(6): p. 906-917.
111. Catherman, A.D., O.S. Skinner, and N.L. Kelleher, *Top Down proteomics: Facts and perspectives.* Biochemical and Biophysical Research Communications, 2014. **445**(4): p. 683-693.
112. Beardsley, R.L., et al., *Noncovalent protein tetramers and pentamers with "n" charges yield monomers with n/4 and n/5 charges.* Analytical chemistry, 2009. **81**(4): p. 1347-1356.

113. Masson, G.R., M.L. Jenkins, and J.E. Burke, *An overview of hydrogen deuterium exchange mass spectrometry (HDX-MS) in drug discovery.* Expert Opinion on Drug Discovery, 2017. **12**(10): p. 981-994.
114. Marcsisin, S.R. and J.R. Engen, *Hydrogen exchange mass spectrometry: what is it and what can it tell us?* Analytical and Bioanalytical Chemistry, 2010. **397**(3): p. 967-972.
115. Vinciauskaite, V. and Glenn R. Masson, *Fundamentals of HDX-MS.* Essays in Biochemistry, 2023. **67**(2): p. 301-314.
116. Bai, Y., et al., *Primary structure effects on peptide group hydrogen exchange.* Proteins, 1993. **17**(1): p. 75-86.
117. Englander, S.W. and N.R. Kallenbach, *Hydrogen exchange and structural dynamics of proteins and nucleic acids.* Quarterly reviews of biophysics, 1983. **16**(4): p. 521-655.
118. Xiao, H., et al., *Mapping protein energy landscapes with amide hydrogen exchange and mass spectrometry: I. A generalized model for a two-state protein and comparison with experiment.* Protein Science, 2005. **14**(2): p. 543-557.
119. Guttman, M., et al., *Analysis of Overlapped and Noisy Hydrogen/Deuterium Exchange Mass Spectra.* Journal of the American Society for Mass Spectrometry, 2013. **24**(12): p. 1906-1912.
120. Zhang, Z. and D.L. Smith, *Determination of amide hydrogen exchange by mass spectrometry: A new tool for protein structure elucidation.* Protein Science, 1993. **2**(4): p. 522-531.
121. Shen, Y., et al., *Differential Hydrogen/Deuterium Exchange during Proteoform Separation Enables Characterization of Conformational Differences between Coexisting Protein States.* Anal Chem, 2019. **91**(6): p. 3805-3809.
122. Fitzgerald, D.J., et al., *Protein complex expression by using multigene baculoviral vectors.* Nature methods, 2006. **3**(12): p. 1021-1032.
123. Olinares, P.D.B., et al., *A Robust Workflow for Native Mass Spectrometric Analysis of Affinity-Isolated Endogenous Protein Assemblies.* Analytical Chemistry, 2016. **88**(5): p. 2799-2807.
124. Kopicki, J.-D., et al., *Opening opportunities for Kd determination and screening of MHC peptide complexes.* Communications Biology, 2022. **5**(1): p. 488.
125. Dülfer, J., et al., *Glycan-induced protein dynamics in human norovirus P dimers depend on virus strain and deamidation status.* Molecules, 2021. **26**(8): p. 2125.
126. Masson, G.R., et al., *Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments.* Nature Methods, 2019. **16**(7): p. 595-602.
127. Yu, D., et al., *AlphaPulldown—a python package for protein–protein interaction screens using AlphaFold-Multimer.* Bioinformatics, 2022. **39**(1): p. 749.
128. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.
129. Evans, R., et al., *Protein complex prediction with AlphaFold-Multimer. bioRxiv [Preprint](2021)*. October.
130. Zhang, Y., J. Carlos de la Torre, and G.B. Melikyan, *Human LAMP1 accelerates Lassa virus fusion and potently promotes fusion pore dilation upon forcing viral fusion with non-endosomal membrane.* PLOS Pathogens, 2022. **18**(8): p. e1010625.
131. Zhang, Y., et al., *Fusogenic structural changes in arenavirus glycoproteins are associated with viroporin activity.* PLOS Pathogens, 2023. **19**(7): p. e1011217.

132. Shtykova, E.V., et al., *Structural Analysis of Influenza A Virus Matrix Protein M1 and Its Self-Assemblies at Low pH.* PLoS ONE, 2013. **8**(12): p. e82431.
133. Xu, X., et al., *Cryo-EM structures of Lassa and Machupo virus polymerases complexed with cognate regulatory Z proteins identify targets for antivirals.* Nature Microbiology, 2021. **6**(7): p. 921-931.
134. Iwasaki, M., et al., *Efficient Interaction between Arenavirus Nucleoprotein (NP) and RNA-Dependent RNA Polymerase (L) Is Mediated by the Virus Nucleocapsid (NP-RNA) Template.* Journal of Virology, 2015. **89**(10): p. 5734-5738.
135. Kerber, R., et al., *Cross-species analysis of the replication complex of Old World arenaviruses reveals two nucleoprotein sites involved in L protein function.* Journal of virology, 2011. **85**(23): p. 12518-12528.
136. Pyle, J.D. and S.P.J. Whelan, *Isolation of Reconstructed Functional Ribonucleoprotein Complexes of Machupo Virus.* Journal of Virology, 2021. **95**(22): p. e01054-21.
137. König, R., et al., *Human host factors required for influenza virus replication.* Nature, 2010. **463**(7282): p. 813-817.
138. Reilly, P.T., et al., *Cracking the ANP32 whips: Important functions, unequal requirement, and hints at disease implications.* BioEssays, 2014. **36**(11): p. 1062-1071.
139. Carrique, L., et al., *Host ANP32A mediates the assembly of the influenza virus replicase.* Nature, 2020. **587**(7835): p. 638-643.
140. Wang, F., et al., *The C-terminal LCAR of host ANP32 proteins interacts with the influenza A virus nucleoprotein to promote the replication of the viral RNA genome.* Nucleic Acids Research, 2022. **50**(10): p. 5713-5725.
141. Bassetto, M., et al., *Structural biology in antiviral drug discovery.* Current Opinion in Pharmacology, 2016. **30**: p. 116-130.
142. Duelfer, J., et al., *Structural mass spectrometry goes viral.* Advances in virus research, 2019. **105**: p. 189-238.
143. Mühlmann, M., et al., *Optimizing recombinant protein expression via automated induction profiling in microtiter plates at different temperatures.* Microb Cell Fact, 2017. **16**(1): p. 220.
144. Chi, E.Y., et al., *Physical stability of proteins in aqueous solution: mechanism and driving forces in nonnative protein aggregation.* Pharmaceutical research, 2003. **20**: p. 1325-1336.
145. VanAernum, Z.L., et al., *Rapid online buffer exchange for screening of proteins, protein complexes and cell lysates by native mass spectrometry.* Nature Protocols, 2020. **15**(3): p. 1132-1157.
146. Olinares, P.D., et al., *A Robust Workflow for Native Mass Spectrometric Analysis of Affinity-Isolated Endogenous Protein Assemblies.* Anal Chem, 2016. **88**(5): p. 2799-807.
147. Gavriilidou, A.F.M., B. Gülbakan, and R. Zenobi, *Influence of Ammonium Acetate Concentration on Receptor–Ligand Binding Affinities Measured by Native Nano ESI-MS: A Systematic Study.* Analytical Chemistry, 2015. **87**(20): p. 10378-10384.
148. Susa, A.C., et al., *Submicrometer Emitter ESI Tips for Native Mass Spectrometry of Membrane Proteins in Ionic and Nonionic Detergents.* J Am Soc Mass Spectrom, 2018. **29**(1): p. 203-206.
149. Saikusa, K., et al., *Native Mass Spectrometry of Protein and DNA Complexes Prepared in Nonvolatile Buffers.* Journal of the American Society for Mass Spectrometry, 2020. **31**(3): p. 711-718.

