Insights into the structure of Lloviu cuevavirus nucleoprotein–RNA complex through cryo-EM

胡 上帆

CONTENTS

page

Abstract 3

Abbreviations 4

Introduction 6

1.1. Filovirus disease distribution and classification. 7
1.2. Emergence, distribution, and pathogenicity of the Lloviu virus. 9
1.3. Structural proteins of the Lloviu viruses. 10
1.4. Lloviu virus replication cycle. 12
1.5. Structure and assembly of filovirus nucleocapsid. 14
1.6. Purpose of this study. 16

Materials and Methods 17

Results 27

3.1. Examination of the helical structure of the LLOV NP–RNA complex. 28
3.2. Architecture of the LLOV NP–RNA complex. 30
3.3. Filovirus NP–RNA protomers exhibit similar structures but different helical arrangements in the helix. 32
3.4. RNA-binding region of LLOV NP. 33
3.5. LLOV NP-NP interactions in the helix. 37
3.6. Residues essential for helical assembly and vRNA synthesis. 42

Discussion 46

Bibliography 50

Acknowledgements 57

Abstract

Lloviu cuevavirus (LLOV) is a novel filovirus detected in Schreiber's bats in Europe and was the first identified member of the *Filoviridae* family outside the *Ebolavirus* (EBOV) and *Marburgvirus* (MARV) genera. Although EBOV and MARV cause severe hemorrhagic fever in humans and nonhuman primates, the pathogenicity of LLOV in humans remains unclear. Recent studies using recombinant LLOV infectious particles and isolated LLOV have confirmed the zoonotic potential of the virus *in vitro*, raising public health concerns. However, the virological and molecular characteristics of LLOV remain unclear.

The LLOV nucleoprotein (NP) encapsidates the single-stranded, viral genomic RNA (vRNA) to form a helical NP–RNA complex, which constitutes a nucleocapsid along with viral proteins VP24, VP30, VP35, and polymerase L. RNA encapsidation and NP oligomerization are required for the formation of the helical NP–RNA complex, yet their molecular mechanisms are not well understood. Recent cryo-electron microscopy (cryo-EM) single-particle analyses have resolved the high-resolution structures of recombinant EBOV NP (residues 1–450)–RNA and MARV NP (residues 1–395)–RNA complexes. However, further structural details of the filovirus nucleocapsid are required to fully understand the filovirus replication cycle.

In this study, to further delineate the structural basis for the formation of the helical LLOV NP–RNA complex, I determined two structures of the LLOV NP–RNA helical complex, comprising a full-length (residues 1–749) and a C-terminal truncated (residues 1–450) NP, using cryo-EM single-particle analysis. The two helical structures were identical, confirming the dominant role of the NP N-terminal region during helical assembly. The LLOV NP–RNA protomers displayed a structure similar to that of EBOV and MARV, however, the spatial arrangements in the helix differed. Structure-based mutational analysis was used to determine the residues involved in helical assembly and vRNA synthesis. The results deepen the understanding of filovirus nucleocapsid formation and may be further utilized toward the design of broad-spectrum antiviral drugs.

Abbreviations

Lloviu cuevavirus	LLOV
Ebola virus	EBOV
Marburg virus	MARV
Nucleoprotein	NP
Viral genomic RNA	vRNA
Cryo-electron microscopy	Cryo-EM
Marburgvirus disease	MVD
Ebola virus disease	EVD
Xīlǎng virus	XILV
Huángjiāo virus	HUJV
Měnglà virus	MLAV
Oberland virus	OBLV
Tapajós virus	TAPV
Viral protein 35,	VP35
Viral protein 24	VP24
Viral protein 30	VP30
Glycoprotein	GP
RNA-dependent RNA polymerase	L
Human macrophage galactose-type C-type lectin	hMGL
Dendritic cell-specific ICAM-3-grabbing nonintegrin	DC-SIGN
Niemann Pick C1	NPC1
Endosomal sorting complex required for transport	ESCRT
Transmission electron microscopy	TEM
Full-length NP	NP(full)
C-terminally truncated NP	NP(1–450)
Human embryonic kidney 293T	HEK 293T
Root-mean-square deviation	RMSD

Host RNA	hRNA
Complementary RNA	cRNA
Protein data bank	PDB
Electron microscopy data	EMD

Chapter 1

Introduction

1.1. Filovirus disease, distribution and classification.

Filoviruses such as the Marburg virus (MARV) and Ebola virus (EBOV) belong to the family *Filoviridae* [1]. This group of single-stranded RNA viruses can cause severe hemorrhagic fever in humans and nonhuman primates with a high fatality rate (up to 90%) [2, 3]. Since the first case of Marburgvirus disease (MVD) was detected in 1967 [4], over 40 outbreaks of filovirus diseases, including MVD and Ebola virus disease (EVD), have been recorded, predominantly in central Africa [5]. Most outbreaks typically occur in remote areas of central Africa and can be contained rapidly through isolation of infected individuals and contact tracing. However, the largest outbreak of EVD in West Africa from 2013 to 2016 spread to ten countries, causing over 28,000 infections and 11,000 fatalities [5,6], highlighting the need for preventive measures and possible treatments to control filoviruses.

Filoviruses were initially considered to exist exclusively in Africa, yet recent advancements in RNA sequencing technologies have revealed that they are more widely distributed, including in Asia [Xīlǎng virus (XILV), Huángjiāo virus (HUJV) and Měnglà virus (MLAV)], Europe [Lloviu virus (LLOV) and Oberland virus (OBLV)], and South America [Tapajós virus (TAPV)] (Table. 1) [7-10]. The genetic diversity and zoonotic potential of these newly identified filoviruses present a significant challenge for the development of virus-targeting prophylactic and therapeutic countermeasures.

Table 1. Distribution, natural host, and pathogenicity in humans of the family *Filoviridae*

Genus	Species	Endemic region	Natural host	Pathogenicity in humans
Ebolavirus	*Zaire ebolavirus*	Africa	Fruit bat?	High
	Sudan ebolavirus	Africa	Unknown	High
	Taï Forest ebolavirus	Africa	Unknown	High
	Reston ebolavirus	Asia	Fruit bat?	Asymptomatic
	Bundibugyo ebolavirus	Africa	Unknown	High
	Bombali ebolavirus	Africa	Free-tailed bat	Unknown
Marburgvirus	*Marburg virus*	Africa	Fruit bat	High
	Ravn virus	Africa	Fruit bat	High
Cuevavirus	*Lloviu cuevavirus*	Europe	Free-tailed bat	Unknown
Striavirus	*Xīlǎng striavirus*	Asia	Frogfish	Unknown
Oblavirus	*Oblavirus percae*	Europe	Perch	Unknown
Thamnovirus	*Huángjiāo thamnovirus*	Asia	Filefish	Unknown
	Thamnovirus percae	Europe	Perch	Unknown
	Thamnovirus kanderense	Europe	Perch	Unknown
Dianlovirus	*Měnglà dianlovirus*	Asia	Fruit bat	Unknown
Not assigned	*Tapajós virus*	South America	Viper	Unknown

1.2. The emergence, distribution and pathogenicity of the Lloviu virus

In 2002, a mass-mortality incident occurred in Schreiber's bat (*Miniopterus schreibersii*) colonies in the caves of France, Spain, and Portugal. Bat carcasses from the Cueva del Lloviu cave (Asturias, Spain) were collected, and a nearly complete LLOV genomic sequence was identified from bat carcasses using deep sequencing and PCR technologies [7]. At the sequence level, LLOV is clearly distinct from EBOV and MARV; therefore, a new filovirus genus, *Cuevavirus* [11], comprising only LLOV, was proposed. The LLOV was the first filovirus not imported from an endemic area of Africa.

After more than a decade, the viral genome was again detected in Schreiber's bats in 2016 in northeast Hungary, along with two additional —possibly associated— mass mortality events from 2013 to 2016, suggesting its wide distribution and continuous circulation in Europe [12]. Recently, two infectious LLOV were successfully isolated from the blood samples of naturally infected Schreiber bats in Hungary and Italy (Fig. 1) [13, 14]. While bats are suspected to be natural hosts of EBOV, the recent isolation of LLOV, as well as the isolation of MARV and Ravn viruses from bats [15], has confirmed the role of bats as natural hosts for filoviruses.

Although the pathogenicity of LLOV in humans remains unknown, studies using recombinant LLOV infectious particles and LLOV isolates have shown that LLOV can infect and replicate in human and nonhuman primate-derived cell lines and primary human macrophages *in vitro* [13, 16]. Based on these data, LLOV is considered a potential zoonotic virus in humans.

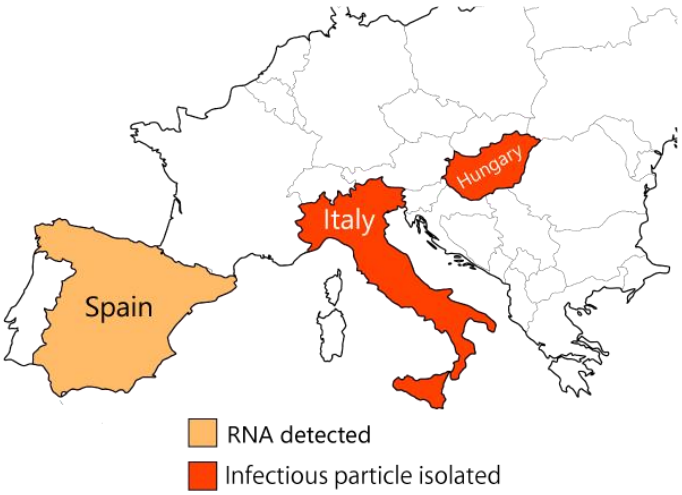


Fig 1. Geographic distribution of LLOV. The map shows the currently known distribution of the LLOV. The LLOV RNA was detected in three countries (colored orange and red). The countries from which LLOV infectious particles were isolated are colored red.

1.3. Structural proteins of the Lloviu virus

Similar to other members of the *Filoviridae* family, LLOV possesses filamentous particles (Fig. 2b) [16, 17]. Seven viral proteins including nucleoprotein (NP), polymerase cofactor (viral protein 35, VP35), matrix protein (VP40), glycoprotein (GP), transcriptional activator (VP30), nucleocapsid-associated protein (VP24), and RNA-dependent RNA polymerase (L) were encoded in its non-segmented, single-stranded, and negative-sense RNA genome (Fig. 2a) [1, 7, 18]. The helical nucleocapsid, containing genomic RNA, NP, VP35, VP30, VP24, and L, resides in the central axis of the filamentous particles (Fig. 2b) [1]. VP40 associates with the lipid envelope in a butterfly-like dimeric form, with two N-terminal domains interacting with each other, and C-terminal basic patches binding to the lipid membrane to form an inner layer underneath the host-derived envelope [19]. The viral particles are covered with GPs composed of GP1 and GP2, forming spike structures on the surface [18, 20, 21].

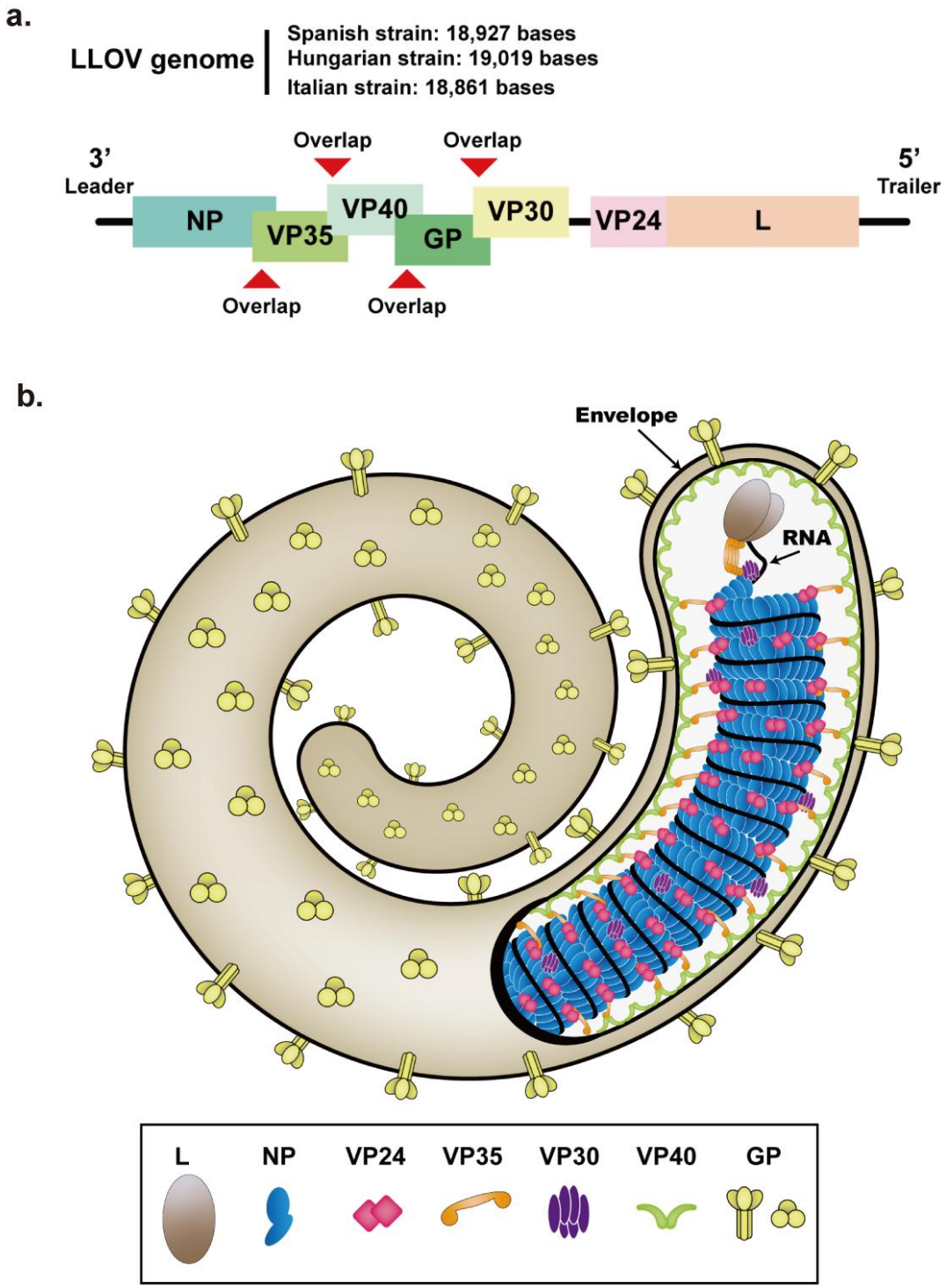


Fig 2. Genome and overall particle structure of LLOV.

(a) Schematic of the LLOV genome. (b) Schematic of LLOV architecture. The host-derived lipid envelope (brown outer layer) is covered with GP trimers. VP40 interacts with the envelope in a dimeric form, creating a lining underneath the envelope. The nucleocapsid, which is composed of the NP–RNA helix, VP35, VP24, VP30, and L, resides on the central axis of the filamentous particle.

