

CONFORMATIONAL STATES AND IMMUNOLOGICAL TARGETING
OF THE ARENAVIRAL GLYCOPROTEIN COMPLEX:
IMPLICATIONS FOR THERAPEUTIC DESIGN



WENG MING NG

MEDICAL SCIENCES DOCTORAL TRAINING CENTRE
NEW COLLEGE
UNIVERSITY OF OXFORD

A thesis submitted in fulfilment of the requirements for the degree of *Doctor of Philosophy*

Trinity Term 2021

ABSTRACT

Envelope-displayed viral glycoproteins mediate host cell entry and are a key determinant of species tropism. As the glycoproteins are the sole antigens exposed on the mature virion surface, they are also primary targets of the host humoral immune response that arises during infection and immunisation. Macromolecular structure determination of these therapeutically important targets constitutes an important approach to obtain molecular-level insights into viral protein structure and function.

The emergence and re-emergence of Old World (OW) and New World (NW) mammarenaviruses (family *Arenaviridae*) from rodent reservoirs persistently threatens human health. The GP1 subunit of the envelope glycoprotein spike complex (GP) is responsible for receptor attachment during host cell entry. Here, I report the GP1 structure of OW Loei River virus (LORV) and the GP1 structures of NW Junín virus (JUNV) and NW Machupo virus (MACV) in complex with neutralising antibodies (nAbs). Comparative analyses of the GP1 glycoprotein architectures of LORV and previously characterised mammarenaviruses expands our understanding of the structural classes sampled by the GP1 glycoproteins and provide molecular blueprint for immunogen design that is capable of mimicking the mature OW mammarenaviral GP. Glycan mapping analysis further revealed a pronounced glycan-mediated masking of the LORV GP, rationalising the observed immune-evasive properties of OW mammarenaviruses. Furthermore, structural characterisation of the nAb epitopes on JUNV and MACV GP1s provides a template for understanding the antigenic surface of the NW GP1 and augments efforts to develop nAb-based prophylactics that target the NW arenavirus surface.

Rabies lyssavirus (RABV; family *Rhabdoviridae*) is the causative agent of rabies, which causes approximately 60,000 deaths annually. NAb-based therapeutics constitute an important part of the rabies treatment. Structure elucidation of the higher-order assembly of RABV glycoprotein (RV-G) allows identification and characterisation of the full antigenic surface of the spike, including the quaternary nAb epitopes. Here, I characterised the structure of RV-G in complex with two nAbs. The structural data revealed the architecture of the first full-length lyssavirus glycoprotein structure in prefusion state and defined nAb-targeting epitopes in the context of a trimeric spike. This work serves as a template for structure-guided engineering of next generation RABV vaccines. The structural insights on the nAb epitopes may also augment the optimisation of pre-existing antibody-based post-exposure prophylactics.

Through crystallographic and cryogenic electron microscopy analysis, the work included in this thesis expand our appreciation of the conformational landscape covered by the glycoproteins of mammarenaviruses and lyssaviruses. The data presented will empower strategies focused on the development of immunogens that resemble the glycoprotein spike presented on the mature virion and capable of eliciting protective immune response.

ACKNOWLEDGEMENTS

I would like to thank my supervisors Thomas Bowden and Juha Huiskonen for being incredibly supportive throughout the entire duration of my DPhil. Their guidance and helpful discussions have helped me to become a junior researcher who is increasingly more independent. I also thank the current and former members from the Bowden and Huiskonen groups for their teaching, assistance, and patience in doing experiments with me. This gratitude goes to Guido Paesen, Antra Zeltina, Ilona Rissanen, Rhys Pryce, Alice Stelfox, Robert Stass, Elizabeth Allen, Victoria Avanzato, Mengqiu Li, and Serban Ilca. I must also specially thank Max Renner who patiently listened to all my scientific and non-scientific ramblings for the past three years.

I would also like to thank the staff members of STRUBI and OPIC for their managerial and administrative supports. My DPhil and experiments would not have been possible without the supports and encouragements from Margaret Jones, Weixian Lu, Oliver Hayward, and Tom Walter. I further thank the OPIC team, especially Dimple Karia, Loic Carrique, and Helen Duyvesteyn for their helps in EM experiments. I would also like to acknowledge Karl Harlos and Guido Paesen who kindly flash-cooled my precious protein crystals. Many thanks also to the admin team, Agata Krupa, Melissa Matthews, Claire Morris, and Jun Dong who made sure my DPhil work was untroubled by administrative work. I also thank the University of Oxford for having awarded me the Clarendon Scholarship and Wellcome for having funded my DPhil.

My DPhil would not be complete without the cooperative efforts from the collaborators. I thank Katie Doores and her team members Jeffrey Seow and Stefanie Krumm at King's College London for providing antibody binding and neutralisation data. I also thank Daniel Pinschewer and his team member Mehmet Sahin at University of Basel for conducting animal infection studies and performing isolation of antigen-specific B cells. I also thank Alexander Douglas and his team members Sofiya Fedosyuk and Luke Thorley at the Jenner Institute for providing samples of rabies glycoproteins. I look forward to more fruitful collaborations with these brilliant scientists.

I dedicate this thesis to my parents, Tek Hiung and Chee Kuen. Their supports, patience, encouragements, and pride in me are priceless and incomparable.

TABLE OF CONTENTS

ABSTRACT.....	II
ACKNOWLEDGEMENTS.....	III
LIST OF FIGURES.....	VIII
LIST OF TABLES.....	X
LIST OF ABBREVIATIONS.....	XI
1 INTRODUCTION.....	1
1.1 The Family <i>Arenaviridae</i> : An Overview.....	1
1.2 Mammarenaviruses.....	5
1.2.1 <i>The Impacts of Mammarenaviruses</i>	5
1.2.2 <i>Epidemiology and Clinical Features</i>	9
1.2.3 <i>Genome Organisation and Gene Expression</i>	12
1.2.4 <i>Structure and Components of the Virion</i>	15
1.2.5 <i>Lifecycle and Replication</i>	23
1.2.5.1 <i>Host Cell Entry</i>	23
1.2.5.2 <i>Endocytosis and pH-Dependent Fusion</i>	26
1.2.5.3 <i>Transcription, Translation, and Genome Replication</i>	29
1.2.5.4 <i>Virion Assembly and Budding</i>	31
1.2.6 <i>Therapeutics</i>	33
1.3 Mammarenaviral Glycoprotein Complex (GP).....	36
1.3.1 <i>Structure and Conformations</i>	36
1.3.2 <i>Shielding of the Protein Surface by N-linked Glycosylation</i>	49
1.4 Rabies Virus.....	52
1.4.1 <i>Rabies Disease</i>	52
1.4.2 <i>RABV Lifecycle</i>	54
1.4.3 <i>Rabies Virus Glycoprotein (RV-G)</i>	57
1.5 Thesis Aims.....	63
2 MATERIALS AND METHODS.....	65
2.1 Molecular Biology and Microbiology.....	65
2.1.1 <i>Complementary DNA (cDNA) Templates</i>	65
2.1.2 <i>Cloning of cDNA into Expression Vectors</i>	66