150. Gan, J., et al., *Native mass spectrometry of recombinant proteins from crude cell lysates.* Analytical chemistry, 2017. **89**(8): p. 4398-4404.
151. Leyrat, C., et al., *Structure of the vesicular stomatitis virus N0-P complex.* PLoS pathogens, 2011. **7**(9): p. e1002248.
152. Pattis, J.G. and E.R. May, *Markov State Model of Lassa Virus Nucleoprotein Reveals Large Structural Changes during the Trimer to Monomer Transition.* Structure, 2020. **28**(5): p. 548-554.e3.
153. Fujii, Y., et al., *Selective incorporation of influenza virus RNA segments into virions.* Proceedings of the National Academy of Sciences, 2003. **100**(4): p. 2002-2007.
154. Raymond, D.D., et al., *Structure of the Rift Valley fever virus nucleocapsid protein reveals another architecture for RNA encapsidation.* Proceedings of the National Academy of Sciences, 2010. **107**(26): p. 11769-11774.
155. Wichgers Schreur, P.J. and J. Kortekaas, *Single-Molecule FISH Reveals Non-selective Packaging of Rift Valley Fever Virus Genome Segments.* PLOS Pathogens, 2016. **12**(8): p. e1005800.
156. Bermúdez-Méndez, E., et al., *Visualizing the ribonucleoprotein content of single bunyavirus virions reveals more efficient genome packaging in the arthropod host.* Communications Biology, 2021. **4**(1): p. 345.
157. Pyle, J.D. and S.P. Whelan, *Isolation of reconstructed functional ribonucleoprotein complexes of Machupo virus.* Journal of virology, 2021. **95**(22): p. e01054-21.
158. Xu, M., et al., *Bunyaviral N Proteins Localize at RNA Processing Bodies and Stress Granules: The Enigma of Cytoplasmic Sources of Capped RNA for Cap Snatching.* Viruses, 2022. **14**(8): p. 1679.
159. Mir, M.A., et al., *Storage of cellular 5′ mRNA caps in P bodies for viral cap-snatching.* Proceedings of the National Academy of Sciences, 2008. **105**(49): p. 19294-19299.
160. Hopkins, K.C., et al., *Virus-induced translational arrest through 4EBP1/2-dependent decay of 5′-TOP mRNAs restricts viral infection.* Proceedings of the National Academy of Sciences, 2015. **112**(22): p. E2920-E2929.
161. Hopkins, K.C., et al., *A genome-wide RNAi screen reveals that mRNA decapping restricts bunyaviral replication by limiting the pools of Dcp2-accessible targets for cap-snatching.* Genes & development, 2013. **27**(13): p. 1511-1525.
162. Baird, N.L., J. York, and J.H. Nunberg, *Arenavirus Infection Induces Discrete Cytosolic Structures for RNA Replication.* Journal of Virology, 2012. **86**(20): p. 11301-11310.
163. Knopp, K.A., et al., *Single Nucleoprotein Residue Modulates Arenavirus Replication Complex Formation.* mBio, 2015. **6**(3): p. e00524-15.
164. Ng, A.K.-L., et al., *Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design.* The FASEB journal, 2008. **22**(10): p. 3638.
165. Ng, A.K., et al., *Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design.* Faseb j, 2008. **22**(10): p. 3638-47.
166. Perez, M., R.C. Craven, and J.C. de la Torre, *The small RING finger protein Z drives arenavirus budding: Implications for antiviral strategies.* Proceedings of the National Academy of Sciences, 2003. **100**(22): p. 12978-12983.

167. Shtanko, O., et al., *A role for the C terminus of Mopeia virus nucleoprotein in its incorporation into Z protein-induced virus-like particles.* Journal of virology, 2010. **84**(10): p. 5415-5422.
168. Capul, A.A., J.C.d.l. Torre, and M.J. Buchmeier, *Conserved Residues in Lassa Fever Virus Z Protein Modulate Viral Infectivity at the Level of the Ribonucleoprotein.* Journal of Virology, 2011. **85**(7): p. 3172-3178.
169. Rodríguez, L., et al., *Human respiratory syncytial virus matrix protein is an RNA-binding protein: binding properties, location and identity of the RNA contact residues.* Journal of general virology, 2004. **85**(3): p. 709-719.
170. Groseth, A., et al., *Efficient Budding of the Tacaribe Virus Matrix Protein Z Requires the Nucleoprotein.* Journal of Virology, 2010. **84**(7): p. 3603-3611.
171. Shtanko, O., et al., *ALIX/AIP1 is required for NP incorporation into Mopeia virus Z-induced virus-like particles.* Journal of virology, 2011. **85**(7): p. 3631-3641.
172. Winter, S.L., et al., *The Ebola virus VP40 matrix layer undergoes endosomal disassembly essential for membrane fusion.* The EMBO Journal, 2023. **42**(11): p. e113578.
173. Safo, M.K., et al., *Crystal Structures of Influenza A Virus Matrix Protein M1: Variations on a Theme.* PLoS ONE, 2014. **9**(10): p. e109510.
174. Harris, A., et al., *The crystal structure of the influenza matrix protein M1 at neutral pH: M1–M1 protein interfaces can rotate in the oligomeric structures of M1.* Virology, 2001. **289**(1): p. 34-44.
175. Fontana, J. and A.C. Steven, *At low pH, influenza virus matrix protein M1 undergoes a conformational change prior to dissociating from the membrane.* Journal of virology, 2013. **87**(10): p. 5621-5628.
176. Fontana, J., et al., *Structural changes in Influenza virus at low pH characterized by cryo-electron tomography.* Journal of virology, 2012. **86**(6): p. 2919-2929.
177. Stauffer, S., et al., *Stepwise Priming by Acidic pH and a High K⁺ Concentration Is Required for Efficient Uncoating of Influenza A Virus Cores after Penetration.* Journal of Virology, 2014. **88**(22): p. 13029-13046.
178. Günther, M., et al., *Interaction of host cellular factor ANP32B with matrix proteins of different paramyxoviruses.* Journal of General Virology, 2020. **101**(1): p. 44-58.
179. Bauer, A., et al., *ANP32B Is a Nuclear Target of Henipavirus M Proteins.* PLOS ONE, 2014. **9**(5): p. e97233.
180. Wang, Y., et al., *ANP32A and ANP32B are key factors in the Rev-dependent CRM1 pathway for nuclear export of HIV-1 unspliced mRNA.* Journal of Biological Chemistry, 2019. **294**(42): p. 15346-15357.
181. García-Sastre, A., *Ten strategies of interferon evasion by viruses.* Cell host & microbe, 2017. **22**(2): p. 176-184.
182. Niebling, S., et al., *Biophysical Screening Pipeline for Cryo-EM Grid Preparation of Membrane Proteins.* Frontiers in Molecular Biosciences, 2022. **9**.
183. Marty, M.T., et al., *Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles.* Analytical Chemistry, 2015. **87**(8): p. 4370-4376.
184. Tyanova, S., T. Temu, and J. Cox, *The MaxQuant computational platform for mass spectrometry-based shotgun proteomics.* Nature Protocols, 2016. **11**(12): p. 2301-2319.
185. Jumper, J., *Highly accurate protein structure prediction with AlphaFold. Nature https://doi. org/10.1038/s41586-021-03819-2 (2021).*

186. Bryant, P., G. Pozzati, and A. Elofsson, *Improved prediction of protein-protein interactions using AlphaFold2.* Nature communications, 2022. **13**(1): p. 1-11.
187. Malhotra, S., et al., *Assessment of protein–protein interfaces in cryo-EM derived assemblies.* Nature communications, 2021. **12**(1): p. 1-12.