1.4. Lloviu virus replication cycle

The LLOV replication cycle in host cells shares many similarities with that of EBOV and MARV. This involves multiple steps, including entry, uncoating, genome transcription and replication, nucleocapsid assembly, virion formation, and budding (Fig. 3). Host cell entry is the first step in the viral life cycle. The entry of LLOV particles into the cell is mediated by the interaction of GP1 with various potential cell receptors/co-receptors, such as human macrophage galactose-type C-type lectin (hMGL) and dendritic cell-specific ICAM-3-grabbing nonintegrin (DC-SIGN) [16, 17]. Following attachment, the viral particles enter the cell mainly through macropinocytosis and are subsequently transported to late endosomes [22, 23]. In endosomes, low pH conditions and cathepsin L (host cysteine proteases) process the cleavage step of GP and facilitate its binding to the second receptor, endosomal membrane protein Niemann Pick C1 (NPC1) [24]. The GPs of all four mammalian filoviruses (EBOV, MARV, LLOV, and MLAV) can use human NPC1 as their entry receptor, suggesting that the use of NPC1 during viral entry is a universal property of filoviruses [9, 24-27]. The GP1–NPC1 interaction leads to fusion between the viral envelope and endosomal membrane, which is mediated by GP2 [25-27]. After membrane fusion, the nucleocapsid is released into the cytoplasm, where the viral RNA-dependent RNA polymerase consisting of L and VP35 transcribes the viral genes into mRNAs [28, 29]. For EBOV and LLOV, four viral proteins, NP, VP35, L, and VP30, are required for the transcription and replication of viral genomic RNA (vRNA) [30-32]. In contrast, three viral proteins, NP, VP35, and L, are sufficient for MARV vRNA transcription and replication [33]. The vRNA synthesis and nucleocapsid assembly occur in “viral factories”, such as inclusion bodies [30, 34,35]. VP24 binding negatively regulates vRNA synthesis and induces mature nucleocapsid formation [36-38]. The mature nucleocapsid is then transported via multivesicular bodies and the actin network to the budding site at the plasma membrane [39, 40]. Progeny virion budding is the final step of the replication cycle and is mainly driven by VP40 via the cellular endosomal sorting complex required for the transport (ESCRT) pathway [41, 42].

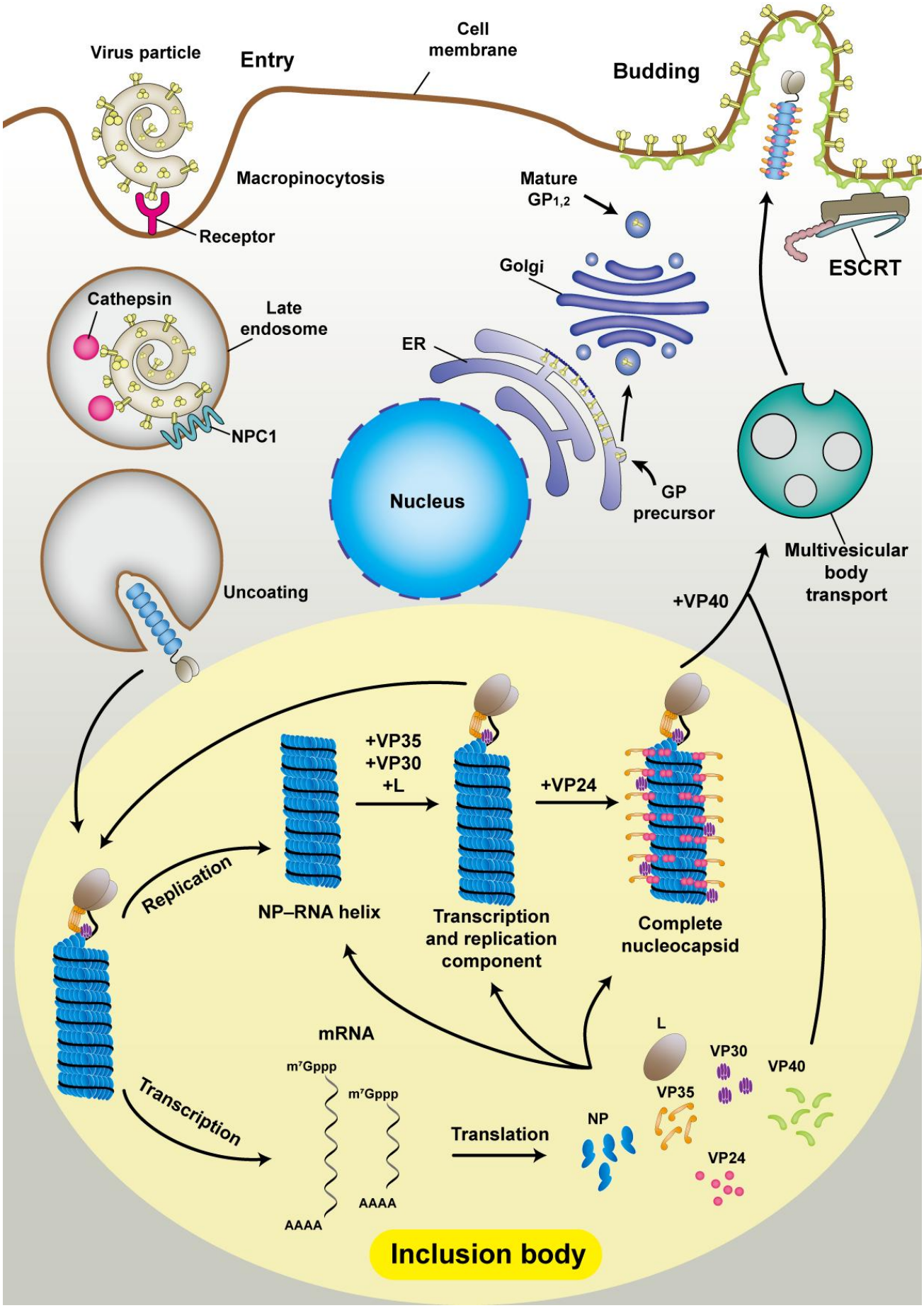

Virus particle
Entry
Cell membrane
Budding
Macropinocytosis
Receptor
Mature GP1,2
Golgi
ESCRT
Cathepsin
Late endosome
NPC1
ER
Nucleus
GP precursor
Multivesicular body transport
Uncoating
+VP40
+VP35
+VP30
+L
+VP24
Replication
NP–RNA helix
Transcription and replication component
Complete nucleocapsid
mRNA
m7Gppp
m7Gppp
AAAA
AAAA
Transcription
Translation
L
VP30
VP40
VP35
NP
VP24
Inclusion body

Fig 3. Schematic diagram of the LLOV replication cycle.

Initially, GP attaches to cellular receptors, facilitating the entry of viral particles into cells through macropinocytosis. Subsequently, the viral particles are trafficked to the late endosome. In the late endosome, GP undergoes cleavage by cathepsins, binds to NPC1, and undergoes conformational changes, resulting in the release of nucleocapsids into the cytoplasm via fusion between the viral envelope and the endosomal membrane. Both vRNA synthesis and nucleocapsid assembly occur in specific liquid organelles called inclusion bodies. The vRNA in the nucleocapsid is transcribed and replicated to generate mRNA and vRNA, respectively. The NP encapsidates the newly synthesized vRNA and forms the NP–RNA helix in the initial stage of nucleocapsid formation. Following the binding of VP35, VP30, and L, the functional nucleocapsid responsible for transcription and replication of the viral genome is assembled. Binding of VP24 to the functional nucleocapsid results in the formation of a structurally mature nucleocapsid. VP40 then transports the mature nucleocapsid to the budding site through the multivesicular body pathway and interacts with ESCRT to enable budding of the progeny virions.

1.5. Structure and assembly of the filovirus nucleocapsid

The NP is the main component of the filovirus nucleocapsid [1, 43]. NP homo-oligomerization, along with RNA binding, produces helical NP–RNA structures. The helical NP–RNA complex acts as a scaffold for binding other viral proteins to form functional nucleocapsids [32, 33]. Transmission electron microscopy (TEM) analyses have shown that the N-terminal domain of NP (EBOV residues 1–450 and MARV residues 1-395) is sufficient to form helical NP–RNA complexes [44, 45]. Cryo-electron microscopy (cryo-EM) single-particle analysis of the EBOV NP (1–450 residues)–RNA complexes and MARV NP (1-394 residues)–RNA complexes revealed that the N-terminal domain of an NP molecule can be divided into an N-terminal arm, N-terminal lobe, and C-terminal lobe (Fig. 4a, b, d, e) [46-48]. An RNA segment composed of six nucleotides was encapsidated in the positively charged cleft between the N-terminal- and C-terminal lobes (Fig. 4b, c, e, f) [46-48]. The C-terminal domain of the NP contains an intrinsically disordered linker region and a C-terminal tail (Fig. 4a, d) [49]. The high flexibility of the C-terminal domain disrupts the rigid helical structure and leads to the formation of loosely coiled helices. Because of that, structural

analyses using cryo-EM have been conducted on C-terminal truncated NPs, which explains the lack of the whole structure of NP [45].

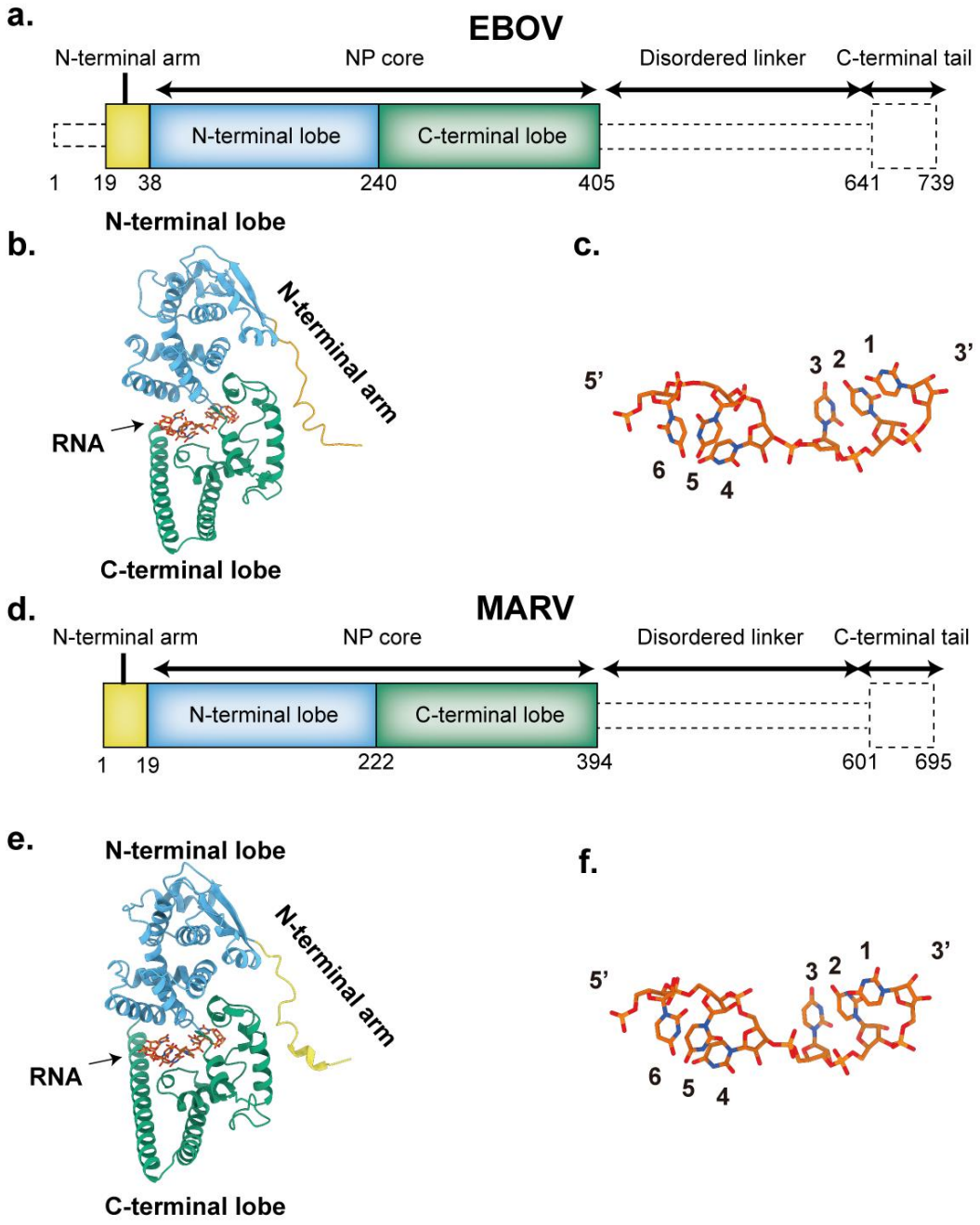


Fig 4. Overall cryo-EM structure of the filovirus NP–RNA complex and its atomic model.

(a, d) Schematic diagram of the domain architecture of EBOV and MARV NP. Each domain is represented by a unique color, and the unsolved region is represented by a dotted outline without color. (b, e) Atomic models of the EBOV NP(1–450 residues)–RNA complex (PDB-ID: 5Z9W) and MARV NP (1-394 residues)–RNA complex (PDB-ID: 7F1M). (c, f) RNA segments encapsidated in EBOV and MARV NP.

1.6. Purpose of this study

Research on filovirus biology has been severely delayed, because all mammalian filoviruses are classified as biosafety level-four (BSL-4) pathogens and cannot be handled outside of BSL-4 facilities. In this study, to advance the understanding of filovirus replication mechanisms, I employed a mammalian expression system in BSL-2 and determined two helical LLOV NP–RNA structures comprising a full-length NP and a C-terminal truncated NP using cryo-EM single particle analysis. Structure-based mutational analyses have determined various key residues involved in helical assembly as well as viral transcription and replication. The findings could contribute to the development of broad-spectrum structure-based antiviral therapies.

Chapter 2

Materials and Methods

2.1. Cells

Human embryonic kidney 293T (HEK 293T) cells were grown in Dulbecco's Modified Eagle's medium supplemented with 10% fetal calf serum (FCS) and 1% penicillin/streptomycin (Nacalai Tesque). The FCS was inactivated at 56 °C, for 30 min before use. The cells were incubated at 37 ℃, and supplied with 5% CO_2.

2.2. Plasmids

All LLOV plasmids (GenBank accession code YP004928135.1) encoded wild-type proteins (pCAGGS-NP, pCAGGS-VP35, pCAGGS-VP30, and pCAGGS-L). The pCAGGS-NP(1–450) that expresses residues 1–450 of the LLOV NP was constructed from pCAGGS-NP by PCR using primers with digestion enzyme sites. The Kozak sequence was added before the initiation codon. The LLOV NP mutants were generated using recombinant PCR and cloned into the pCAGGS vector. A C-terminal myc tag was added via PCR. Primers used in this study are listed in Table S3. All the constructs were sequenced to confirm the absence of unwanted mutations.