2.1.2.1	<i>Primer Specifications</i>	66
2.1.2.2	<i>Expression and Transfer Vectors</i>	67
2.1.3	<i>Polymerase Chain Reaction (PCR)</i>	68
2.1.4	<i>Agarose Gel Electrophoresis</i>	69
2.1.5	<i>Restriction Enzyme Digestion</i>	69
2.1.6	<i>DNA Ligation</i>	70
2.1.7	<i>Transformation of Chemically-Competent E. coli</i>	70
2.1.8	<i>Plasmid Amplification and Purification</i>	71
2.1.9	<i>Gene Mutagenesis Using Splicing by Overlap Extension PCR</i>	72
2.2	<i>Mammalian Tissue Culture</i>	73
2.2.1	<i>Cultivation of Mammalian Cell Lines</i>	73
2.2.2	<i>Production of Stable Cell Lines</i>	74
2.2.3	<i>Storage and Revival of Cell Lines</i>	74
2.3	<i>Recombinant Protein Expression and Purification</i>	75
2.3.1	<i>Small-Scale Protein Expression</i>	75
2.3.2	<i>SDS-PAGE</i>	76
2.3.3	<i>Western Blot</i>	76
2.3.4	<i>Large-Scale Protein Expression</i>	77
2.3.5	<i>Cell Supernatant Harvesting and Diafiltration</i>	78
2.3.6	<i>Immobilised Metal Ion Affinity Chromatography (IMAC)</i>	78
2.3.7	<i>Enzymatic Deglycosylation</i>	79
2.3.8	<i>Size Exclusion Chromatography (SEC)</i>	79
2.3.9	<i>Enzyme-Linked Immunosorbent Assay (ELISA)</i>	80
2.4	<i>Virus-Like Particle (VLP) Production and Purification</i>	81
2.4.1	<i>VLP Stable Cell Line Culture</i>	81
2.4.2	<i>VLP Purification</i>	82
2.5	<i>Structural Biology Methods</i>	82
2.5.1	<i>Protein Crystallisation</i>	82
2.5.2	<i>Cryo-Protection and Flash Cooling of Protein Crystals</i>	83
2.5.3	<i>X-Ray Data Collection and Processing</i>	83
2.5.4	<i>Cryogenic Electron Microscopy</i>	84
2.5.4.1	<i>Grid Preparation</i>	84
2.5.4.2	<i>Data Collection and Processing</i>	85
2.6	<i>Structure-Based Phylogenetic Analysis (SBPA)</i>	86

3 DIFFERENTIAL CONFORMATIONS OF OLD AND NEW WORLD

ARENAVIRAL GP1 GLYCOPROTEINS.....	87
3.1 Introduction.....	87
3.2 Protein Production and Structure Determination of Old World Loei River Virus (LORV) GP1.....	89
3.2.1 <i>Construct Design and Small-Scale Transient Expression</i>	89
3.2.2 <i>Large-Scale Transient Expression and Purification</i>	91
3.2.3 <i>LORV GP1 Crystallisation Experiment</i>	91
3.2.4 <i>X-Ray Diffraction Data Collection and Processing</i>	92
3.3 Structure Analysis.....	93
3.3.1 <i>Structure Overview of LORV GP1</i>	93
3.3.2 <i>Structure Comparison of LORV GP1 with Characterised Arenaviral GP1</i>	97
3.3.3 <i>Structural Analysis of Characterised Receptor-Binding Sites</i>	100
3.3.4 <i>LORV GP1 Is Highly Glycosylated</i>	106
3.4 Structure-Based Phylogenetic Analysis (SBPA) of the Arenaviral GP1.....	108
3.4.1 <i>Introduction</i>	108
3.4.2 <i>Background of SBPA</i>	109
3.4.3 <i>SBPA of the Arenaviral GP1</i>	112
3.5 Summary and Discussion.....	115

4 ANTIBODY-MEDIATED NEUTRALISATION OF NEW WORLD JUNÍN

VIRUS AND MACHUPO VIRUS.....	119
4.1 Introduction.....	119
4.2 Protein Production and Structure Determination of New World Junín Virus (JUNV) and Machupo Virus (MACV) GP1s in Complex with Fab Fragments.....	122
4.2.1 <i>Large-Scale Transient Protein Expression and Purification</i>	122
4.2.2 <i>Crystallisation of GP1–Fab Complexes</i>	123
4.2.3 <i>X-Ray Diffraction Data Collection and Processing</i>	123
4.3 Structure Analysis.....	126
4.3.1 <i>Structural Basis of JUN1-Mediated Neutralisation of JUNV GP1</i>	126
4.3.2 <i>Structural Basis for MACV GP1 Recognition by nAb MAC1</i>	135
4.3.3 <i>Comparison of Anti-MACV GP1 nAbs MAC1 and CR1-07</i>	143

4.4 Summary and Discussion.....	144
5 STRUCTURE ELUCIDATION OF RABIES VIRUS GLYCOPROTEIN.....	150
5.1 Introduction.....	150
5.2 Single-Particle Cryo-EM Analysis of RV-G in complex with the Fab Fragments of nAbs 17C7 and 1112-1.....	152
5.2.1 <i>Protein Production and Purification</i>	152
5.2.2 <i>Cryo-EM Grid Preparation and Data Collection</i>	153
5.2.3 <i>Cryo-EM Single-Particle Data Processing</i>	155
5.2.3.1 <i>Preliminary Glacios Data Set Processing</i>	157
5.2.3.2 <i>High-Resolution Titan Krios Data Set Processing</i>	168
5.3 Summary and Discussion.....	175
6 CONCLUDING REMARKS AND FUTURE WORK.....	179
6.1 Mammarenavirus.....	179
6.2 Rabies Virus.....	185
7 REFERENCES.....	188
8 APPENDIX.....	226
8.1 cDNA and Amino Acid Sequences of Fab Fragments.....	226
8.2 List of Publications.....	232