7 Supplement

Table 20: Experimental mass determination of all measured species

All measured mass species are listed together with the theoretical mass (M_{th}), experimental mass (M_{exp}), standard deviation from the experimental mass (s) and FWHM. M_{exp} were determined from 3 independent MS measurements.

Mass species	M_{th} /Da	M_{exp} /Da	s (M_{exp}) /Da	FWHM /Da
NP	189168	189557	25	784
Z	11187	11347	38	12
R52A	62971	63900	82	433
R52A-Z	74158	75473	232	519
NP_1-$RNA12_1$	66840	67033	47	313
NP_2-$RNA12_2$	133680	134033	47	509
NP_3-$RNA12_3$	200520	201167	125	638
NP_3-$RNA12_3$	192952	193633	94	583
NP_4-$RNA12_4$	267360	268250	50	709
NP_5-$RNA12_5$	334200	335780	n/d	897
NP_6-$RNA12_6$	401040	393181	n/d	1116
NP_3-$RNA9_1$	192007	193000	82	1077
NP_3-Z_1	200355	201000	82	1258
NP_3-Z_2	211542	212367	125	330
NP_3-Z_3	222729	223867	170	319
NP_1-$RNA9_1$	65895	66085	22	784
NP_1-Z_1-$RNA9_1$	77082	77377	56	12
NP_3-$RNA18_1$	194983	195273	9	195
NP_1-$RNA18_1$	68871	68970	29	95
NP_2-$RNA18_1$	131927	132123	25	192
NP_3-$RNA18_2$	200798	201090	24	300
NP_1-$RNA25_1$	71025	71163	5	113
NP_2-$RNA25_1$	134081	134293	9	190
NP_3-$RNA25_2$	205106	205443	12	250
NP_4-$RNA25_2$	268162	268625	15	376
NP_3-$RNA27_1$	197732	198187	41	250
NP_1-$RNA27_1$	71620	71793	26	166
NP_2-$RNA27_1$	134676	134950	36	237
NP_3-$RNA27_2$	206296	206693	33	347
NP_4-$RNA27_2$	269352	269923	45	506
L-C	256893	255975	68	297
L-407	254500	254737	31	217
L-C-Z	268080	267653	177	636
L_C-Z-Z	279267	278785	55	207
L_C_RNA5'3'1-18	268386	267930	58	754
L_C_RNA5'3'1-18-Z	279573	279583	76	1000
NP1_RNA3'1-18	68763	69337	170	486
NP_2_RNA3'1-18_2	137526	138010	1	368
L_C_3'1-7-r20*$_2$		270223	17	565
L_C_3'1-7-$r20_1$	265.457	264790	15	547
NP_1+3'1-7-r20*		67757	31	167

Table 21: RNA sequences used in this study

The names, length and sequence are listed. The sequence is listed from the 5’ to the 3’ end.

RNA	Sequence (5’-3’)
9 nt RNA	UAGGAAUCU
12 nt RNA	ACACAAAGACCC
18 nt RNA	CGGACACACAAAAAGAAA
25 nt RNA	GCCUAGGAUCCACUGUGCGUGUUGU
27 nt RNA (LASV 3'1-7-r20)	ACCAACACACCAACAACAACCUGUGCG
LASV 3'1-18	GCGUGUCACCUAGGAUCC
LASV 5'1-18	CGCACCGGGGAUCCUAGG

Purification of ANP32A

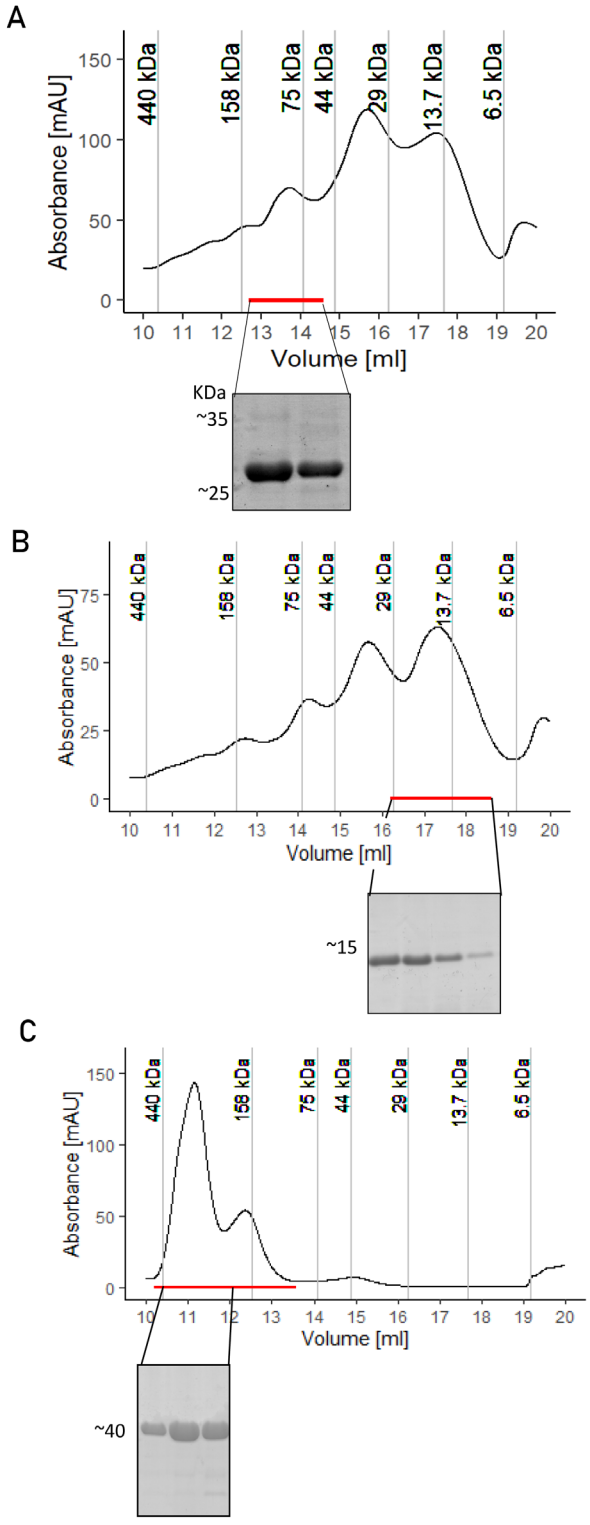

Figure 47: Purification of ANP32A and truncated versions

(A) Elution profile from SEC and SDS-PAGE of the respective fractions of the ANP32A purification. ANP32A was expressed as a fusion protein with GST-tag and purified via affinity chromatography and SEC. (B) ANP32A$_{LRR}$ domain was expressed and purified analogously to WT ANP32A. Elution profile and SDS PAGE of the ANP32$_{LRR}$ domain are visible. (C) ANP32A$_{LCAR}$ domain was expressed and purified as a fusion protein with GST-tag. Elution profile from SEC and SDS PAGE of the ANP32A$_{LCAR}$ domain. The data from this figure was produced by Elisabeth Eckelmann