Table 2. Primers used in this study

Name	Forward /Reverse	Oligo sequence (5′ to 3′)	Purpose
XmaI_kozak_LLOVNP_F	Forward	GATCCCCGGGGCCGCCACCATGAATCGTTATCTTGGTCATGG	insert
LLOVNP(1–450)_NheI_R	Reverse	GATCGCTAGCTTAAGTGTAGGAGGATGACTCTTTG	insert
LLOV-NP_fwd	Forward	GATCCTCGAGTTATGAATCGTTATCTTGGTCATGGTAC	insert
LLOV-NP-M222A-fwd	Forward	GGCATGCACATGGCGGCTGGGCATGATGCTAATGATGC	mutation introduction
LLOV-NP-M222A-rev	Reverse	ATGCCCAGCCGCCATGTGCATGCCCTGGTG	mutation introduction
LLOV-NP-M222E-fwd	Forward	GGCATGCACATGGAAGCTGGGCATGATGCTAATGATGC	mutation introduction
LLOV-NP-M222E-rev	Reverse	ATGCCCAGCTTCCATGTGCATGCCCTGGTG	mutation introduction
LLOV-NP-H102-fwd	Forward	CATTATTTAAAAGGAGCGGGTATCCACTTCGAAATCCAGCATCG	mutation introduction
LLOV-NP-H102-rev	Reverse	GTGGATACCCGCTCCTTTTAAATAATGAGCTGCAGGAC	mutation introduction

LLOV-NP-R37A-fwd	Forward	CATACCATTGTAGCGAAGAAGAGCATCCCGTTGTTTGAG	mutation introduction
LLOV-NP-R37A-rev	Reverse	CGGGATGCTCTTCTTCGCTACAATGGTATGATCTACATTTAACCC	mutation introduction
LLOV-NP-I24E-fwd	Forward	GAATTACATGGTGAACTAAGCCTCGGGTTAAATGTAGATC	mutation introduction
LLOV-NP-I24E-rev	Reverse	CCCGAGGCTTAGTTCACCATGTAATTCACTCAGGTTGG	mutation introduction
LLOV-NP-H110A-fwd	Forward	CACTTCGAAATCCAGGCGCGTGATAATGTTGATCACATC	mutation introduction
LLOV-NP-H110A-rev	Reverse	CAACATTATCACGCGCCTGGATTTCGAAGTGGATACCATG	mutation introduction
LLOV-NP-370-382-fwd	Forward	CTTGGACTGGATGCGCAAGAAAGGGCGATCCTTGCAACATTCCACTCCGCGAAAAATGAGATCAAC	mutation introduction
LLOV-NP-370-382-rev	Reverse	GTTGATCTCATTTTTCGCGGAGTGGAATGTTGCAAGGATCGCCCTTTCTTGCGCATCCAGTCCAAG	mutation introduction
LLOV-NP-I24E-bbone_fwd	Forward	GATCCTCGAGTTCTGTCTCATCATTTTGGCAAAGAATTCCTTGAGCGTCTCTCATG	mutation introduction
LLOV-NP-Cmyc_n_rev	Reverse	AATCGCTAGCAACTACAGATCCTCTTCTGAGATGAGTTTTTGTTCGTTATTTATCAAGAAAAGGTACCGTTCC	insert
LLOV_557F	Forward	AATATCCCACCCAATGGCAG	sequence
LLOV_1220F	Forward	AGGCCCTCAATGAAGAAAAG	sequence

2.3. Expression and purification of the NP–RNA complex

For the LLOV NP(full)–RNA complex and its mutant, HEK 293T cells were seeded in two 10 cm dishes (4×10^6 cells/dish) and incubated at 37 ℃ for 24 h. The cells were transfected with 15 μg/dish pCAGGS-LLOV NP(full), using 45 μl/dish polyethylenimine (PEI) MAX (Polysciences) and 1 ml/dish Opti-MEM (Gibco). Two days after transfection, the cells were collected in 1× Dulbecco's Phosphate Buffered Saline (D-PBS) (-) (Nacalai Tesque), containing 2.68 mM KCl, 136.89 mM NaCl, 8.1 mM Na_2HPO_4, 1.46 mM KH_2PO_4, and further lysed with 0.1% (v/v) Nonidet P-40 (Wako) in Tris-HCl buffer [10 mM Tris-HCl, pH 8.0 (Invitrogen); 150 mM NaCl (Nacalai Tesque); 25 mM $MgCl_2$ (Wako); 5 mM $CaCl_2$ (Wako); 1 mM Dithiothreitol (Nacalai Tesque)], supplemented with protease inhibitor cocktail tablets (EDTA-free, Roche) and 10 mM ribonucleoside-vanadyl complex (New England Biolabs, NEB). The lysate was centrifuged at 3200 ×

g at 4 ℃ for 10 min to remove insoluble compounds. The supernatant was loaded onto a discontinuous 25–40% (w/w) CsCl (Wako) gradient and centrifuged at 191,600 × g at 4 °C for 5 h. Fractions containing the NP–RNA complex were collected and centrifuged at 48,500 × g at 4 °C for 1 h. The pellets were suspended in the Tris-HCl buffer and stored at 4 ℃.

For the LLOV NP(1–450)–RNA complex, the HEK 293T cells were seeded in four 15 cm dishes (1×10^7 cells/dish), and incubated at 37 ℃ for 24 h. The cells were transfected with 30 μg/dish pCAGGS-LLOV NP(1–450), using 120 μl/dish PEI MAX and 2 ml/dish Opti-MEM. Three days post-transfection, NP–RNA complexes were purified as per the method described above.

2.4. Negative-stain TEM

Aliquots (5 μl) of purified samples were placed on a carbon-coated copper grid (Thick & Thin Bar Sq 600 mesh Cu, Gilder), and hydrophilized through glow-discharge, prior to staining with 2% uranyl acetate. Images were obtained using an HT-7700 transmission electron microscope (Hitachi High-Tech) operated at 80 kV with an XR81-B CCD camera.

2.5. Cryo-EM specimen preparation.

For the LLOV NP(full)–RNA complex, a 2.5 μl-aliquot of purified NP–RNA complex solution at a concentration of 1 mg/ml, was applied to Quantifiol R1.2/1.3 200 mesh grids and blotted for 7 s, with a humidity of 100% at 4 ℃, before being rapidly frozen with liquid ethane using the Vitrobot Mark IV system (Thermo Fisher Scientific). Images of the frozen sample were acquired as movies on a Titan Krios cryo-TEM (Thermo Fisher Scientific) operated at 300 kV and equipped with a Falcon3EC direct electron detector (Thermo Fisher Scientific) and a Cs corrector (CEOS GmbH) at the Institute for Protein Research, Osaka University.

For the LLOV NP(1–450)–RNA complex, a 3 μl-aliquot of purified NP–RNA complex solution at a concentration of 0.3 mg/ml was applied to Quantifiol R1.2/1.3 300 mesh grids and blotted for 6 s, with a humidity of 100% at 4 ℃, before being rapidly frozen in liquid ethane using the Vitrobot Mark IV system. Images were acquired using a Glacios cryo-TEM (Thermo Fisher Scientific) operated at 200 kV with a Falcon4 direct electron detector (Thermo Fisher Scientific) in electron counting mode at the Institute for Life and Medical Sciences, Kyoto University.

The detailed imaging conditions for both samples are listed in Table 3.

Table 3. Cryo-EM data collection, refinement, and validation statistics.

	LLOV NP(full)	LLOV NP(1-450)
Data collection and processing		
Magnification	59000x	150000x
Voltage (kV)	300	200
Electron exposure (e-/Å^2)	40	40
Defocus range (μm)	-0.8 to -1.8	-0.4 to -2.0
Calibrated pixel size (Å)	1.13	0.925
File Format	MRC	EER
Symmetry imposed	C1 helical	C1 helical
Initial segment images (no.)	62,855	442,067
Final segment images (no.)	26,797	399,761
Map resolution (Å)	3.04	3.07
FSC threshold	0.143	0.143
Map resolution range (Å)	2.9-5.6	2.9-3.2
Refinement		
Model resolution (Å)	3.11	3.18
FSC threshold	0.5	0.5
Map sharpening B factor (Å^2)	-147.231	-126.432
Model composition		
Non-hydrogen atoms	3180	3180
Protein residues	390	390
R.m.s deviations		
Bond lengths (Å)	0.006	0.006
Bond angles (°)	1.074	1.208
Validation		
Molprobity score	2.17	2.46
Clashscore	6.33	12.06
Poor rotamers (%)	6.36	8.51
Ramachandran plot		
Favored (%)	96.63	97.14
Allowed (%)	3.37	2.86
Disallowed (%)	0	0
Helical configuration		
Rise (Å)	3.19	3.16
Twist (°)	-15.11	-15.11
Asymmetric unit	Protomer	Protomer
Handedness	Left Handed	Left Handed
Inner diameter (Å)	≈195	≈195
Outer diameter (Å)	≈260	≈260

2.6. Image processing

Beam-induced motion was corrected using RELION, and the contrast transfer function parameters were estimated using CTFFIND-4.1.14 [50] for all micrographs. The helical structure image processing steps for the LLOV NP(full)–RNA and NP(1–450)–RNA complexes were performed using the software packages RELION3.0 and RELION3.1 [51], respectively. The coordinates of the start and end points of the helices in each micrograph were recorded manually. Then, helical segments were extracted from the helices, using a box size of 448 pixels at an unbinned pixel size and an interbox distance of 43 Å; the 2x binned images were subjected to 2D classification.

For the LLOV NP(full)–RNA complex dataset, a total of 62,855 segments were selected after two rounds of 2D classification from 3,122 micrographs and further subjected to 3D classification with a featureless cylinder generated by the relion_helix_toolbox [52] as the initial model. Subsequently, a 3D classification was performed using a symmetry search. The search range for twist was between -14.5° and -15.5°. The search range for rise was between 2.7 and 3.7 Å. After the first 3D classification, the model displayed clear features of a helical structure. A segment stack belonging to this class (60,818 segments) was selected and subjected to 3D refinement. Furthermore, 3D classification without alignment resulted in a better 3D class average, with 44.1% of the segments displaying clear secondary protein structural features. This subset was subjected to a second 3D refinement with a symmetry search. The search range for twist was between -14.9° and -15.3°. The search range for rise was between 3.0 and 3.5 Å. The helical symmetry parameter converged to a rise = 3.19 Å and twist = -15.11°. The resulting segment set was further processed using CtfRefine, Bayesian polishing, and CtfRefine before the third 3D refinement process to obtain the final reconstruction. The resolution estimate of the final reconstruction was based on the Fourier shell correlation (FSC) [53] between independently refined half-sets, which yielded a resolution of 3.0 Å.

The LLOV NP(1–450)–RNA complex dataset was processed according to the methods described above. After the first 3D classification step, CtfRefine and Bayesian polishing were performed, and a mask was created and used in the second 3D classification step. After the second 3D classification without alignment, 90.4 % of the segments from the resulting 3D class average that displayed the clearest secondary protein structures were re-extracted with a refined pixel size of 0.925 Å/pix. The reference map and mask were also rescaled to the same pixel size and subjected to

a second 3D refinement along with the re-extracted segments. The CtfRefine and Bayesian polishing were conducted before the final 3D refinement step with a helical symmetry local search. The search range for twist was between -14.9° and -15.3°. The search range for rise was between 3.0 and 3.5 Å. The resolution of the final construction was estimated as 3.1 Å. The helical symmetry parameter converged to a rise = 3.16 Å and twist = -15.11°. The detailed imaging processes for NP(full) and NP(1–450) are shown in Fig. 5.

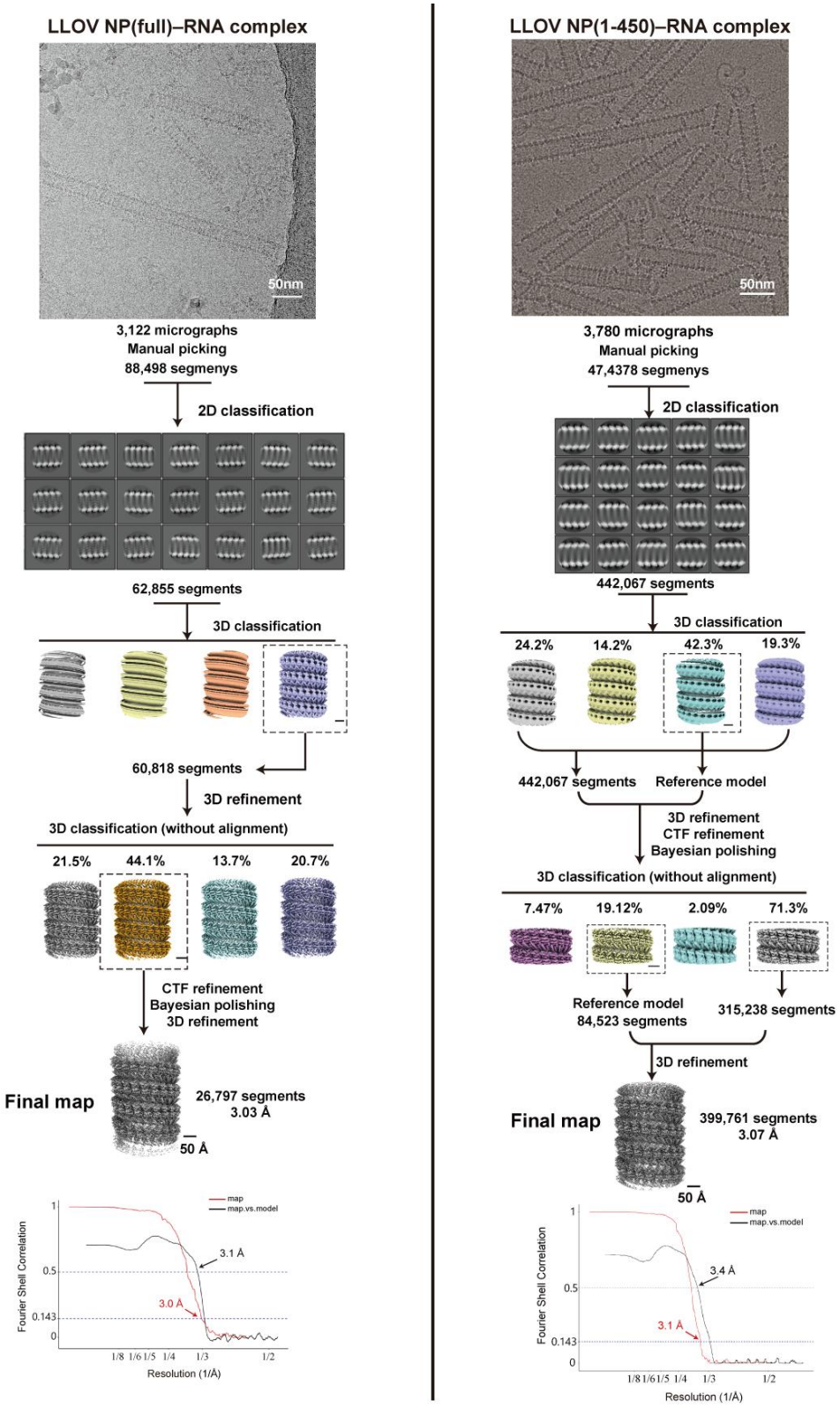


Fig 5. Single-particle cryo-EM workflow for LLOV NP(full) and NP(1–450)–RNA complex.

The cryo-EM images were processed by 2D classification and 3D refinement to construct the final cryo-EM map. The resolution of the final map of LLOV NP(full)–RNA complex and NP(1–450)–RNA complex was determined at 3.0 Å and 3.1 Å resolution, respectively.

2.7. Model building and refinement

For the LLOV NP(full)–RNA complex structure, the atomic model of the EBOV NP–RNA complex (NP residues: 19–407, PDB-ID: 5Z9W) [46] was used as the initial model and fitted as a rigid body into the cryo-EM map using UCSF Chimera (ver 1.13.1) [54]. Next, the EBOV residues were replaced with LLOV residues and the model was adjusted manually using COOT [55, 56]. The stereochemistry of peptide bonds was monitored using a Ramachandran plot. The N- and C-terminal residues 1–19 and 407–749, respectively, were not modeled because of the lack of interpretable map features. The atomic models were then refined against the corresponding maps in the presence of eight surrounding models using phenix.real_space_refine within PHENIX software [57] in the presence of eight surrounding subunits, including an extended RNA strand. The resulting models were selected based on the clash score, Ramachandran outliers, and MolProbity score, and the final model was presented using UCSF Chimera and ChimeraX [54, 58].

For the LLOV NP(1–450)–RNA complex structure, the atomic model of the LLOV NP(full)–RNA complex was used as the initial model, and the model was refined according to the method described above. The detailed model statistics are presented in Table 3.

2.7. Minigenome assay

All LLOV minigenome assays were performed as previously described [31]. Briefly, HEK 293T cells were transfected with the PolII-driven LLOV minigenome (pCAGGS-LLOV-vRNA-hrluc; 125 ng), pCAGGS-LLOV-L (500 ng), pCAGGS-LLOV-VP30 (37.5 ng), and pCAGGS-LLOV-VP35 (62.5 ng), as well as the myc-tagged NP mutants or NP-WT (62.5 ng), using Transit LT-1, according to the manufacturer's instructions. For normalization, the cells were additionally transfected with pCAGGS-luc2 (12.5 ng), and polymerase was omitted (-L) as a negative control. Two days post-transfection, the cells were lysed in 1× Lysis Juice (PJK) for 10 min, and cleared by centrifugation for 3 min at 10,000 x g. Reporter activity was measured by adding 40 µl cleared lysate to either 40 µl Beetle Juice (PJK) or 40 µl *Renilla* Glo Juice in opaque 96-well plates. Luciferase activity was measured using the Glomax Multi Microplate Reader (Promega). The reporter activity obtained for *Renilla* luciferase (minigenome reporter) was normalized to firefly luciferase activity (control luciferase). Only samples that showed NP expression levels between 40% and 250% of

NP-myc (+L) were included in the reporter activity calculations. Statistical analysis was conducted by one-way ANOVA with Dunnett's post hoc test. The experiments were performed in pentaplicate (n = 5). Significance was set at $p < 0.05$ (*) and $p < 0.001$ (***).