LIST OF FIGURES

Fig. 1.1: Maximum likelihood phylogenetic tree of the complete L amino acid sequences of 50 arenaviruses assigned to the family <i>Arenaviridae</i>	3
Fig. 1.2: Maximum likelihood phylogenetic trees of the L amino acid sequences of the genus <i>Mammarenavirus</i>	4
Fig. 1.3: Geographic distribution of human pathogenic mammarenaviruses.....	8
Fig. 1.4: Arenavirus genome replication and gene transcription.....	12
Fig. 1.5: Structure overview of LASV virion and GP spike.....	17
Fig. 1.6: Structure overview of the mammarenaviral Z, NP, and L proteins.....	22
Fig. 1.7: Host cell entry pathways utilised by OW and NW mammarenaviruses.....	27
Fig. 1.8: Schematic representation of mammarenavirus gene transcription, replication, and translation.....	30
Fig. 1.9: Model of arenavirus budding.....	32
Fig. 1.10: Structure overview of NW mammarenaviral GP1s.....	39
Fig. 1.11: Structure overview of OW mammarenaviral GP1s.....	42
Fig. 1.12: Structure overview of OW and NW mammarenaviral GP2s.....	47
Fig. 1.13: Compositional analysis and structure-based mapping of LASV GP N-linked glycans.....	52
Fig. 1.14: RABV lifecycle.....	56
Fig. 1.15: An overview of the structures of class III viral membrane fusion proteins.....	58
Fig. 1.16: An overview of reported RV-G ectodomain structures at pH 8.0 and pH 6.5.....	60
Fig. 1.17: Representation of antigenic sites on the RV-G.....	62
Fig. 2.1: Plasmid maps.....	67
Fig. 2.2: A schematic outline of site-directed mutagenesis accomplished by OE-PCR.....	72
Fig. 3.1: A maximum-likelihood phylogeny of 35 arenaviral GP amino acid sequences dividing the genus into OW and NW (clades A to D) serogroups.....	88
Fig. 3.2: Protein construct design and expression test of LORV GP1.....	90
Fig. 3.3: Purification and crystallisation of LORV GP1.....	92
Fig. 3.4: Structure and organisation of LORV GP1.....	96
Fig. 3.5: Structure comparison of LORV GP1 with characterised mammarenaviral GP1s.....	99
Fig. 3.6: Comparison of LAMP1 and α -DG binding residues on the GP2-associated LASV GP1 and the GP2-free LORV GP1.....	101

Fig. 3.7: Structural comparison of strand $\beta 5$ and loop L7 of GP2-free LORV, LASV, and MORV GP1s.....	104
Fig. 3.8: Mapping of LORV N-linked glycosylation sites onto the structure of trimeric LASV GP1–GP2.....	108
Fig. 3.9: Structure-based phylogenetic tree showing the conformations adopted by arenaviral GP1 attachment glycoproteins.....	115
Fig. 4.1: Purification and crystallisation of JUNV GP1–Fab JUN1 and MACV GP1–Fab MAC1 complexes.....	124
Fig. 4.2: Isolation and characterisation of JUNV and MACV GP1-reactive mAbs.....	127
Fig. 4.3: Structural basis for Fab JUN1 recognition of JUNV GP1.....	129
Fig. 4.4: Amino acid sequence alignments of Fab JUN1 and Fab MAC1 with respect to their closest corresponding germline sequences.....	131
Fig. 4.5: Tyrosine-mediated JUNV GP1 and MACV GP1 recognition: a common mode of mAb neutralisation.....	132
Fig. 4.6: Schematic diagram of contacts formed in the JUNV GP1–Fab JUN1 interface.....	133
Fig. 4.7: Visualisation of crystallographically ordered N-linked glycans on JUNV GP1.....	135
Fig. 4.8: Structural basis for Fab MAC1 recognition of MACV GP1.....	136
Fig. 4.9: Comparison of MACV GP1–Fab MAC1 and MACV GP1–Fab CR1-07 complex interfaces.....	138
Fig. 4.10: Structure overlay analysis of MACV GP1 with Fabs JUN1, OD01, GD01, and CR1-28.	139
Fig. 4.11: Schematic diagram of contacts formed in the MACV GP1–Fab MAC1 interface..	141
Fig. 4.12: Visualisation of crystallographically ordered N-linked glycans on MACV GP1...	142
Fig. 4.13: Footprints of hTfR1, Fab MAC1, and Fab CR1-07 on MACV GP1.....	144
Fig. 4.14: Models of JUNV GP1–Fab JUN1, MACV GP1–Fab MAC1, and MACV GP1–hTfR1 complexes in the context of the expected higher-order trimeric GP assembly...	147
Fig. 5.1: Purification of full-length trimeric RV-G in complex with Fabs 17C7 and 1112-1..	153
Fig. 5.2: Initial cryo-EM analysis of full-length RV-G in complex with the Fab of a neutralising mAb, termed RVC20.....	159
Fig. 5.3: Initial cryo-EM analysis of full-length RV-G in complex with the Fab of a neutralising mAb, termed 17C7.....	161
Fig. 5.4: Initial cryo-EM analysis of full-length RV-G in complex with Fabs RVC20 and 17C7.....	162

Fig. 5.5: Competition between IgG 1112-1 and Fab 17C7 or Fab RVC20 for binding to RV-G spike determined by ELISA.....	164
Fig. 5.6: Initial cryo-EM analysis of full-length RV-G in complex with Fabs 1112-1 and 17C7.....	165
Fig. 5.7: Initial cryo-EM analysis of full-length RV-G in complex with the Fab of a neutralising mAb, termed E559.....	167
Fig. 5.8: Cryo-EM data processing workflow for the RV-G–Fab 17C7–Fab 1112-1 complex.....	170
Fig. 5.9: Initial cryo-EM analysis of the higher-order assembly of the RV-G–(Fab 17C7) ₃ –(Fab 1112-1) ₂ complex.....	173
Fig. 5.10: Rigid fitting of RV-G, VSV-G, and CHAV-G ectodomain crystal structures into the cryo-EM-derived 5-Fab bound full-length trimeric RV-G map.....	174
Fig. 6.1: Preparation of LASV GP-displayed VLPs lays a groundwork for cryo-ET investigation of mAb recognition of LASV GP	184