Exemplary settings for nMS

Q Exactive™ UHMR Hybrid Quadrupole Orbitrap™

20240109 LS03 0.3 uM L-C +5 uM 3' rPur + 1.5 uM NP #117 RT: 0.0215
Total Ion Current: 1546390.50
Scan Low Mass: 1000.00
Scan High Mass: 15000.00
Scan Start Time (min): 2.38
Scan Number: 117
Base Peak Intensity: 20184.45
Base Peak Mass: 7788.56
Scan Mode: FTMS + p NSI sid=25.00 Full ms [1000.00-15000.00]
Q Exactive UHMR Orbitrap Data:
==========
Multiple Injection: ff
Multi Inject Info: IT=50;50
AGC: Off
Micro Scan Count: 10
Scan Segment: 0
Scan Event: 0
Master Index: 0
Charge State: 0
Monoisotopic M/Z: 0.0000
Ion Injection Time (ms): 100.000
Max. Ion Time (ms): 5.00
FT Resolution (m/z=200): 8750
FT Resolution (m/z=400): 6250
MS2 Isolation Width: 0.00
MS2 Isolation Offset: 0.00
AGC Target: 1000000
HCD Energy:
Analyzer Temperature: 28.54
=== Mass Calibration: ===:
Conversion Parameter B: 67840718.0293
Conversion Parameter C: -5960501.6528
Temperature Comp. (ppm): -4.03
RF Comp. (ppm): 0.00
Space Charge Comp. (ppm): -0.13
Resolution Comp. (ppm): 0.07
Number of Lock Masses: 0
Lock Mass #1 (m/z): 0.0000
Lock Mass #2 (m/z): 0.0000
Lock Mass #3 (m/z): 0.0000
LM Search Window (ppm): 0.0
LM Search Window (mmu): 0.0
Number of LM Found: 0
Last Locking (sec): 0.0
LM m/z-Correction (ppm): 0.00
=== Ion Optics Settings: ===:
S-Lens RF Level: 200.00
S-Lens Voltage (V): 21.00
Skimmer Voltage (V): 15.00

Inject Flatapole Offset (V): 5.00
Bent Flatapole DC (V): 2.00
MP2 and MP3 RF (V): 900.00
Gate Lens Voltage (V): 1.80
C-Trap RF (V): 2950.0
==== Diagnostic Data: ====:
APD: On
Dynamic RT Shift (min): 0.00
Intens Comp Factor: 0.6323
Res. Dep. Intens: 1.000
CTCD NumF: 0
CTCD Comp: 1.000
CTCD ScScr: 0.000
RawOvFtT: 138974.2
LC FWHM parameter: 30.0
Rod: 0
PS Inj. Time (ms): 0.000
AGC PS Mode: 0
AGC PS Diag: 0
HCD Energy eV: 0.000
AGC Fill: 1.00
Injection t0: 0.000
t0 FLP: 0.00
Access Id: 0
Analog Input 1 (V): 0.000
Analog Input 2 (V): 0.000
RF-Only: On
Extended Trapping: -1.000
Direct Mass: AIC Density Target (%): 0.00
Direct Mass: Manual Inject Time (ms): 0.00

QTof2

Aquisition Experiment Report

File:d:\mydocuments\lassa virus project\data\qtof2\20210930\20210930 ls06.raw
Header
Acquired File Name: 20210930 LS06
Acquired Date: 30-Sep-2021
Acquired Time: 15:12:36
Job Code:
Task Code:
User Name:
Laboratory Name:
Instrument: QTOF
Conditions:
Submitter:
SampleID:
Bottle Number:
Description: 18.2 uM SFTS CBD_His + 25 uM m7GTP_Cy5 msms 249

Instrument Calibration:
Calibration File:
Parameters

MS1 Static: None
MS1 Scanning:
Mass: 100 Da to 20000 Da.
Resolution: 0.0/0.0
Ion Energy: 0.0
Reference File: Csiesi
Acquisition File: 20210930 LS01
MS1 Scan Speed Compensation: None
MS2 Static: None
MS2 Scanning:
Mass: 100 Da to 20000 Da.
Resolution: 0.0/0.0
Ion Energy: 0.0
Reference File: Csiesi
Acquisition File: 20210930 LS01
MS2 Scan Speed Compensation: None
Calibration Time: 11:34:46
Calibration Date: 30-Sep-2021
Coefficients
MS1 Static: None
MS2 Static: None
Function 1: 1.000748089786*x +-0.049923547139, Root Mass

Parameters for C:\MassLynx\Qtof\tuneexp.exp

Temperature Correction Is Disabled
TDC Gain Control 0.0
TDC Amp Edge Control 0.0
Using 4 GHz TDC NO
Using TTP 4 GHz TDC NO

Instrument Parameters - Function 1:

Polarity	ES+	
Calibration	Dynamic 2	
Capillary	1.45	1.45
Cone 125	228	
Extractor	10	100
RF Lens	0.10	88.77
Source Temp (°C)	80	79
Desolvation Temp (°C)	20	1020
LM Resolution	0.0	
HM Resolution	0.0	
Collision Energy	50.0	89.0
Ion Energy	1.0	
Steering	0.44	-2.03
Entrance	65.0	-63.7
Pre-filter	5.0	84.9
Transport	15.0	-15.7
Aperture2	18.0	-18.5
Acceleration	200	-199
Focus 0	0	
Tube Lens	80	-79
Offset1	-0.9	-64.3
Offset2	0.0	-64.7

Pusher 980 924
TOF (kV) 9.10 -9.26
Reflectron 35.00 2.08
Pusher Cycle Time (µs) 120
Pusher Frequency (Hz) 8333.3333
Multiplier 550 -552
MCP 1900 1908
Centroid Threshold 0.0
Min Points 2.0
Np Multiplier 0.70
Resolution 4000.0
Lock Mass 0.0000
Mass Window +/- 1.0000
Lteff 1802.5600
Veff 9100.0000
1.0000
TDC Start (mV) 700.0000
TDC Stop (mV) 42.0000
TDC Threshold 0.0000
Pirani Pressure(mbar) 2.83e0
Penning Pressure(mbar) 1.03e-4
Tof Penning Pressure(mbar) 8.31e-7

Function Parameters - Function 1 - TOF MSMS FUNCTION
Set Mass 2049.0
Start Mass 100.0
End Mass 20000.0
Start Time (mins) 0.0
End Time (mins) 300.0
Data Format Continuum
Ion Mode ES Mode
Polarity Positive
Instrument Parameter Filename
C:\MassLynx\DEFAULT.PRO\ACQUDB\TOF_jh_2021 zerovoltage.ipr
Scans To Sum 1000000
Scan Time (sec) 1.0
Interscan Time (sec) 0.1
Use Tune Page Collision Energy YES
Collision Energy (eV) 30.0
Use Tune Page Cone Voltage YES
Cone Voltage (V) 35.0

Function 1
Scans in function: 436
Cycle time (secs): 1.100
Scan duration (secs): 1.000
Inter Scan Delay (secs): 0.100
Start and End Time(mins): 0.000 to 300.000
Ionization mode: ES+
Data type: Accurate Mass
Function type: TOF Daughter
Mass range: 100 to 20000
Collision Energy: 0.0