2.8. Western blotting

To determine NP expression levels, 4x SDS-PAGE buffer [125 mM Tris-HCl (pH 6.8), 40% glycerin, 8% SDS, 20% beta-mercaptoethanol, and 10 µg/ml bromophenol blue] was added to the remaining lysates from the luciferase assay, and further boiled for 10 min at 99 °C, before the samples were subjected to SDS-PAGE and western blotting. Nitrocellulose membranes were blocked with 7% skimmed milk powder in D-PBS (-) and stained with anti-c-myc (rabbit; #PA1-981; Thermo Fisher Scientific) and anti-rabbit (goat anti-rabbit; #111-655-144; Dianova) antibodies for the detection of NP-myc. Anti-GAPDH (mouse; #sc47724; Santa Cruz) and anti-mouse (goat anti-mouse; #926-68070; Li-cor) antibodies were used as loading controls. Western Blots were visualized using the Li-Cor Odyssey imaging system, and band intensity was determined using ImageStudio Lite software (Li-Cor).

Chapter 3

Results

3.1. Examination of the helical structure of LLOV NP–RNA complex

To reveal the rigid, straight and homogeneous helical structure as a representative of virus nucleocapsid core within virus particle, I first performed protein structure prediction and negative-stain TEM to examine the structure of the LLOV NP–RNA complex.

Protein structure prediction using PSIPRED 4.0 [59] suggested that LLOV NP contains two relatively ordered regions: an N-terminal core comprising residues 1-410 and a C-terminal tail comprising residues 660-749 (Fig. 6a). The two ordered regions are linked by a disordered linker region, implying that LLOV NP may share a similar structure with EBOV and MARV NP [46, 48]. Next, to examine the structure of the LLOV NP–RNA complex, I expressed a full-length and C-terminal truncated NP consisting of its N-terminal 1–450 residues [NP(full) and NP(1–450), respectively) in HEK 293T cells. The NP–RNA complexes were purified using CsCl density-gradient ultracentrifugation, and visualized by negative-stain TEM. Both the NP(full)–RNA and NP(1–450)–RNA complexes maintained straight, rigid, and homogeneous helical structures (Fig. 6b, c). In contrast, the NP(full)–RNA complex displayed a more flexible structure, with several helices uncoiled during the purification process (Fig. 6b, c). A number of irregular protrusions were observed on the surface of the NP(full)–RNA complex, whereas these were not observed on the NP(1–450)–RNA complex (Fig. 6d, e). These protrusions may represent a disordered C-terminal region of the NP. To determine the structure of the LLOV NP–RNA complex and assess the impact of the NP C-terminal region on helical formation, both the NP(full)–RNA and NP(1–450)–RNA complexes were used for further cryo-EM single particle analysis.

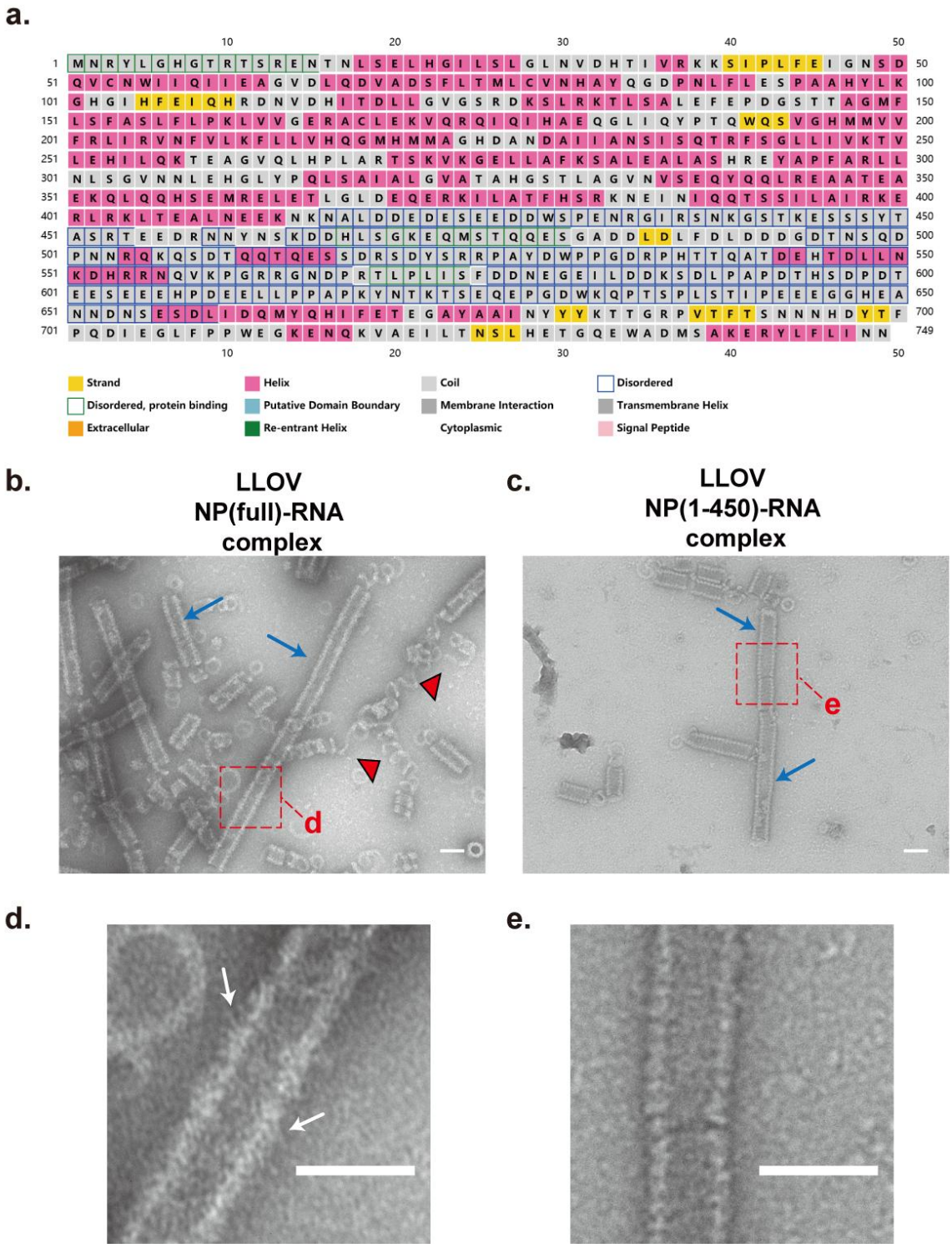
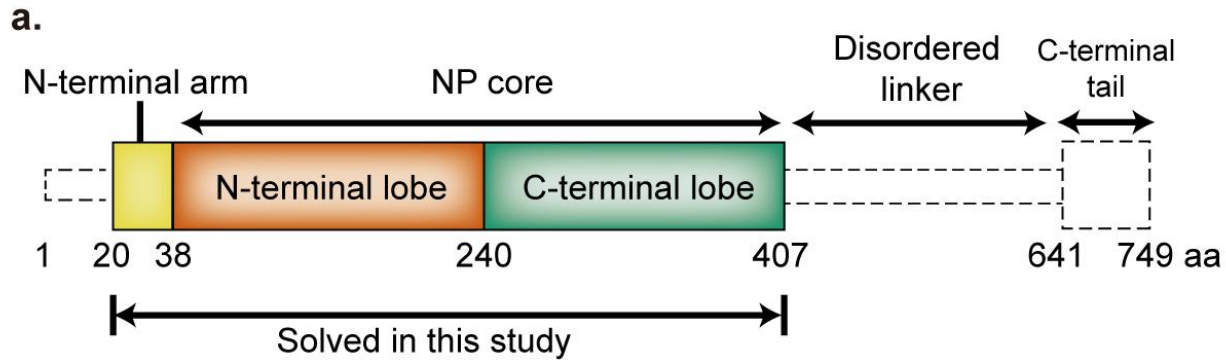


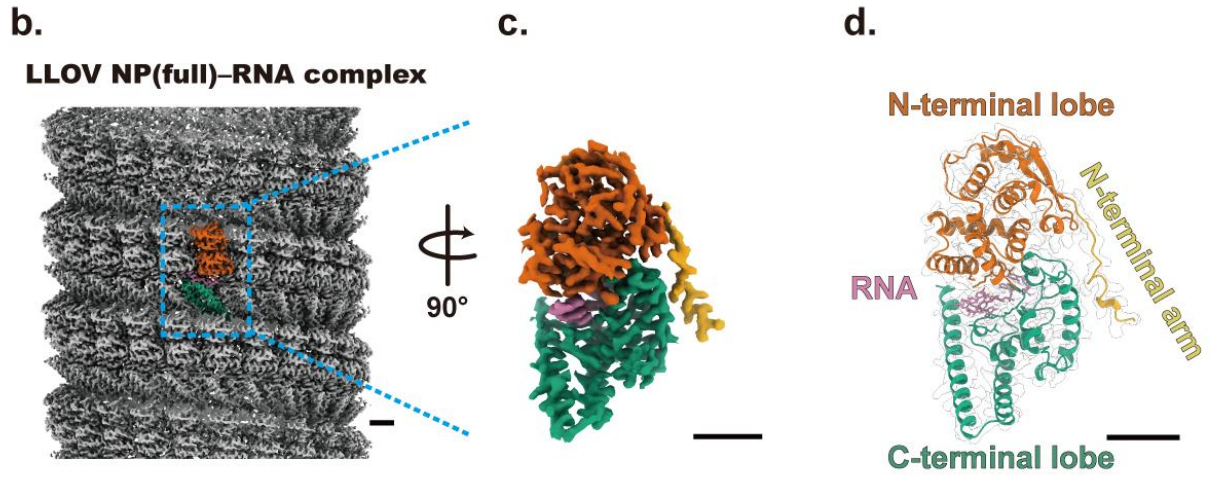
Fig 6. Morphology of LLOV helical NP–RNA complex.

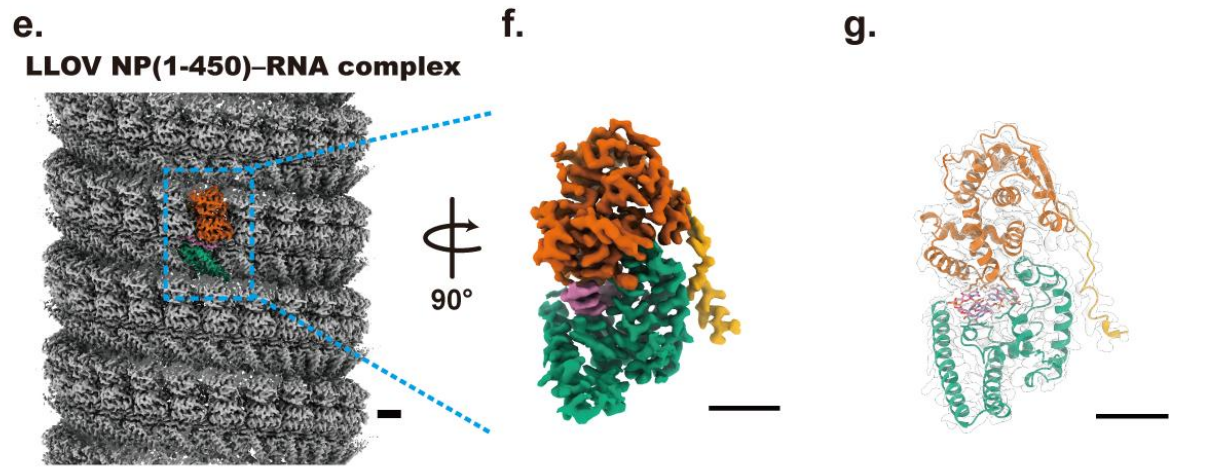
(a) LLOV NP secondary structure prediction using PSIPRED 4.0 [59]. (b, c) Images of LLOV NP(full) and NP(1–450)–RNA complexes visualized by negative-stain TEM. The loosely uncoiled region is marked by red arrowheads and the rigid region is marked by blue arrows. (d, e) Close-up view of the NP–RNA complex. Heterogeneous protrusions on the surface of the NP(full)–RNA complex are marked by white arrows. Scale bars, 50 nm.

3.2. Overall architecture of the LLOV NP–RNA complex.

Cryo-EM single-particle analysis determined the structures of the NP(full)–RNA and NP(1–450)–RNA complexes at overall resolutions of 3.0 Å and 3.1 Å, respectively (Fig. 5). Both the NP(full)–RNA and NP(1–450)–RNA complexes were left-handed helices with 260 Å and 195 Å in outer and inner diameters, respectively, and contained 23.8 subunits per turn (Fig. 7b, e, Table. 3). The helical parameters of NP(full)–RNA (rise = 3.19 Å, twist = -15.11°) and NP(1–450)–RNA complexes (rise = 3.16 Å, twist = -15.11°) are quite similar, indicating that the impact of C-terminal truncation on the helical formation is negligible (Table. 3). Based on the two cryo-EM maps of the NP–RNA complex, NP–RNA protomers were isolated, and their atomic models were generated (Fig. 7b-g). Only the N-terminal residues 20–407 of NP were modelled in both cryo-EM maps (Fig. 7a). The lack of map density may be due to the heterogeneous structure and spatial arrangement of a disordered region within 19 N-terminal and 342 C-terminal residues. The N-terminal region (residues 20–407) was divided into two domains: the N-terminal arm (residues 20–38), and the N-terminal core, comprising the N-terminal lobe (residues 39–240) and the C-terminal lobe (residues 241–407) (Fig. 7a, d). A host-derived single-stranded RNA segment was encapsidated between the N-terminal- and C-terminal lobes of each NP (Fig. 7i). Structural comparison between the NP(full) and NP(1–450) protomers revealed a high degree of similarity (Fig. 7h), with only a 0.5 Å backbone Root Mean Square Deviation (RMSD), suggesting the C-terminal region had minimal impact on the folding of NP.


a.
N-terminal arm
NP core
Disordered linker
C-terminal tail
N-terminal lobe
C-terminal lobe
1
20
38
240
407
641
749 aa
Solved in this study


b.
LLOV NP(full)–RNA complex
90°
c.
d.
N-terminal lobe
N-terminal arm
RNA
C-terminal lobe


e.
LLOV NP(1-450)–RNA complex
90°
f.
g.

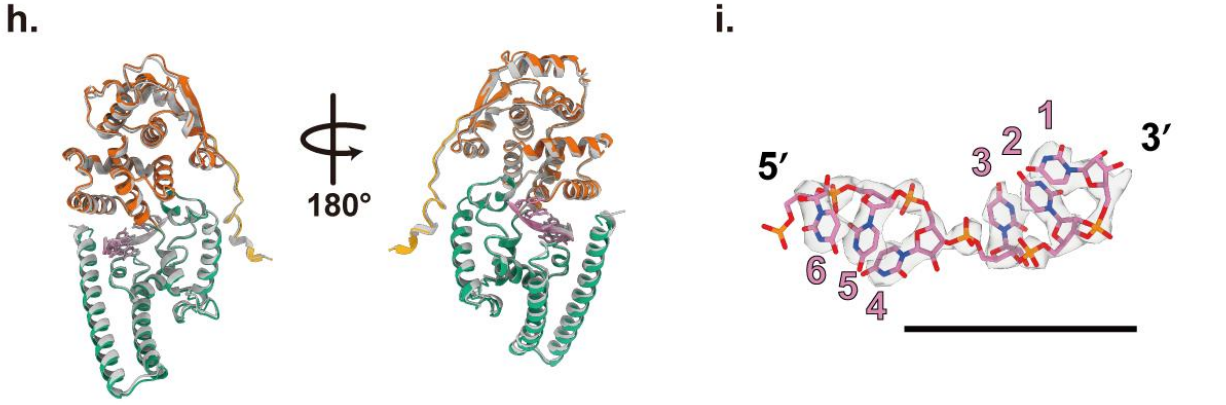

h.
180°
i.
5′
3′
1
2
3
4
5
6

Fig 7. Overall cryo-EM structure of the LLOV NP–RNA complex and its atomic model.