LIST OF TABLES

Table 1.1: Summary of mammarenaviral GP1 structures determined to date.....	43
Table 1.2: Summary of mammarenaviral GP2 structures determined to date.....	48
Table 1.3: <i>N</i> -linked glycosylation location in mammarenaviral GP sequences.....	50
Table 2.1: List of studied protein constructs.....	65
Table 2.2: PCR thermocycling conditions.....	69
Table 2.3: Summary of cell types used and their corresponding growth media.....	73
Table 2.4: Summary of the production and usage of stable cell lines.....	74
Table 3.1: Crystallographic data collection and refinement statistics for LORV GP1	94
Table 4.1: Crystallographic data collection and refinement statistics for JUNV GP1–Fab JUN1 and MACV GP1–Fab MAC1 complexes.....	125
Table 4.2: Summary of structurally characterised anti-JUNV and anti-MACV nAbs.....	130
Table 5.1: Cryo-EM data collection and refinement statistics for the RV-G–(Fab 17C7) ₃ –(Fab 1112-1) ₂ complex.....	156

LIST OF ABBREVIATIONS

α/β	alpha/beta
α -DG	alpha-dystroglycan
Å	Ångström (10^{-10} m)
APBS	Adaptive Poisson-Boltzmann Solver
Asn	asparagine
AU	absorbance units
β -OG	octyl-beta-glucoside
BHK	Baby hamster kidney
BIBD	boid inclusion body disease
bnAbs	broadly neutralising antibodies
BSL	Biosafety Level
CD	central domain
CD4/CD8	cluster of differentiation 4 / 8
CDC	Centers for Disease Control and Prevention
cDNA/cRNA	complementary deoxyribonucleic acid / complementary ribonucleic acid
CDR	complementary-determining region
CEPI	Coalition for Epidemic Preparedness Innovations
CHAPS	3-[(3-cholamidopropyl) dimethylammonio]-1- propanesulfonate
CO ₂	carbon dioxide
Cryo-EM	cryogenic electron microscopy
Cryo-ET	cryogenic electron tomography
CTD	C-terminal domain
CTF	contrast transfer function
DC-SIGN	dendritic cell-specific intercellular adhesion molecule-3-grabbing-nonintegrin
DLS	Diamond Light Source
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
ECD	extracellular domain
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
EndoF ₁	endoglycosidase F ₁
Env	envelope protein
ER	endoplasmic reticulum
ESCRT	endosomal sorting complex required for transport
Fab	antigen-binding fragment
FD	fusion domain
FSC	Fourier shell correlation
G	glycoprotein
gB	glycoprotein B
GIF	Gatan Imaging Filter

GlcNAc	N-acetylglucosamine
GP	glycoprotein complex
GP1	glycoprotein 1
GP2	glycoprotein 2
GSFSC	gold-standard Fourier shell correlation
GTPase	guanosine triphosphatase
h	hour
HEK	human embryonic kidney
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HF	haemorrhagic fever
His	histidine
His ₆	hexa-histidine
HSF	Homologous Structure Finder
hTfR1	human transferrin receptor 1
hTIM1	human T-cell immunoglobulin and mucin-domain containing proteins 1
ICTV	International Committee on Taxonomy of Viruses
IFN	interferon
IGR	intergenic region
IMAC	immobilised metal ion affinity chromatography
IV	intraviral
kDa	kilodalton
L	(<i>gene</i>) Large segment / (<i>protein</i>) L polymerase / (<i>domain</i>) Late domain
LAMP1	lysosomal-associated membrane protein 1
LB	Luria broth
LSECTin	liver and lymph node sinusoidal endothelial calcium-dependent lectin
M/Z protein	matrix protein
mAb	monoclonal antibody
MCS	multiple cloning site
MDCK	Madin-Darby Canine Kidney II
min	minute
mGluR2	metabotropic glutamate receptor 2
mRNA	messenger ribonucleic acid
MUSTANG	Multiple Structural Alignment Algorithm
myr	myristoylation
nAb	neutralising antibody
nAChR	nicotinic acetylcholine receptor
NaCl	sodium chloride
NCAM	nerve cell adhesion molecule
NHPs	non-human primates
NHS	National Health Service
N-linked	asparagine-linked
N/NP	nucleoprotein
nt	nucleotide
NMR	nuclear magnetic resonance

NRP2	neurophilin-2
NTD	N-terminal domain
NTP	nucleoside triphosphate
NW	New World
OD	optical density
OE-PCR	overlap extension-polymerase chain reaction
ORF	open-reading frame
OW	Old World
P protein	phosphoprotein
p75NTR	low-affinity nerve growth factor receptor
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethylene glycol
PEI	polyethylenimine
PEP	post-exposure prophylaxis
PHD	pleckstrin homology domain
PHYLP	PHYLogeny Inference Package
px	pixel
RBS	receptor-binding site
RdRp	RNA-dependent RNA polymerase
RIG	rabies immunoglobulin
RIG-I	retinoic acid-inducible gene I
RING	Really Interesting New Gene
RMSD	root mean square deviation
RNA	ribonucleic acid
RNP	ribonucleoprotein
rpm	revolution per minute
RV-G	rabies glycoprotein
s	second
SARS-CoV	severe acute respiratory syndrome-related coronavirus
SBPA	structure-based phylogenetic analysis
scFv	single-chain variable fragment
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
SHP	Structural Homology Program
SKI-1/S1P	subtilisin kexin isozyme-1/site-1 protease
SOC	Super Optimal Broth
SPA	single-particle analysis
SPase	signal peptidase
SSP	stable signal peptide
STA	subtomogram averaging
TEM	transmission electron microscopy
TfR1	transferrin receptor 1

TLS	translation-libration-screw
T _m	melting temperature
TM	transmembrane
Tsg101	tumour susceptibility gene 101
USAMRII	United States Army Medical Research Institute of Infectious Diseases
UTR	untranslated region
V	Volts
VGCCs	voltage-gated calcium channels
VLP	virus-like particles
WHO	World Health Organization

List of virus abbreviations:

AMAV	Amapari virus
CHAV	Chandipura virus (vesiculovirus)
CHAV	Chapare virus (mammarenavirus) – causative agent of Chapare haemorrhagic fevers (CHFF)
DENV	Dengue virus
EBV	Epstein-Barr virus
GTOV	Guanarito virus – causative agent of Venezuelan haemorrhagic fevers (VHF)
HCMV	human cytomegalovirus
HISV-1	Haartman Institute snake virus-1
HIV-1	human immunodeficiency virus type-1
HSV	herpes simplex virus
JUNV	Junín virus – causative agent of Argentinian haemorrhagic fever (AHF)
LASV	Lassa virus – causative agent of Lassa fever (LF)
LATV	Latino virus
LCMV	Lymphocytic choriomeningitis virus
LORV	Loei River virus
LUJV	Lujo virus – causative agent of Lujo haemorrhagic fever (LHF)
MACV	Machupo virus – causative agent of Bolivian haemorrhagic fever (BHF)
MOKV	Mokola virus
MOPV	Mopeia virus
MORV	Morogoro virus
OLV	Oliveros virus
PICV	Pichinde virus
RABV	Rabies virus
RYKV	Ryukyo virus
SABV	Sabia virus – causative agent of Brazilian haemorrhagic fevers (BzHF)
TCRV	Tacaribe virus
VSV	vesicular stomatitis virus
WENV	Wenzhou virus
WWAV	Whitewater Arroyo virus