substance	hazard pictogram	hazard statements	precautionary statements
2-mercaptoethanol		H301 + H331 Toxic if swallowed or if inhaled. H310 Fatal in contact with skin. H315 Causes skin irritation. H317 May cause an allergic skin reaction. H318 Causes serious eye damage. H361d Suspected of damaging the unborn child. H373 May cause damage to organs (Liver, Heart) through prolonged or repeated exposure if swallowed. H410 Very toxic to aquatic life with long lasting effects.	P273 Avoid release to the environment. P280 Wear protective gloves/protective clothing/eye protection/face protection. P301 + P310 IF SWALLOWED: Immediately call a POISON CENTER/doctor. P302 + P352 + P310 IF ON SKIN: Wash with plenty of water. Immediately call a POISON CENTER/doctor. P304 + P340 + P311 IF INHALED: Remove person to fresh air and keep comfortable for breathing. Call a POISON CENTER/doctor. P305 + P351 + P338 IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing.
Acrylamide		H301 - Toxic if swallowed H312 + H332 - Harmful in contact with skin or if inhaled H315 - Causes skin irritation H319 - Causes serious eye irritation H317 - May cause an allergic skin reaction H372 - Causes damage to organs through prolonged or repeated exposure H350 - May cause cancer H340 - May cause genetic defects H361 - Suspected of damaging fertility or the unborn child	P301 + P310 - IF SWALLOWED: Immediately call a POISON CENTRE or doctor/physician P302 + P352 - IF ON SKIN: Wash with plenty of soap and water P304 + P340 - IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing P305 + P351 + P338 - IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing P333 + P313 - If skin irritation or rash occurs: Get medical advice/attention P260 - Do not breathe dust/fume/gas/mist/vapors/spray P202 - Do not handle until all safety precautions have been read and understood
acetic acid (64-19-7)		H226 Flammable liquid and vapor. H314 Causes severe skin burns and eye damage.	P210 Keep away from heat, hot surfaces, sparks, open flames, and other ignition sources. No smoking. P280 Wear protective gloves/protective clothing/eye protection/face protection. P303+P361+P353 IF ON SKIN (or hair): Take off immediately all contaminated clothing. Rinse skin with water. P305+P351+P338+P310 IF IN EYES: Rinse cautiously with water for several minutes. Remove contact

			lenses, if present and easy to do. Continue rinsing. Immediately call a poison center/doctor.
acetone		H225: Highly flammable liquid and vapor H319: Causes serious eye irritation H336: May cause drowsiness or dizziness	P210: Keep away from heat/sparks/open flames/hot surfaces. - No smoking. P233: Keep container tightly closed. P240: Ground/bond container and receiving equipment. P241: Use explosion-proof electrical/ventilating/lighting/equipment. P242: Use only non-sparking tools. P243: Take precautionary measures against static discharge. P261: Avoid breathing dust/fume/gas/mist/vapors/spray. P264+P265: Wash hands and face thoroughly after handling. P271: Use only outdoors or in a well-ventilated area. P280: Wear protective gloves/protective clothing/eye protection/face protection. P303+P361+P353: IF ON SKIN (or hair): Remove/Take off immediately all contaminated clothing. Rinse skin with water/shower. P304+P340: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. P305+P351+P338: IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. P319: Get medical attention if you feel unwell. P337+P317: If eye irritation persists: Get medical advice/attention. P370+P378: In case of fire: Use dry sand, dry chemical, or alcohol-resistant foam for extinction. P403+P233: Store in a well-ventilated place. Keep container tightly closed. P403+P235: Store in a well-ventilated place. Keep cool. P405: Store locked up. P501: Dispose of contents/container in accordance with local/regional/national/international regulations.
acetonitrile		H225 - Highly flammable liquid and vapor H302 - Harmful if swallowed H312 - Harmful in contact with skin H319 - Causes serious eye irritation H332 - Harmful if inhaled	P210 Keep away from heat/sparks/open flames/hot surfaces. - No smoking P280 - Wear protective gloves/ protective clothing/ eye protection/ face protection P301 + P312 - IF SWALLOWED: Call a POISON CENTER or doctor/ physician if you feel unwell P302 + P352 - IF ON SKIN: Wash with plenty of soap and water P304 + P340 - IF INHALED: Remove to fresh air and keep at rest in a position comfortable for breathing P305 + P351 + P338 - IF IN EYES: Rinse

			cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing
ammonium persulfate		H272 - May intensify fire; oxidize H302 - Harmful if swallowed H315 - Causes skin irritation H317 - May cause an allergic skin reaction H319 - Causes serious eye irritation H334 - May cause allergy or asthma symptoms or breathing difficulties if inhaled H335 - May cause respiratory irritation	P210 - Keep away from heat, hot surfaces, sparks, open flames and other ignition sources. No smoking P220 - Keep/Store away from clothing/ combustible materials P261 -Avoid breathing dust/fume/gas/mist/vapors/spray P264 - Wash hands thoroughly after handling P280 - Wear protective gloves/protective clothing/eye protection/face protection P301 + P312 - IF SWALLOWED: Call a POISON CENTER or doctor/physician if you feel unwell P302 + P352 - IF ON SKIN: Wash with plenty of soap and water P304 + P340 - IF INHALED: Remove person to fresh air and keep comfortable for breathing P305 + P351 + P338 - IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing P333 + P313 - If skin irritation or rash occurs: Get medical advice/attention P403 + P233 - Store in a well-ventilated place. Keep container tightly closed P501 - Dispose of contents/ container to an approved waste disposal plant
ammonium hydroxide solution		H314 Causes severe skin burns and eye damage. H335 May cause respiratory irritation. H410 Very toxic to aquatic life with long lasting effects.	P261 Avoid breathing dust/fume/gas/mist/vapors/spray. P271 Use only outdoors or in a well-ventilated area. P273 Avoid release to the environment. P280 Wear protective gloves/protective clothing/eye protection/face protection. P303+P361+P353 IF ON SKIN (or hair): Take off immediately all contaminated clothing. Rinse skin with water. P305+P351+P338 IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do.

			Continue rinsing.
Ampicillin		H315: Causes skin irritation H317: May cause an allergic skin reaction H319: Causes serious eye irritation H334: May cause allergy or asthma symptoms or breathing difficulties if inhaled H335: May cause respiratory irritation	P261: Avoid breathing dust/fume/gas/mist/vapors/spray. P264: Wash thoroughly after handling. P264+P265: Wash hands and face thoroughly after handling. P271: Use only outdoors or in a well-ventilated area. P272: Contaminated work clothing should not be allowed out of the workplace. P280: Wear protective gloves/protective clothing/eye protection/face protection. P284: Wear respiratory protection. P302+P352: IF ON SKIN: Wash with plenty of soap and water. P304+P340: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. P305+P351+P338: IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. P319: Get medical attention if you feel unwell. P321: Specific treatment (see medical advice on this label). P332+P317: If skin irritation occurs: Get medical advice/attention. P333+P313: If skin irritation or rash occurs: Get medical advice/attention. P337+P317: If eye irritation persists: Get medical advice/attention. P342+P316: If experiencing respiratory symptoms: Get medical advice/attention. P362+P364: Take off contaminated clothing and wash it before reuse. P403+P233: Store in a well-ventilated place. Keep container tightly closed. P405: Store locked up. P501: Dispose of contents/container in accordance with local/regional/national/international regulations.
Argon		H280: Contains gas under pressure; may explode if heated. H281: Contains refrigerated gas; may cause cryogenic burns or injury.	P403+P410: Store in a well-ventilated place. Protect from sunlight.
Carbenicillin disodium		H317: May cause an allergic skin reaction. H334: May cause allergy or asthma symptoms or breathing difficulties if inhaled.	P261: Avoid breathing dust/fume/gas/mist/vapors/spray. P272: Contaminated work clothing should not be allowed out of the workplace. P280: Wear protective gloves/protective clothing/eye protection/face protection. P284: Wear respiratory protection. P302+P352: IF ON SKIN: Wash with plenty of soap and water. P304+P340: IF INHALED: Remove victim to fresh air and keep at rest in a