(a) Schematic of the domain architecture of the LLOV NP. Each domain is represented by a unique color, and the unsolved region is represented by a dotted outline without color. (b, e) Cryo-EM maps, (c, f) isolated cryo-EM maps, and (d, g) atomic models of NP(full)–RNA and NP(1–450)–RNA complexes for the asymmetric unit; NP protomer associated with RNA. A single NP–RNA protomer is highlighted in the following colors: RNA strand in purple, NP in red, green, and yellow, according to the color code shown in (a). (h) Structural comparison of the NP(full)–RNA complex (colored as described above) and the NP(1–450)–RNA complex (colored gray). (i) Isolated cryo-EM map and atomic model of RNA associating with an NP protomer. The RNA bases were modeled as uracils. Scale bars, 20 Å.

3.3. Filoviruses NP–RNA protomers exhibit similar structure but different arrangement in helix

Each LLOV, EBOV, and MARV NP protomer encapsidates six nucleotides [46, 48], similar to the paramyxovirus NP [60]. Structural comparison of the LLOV NP protomer with that of the EBOV NP (PDB-ID: 5Z9W) and MARV NP (PDB-ID: 7F1M) revealed 1.0 Å and 0.8 Å backbone RMSDs, respectively, suggesting a common structure of the filovirus NP protomer (Fig. 8a, b) [46, 48]. However, the NP–RNA helices of three filoviruses exhibited differences in helical parameters [LLOV NP(full)–RNA, rise = 3.19 Å, twist = -15.11°; EBOV NP(1–450])–RNA, rise = 3.01 Å, twist = -14.73°; MARV NP(1-395)–RNA, rise = 4.23 Å, twist = -11.81°] [46, 48]. These different helical parameters result in differences in the size of NP–RNA helices (Fig. 8c) and the length of viral particles [16, 61]. The LLOV NP–RNA and EBOV NP–RNA helices have close inner diameters (195 Å and 175 Å, respectively), containing similar amounts of subunits per turn (23.8 subunits and 24.4 subunits per turn, respectively), while the MARV NP–RNA helices have a larger inner diameter (255 Å) with more subunits per turn (30.5 subunits per turn) (Fig. 8c) [46, 48]. Since the genome sizes of these three filoviruses were similar (19,019 bases for LLOV, 18,961 bases for EBOV, and 19,111 bases for MARV), the LLOV and EBOV viral particles needed to be longer to pack the whole genome. As a result, the lengths of EBOV and LLOV viral particles are approximately 1000 nm, while that of MARV is approximately 890 nm [16, 61]. These results suggest that filoviruses of

different genera exhibit similar NP-folding mechanisms, but their NP–RNA helices have different quaternary structures.

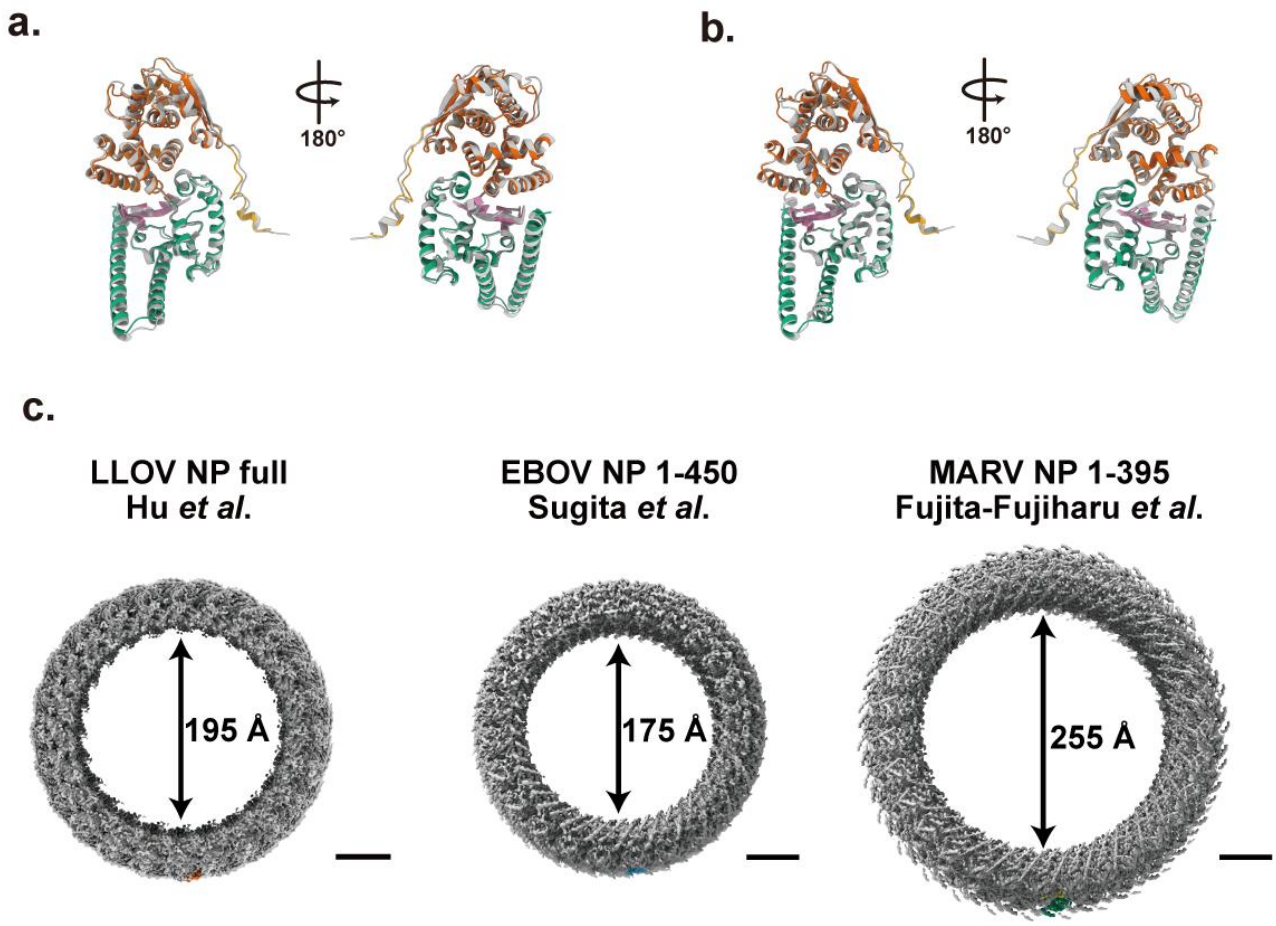


Fig 8. Comparison of the LLOV, EBOV and MARV NP(full)–RNA complexes.

(a, b) Overall structure of the LLOV NP(full)–RNA complex (PDB-ID: 7YPW, colored in this study) superimposed on the (a) EBOV NP(1–450)–RNA complex (PDB-ID: 5Z9W, gray) and (b) MARV NP(1-394)–RNA complex (PDB-ID: 7F1M, gray). (c) Comparison of the inner diameters of the LLOV NP(full)–RNA complex (EMD-34016) with the EBOV NP(1–450)–RNA complex (EMD-6903) and the MARV NP(1-394)–RNA complex (EMD-31420).

3.4. RNA binding region of the LLOV NP

The LLOV NP(full) interacts with the RNA segment through a positively charged cleft formed by the N-terminal and C-terminal lobes (Fig. 9), similar to the C-terminal truncated EBOV and MARV NPs [46-48]. From the top view, six RNA nucleotides were embedded in the cleft in a

"3-bases-inward, 3-bases-outward" configuration (Fig. 10a, b). Hydrogen bonding and electrostatic interactions between the RNA sugar-phosphate backbone and positively charged NP residues play a major role in the NP-RNA interactions. A short 3_{10}-helix (η7, 330-334 residues), which is placed beneath the RNA backbone, likely stabilizes the RNA backbone from C-terminal side through the hydrogen bonds (Fig. 10c, e). The side chains of the positively charged residues K160, K171, R174, K248, R298, and R401 were located within a distance of < 4 Å from the RNA backbone, and pointed toward the sugar-phosphate backbone of the RNA (Fig. 10c, d, f). Sequence alignment of LLOV, EBOV, and MARV NPs revealed a high degree of conservation of these positively charged residues (Fig. 11), suggesting a common RNA-binding mechanism in mammalian filoviruses. Among these positive residues, K160 and K171 are reported to be important for vRNA synthesis and NP–RNA helix formation in EBOV and MARV NPs (K142 and K153 for MARV), whereas R174 and K248 are negligible in NP-RNA interactions [34, 48, 62, 63]. Lastly, LLOV NP interacts with the RNA backbone instead of the RNA bases (Fig. 10d, f), indicating that NP binds RNA in a sequence-independent manner, similar to other negative-sense RNA viruses [45, 46-48, 60, 64-68]. This likely explains why LLOV NP can form NP–RNA complexes in the absence of vRNA.

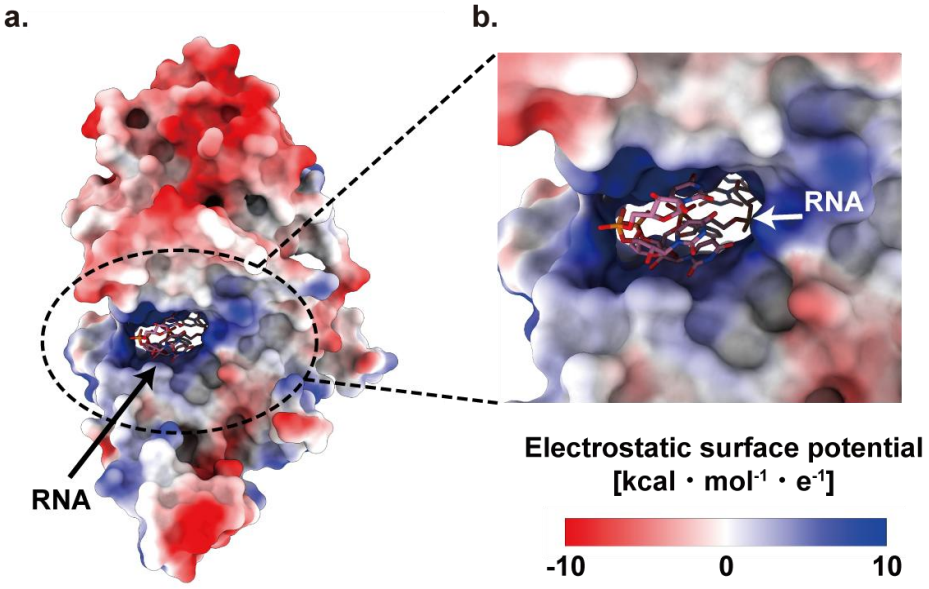


Fig 9. Electrostatic surface of the LLOV NP–RNA protomer.

(a) The electrostatic surface calculated using the atomic model of the LLOV NP(full)–RNA protomer. Scale ranged from -10 (red) to +10 (blue) kcal mol^{-1} e^{-1}. (b) Close-up view of the RNA-binding cleft in the NP.

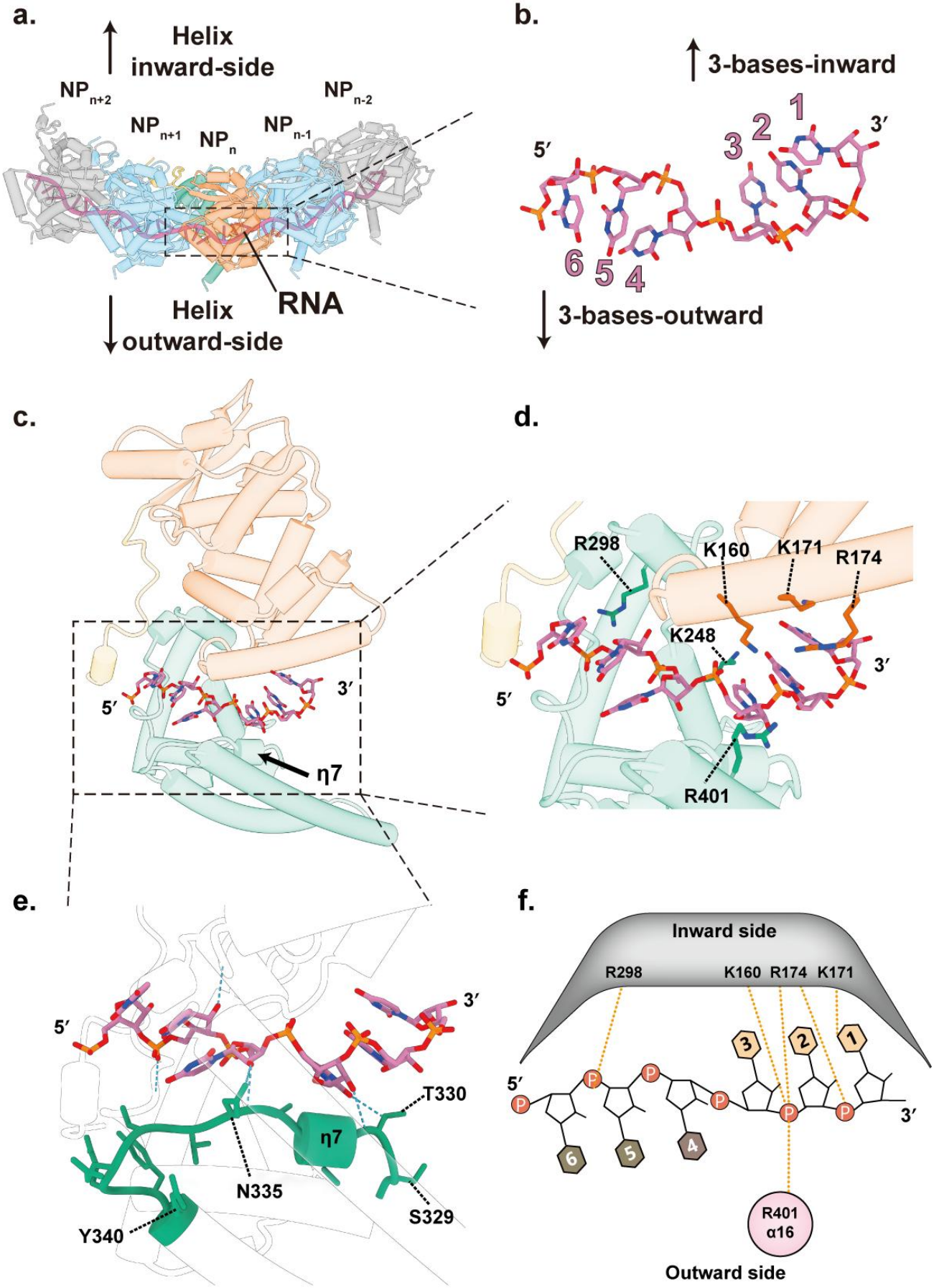

a.
Helix
inward-side
NP n+2
NP n-2
NP n+1
NP n
NP n-1
RNA
Helix
outward-side
b.
3-bases-inward
5′
3′
1
2
3
4
5
6
3-bases-outward
c.
5′
3′
η7
d.
R298
K160
K171
R174
K248
5′
3′
R401
e.
5′
3′
T330
η7
N335
S329
Y340
f.
Inward side
R298
K160
R174
K171
5′
3′
R401
α16
Outward side

Fig 10. NP-RNA interactions in the helical NP(full)–RNA complex.

(a) Top view of five adjacent (full) NP RNA protomers on a strand. (b) Six nucleotides of RNA segments were arranged in a "3-bases-inward, 3-bases-outward" configuration. (c) The overall structure of the protomeric NP(full)–RNA complex in a cylindrical presentation (colored in Figure 6a). (d) Close-up view of the NP-RNA interactions. Positively charged amino acids are shown. (e) Close-up view of the RNA-binding site. The residues possibly forming hydrogen bonds between NP and RNA (distance < 4 Å) were detected by ChimeraX and shown as blue dotted lines. (f) Schematic diagram of NP-RNA interactions. The RNA bases are numbered in the 3′ to 5′ direction. The RNA is based on a helix facing inward (light orange) and outward (light brown). Secondary structure labels were based on the atomic model. Orange dotted lines represent potential interactions between NP positively charged residues and the RNA backbone.