1 INTRODUCTION

1.1 THE FAMILY *ARENAVIRIDAE*: AN OVERVIEW

Established in 1976, the family *Arenaviridae* encompassed rodent-borne enveloped viruses with bisegmented, ambisense single-stranded RNA genome (Fenner, 1976). The genome encodes for surface glycoprotein complex (GP), matrix protein (Z), RNA-dependent RNA polymerase (L), and nucleoprotein (NP) (Hallam et al., 2018b). The family name “arena” was derived from the Latin word ‘*arenosus*’, meaning ‘sandy’, and was given due to the grainy appearance of the ribosomes acquired from host cells as visualised by electron microscope (Tani et al., 2012). The family was monogeneric with a single genus *Arenavirus*.

With the discovery of arenaviruses that infect snakes (e.g. boa constrictors and pythons) in 2012–2014 (Bodewes et al., 2013; Hetzel et al., 2013; Stenglein et al., 2012), the International Committee on Taxonomy of Viruses (ICTV) accepted a proposal to establish a new genus, *Reptarenavirus*, while the genus *Arenavirus* was renamed *Mammarenavirus*, which harbours the “classical” arenaviruses that infect mammalian hosts (Radoshitzky et al., 2015) (**Fig. 1.1**). Since then, members of the genus *Reptarenavirus* have increased to five species, namely California, Giessen, Golden, Ordinary, and Rotterdam reptarenaviruses. These reptarenaviruses are found to be the culprits of boid inclusion body disease (BIBD) that is associated with eventual death of affected snakes, and have been confirmed in captive snakes in Asia, Australia, Europe, and North America (Argenta et al., 2020). On the other hand, the genus *Mammarenavirus* consists of 39 distinct rodent and bat-borne (Downs et al., 1963; Radoshitzky et al., 2015) species according to ICTV taxonomy (**Fig. 1.2**), although the number of undiscovered and newly discovered arenaviruses that have yet to be included remains uncertain. Mammarenaviruses are further divided into Old World (OW) and New World (NW)

serogroups based on their genetic, geographical, and epidemiological relationship (Charrel et al., 2008; Radoshitzky et al., 2015).

The discovery of Haartman Institute snake virus-1 (HISV-1) in 2015 marks a departure from the genome and gene organisation of the established arenaviruses (Hepojoki et al., 2015). While HISV-1, like reptarenaviruses, infects snakes, its non-canonical RNA genome lacks an open-reading frame (ORF) for the Z protein (Hepojoki et al., 2018), reflective of a complicated snake arenaviral biology and the widespread recombination and reassortment of arenaviral genome in alethinophidian host (Stenglein et al., 2015). Identification of HISV-1 led to the formation of a third genus *Hartmanivirus* (**Fig. 1.1**) which houses alethinophid arenaviruses that lack Z coding frame (Maes et al., 2018). The genus currently consists of 4 species, including Haartman, Muikkunen, Schoolhouse, and Zurich hartmaniviruses. So far it is unclear whether hartmaniviruses contribute to BIBD pathogenesis.

The fourth genus *Antennavirus* with two new species was established in 2019 (**Fig. 1.1**) following the identification of arenaviruses that infect frogfish, namely Striated and Hairy antennaviruses (originally referred to as Wēnlǐng frogfish arenavirus 1 and 2, respectively) (Abudurexiti et al., 2019; Shi et al., 2018). Unlike mammarenaviruses, reptarenaviruses, and hartmaniviruses, antennaviruses have tri-segmented RNA genomes, with GP and NP encoded on separate segments (Shi et al., 2018).

Finally, the family *Arenaviridae* with its four genera *Mammarenavirus*, *Reptarenavirus*, *Hartmanivirus*, and *Antennavirus* is classified under the order *Bunyavirales* (**Fig. 1.1**). As the number of negative and ambisense segmented RNA viruses increases steadily, *Bunyaviridae* (previously a family) has been re-ranked as the order *Bunyavirales* to accommodate all related viruses (ICTV TaxoProp 2016.030a-vM). As per ICTV taxonomy, the order now contains twelve large and diverse family members, including *Arenaviridae*. As my DPhil research

primarily focused on the glycoproteins of mammarenaviruses, the following sections in this chapter discuss them in detail.