			position comfortable for breathing. P321: Specific treatment (see medical advice on this label). P333+P313: If skin irritation or rash occurs: Get medical advice/attention. P342+P316: If experiencing respiratory symptoms: Get medical advice/attention. P362+P364: Take off contaminated clothing and wash it before reuse. P501: Dispose of contents/container in accordance with local/regional/national/international regulations.
Cesium Iodide		H315 Causes skin irritation. H317 May cause an allergic skin reaction. H319 Causes serious eye irritation. H335 May cause respiratory irritation. H410 Very toxic to aquatic life with long lasting effects.	P280 Wear protective gloves. P305+P351+P338 IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing.
Chloramphenicol		H317: May cause an allergic skin reaction. H318: Causes serious eye damage. H350: May cause cancer. H351: Suspected of causing cancer. H360: May damage fertility or the unborn child. H361: Suspected of damaging fertility or the unborn child.	P203: Keep container tightly closed. P261: Avoid breathing dust/fume/gas/mist/vapors/spray. P264+P265: Wash hands and face thoroughly after handling. P272: Contaminated work clothing should not be allowed out of the workplace. P280: Wear protective gloves/protective clothing/eye protection/face protection. P302+P352: IF ON SKIN: Wash with plenty of soap and water. P305+P354+P338: IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. P317: Get medical advice/attention. P318: IF exposed or concerned: Get medical advice/attention. P321: Specific treatment (see medical advice on this label). P333+P313: If skin irritation or rash occurs: Get medical advice/attention. P362+P364: Take off contaminated clothing and wash it before reuse. P405: Store locked up. P501: Dispose of contents/container in accordance with local/regional/national/international regulations.
Dimethyl sulfoxide		H315: Causes skin irritation. H319: Causes serious eye irritation. H335: May cause respiratory irritation.	P261: Avoid breathing dust/fume/gas/mist/vapors/spray. P264: Wash thoroughly after handling. P264+P265: Wash hands and face thoroughly after handling. P271: Use only outdoors or in a well-ventilated area. P280: Wear protective

			gloves/protective clothing/eye protection/face protection. P302+P352: IF ON SKIN: Wash with plenty of soap and water. P304+P340: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. P305+P351+P338: IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. P319: Get medical attention if you feel unwell. P321: Specific treatment (see medical advice on this label). P332+P317: If skin irritation occurs: Get medical advice/attention. P337+P317: If eye irritation persists: Get medical advice/attention. P362+P364: Take off contaminated clothing and wash it before reuse. P403+P233: Store in a well-ventilated place. Keep container tightly closed. P405: Store locked up. P501: Dispose of contents/container in accordance with local/regional/national/international regulations.
Dithiothreitol		H302: Harmful if swallowed. H315: Causes skin irritation. H319: Causes serious eye irritation. H335: May cause respiratory irritation.	P261: Avoid breathing dust/fume/gas/mist/vapors/spray. P264: Wash thoroughly after handling. P264+P265: Wash hands and face thoroughly after handling. P270: Do not eat, drink, or smoke when using this product. P271: Use only outdoors or in a well-ventilated area. P280: Wear protective gloves/protective clothing/eye protection/face protection. P301+P317: IF SWALLOWED: Get medical help. P302+P352: IF ON SKIN: Wash with plenty of soap and water. P304+P340: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. P305+P351+P338: IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. P319: Get medical attention if you feel unwell. P321: Specific treatment (see medical advice on this label). P330: Rinse mouth. P332+P317: If skin irritation occurs: Get medical advice/attention. P337+P317: If eye irritation persists: Get medical advice/attention. P362+P364: Take off contaminated clothing and wash it before reuse. P403+P233: Store in a well-ventilated place. Keep container tightly closed. P405: Store locked up. P501: Dispose of contents/container

			in accordance with local/regional/national/international regulations.
Ethylenediaminetetraacetic acid (EDTA)		H302: Harmful if swallowed. H312: Harmful in contact with skin. H315: Causes skin irritation. H319: Causes serious eye irritation. H332: Harmful if inhaled. H335: May cause respiratory irritation. H373: May cause damage to organs through prolonged or repeated exposure. H412: Harmful to aquatic life with long-lasting effects.	P260: Do not breathe dust/fume/gas/mist/vapors/spray. P261: Avoid breathing dust/fume/gas/mist/vapors/spray. P264: Wash thoroughly after handling. P264+P265: Wash hands and face thoroughly after handling. P270: Do not eat, drink, or smoke when using this product. P271: Use only outdoors or in a well-ventilated area. P273: Avoid release to the environment. P280: Wear protective gloves/protective clothing/eye protection/face protection. P301+P317: IF SWALLOWED: Get medical help. P302+P352: IF ON SKIN: Wash with plenty of soap and water. P304+P340: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. P305+P351+P338: IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. P317: Get medical advice/attention. P319: Get medical attention if you feel unwell. P321: Specific treatment (see medical advice on this label). P330: Rinse mouth. P332+P317: If skin irritation occurs: Get medical advice/attention. P337+P317: If eye irritation persists: Get medical advice/attention. P362+P364: Take off contaminated clothing and wash it before reuse. P403+P233: Store in a well-ventilated place. Keep container tightly closed. P405: Store locked up. P501: Dispose of contents/container in accordance with local/regional/national/international regulations.
Ethidium bromide		H302: Harmful if swallowed. H330: Fatal if inhaled. H341: Suspected of causing genetic defects.	P201: Obtain special instructions before use. P202: Do not handle until all safety precautions have been read and understood. P260: Do not breathe dust/fume/gas/mist/vapors/spray. P284: Wear respiratory protection. P301+P312: IF SWALLOWED: Call a POISON CENTER or doctor/physician if you feel unwell. P304+P340+P310: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. Immediately call a POISON CENTER or doctor/physician.

Formic acid		H226 Flammable liquid and vapour H290 May be corrosive to metals H302 Harmful if swallowed H314 Causes severe skin burns and eye damage H331 Toxic if inhaled	P210 Keep away from heat. No smoking. P280 Wear protective gloves/protective clothing/eye protection/face protection. P303+P361+P353 IF ON SKIN (or hair): take off immediately all contaminated clothing. Rinse skin with water/shower. P304+P340 IF INHALED: Remove person to fresh air and keep comfortable for breathing. P305+P351+P338 IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. P310 Immediately call a POISON CENTER/doctor.
Guanidine Hydrochloride (GndHCl)		H302 - Harmful if swallowed H332 - Harmful if inhaled H315 - Causes skin irritation H319 - Causes serious eye irritation	P301 + P330 + P331 - IF SWALLOWED: rinse mouth. Do NOT induce vomiting P312 - Call a POISON CENTER or doctor/ physician if you feel unwell P304 + P340 - IF INHALED: Remove to fresh air and keep at rest in a position comfortable for breathing P302 + P352 - IF ON SKIN: Wash with plenty of soap and water P337 + P313 - If eye irritation persists: Get medical advice/ attention P280 - Wear protective gloves/ protective clothing/ eye protection/ face protection
Hydrochloric acid		H314: Causes severe skin burns and eye damage. H331: Toxic if inhaled.	P260: Do not breathe dust/fume/gas/mist/vapors/spray. P261: Avoid breathing dust/fume/gas/mist/vapors/spray. P264: Wash thoroughly after handling. P271: Use only outdoors or in a well-ventilated area. P280: Wear protective gloves/protective clothing/eye protection/face protection. P301+P330+P331: IF SWALLOWED: Rinse mouth. Do NOT induce vomiting. P302+P361+P354: IF ON SKIN: Take off immediately all contaminated clothing. Rinse skin with water/shower. P304+P340: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. P305+P354+P338: IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. P316: Get emergency medical help for eye contact. P321: Specific treatment (see medical advice on this label). P363: Wash contaminated clothing before reuse. P403+P233: Store in a well-ventilated place. Keep container tightly closed. P405: Store locked up. P501: Dispose of contents/container in accordance with local/regional/national/international regulations.