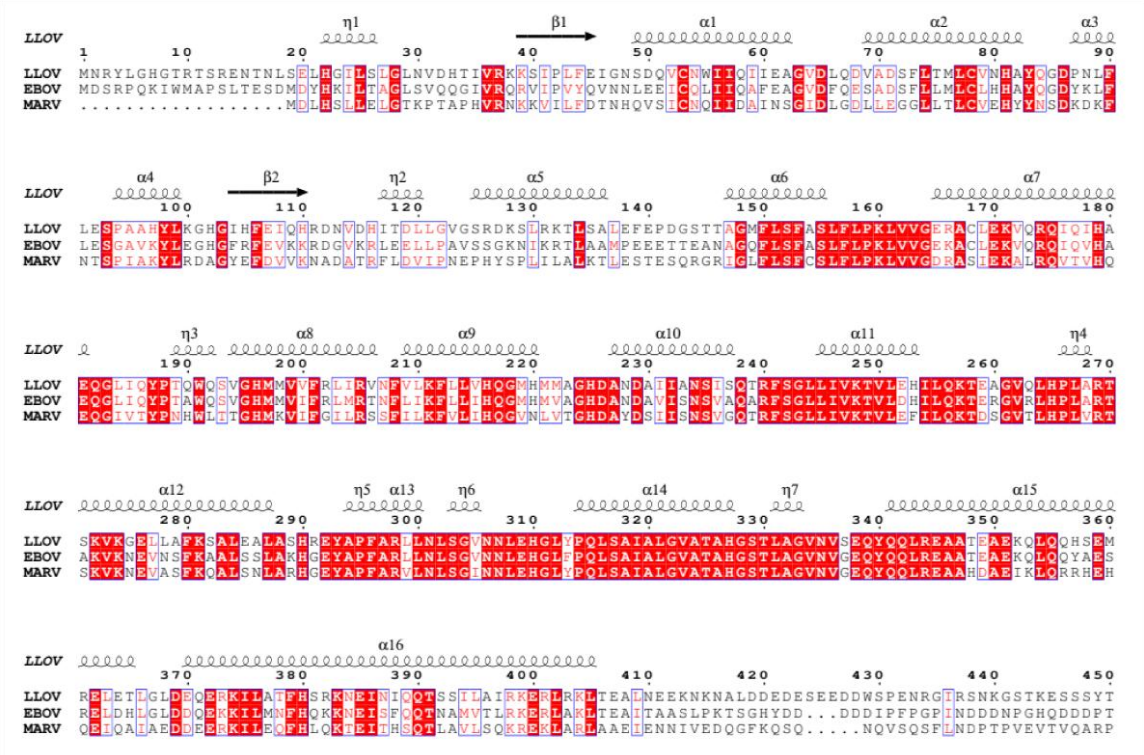


Fig 11. Secondary structure assignment of the LLOV NP and the sequence alignment with the EBOV and MARV NPs.

Sequences of NPs from *Lloviu cuevavirus* (GenBank: YP_009428135.1), *Zaire ebolavirus* (GenBank: AAD14590.1), and *Marburg virus* (GenBank: CAA82536.1) were aligned using ClastalW 2.1 [69], and presented using ESPript 3.1 [70]. Conserved residues and similar residues are boxed in red and

white, respectively. The secondary structure assignment of the atomic model for LLOV NP(20–407) was predicted using phenix_secondary_structure restraints [57] and is shown at the top of the sequence alignment results.

3.5. LLOV NP-NP interactions in the helix

The NP-NP interactions play a significant role in the formation of the NP–RNA helix, along with NP-RNA interactions. Based on the spatial position of the NP–RNA protomers in the helix, NP-NP interactions can be classified into two groups: interactions between neighboring NPs on the strand (intra-strand NP-NP interactions) and associations between helical strands (inter-strand NP-NP interactions) (Fig. 12a, b).

Intra-strand NP-NP interactions are mediated by both hydrophobic and electrostatic interactions involving multiple residues and domains (Fig. 12b). The N-terminal arm plays an essential role in NP oligomerization [45, 46-48, 62, 63]. The atomic model shows that the helix η1 in the N-terminal arm of NP_n binds to the hydrophobic pocket located in the C-terminal lobe of NP_{n+1} (Fig. 11c, and Fig. 13a, b, d). Residues located on or near helix η1, such as L21, H22, I24, L25, and L27, likely participate in the interaction between the N-terminal arm and the hydrophobic pocket (Fig. 12c). A loop consisting of residues 222–226 on the N-terminal lobe of NP_{n+1} forms a screw-like structure, which fits into another nut-like conformation composed of four benzene rings: H33 on the N-terminal arm, H102 and F208 on the N-terminal lobe, and H290 on the C-terminal lobe of NP_n (Fig. 12d, Fig. 13c). Specifically, the M222 side chain of NP_{n+1} fits into the opening of this nut-like structure, possibly stabilizing the contact region between the N-terminal arm and N-terminal lobe (Fig. 12d, Fig. 13c). Interactions between the two C-terminal lobes of adjacent NPs mainly occur between α15 of NP_n and α16 of NP_{n+1} (Fig. 12e). The atomic model shows that these interactions are composed of electrostatic (H357/R361 on α15 and D369/E372 on α16) and hydrophobic interactions (L354 on α15 and I375/L368/L376 on α16), which may strengthen the stability of the helical strands (Fig. 12e).

Inter-strand NP-NP interactions are mainly mediated by electrostatic interactions between the C-terminal lobe of the upper NP and the N-terminal lobe of the lower NP (Fig. 12a, b, f). In the model, any arbitrary NP_n of NP(full) formed pairs with NP_{n-24} but did not associate with NP_{n-23} or

NP_{n-25} in the helix (Fig. 14). The same was true for the helix composed of LLOV NP(1–450), but not for the helix composed of the C-terminal truncated EBOV NP, where NP_n interacted with both NP_{n-23} and NP_{n-24} (Fig. 14) [46]. The different arrangements of NPs in the helix between LLOV and EBOV may correspond to the stability of the helical structures.

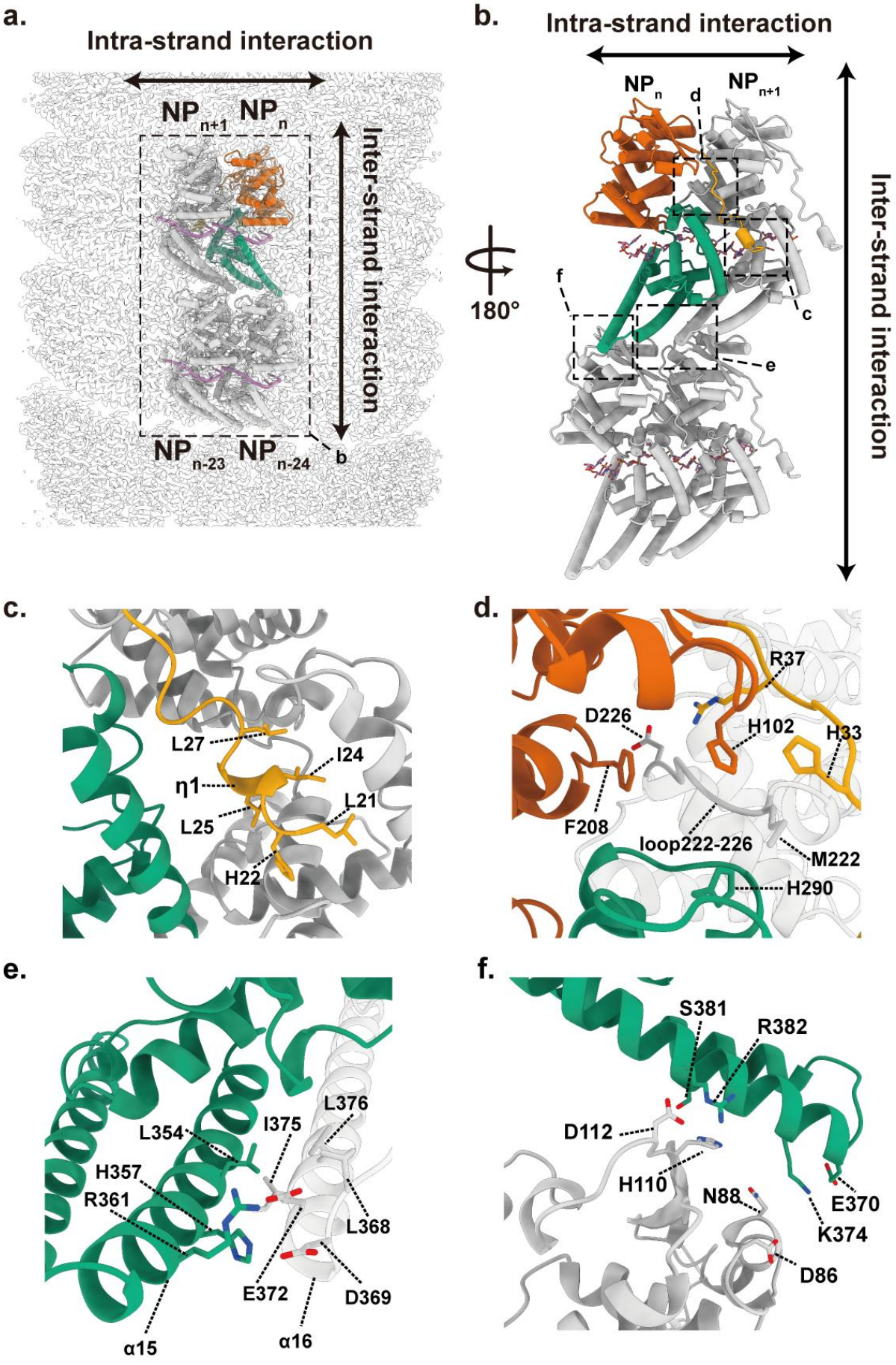

a.
Intra-strand interaction
NP n+1
NP n
Inter-strand interaction
NP n-23
NP n-24
b
b.
Intra-strand interaction
NP n
d
NP n+1
180°
f
c
e
Inter-strand interaction
c.
L27
I24
η1
L21
L25
H22
d.
R37
D226
H102
H33
F208
loop222-226
M222
H290
e.
L376
L354
I375
H357
R361
L368
E372
D369
α15
α16
f.
S381
R382
D112
H110
N88
E370
K374
D86

Fig 12. NP-NP interactions in the helical NP(full)–RNA complex.

(a) Overall and (b) close-up views of four adjacent NP–RNA protomers in a helix. The NP_n was colored according to the colors shown in Figure 6a. The larger and smaller rectangles with dotted lines highlight the close-up of (c, d, e) intra-strand NP-NP interactions between NP_n and NP_{n+1}, and (f) inter-strand NP-NP interactions between NP_n and NP_{n-24}.

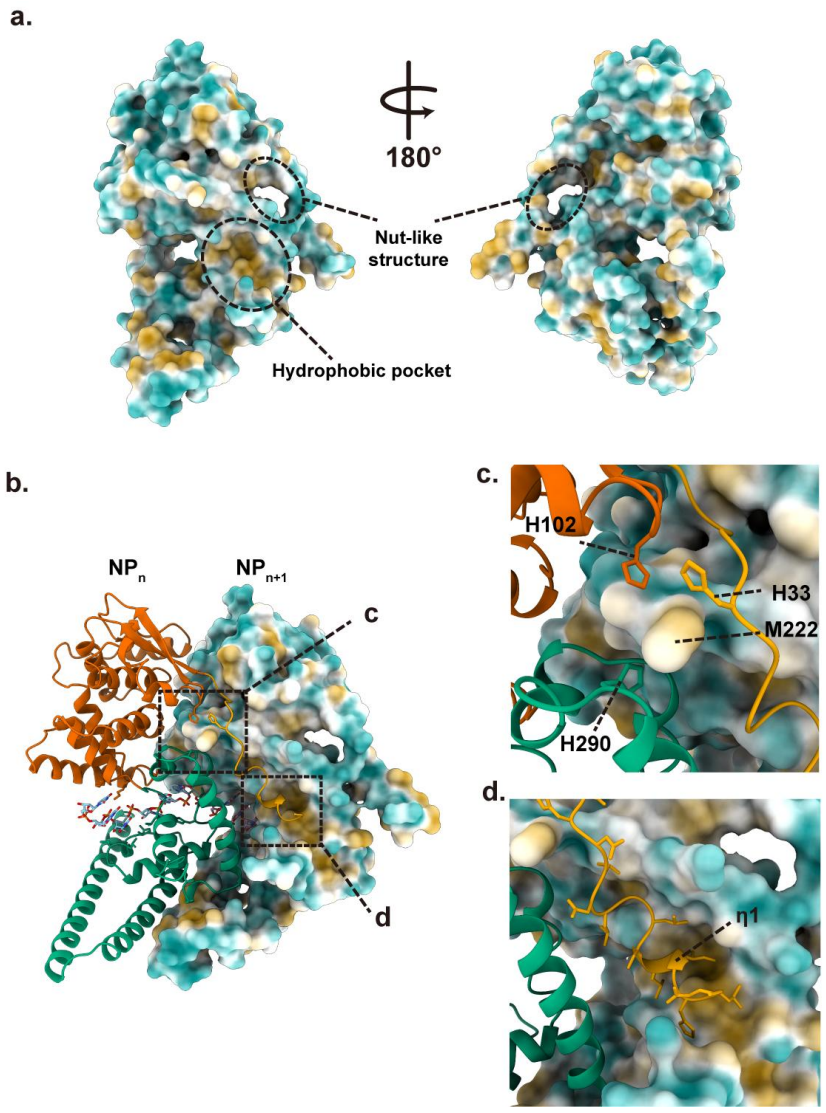


Fig 13. Molecular lipophilicity potential maps of the LLOV NP.

(a) Hydrophilic regions of the LLOV NP are colored in dark cyan, and the hydrophobic domains are colored as dark gold rods. (b) Overall view of the intra-strand interaction between NP_n (ribbon) and NP_{n+1}, with the surface colored as in (a). (c) Close-up view of the interaction between the nut structure of NP_n and the M222 screw structure of its adjacent NP (NP_{n+1}). (d) A close-up view of the interaction between the N-terminal arm (yellow) of NP_n and the hydrophobic pocket of its adjacent

NP (NP_{n+1}).