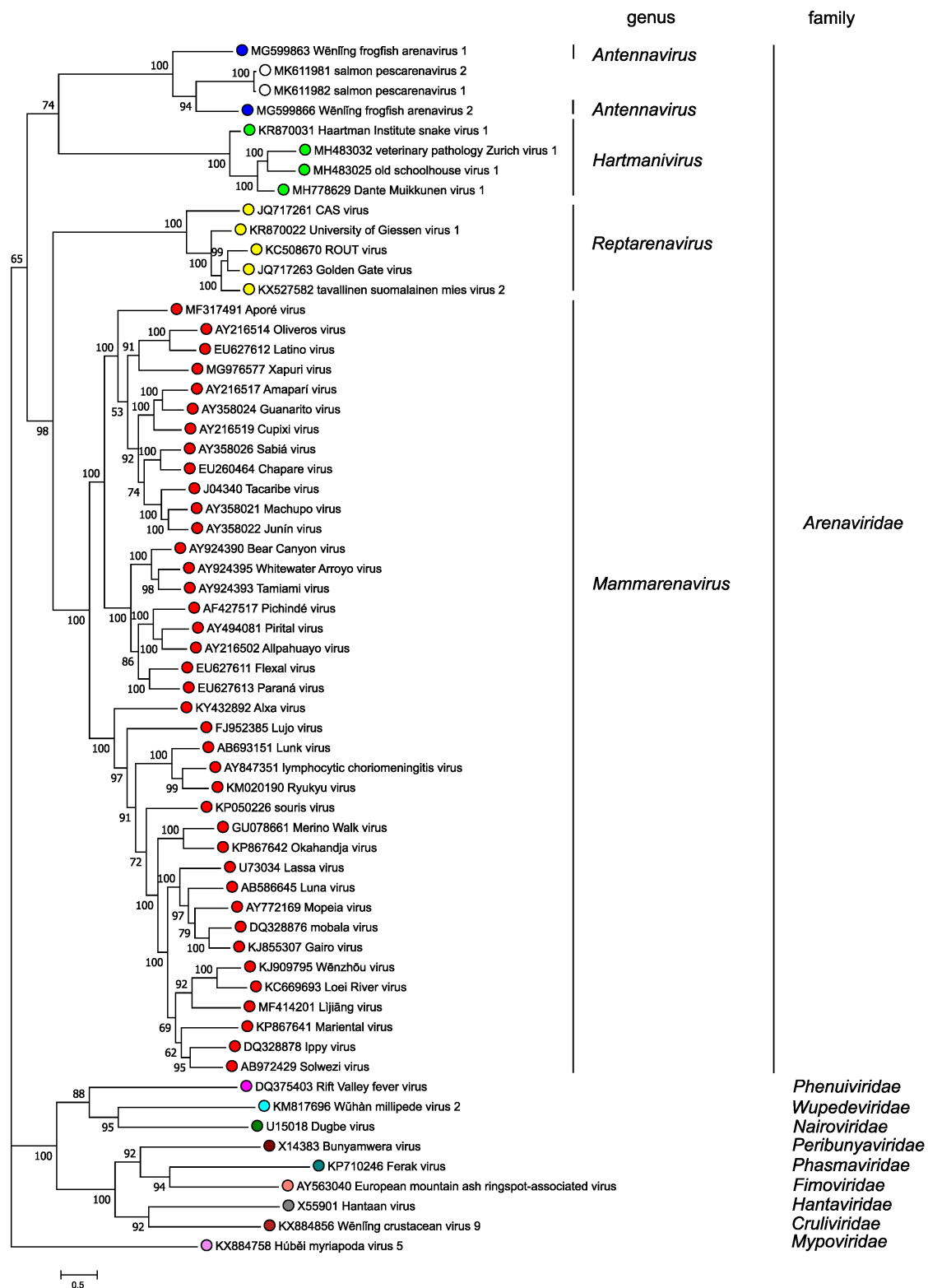


Fig. 1.1. Maximum likelihood phylogenetic tree of the complete L amino acid sequences of 50 arenaviruses assigned to the family *Arenaviridae*. The genera *Antennavirus* (blue dots;

white dots for unclassified viruses within the genus), *Hartmanivirus* (green dots), *Mammarenavirus* (red dots), and *Reptarenavirus* (yellow dots), along with representative viruses of other bunyavirus families are shown. The maximum likelihood tree with 1,000 bootstrap replicates was produced using RAXML (v. 8) and visualised using FigTree. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap is shown next to branch nodes. This figure is reproduced from the ICTV webpage: <https://talk.ictvonline.org/ictv-reports/ictv-online-report/negative-sense-rna-viruses/w/arenaviridae/>, where this work is licensed under a Creative Commons Attribution License (CC BY-SA 4.0).

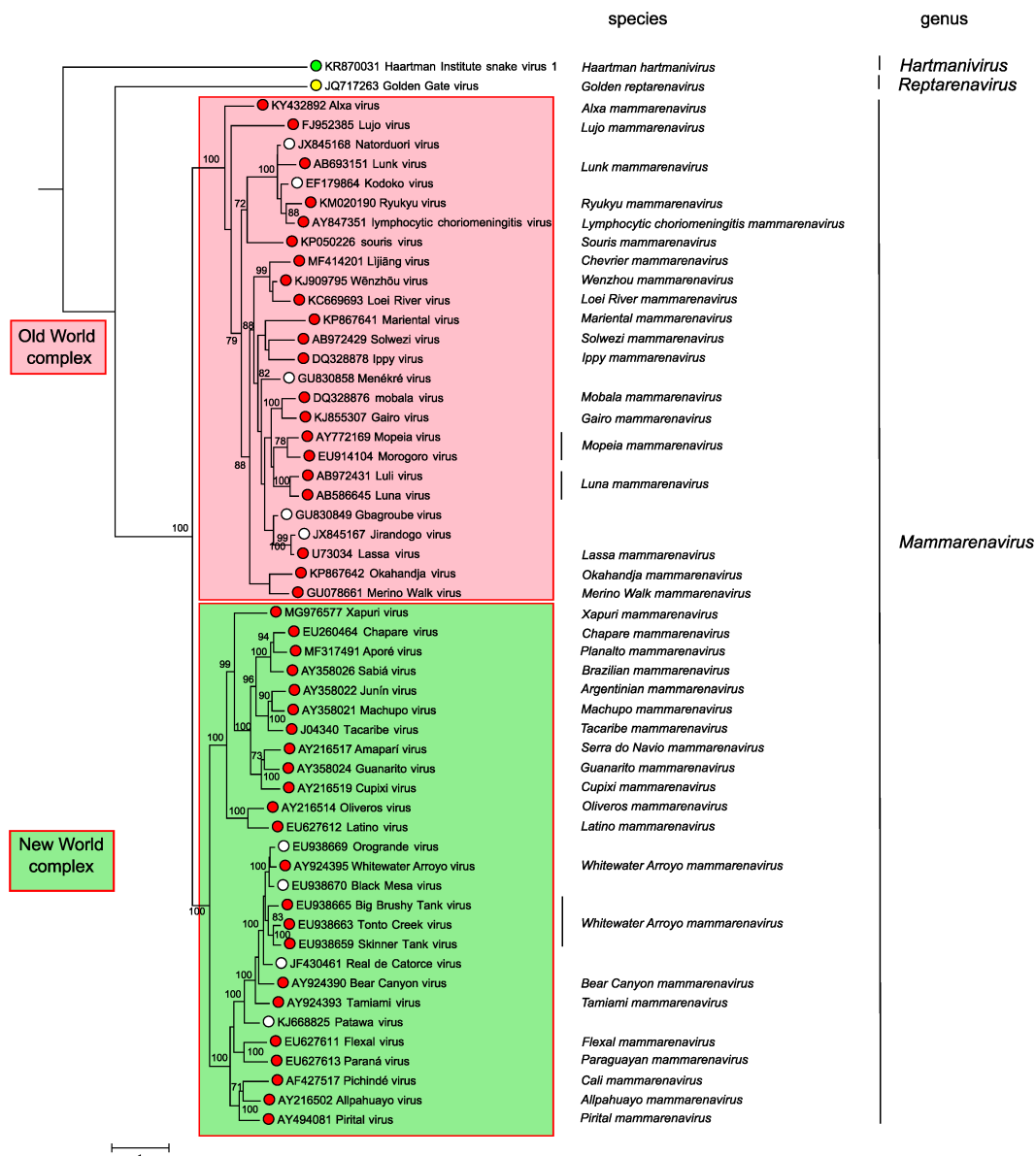


Fig. 1.2. Maximum likelihood phylogenetic trees of the L amino acid sequences of the genus *Mammarenavirus*. The phylogeny includes sequences of 44 mammarenaviruses assigned to 39 species (red dots) and 9 unclassified mammarenaviruses (white dots). Representative viruses of the genera *Hartmanivirus* and *Reptarenavirus* are also included (green and yellow dots, respectively). Viruses belonging to the OW and NW complexes are indicated by background shading of red and green, respectively. The tree was generated using