HEPES		H315: Causes skin irritation. H319: Causes serious eye irritation. H335: May cause respiratory irritation.	P261: Avoid breathing dust/fume/gas/mist/vapors/spray. P264: Wash hands thoroughly after handling. P270: Do not eat, drink, or smoke when using this product. P271: Use only outdoors or in a well-ventilated area. P280: Wear protective gloves/protective clothing/eye protection/face protection. P301+P312: IF SWALLOWED: Call a POISON CENTER or doctor/physician if you feel unwell. P302+P352: IF ON SKIN: Wash with plenty of soap and water. P304+P312: IF INHALED: Call a POISON CENTER or doctor/physician if you feel unwell. P304+P340: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. P305+P351+P338: IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. P312: Call a POISON CENTER or doctor/physician if you feel unwell. P321: Specific treatment (see medical advice on this label). P322: Specific measures (see medical advice on this label). P330: Rinse mouth. P332+P313: If skin irritation occurs: Get medical advice/attention. P337+P313: If eye irritation persists: Get medical advice/attention. P362: Take off contaminated clothing and wash before reuse. P363: Wash contaminated clothing before reuse. P403+P233: Store in a well-ventilated place. Keep container tightly closed. P405: Store locked up. P501: Dispose of contents/container in accordance with local/regional/national/international regulations.
Kanamycin sulfate		H360: May damage fertility or the unborn child.	P203: Keep container tightly closed. P280: Wear protective gloves/protective clothing/eye protection/face protection. P318: IF exposed or concerned: Get medical advice/attention. P405: Store locked up. P501: Dispose of contents/container in accordance with local/regional/national/international regulations.
Lysozyme (egg white)		H334: May cause allergy or asthma symptoms or breathing difficulties if inhaled.	P261: Avoid breathing dust/fume/gas/mist/vapors/spray. P284: Wear respiratory protection. P304+P340: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. P342+P316: If experiencing

			respiratory symptoms: Get medical advice/attention. P501: Dispose of contents/container in accordance with local/regional/national/international regulations.
Methanol		H225: Highly flammable liquid and vapor. H301: Toxic if swallowed. H311: Toxic in contact with skin. H331: Toxic if inhaled. H370: Causes damage to organs.	P210: Keep away from heat/sparks/open flames/hot surfaces. - No smoking. P233: Keep container tightly closed. P240: Ground/bond container and receiving equipment. P241: Use explosion-proof electrical/ventilating/lighting/equipment. P242: Use only non-sparking tools. P243: Take precautionary measures against static discharge. P260: Do not breathe dust/fume/gas/mist/vapors/spray. P261: Avoid breathing dust/fume/gas/mist/vapors/spray. P264: Wash hands thoroughly after handling. P270: Do not eat, drink, or smoke when using this product. P271: Use only outdoors or in a well-ventilated area. P280: Wear protective gloves/protective clothing/eye protection/face protection. P301+P316: IF SWALLOWED: Get emergency medical help. P302+P352: IF ON SKIN: Wash with plenty of soap and water. P303+P361+P353: IF ON SKIN (or hair): Remove/Take off immediately all contaminated clothing. Rinse skin with water/shower. P304+P340: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. P308+P316: IF exposed or concerned: Get medical advice/attention. P316: Get emergency medical help. P321: Specific treatment (see medical advice on this label). P330: Rinse mouth. P361+P364: Take off immediately all contaminated clothing and wash it before reuse. P370+P378: In case of fire: Use dry sand, dry chemical, or alcohol-resistant foam for extinction. P403+P233: Store in a well-ventilated place. Keep container tightly closed. P403+P235: Store in a well-ventilated place. Keep cool. P405: Store locked up. P501: Dispose of contents/container in accordance with local/regional/national/international regulations.
Monopotassium phosphate		H315: Causes skin irritation. H319: Causes serious eye irritation.	P264: Wash hands thoroughly after handling. P280: Wear protective gloves/protective clothing/eye

			protection/face protection. P305+P351+P338: IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. P321: Specific treatment (see medical advice on this label). P332+P313: If skin irritation occurs: Get medical advice/attention. P337+P313: If eye irritation persists: Get medical advice/attention.
Nitrogen (liquid)		H280: Contains gas under pressure; may explode if heated. H281: Contains refrigerated gas; may cause cryogenic burns or injury.	P282: Wear cold insulating gloves/face shield/eye protection. P336+P317: If eye irritation persists: Get medical advice/attention. P403: Store in a well-ventilated place. P410+P403: Protect from sunlight. Store in a well-ventilated place.
Protease Inhibitor Cocktail cOmplete™, EDTA-free		H314: Causes severe skin burns and eye damage.	P260: Do not breathe dust/fume/gas/mist/vapors/spray. P280: Wear protective gloves/protective clothing/eye protection/face protection. P301+P330+P331: IF SWALLOWED: Rinse mouth. Do NOT induce vomiting. P303+P361+P353: IF ON SKIN (or hair): Remove/Take off immediately all contaminated clothing. Rinse skin with water/shower. P304+P340+P310: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. Immediately call a POISON CENTER or doctor/physician. P305+P351+P338+P310: IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. Immediately call a POISON CENTER or doctor/physician.
Pepsin immobilized on Resin beads		H334: May cause allergy or asthma symptoms or breathing difficulties if inhaled.	P261: Avoid breathing dust/fume/gas/mist/vapors/spray. P284: In case of inadequate ventilation, wear respiratory protection. P304+P340: IF INHALED: Remove person to fresh air and keep comfortable for breathing. P342+P311: If experiencing respiratory symptoms: Call a POISON CENTER or doctor/physician. P501: Dispose of contents/container to an approved waste disposal plant.w
Pierce® LTQ Velos ESI Positive Ion Calibration Solution	  	H225: Highly flammable liquid and vapor. H301 + H311 + H331: Toxic if swallowed, in contact with skin, or if inhaled. H319: Causes serious eye irritation. H370: Causes damage to organs.	P280: Wear protective gloves/protective clothing/eye protection/face protection. P210: Keep away from heat, hot surfaces, sparks, open flames, and other ignition sources. No smoking. P241: Use explosion-proof electrical/ventilating/lighting/equipment. P260: Do not breathe dust/fume/gas/mist/vapors/spray. P304+P340: IF INHALED: Remove person to fresh air and keep comfortable for breathing. P301+P310: IF SWALLOWED:

			Immediately call a POISON CENTER or doctor/physician. P303+P361+P353: IF ON SKIN (or hair): Remove/Take off immediately all contaminated clothing. Rinse skin with water/shower. P403+P235: Store in a well-ventilated place. Keep cool. P501: Dispose of contents/container to an approved waste disposal plant
Sodium dodecyl sulfate (SDS)		H228: Flammable solid. H302: Harmful if swallowed. H315: Causes skin irritation. H318: Causes serious eye damage. H319: Causes serious eye irritation. H332: Harmful if inhaled. H335: May cause respiratory irritation. H412: Harmful to aquatic life with long-lasting effects.	P210: Keep away from heat, hot surfaces, sparks, open flames, and other ignition sources. No smoking. P240: Ground/bond container and receiving equipment. P241: Use explosion-proof electrical/ventilating/lighting/equipment. P261: Avoid breathing dust/fume/gas/mist/vapors/spray. P264: Wash hands thoroughly after handling. P264+P265: Wash hands and face thoroughly after handling. P270: Do not eat, drink, or smoke when using this product. P271: Use only outdoors or in a well-ventilated area. P273: Avoid release to the environment. P280: Wear protective gloves/protective clothing/eye protection/face protection. P301+P317: IF SWALLOWED: Get medical help. P302+P352: IF ON SKIN: Wash with plenty of soap and water. P304+P340: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. P305+P351+P338: IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. P305+P354+P338: IF ON SKIN (or hair): Rinse with plenty of water. P317: Get medical advice/attention. P319: Get medical attention if you feel unwell. P321: Specific treatment (see medical advice on this label). P330: Rinse mouth. P332+P317: If skin irritation occurs: Get medical advice/attention. P337+P317: If eye irritation persists: Get medical advice/attention. P362+P364: Take off contaminated clothing and wash it before reuse. P370+P378: In case of fire: Use dry sand, dry chemical, or alcohol-resistant foam for extinction. P403+P233: Store in a well-ventilated place. Keep container tightly closed. P405: Store locked up. P501: Dispose of contents/container in accordance with local/regional/national/international