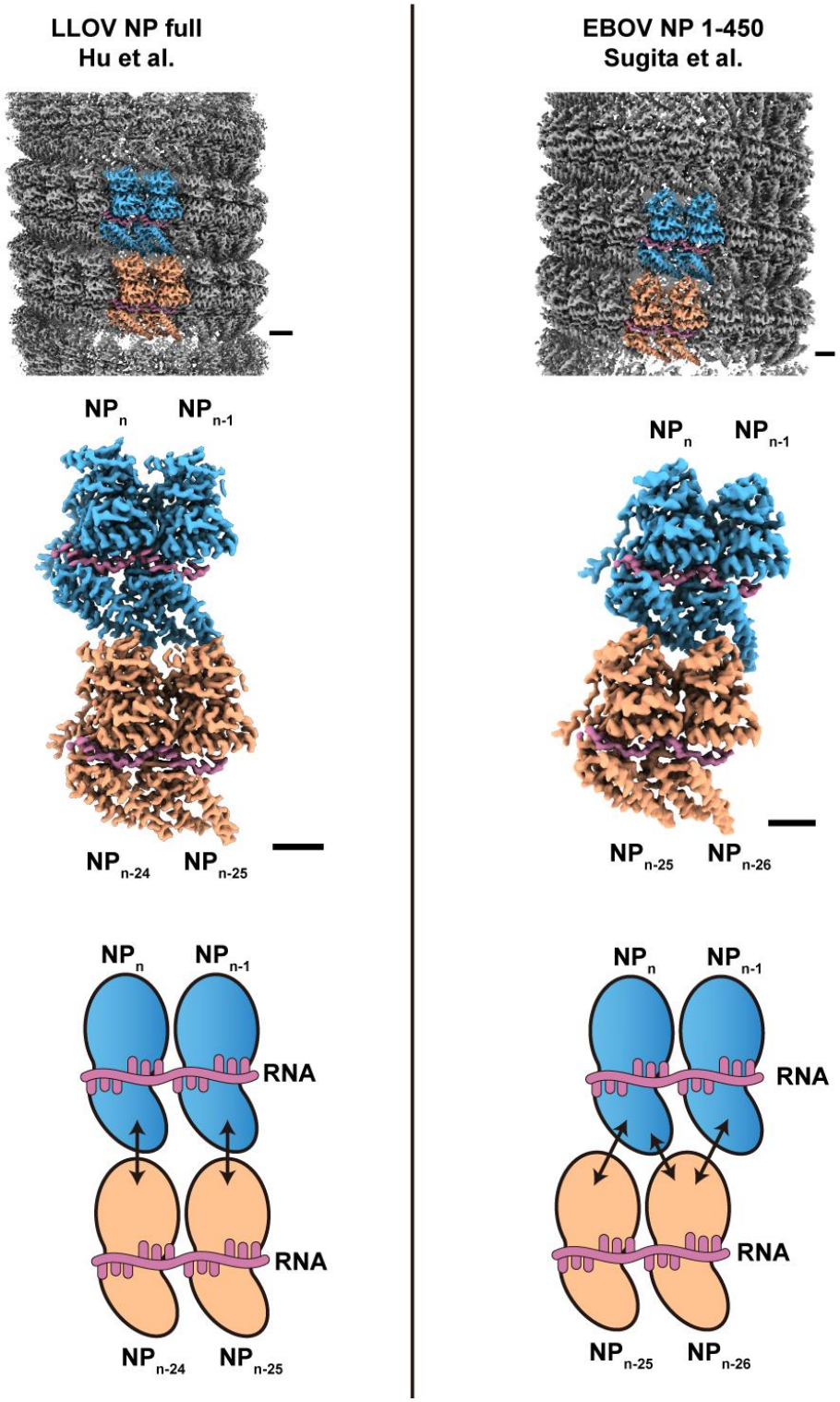


Fig 14. Structural comparison of the filoviruses' NP helical assemblies.

Overall and close-up isopotential surface views of three filovirus helical NP assemblies (LLOV NP(full)–RNA complex from this study, and EBOV NP(1–450)–RNA complex [EMD-6903] [46]). Different helical arrangements are shown in the schematic diagrams. Only the core domain and the

RNA (purple) are depicted. The interactions between NPs are represented by black arrows.

3.6. Essential residues for the helical assembly and vRNA synthesis

The residues involved in helical assembly are promising targets for further drug development. To identify the residues required for helical assembly, I performed a structure-based mutational analysis of the residues participating in molecular interactions in the helix. Since the residues involved in NP-RNA interactions, which are highly conserved among mammalian filoviruses, have already been identified (Fig. 11) [48], I mainly focused on the residues involved in intra-strand and inter-strand NP-NP interactions (Table. 4). Western blot analysis revealed that the expression levels of all NP(full) mutants were comparable to those of the wild-type NP(full) (Fig. 15). Using these NP mutants, negative-stain TEM was performed to evaluate their ability to form helical structures (Fig. 16a). A minigenome assay was also conducted to assess the effects of the mutation on transcription and replication (Fig. 16b).

Mutation on helix η1 in the N-terminal arm of the NP I24E mutant (Fig. 12c), significantly decreased transcription and replication activity, with that being less than 1% of the wild-type NP (Fig. 16b). Likewise, the NP I24E mutants were the only mutants that rarely formed NP–RNA strands and helices (Fig. 16a), indicating a remarkable decrease in oligomerization activity. This finding is consistent with those for the corresponding mutants EBOV I24E and MARV L6E [48]. Except for NP I24E, the NP mutants remained capable of forming stable NP–RNA strands. However, none of the NP mutants formed rigid NP–RNA helices, unlike the wild-type NP (Fig. 16a). The R37A, H102A, and M222A mutations, which were in or close to the nut- and screw-like structures, (Fig. 12d), resulted in irregular and loose helical conformations, with more than 50% reduction in transcription and replication activity (Fig. 16). These results suggest that interactions between nut- and screw-like structures play an important role in helical assembly and vRNA synthesis. Of note, the NP M222E mutant formed mildly loose but regular helical structures and maintained its transcriptional and replication activities (Fig. 16), indicating that the M222 side chain could interact with the nut structure H33/H102/F208/H290 through either hydrophobic or electrostatic interactions. Because the corresponding residues, EBOV V222 and MARV V204 (Fig. 11), are both hydrophobic, the LLOV NP M222A mutant may lose its activity due to its short side chain. In addition, the

corresponding mutants, EBOV R37A and MARV R19A, exhibited reduced transcription and replication activity [48]; however, their impact on helical assembly was negligible.

The NP H110E and E370A/K374A/R382A mutants (Fig. 12f), corresponding to inter-strand NP-NP interactions, barely formed any regular helical structures, which appeared rather flexible and irregularly uncoiled (Fig. 16a). The transcription and replication activities of these mutants were reduced to less than 50% compared with wild-type NP (Fig. 16b), indicating that inter-strand electrostatic interactions are not only important for the maintenance of the helical structure, but also for viral RNA synthesis.

Table 4. Target residues for the LLOV NP mutational analysis

Residues	Location	Contact region
I24E	N-terminal arm	Hydrophobic pocket of neighbouring NP
R37A	N-terminal arm	N-terminal lobe loop 220-229 of neighbouring NP
H102A	N-terminal lobe	N-terminal lobe loop 220-229 of neighbouring NP
M222A	N-terminal lobe loop 220-229	Hydrophobic ring of neighbouring NP
M222E	N-terminal lobe loop 220-229	Hydrophobic ring of neighbouring NP
H110E	N-terminal lobe	C-terminal lobe α16 of Upper NP
E370A·K374A·R382A	C-terminal lobe α16	N-terminal lobe of lower NP

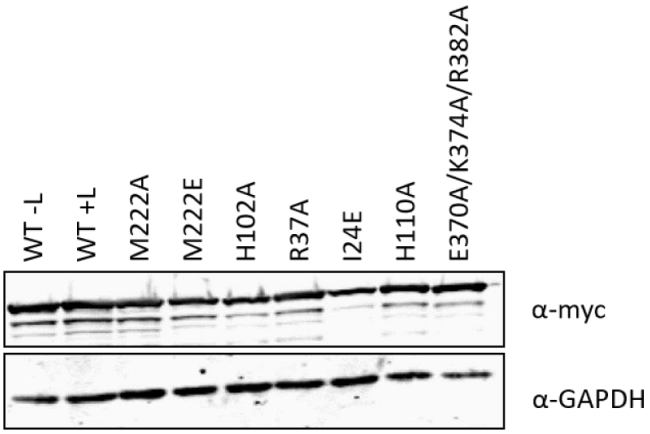


Fig 15. Expression levels of respective LLOV NP mutants in cells.

(a) Western blot analysis of cells transfected with plasmids expressing the LLOV nucleocapsid components. A myc tag was fused to the C terminus of native and mutant LLOV NPs. LLOV NP-myc and GAPDH were detected using specific antibodies. All experiments were performed in pentaplicate.

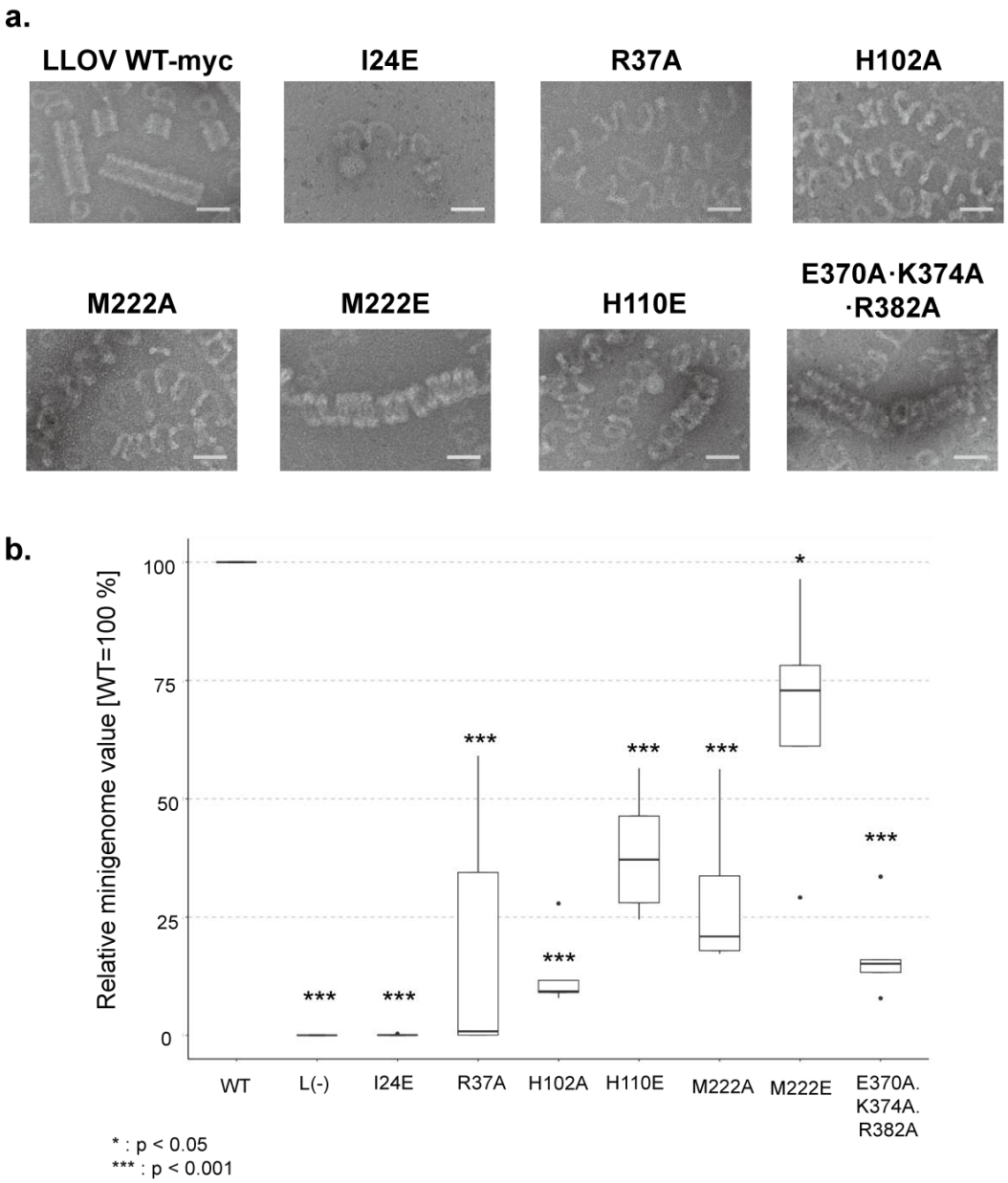


Fig 16. Structure-based mutational analysis of the LLOV NP.

(a) Images of the negative-stain TEM, purified helical structure of the LLOV NP(full) mutant–RNA complex. A myc tag was fused to the C terminus of native and mutant LLOV NPs. Scale bars, 50 nm. (b) Assessment of the function of LLOV NP-myc mutants in genome transcription and replication evaluated using minigenome assays. The experiments were performed in pentaplicate (n = 5). Significance was set at $p < 0.05$ (*) and $p < 0.001$ (***).

Chapter 4

Discussion

The LLOV has been detected in three European countries, with infectious particles isolated from two of these countries [13, 14], indicating a wide distribution of this virus in Europe. Considering the zoonotic potential of LLOV, there is an urgent need to better understand the underlying molecular mechanisms of LLOV. In this study, using single-particle cryo-EM, I determined the high-resolution structures of the LLOV NP(full)–RNA and C-terminal truncated NP(1–450)–RNA complexes (Fig. 7b, e). Based on the structures, atomic models of the two LLOV NP–RNA protomers were constructed and showed identical structures (Fig. 7d, g, h). By combining the cryo-EM maps and atomic models, I identified amino acid residues on the NP that are potentially involved in intra- and inter-strand NP-NP interactions in the helix (Fig. 12). Structure-based mutational analyses highlighted several important amino acid residues involved in helical assembly and vRNA synthesis (Fig. 16). Collectively, this study revealed the common structural features of mammalian filovirus NP–RNA complexes and may therefore aid toward the development of structure-based pan-filovirus antiviral agents.

To date, cryo-EM structural studies of the mononegaviral NP–RNA complex, such as filoviruses and paramyxoviruses [46-48, 64, 68], have always employed C-terminal truncated NPs due to the intrinsically disordered region interfering with cryo-EM single-particle analysis. Consequently, the impact of C-terminal truncation on the folding of the NP–RNA protomer and helical assembly of the NP–RNA complex remains unknown. This study is the first to determine the helical structure of both the NP(full)–RNA and C-terminal truncated NP–RNA complexes (Fig. 7b, e), demonstrating that C-terminal truncation has a negligible effect on both the folding of the NP–RNA protomer and the helical assembly of the NP–RNA complex. However, it is still challenging to explain why only the LLOV NP(full)–RNA helix, but not the NP–RNA complexes of EBOV and MARV [45-48], could maintain a rigid helical structure. The unresolved C-terminal regions of filoviruses from different genera may have different features and play various roles in nucleocapsid formation. Further studies using full-length NP and other nucleocapsid components are warranted.

The LLOV NP-RNA interaction was mainly mediated by electrostatic interactions between the NP core and RNA backbone (Fig. 10f). The sequence-independent NP-RNA interaction mechanism explains why LLOV NP could form NP–RNA complexes without vRNA and why the minigenome assay could be established. Structural comparison of LLOV NP, EBOV NP (PDB-ID: 5Z9W), and MARV NP (PDB-ID: 7F1M) revealed that NPs share a common RNA-binding cleft structure and

that NP-RNA interactions adopt several highly conserved or similar positively charged residues, such as K160, K171, R174, R298, and R401 (K142, K153, R156, R280, and K383 for MARV NP) (Fig. 10d, Fig. 11). These results suggest that filovirus NPs employ a common RNA-binding mechanism.

The mechanisms by which filoviruses recognize and distinguish between host RNA and vRNA are not well understood. In artificial protein expression systems, filovirus NPs randomly bind to host RNAs (hRNAs) in the cytoplasm to form an NP–hRNA complex [46-48]. However, it is difficult to determine the role of NP–hRNA complexes in the viral replication cycle. Because of the inability to independently recognize RNA, NP may be assisted by other nucleocapsid components, such as VP35 and VP30, to facilitate the formation of the NP–vRNA complex, which can be used for vRNA synthesis. Moreover, during viral genome replication, a full-length complementary copy of the negative-sense RNA genome (cRNA), that is, the positive-sense antigenome, is synthesized to serve as a template for the production of more genomes [32, 33]. Similar to the genome, the antigenome is also encapsidated by the nucleocapsid components, indicating the existence of the NP–cRNA complex. However, the NP–cRNA complex participates in viral genome replication and is not packaged into progeny viral particles. Although this mechanism remains to be elucidated, the NP–cRNA complex may not form structurally mature nucleocapsids that are ready for transport and packaging into progeny virus particles, unlike NP-vRNA complexes.

The formation of a helical NP–RNA complex requires NP oligomerization and RNA binding. A recent fluorescence analysis using an RNA-binding-deficient EBOV NP showed that RNA binding is not required for the formation of inclusion bodies [71]. The N-terminal arm plays a critical role in NP oligomerization and interactions between neighboring NPs [45, 46-48, 62, 63]. This is consistent with my finding that a single-residue mutation in a short helix located on the N-terminal arm abolished the formation of helical structures, as well as NP–RNA multimers (Fig. 15 a). These results underscore how the interaction between the N-terminal arm and the hydrophobic pocket may initiate NP oligomerization and subsequently, the formation of the NP–RNA complex. In EBOV and MARV, the N-terminus of VP35 can bind to the NP hydrophobic pocket, resulting in an RNA-free NP-VP35 complex [62, 63]. Because LLOV NP also possesses a hydrophobic pocket similar to that of EBOV and MARV NPs, and most hydrophobic residues in this region are conserved, the mechanism of the NP-VP35 interaction in LLOV is likely similar to that in EBOV and MARV. Interestingly, NP has two VP35-binding regions: a hydrophobic pocket on the inner side of the NP–RNA helix and an

intrinsically disordered region encompassing residues 481–500 on the outer side of the NP–RNA helix [72]. The two stages of the NP-VP35 interaction suggest that VP35 may associate with NP before the formation of the NP–RNA complex and likely contribute to its assembly.