the same software and parameters as the one presented in *Fig 1.1*. This figure is reproduced from the ICTV webpage: https://talk.ictvonline.org/ictv-reports/ictv_online_report/negative-sense-rna-viruses/w/arenaviridae/1117/genus-mammarenavirus, where this work is licensed under a Creative Commons Attribution License (CC BY-SA 4.0).

1.2 MAMMARENAVIRUSES

1.2.1 The Impacts of Mammarenaviruses

The genus *Mammarenavirus* constitutes a considerable number of diverse rodent-borne viruses (Hallam et al., 2018b). The interaction between these viruses and human populations has dramatically increased due to the extensive overlapping of human dwellings and rodent host reservoirs over the past decades (Blasdell et al., 2016). Largely endemic in the African continent and the Americas (Hallam et al., 2018b) (**Fig. 1.3**), the known geographical distribution of mammarenaviruses has expanded following the recent discovery of novel species in China, Thailand, and Cambodia (Blasdell et al., 2016; Li et al., 2015; Wu et al., 2018). The increased frequency of viral outbreaks, combined with the continuous discovery of new arenaviruses with unknown zoonotic potential, present a major risk and burden to public health and to the socio-economic development of affected countries.

Many mammarenaviruses are highly pathogenic and capable of causing fatal human haemorrhagic fever (HF) upon zoonosis to humans (Paessler and Walker, 2013). While most mammarenaviruses circulate within their respective endemic regions, lymphocytic choriomeningitis virus (LCMV) constitutes the most widely prevalent mammarenavirus as it is carried and transmitted by common house mouse, *Mus musculus* (Lavergne et al., 2016). In West Africa, Lassa virus (LASV) is spread by the Natal multimammate mouse (*Mastomys natalensis*) and plagues Guinea, Sierra Leone, Nigeria, and Liberia (Safronetz et al., 2013). Fortunately, the spread of LASV outside of these regions has been limited. LASV recorded its largest known outbreak in 2018 when the virus spread to eighteen Nigerian states (Zhao et al., 2020). In 2020–2021, Nigeria continued to face a large Lassa fever outbreak with upwards of

20% case fatality while the COVID-19 pandemic, caused by severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), wreaked havoc across the world (Bagechi, 2020). Another deadly mammarenavirus, Lujo virus (LUJV), was responsible for a limited but highly fatal outbreak (80% case mortality rate) in South Africa in 2008 (Simulundu et al., 2016).

In South America, virulent Junín virus (JUNV), Machupo virus (MACV), Guanarito virus (GOTV), Sabia virus (SABV), and Chapare virus (CHAV) circulate in multiple hot spots (Paessler and Walker, 2013). High mortality rates following the rapid onset of severe HF underscore the public health impacts of arenaviral infections and the wide scale need for effective vaccines and therapeutics. For example, MACV and CHAV, aetiological agents of Bolivian and Chapare haemorrhagic fevers (BHF and CHHF), are both endemic in Bolivia with case fatality rates reported up to 35% and 60%, respectively (Escalera-Antezana et al., 2020; Patterson et al., 2014). Alarming, a CHAV outbreak near La Paz, Bolivia in 2019 has been confirmed to be facilitated by human-to-human transmission (Escalera-Antezana et al., 2020). Mammarenaviruses also pose a threat to North America. Indeed, Whitewater Arroyo virus (WWAV) was tentatively associated with three human fatalities in California from 1999 to 2000 (Enserink, 2000). WWAV has since then been recovered from woodrats in New Mexico, Utah, Oklahoma, and Texas (Fulhorst et al., 2001).

As viral surveillance initiatives continue to develop, the diversity of the family *Arenaviridae* is anticipated to grow dramatically. The recent identification of novel arenaviruses in China, Thailand, and Cambodia underscores the wider-than-expected geographical and host range of mammarenaviruses, and the existence of more undiscovered arenaviruses around the world (Blasdell et al., 2016; Li et al., 2015; Wu et al., 2018). A recent seroprevalence study revealed that Wenzhou virus (WENV), first identified in Zhejiang Province, China in 2014, may sporadically infect humans (Guo et al., 2020). Indeed, a separate study showed that patients from Cambodia with fever and respiratory illnesses were tested

positive for WENV, indicating that the virus may be spreading more widely among humans in Asia (Blasdell et al., 2016). Evidently, a genetic variant of WENV named Cardamones was identified in Cambodia in 2016, and has been associated with patients presenting with dengue-like and influenza-like symptoms (Blasdell et al., 2016). Similarly, isolation of Loei River virus (LORV) from *Bandicota* rat in Thailand and Ryukyo virus (RYKV) from Ryukyo mouse in China highlights the growing concerns for arenaviral epidemiology and zoonotic transmission in Asia (Blasdell et al., 2016; Wu et al., 2018).

Zoonotic mammarenaviruses are also economically relevant pathogens. JUNV, the causative agent for Argentine haemorrhagic fever (AHF), poses a pronounced threat to the Argentinian economy (Grant et al., 2012). The pampas (plains for crops and livestock), where JUNV circulates, are vital to the nation's agricultural industry. The health impact of AHF on the agricultural workers, who are often rural populations in endemic areas, directly translates to the economic damage (Burrell et al., 2017). Furthermore, the possibility that pathogenic mammarenaviruses such as LASV, JUNV, and MACV could be used as biological warfare agents has also raised the political profile of these viruses (Borio et al., 2002). Indeed, in line with the politico-economic motivation, the United States Army Medical Research Institute of Infectious Diseases (USAMRII) and the Argentine Ministry of Health and Social Action worked together to develop a vaccine utilising an attenuated JUNV strain (Candid#1) which proved to be successful against JUNV with an estimated effectiveness of 95% in humans (Maiztegui et al., 1998). Along with other deadly viral diseases such as COVID-19, Rift Valley fever, Ebola virus disease, and henipaviral diseases, Lassa fever has been included in World Health Organization's (WHO) list of priority diseases for research and development in emergency contexts, owing to the great public health risk and epidemic potential of LASV (Mehand et al., 2018).