			regulations.
Sodium hydroxide (NaOH)		H314: Causes severe skin burns and eye damage.	P260: Do not breathe dust/fume/gas/mist/vapors/spray. P264: Wash hands thoroughly after handling. P280: Wear protective gloves/protective clothing/eye protection/face protection. P301+P330+P331: IF SWALLOWED: Rinse mouth. Do NOT induce vomiting. P302+P361+P354: IF ON SKIN: Take off immediately all contaminated clothing. Rinse skin with water/shower. P304+P340: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. P305+P354+P338: IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. P316: Get emergency medical help for eye contact. P321: Specific treatment (see medical advice on this label). P363: Wash contaminated clothing before reuse. P405: Store locked up. P501: Dispose of contents/container in accordance with local/regional/national/international regulations.
Sodium chloride		H319: Causes serious eye irritation.	P264+P265: Wash hands and face thoroughly after handling. P280: Wear protective gloves/protective clothing/eye protection/face protection. P305+P351+P338: IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. P337+P317: If eye irritation persists: Get medical advice/attention.
Tris base		H315: Causes skin irritation. H319: Causes serious eye irritation. H335: May cause respiratory irritation.	P261: Avoid breathing dust/fume/gas/mist/vapors/spray. P264: Wash thoroughly after handling. P264+P265: Wash hands and face thoroughly after handling. P271: Use only outdoors or in a well-ventilated area. P280: Wear protective gloves/protective clothing/eye protection/face protection. P302+P352: IF ON SKIN: Wash with plenty of soap and water. P304+P340: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. P305+P351+P338: IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. P319: Get medical attention if you feel unwell. P321: Specific treatment (see medical advice on this label).

			P332+P317: If skin irritation occurs: Get medical advice/attention. P337+P317: If eye irritation persists: Get medical advice/attention. P362+P364: Take off contaminated clothing and wash it before reuse. P403+P233: Store in a well-ventilated place. Keep container tightly closed. P405: Store locked up. P501: Dispose of contents/container in accordance with local/regional/national/international regulations.
Triton X-100		H302: Harmful if swallowed. H315: Causes skin irritation. H318: Causes serious eye damage. H319: Causes serious eye irritation. H400: Very toxic to aquatic life. H410: Very toxic to aquatic life with long-lasting effects. H411: Toxic to aquatic life with long-lasting effects. H412: Harmful to aquatic life with long-lasting effects.	P264: Wash thoroughly after handling. P264+P265: Wash hands and face thoroughly after handling. P270: Do not eat, drink, or smoke when using this product. P273: Avoid release to the environment. P280: Wear protective gloves/protective clothing/eye protection/face protection. P301+P317: IF SWALLOWED: Get medical help. P302+P352: IF ON SKIN: Wash with plenty of soap and water. P305+P351+P338: IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. P305+P354+P338: IF ON SKIN (or hair): Rinse with plenty of water. P317: Get medical advice/attention. P321: Specific treatment (see medical advice on this label). P330: Rinse mouth. P332+P317: If skin irritation occurs: Get medical advice/attention. P337+P317: If eye irritation persists: Get medical advice/attention. P362+P364: Take off contaminated clothing and wash it before reuse. P391: Collect spillage. P501: Dispose of contents/container in accordance with local/regional/national/international regulations.
Uranyl acetate		H300+H330: Fatal if swallowed or if inhaled. H300: Fatal if swallowed. H330: Fatal if inhaled. H373: May cause damage to organs through prolonged or repeated exposure. H411: Toxic to aquatic life with long-lasting effects.	P260: Do not breathe dust/fume/gas/mist/vapors/spray. P264: Wash thoroughly after handling. P270: Do not eat, drink, or smoke when using this product. P271: Use only outdoors or in a well-ventilated area. P273: Avoid release to the environment. P284: Wear respiratory protection. P301+P316: IF SWALLOWED: Get emergency medical help. P304+P340: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. P316: Get emergency medical help. P319: Get medical attention if you feel unwell.

			P320: Specific treatment is urgent (see supplemental first aid instructions on this label). P321: Specific treatment (see medical advice on this label). P330: Rinse mouth. P391: Collect spillage. P403+P233: Store in a well-ventilated place. Keep container tightly closed. P405: Store locked up. P501: Dispose of contents/container in accordance with local/regional/national/international regulations.
N,N,N',N'-Tetramethyl ethylenediamine		H225: Highly flammable liquid and vapor. H301: Toxic if swallowed. H314: Causes severe skin burns and eye damage. H318: Causes serious eye damage. H331: Toxic if inhaled.	P210: Keep away from heat, hot surfaces, sparks, open flames, and other ignition sources. No smoking. P280: Wear protective gloves/protective clothing/eye protection/face protection. P301+P310: IF SWALLOWED: Immediately call a POISON CENTER or doctor/physician. P303+P361+P353: IF ON SKIN (or hair): Remove/Take off immediately all contaminated clothing. Rinse skin with water/shower. P304+P340: IF INHALED: Remove victim to fresh air and keep at rest in a position comfortable for breathing. P403+P235: Store in a well-ventilated place. Keep cool. P501: Dispose of contents/container to an approved waste disposal plant.

Acknowledgements

First, I would like to thank my direct supervisors Dr. Maria Rosenthal and Prof. Dr. Charlotte Uetrecht for giving me the opportunity to do the PhD with this shared project and their excellent support. It was great to do this shared project in two groups from which I learned a lot from two different perspectives. Additionally, giving me the time to develop both professionally and personally. I would also like to thank my third supervisor, Prof. Dr Kay Grünewald from the University of Hamburg. I would also like to thank Prof. Hartmut Schlüter at the UKE, for using the mass spectrometers at UKE and the second reviewer for my thesis and Jan Kosinski for providing the structure predictions.

Special thanks go to the group of Structural Virology at BNITM for their help and great working atmosphere. To my PhD student soulmates Kristina, Janna, Selina and Saskia for the great time. Harry for helping me with all the microscopy nonsense :) and the Bermuda nights. Dominik and Silke for all their support. My special thanks also to the lab bosses Carola and Morlin for the great time and big support.

I would also thank the whole virology department for the nice working atmosphere and the fruitful discussions.

On the other part of Hamburg, many thanks to the group of Dynamics of Viral Structures at CSSB and LIV. Here I learned everything about mass spec. Special thanks to Janine who keeps the shop open and running in every aspect. It was great working with you and doing partial derivatives! Jasmin, for bringing HDX-MS to me and Boris showing me nMS for the first time! Many thanks also to the other PhD student soulmates Lars, Marius, Jonas and Kira for the great time and funny moments. Also, a big thank you to all the other members of the group.

And of course, a big thank you to all my friends and family, especially Anne for all your support!

Declaration of authorship

Declaration of authorship | Eidesstattliche Versicherung

I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids. The submitted written document corresponds to the file on the electronic storage medium. I further declare that this thesis has not been presented previously to another examination board.

Hiermit versichere ich an Eides statt, die vorliegende Dissertation selbst verfasst und keine anderen als die angegebenen Hilfsmittel benutzt zu haben. Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium. Ich versichere, dass diese Dissertation nicht in einem früheren Promotionsverfahren eingereicht wurde

05.04.2024