Inter-strand NP-NP interactions are only mediated by electrostatic interactions. These interactions are sensitive to the changes in the ionic strength and pH of the solution buffer, which may be why NP–RNA helix is easily uncoiled during purification. Pairs of the charged residues could maintain electrostatic interactions over a range of distances, resulting in a flexible helical structure, which could explain the flexibility of the filoviruses particles.

Among the mammalian filoviruses, LLOV appears to be more similar to EBOV. This includes similarities in the replication promoter structure [73], the requirement of VP30 for minigenome transcription [30], and the use of VP24 to inhibit interferon signaling [74]. The studies also demonstrated that the structure of LLOV NP–RNA helix is similar to that of EBOV instead of MARV, which is consistent with the previous studies. The two distinct patterns of NP–RNA helical structures may correspond to the different strategies for genome transcription and replication adopted by EBOV, LLOV and MARV.

In summary, this study presents two cryo-EM structures of LLOV NP–RNA complexes. Structural studies on various filoviruses should aid in a comprehensive understanding of the replication cycle of filoviruses and contribute to the development of broad-spectrum structure-based anti-filoviral therapies.

Bibliography

[1]. Kuhn J H, Amarasinghe G K, Perry D L (2021) Fields Virology, 7th ed., Vol.1: Emerging viruses, ed. Howley P M & Knipe D M, Ch. 11, 449-503 (Wolters Kluwer).

[2]. Rougeron V, Feldmann H, Grard G, Becker S & Leroy E M (2015) Ebola and Marburg haemorrhagic fever. Journal of Clinical Virology 64, 111-119.

[3]. Messaoudi I, Amarasinghe G K & Basler C F (2015) Filovirus pathogenesis and immune evasion: insights from Ebola virus and Marburg virus. Nat Rev Microbiol 13, 663-676.

[4]. Siegert R, Shu H L., Slenczka W, Peters D & Müller G (1967) On the etiology of an unknown human infection originating from monkeys [German]. Dtsch Med Wochenschr 92, 2341-2343.

[5]. Languon, S. & Quaye, O. Filovirus Disease Outbreaks: A Chronological Overview. Virology (Auckl) 10, 1178122X19849927 (2019).

[6]. Jacob S T et al. (2020) Ebola virus disease. Nat Rev Dis Primers 6, 13.

[7]. Negredo A et al. (2011) Discovery of an Ebolavirus-Like Filovirus in Europe. PLOS Pathogens 7, e1002304.

[8]. Shi M et al. (2018) The evolutionary history of vertebrate RNA viruses. Nature 556, 197-202.

[9]. Yang X.-L et al. (2019) Characterization of a filovirus (Měnglà virus) from Rousettus bats in China. Nat Microbiol 4, 390-95.

[10]. Horie M (2021) Identification of a novel filovirus in a common lancehead (Bothrops atrox (Linnaeus, 1758)). J Vet Med Sci 83, 1485-1488.

[11]. Kuhn, J. H. et al. ICTV Virus Taxonomy Profile: Filoviridae. J Gen Virol 100, 911-912 (2019).

[12]. Kemenesi, G. et al. Re-emergence of Lloviu virus in Miniopterus schreibersii bats, Hungary,

2016. Emerging Microbes & Infections 7, 1-4 (2018).

[13]. Kemenesi, G. et al. Isolation of infectious Lloviu virus from Schreiber's bats in Hungary. Nat Commun 13, 1706 (2022).

[14]. Tóth, G. E. et al. Isolation and genome characterization of Lloviu virus from Italian Schreibers's bats. Sci Rep 13, 11310 (2023).

[15]. Towner, J. S. et al. Isolation of Genetically Diverse Marburg Viruses from Egyptian Fruit Bats. PLoS Pathog 5, e1000536 (2009).

[16]. Hume, A. J. et al. Recombinant Lloviu virus as a tool to study viral replication and host responses. PLoS Pathog 18, e1010268 (2022).

[17]. Maruyama, J. et al. Characterization of the Envelope Glycoprotein of a Novel Filovirus, Lloviu Virus. Journal of Virology 88, 99–109 (2014).

[18]. Elliott L H, Kiley M P & McCormick J B (1985) Descriptive analysis of Ebola virus proteins. Virology 147, 169-176.

[19]. Bornholdt Z A et al. (2013) Structural rearrangement of ebola virus VP40 begets multiple functions in the virus life cycle. Cell 154, 763-774.

[20]. Lee J E et al. (2008) Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. Nature 454, 177-182.

[21]. Bale S et al. (2012) Structural basis for differential neutralization of ebolaviruses. Viruses 4, 447-470.

[22]. Saeed M F, Kolokoltsov A A, Albrecht T & Davey R A (2010) Cellular entry of ebola virus involves uptake by a macropinocytosis-like mechanism and subsequent trafficking through early and late endosomes. PLoS Pathog 6, e1001110.

[23]. Nanbo A et al. (2010) Ebolavirus is internalized into host cells via macropinocytosis in a viral glycoprotein-dependent manner. PLoS Pathog 6, e1001121.

[24]. Ng, M. et al. Cell entry by a novel European filovirus requires host endosomal cysteine proteases and Niemann-Pick C1. Virology 468-470, 637-646 (2014).

[25]. Carette J E et al. (2011) Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. Nature 477, 340-343.

[26]. Wang H et al. (2016) Ebola Viral Glycoprotein Bound to Its Endosomal Receptor Niemann-Pick C1. Cell 164, 258-268.

[27]. Wang L L et al. (2021) Identification of Filovirus Entry Inhibitors Targeting the Endosomal Receptor NPC1 Binding Site. Antiviral Res 189, 105059.

[28]. Yuan, B. et al. Structure of the Ebola virus polymerase complex. Nature 610, 394–401 (2022).

[29]. Peng, Q. et al. Molecular mechanism of de novo replication by the Ebola virus polymerase. Nature 622, 603–610 (2023).

[30]. Manhart, W. A. et al. A Chimeric Lloviu Virus Minigenome System Reveals that the Bat-Derived Filovirus Replicates More Similarly to Ebolaviruses than Marburgviruses. Cell Rep 24, 2573-2580.e4 (2018).

[31]. Kämper, L. et al. Assessment of the function and intergenus-compatibility of Ebola and Lloviu virus proteins. Journal of General Virology 100, 760-772 (2019).

[32]. Mühlberger, E., Weik, M., Volchkov, V. E., Klenk, H.-D. & Becker, S. Comparison of the Transcription and Replication Strategies of Marburg Virus and Ebola Virus by Using Artificial Replication Systems. Journal of Virology 73, 2333-2342 (1999).

[33]. Mühlberger E, Lötfering B, Klenk H D & Becker S (1998) Three of the four nucleocapsid

proteins of Marburg virus, NP, VP35, and L, are sufficient to mediate replication and transcription of Marburg virus-specific monocistronic minigenomes. J Virol 72, 8756-8764.

[34]. Wendt, L. et al. Evidence for Viral mRNA Export from Ebola Virus Inclusion Bodies by the Nuclear RNA Export Factor NXF1. J Virol e0090022 (2022) doi:10.1128/jvi.00900-22.

[35]. Hoenen, T. et al. Inclusion bodies are a site of ebolavirus replication. J Virol 86, 11779–11788 (2012).

[36]. Watanabe S, Noda T, Halfmann P, Jasenosky L & Kawaoka Y (2007) Ebola virus (EBOV) VP24 inhibits transcription and replication of the EBOV genome. J Infect Dis 196 Suppl 2, S284-290.

[37]. Hoenen T, Jung S, Herwig A, Groseth A & Becker S (2010) Both matrix proteins of Ebola virus contribute to the regulation of viral genome replication and transcription. Virology 403, 56-66.

[38]. Mateo M et al. (2011) Knockdown of Ebola virus VP24 impairs viral nucleocapsid assembly and prevents virus replication. J Infect Dis 204 Suppl 3, S892-896.

[39]. Dolnik, O., Stevermann, L., Kolesnikova, L. & Becker, S. Marburg virus inclusions: A virus-induced microcompartment and interface to multivesicular bodies and the late endosomal compartment. Eur J Cell Biol 94, 323-331 (2015).

[40]. Schudt, G., Kolesnikova, L., Dolnik, O., Sodeik, B. & Becker, S. Live-cell imaging of Marburg virus-infected cells uncovers actin-dependent transport of nucleocapsids over long distances. Proc Natl Acad Sci U S A 110, 14402-14407 (2013).

[41]. Han Z et al. (2015) ALIX Rescues Budding of a Double PTAP/PPEY L-Domain Deletion Mutant of Ebola VP40: A Role for ALIX in Ebola Virus Egress. J Infect Dis 212 Suppl 2, S138-145.

[42]. Urata S et al. (2007) Interaction of Tsg101 with Marburg virus VP40 depends on the PPPY

motif, but not the PT/SAP motif as in the case of Ebola virus, and Tsg101 plays a critical role in the budding of Marburg virus-like particles induced by VP40, NP, and GP. J Virol 81, 4895–4899.

[43]. Ruigrok, R. W. H., Crépin, T. & Kolakofsky, D. Nucleoproteins and nucleocapsids of negative-strand RNA viruses. Curr Opin Microbiol 14, 504-510 (2011).

[44]. Watanabe, S., Noda, T. & Kawaoka, Y. Functional mapping of the nucleoprotein of Ebola virus. J Virol 80, 3743-3751 (2006).

[45]. Hu, S. & Noda, T. Filovirus helical nucleocapsid structures. Microscopy (Oxf) 72, 178-190 (2023).

[46]. Sugita, Y., Matsunami, H., Kawaoka, Y., Noda, T. & Wolf, M. Cryo-EM structure of the Ebola virus nucleoprotein–RNA complex at 3.6 Å resolution. Nature 563, 137-140 (2018).

[47]. Kirchdoerfer, R. N., Saphire, E. O. & Ward, A. B. Cryo-EM structure of the Ebola virus nucleoprotein–RNA complex. Acta Cryst F 75, 340-347 (2019).

[48]. Fujita-Fujiharu, Y. et al. Structural insight into Marburg virus nucleoprotein–RNA complex formation. Nat Commun 13, 1191 (2022).

[49]. Dziubańska, P. J., Derewenda, U., Ellena, J. F., Engel, D. A. & Derewenda, Z. S. The structure of the C-terminal domain of the Zaire ebolavirus nucleoprotein. Acta Crystallogr D Biol Crystallogr 70, 2420-2429 (2014).

[50]. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J Struct Biol 192, 216-221 (2015).

[51]. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife 7, e42166 (2018).

[52]. He, S. & Scheres, S. H. W. Helical reconstruction in RELION. J Struct Biol 198, 163-176

(2017).

[53]. Liao, H. Y. & Frank, J. Definition and estimation of resolution in single-particle reconstructions. Structure 18, 768-775 (2010).

[54]. Pettersen, E. F. et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25, 1605-1612 (2004).

[55]. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60, 2126-2132 (2004).

[56]. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66, 486-501 (2010).

[57]. Adams, P. D. et al. PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr D Biol Crystallogr 58, 1948-1954 (2002).

[58]. Goddard, T. D. et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci 27, 14-25 (2018).

[59]. Buchan, D. W. A. & Jones, D. T. The PSIPRED Protein Analysis Workbench: 20 years on. Nucleic Acids Res 47, W402-W407 (2019).

[60]. Li, T. & Shen, Q.-T. Insights into Paramyxovirus Nucleocapsids from Diverse Assemblies. Viruses 13, 2479 (2021).

[61]. Bharat, T. A. M. et al. Structural dissection of Ebola virus and its assembly determinants using cryo-electron tomography. Proceedings of the National Academy of Sciences 109, 4275-4280 (2012).

[62]. Leung, D. W. et al. An intrinsically disordered peptide from Ebola virus VP35 controls viral RNA synthesis by modulating nucleoprotein-RNA interactions. Cell Rep 11, 376-389 (2015).

[63]. Liu, B. et al. Structural Insight into Nucleoprotein Conformation Change Chaperoned by VP35

Peptide in Marburg Virus. J Virol 91, e00825-17 (2017).

[64]. Renner, M. et al. Nucleocapsid assembly in pneumoviruses is regulated by conformational switching of the N protein. Elife 5, e12627 (2016).

[65]. Gutsche, I. et al. Structural virology. Near-atomic cryo-EM structure of the helical measles virus nucleocapsid. Science 348, 704-707 (2015).

[66]. Green, T. J., Zhang, X., Wertz, G. W. & Luo, M. Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. Science 313, 357-360 (2006).

[67]. Hastie, K. M. et al. Crystal structure of the Lassa virus nucleoprotein-RNA complex reveals a gating mechanism for RNA binding. Proc Natl Acad Sci U S A 108, 19365-19370 (2011).

[68]. Alayyoubi, M. et al. Structure of the paramyxovirus parainfluenza virus 5 nucleoprotein-RNA complex. Proc Natl Acad Sci U S A 112, E1792-1799 (2015).

[69]. Larkin, M. A. et al. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947-2948 (2007).

[70]. Gouet, P., Robert, X. & Courcelle, E. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res 31, 3320-3323 (2003).

[71]. Bodmer, B. S. et al. Ebola virus inclusion bodies are liquid organelles whose formation is facilitated by nucleoprotein oligomerization. Emerg Microbes Infect 12, 2223727 (2023).

[72]. Miyake, T. et al. Ebola Virus Inclusion Body Formation and RNA Synthesis Are Controlled by a Novel Domain of Nucleoprotein Interacting with VP35. J Virol 94, e02100-19 (2020).

[73]. Hume, A. J. & Möhlberger, E. Distinct genome replication and transcription strategies within the growing filovirus family. J Mol Biol 431, 4290–4320 (2019).

[74]. Feagins, A. R. & Basler, C. F. Lloviu virus VP24 and VP35 proteins function as innate immune antagonists in human and bat cells. Virology 485, 145–152 (2015).

Acknowledgements

This doctoral thesis was based on the following articles:

Hu Shangfan, Fujita-Fujiharu Yoko, Sugita Yukihiko, Wendt Lisa, Muramoto Yukiko, Nakano Masahiro, Hoenen Thomas, Noda Takeshi. **Cryoelectron microscopic structure of the nucleoprotein–RNAcomplex of the European filovirus, Lloviu virus.** PNAS Nexus. 2023 Apr 6;2(4):pgad120. doi: 10.1093/pnasnexus/pgad120. PMID: 37124400; PMCID: PMC10139700.

I would like to thank my supervisors, Prof. Takeshi Noda, Assistant Prof. Masahiro Nakano, Assistant Prof. Yukiko Muramoto, and Associate Prof. Yukihiko Sugita, for their guidance during this process. Their support helped me finish my doctoral degree and grow as a researcher and person.

Second, I thank Prof. Ayato Takada of Hokkaido University for providing the plasmids used in this study. I would also like to thank Prof. Thomas Hoenen and Dr. Lisa Wendt for supplying the plasmids and data for the minigenome assay.

In addition, I would like to thank my colleagues in the laboratory, with special thanks also to Ai Hirabayashi and Yoko Fujita for their support with the electron microscopy technology and data analysis, as well as to Yuki Takamatsu, Sho Miyamoto, Junichi Kajikawa, and Kazuya Houri for their useful advice on the hemorrhagic fever virus research. I am also grateful to everyone for their tremendous aid in my research and daily life. Their advice, experimental expertise, and friendships are invaluable.

Moreover, I would like to thank Drs. Mika Hirose and Takayuki Kato at the Institute for Protein Research, Osaka University, for providing the support in the cryo-EM facility and the Otsuka Toshimi Scholarship for the financial support from 2021–2024.

Finally, I thank my family members, especially my parents and sister, for their financial support and encouragement.