Infectious diseases, including those associated with emerging viruses, constitute a major cause of morbidity and mortality globally, resulting in more than fifteen million deaths annually (Morens et al., 2004). Many of the pathogenic causative agents remain undiagnosed, misdiagnosed, or underdiagnosed due to the lack of understanding of both the ecological and epidemiological patterns of these pathogens and diseases. Therefore, as more emerging arenaviruses are detected, it is of paramount importance to characterise them to bridge the gap of information concerning their human pathogenic potential and the corresponding therapeutic countermeasures.

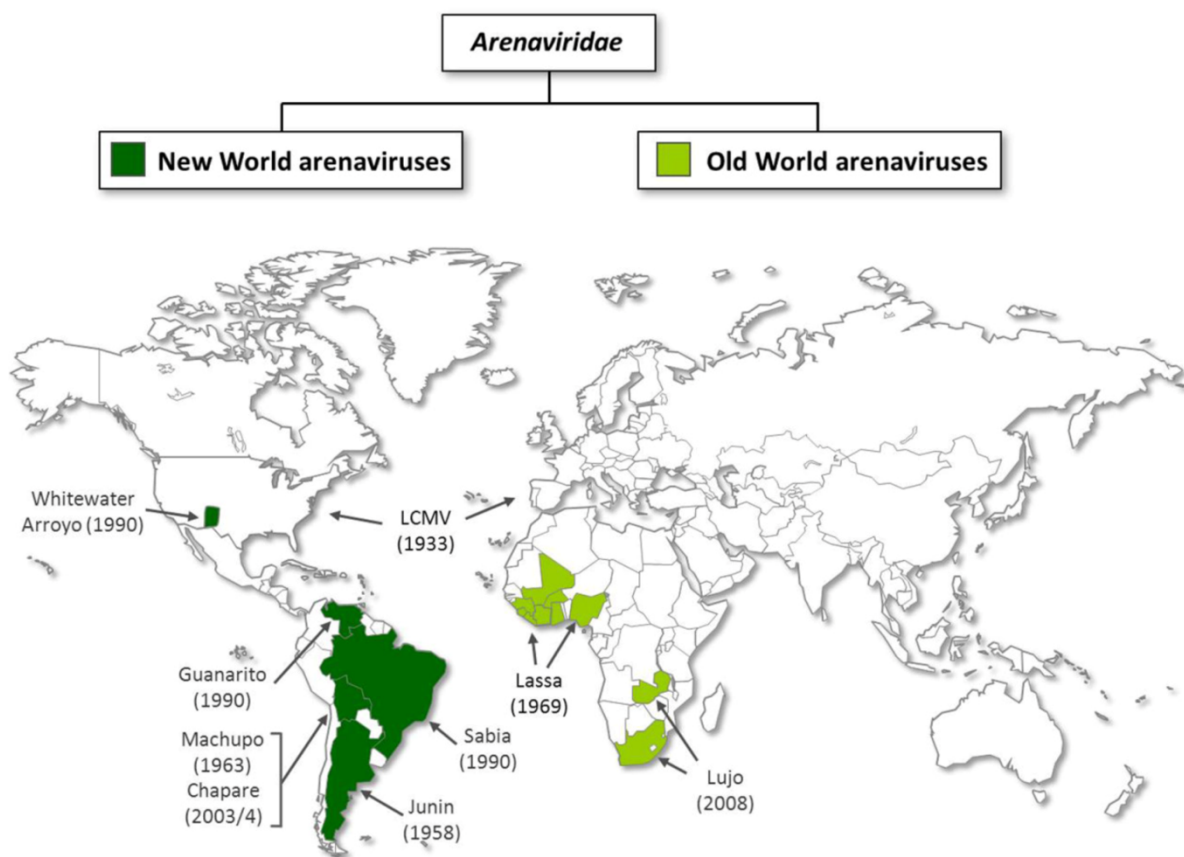


Fig. 1.3. Geographic distribution of human pathogenic mammarenaviruses. This map summarises the distribution of human pathogenic OW and NW mammarenavirus species. The year of the first description is indicated in brackets. This figure is reproduced from *Viruses* (Fehling et al., 2012), where this work is licensed under a Creative Commons Attribution License (CC BY-NC-SA 3.0).

1.2.2 Epidemiology and Clinical Features

As detailed in section 1.2.1, mammarenaviruses represent some of the deadliest human pathogens, including LCMV, LASV, LUJV, JUNV, MACV, SABV, CHAV, and GTOV (Hallam et al., 2018b). They circulate in nature within their respective rodent hosts, with the exception of Tacaribe virus (TCRV), a mammarenavirus non-pathogenic to humans isolated from the *Artibeus* species bats (Neotropical fruit bats) (Downs et al., 1963). Each mammarenavirus is closely associated with a specific rodent host, although some can be carried by a few closely related species. Therefore, geographical distribution of their rodent hosts corresponds to the geographical range of virus dissemination. For instance, LCMV is the only arenavirus to exhibit a worldwide distribution (**Fig. 1.3**) due to the widespread presence of common house mouse, *M. musculus* (Lavergne et al., 2016). According to their genetic and geographical relationships, mammarenaviruses are categorised into OW and NW serogroups (**Fig. 1.2**), in which the former circulates in the African and Asian continents, and the latter are distributed in the Americas (Charrel et al., 2008; Radoshitzky et al., 2015).

Human activities such as agricultural, urban land expansion, and recreational work involving incursions into rodent host habitats promote increased interactions between infected rodents and the human populations (Jones et al., 2013). Interestingly, natural ecological perturbation of the rodent host environment may also result in the spread of mammarenaviruses. Indeed, MACV infection cases increased dramatically following flooding in El Beni region of Bolivia (Charrel et al., 2011). A number of cases have also been attributed to laboratory-acquired infection as arenaviruses are infectious as aerosols (Buchmeier et al., 1974; Pfau, 1996). Therefore, pathogenic mammarenaviruses including LASV, LUJV, JUNV, MACV, CHAV, GOTV, and SABV have been listed as Biosafety Level 4 (BSL-4) pathogens (Charrel et al., 2011). Hospitals also constitute a growing player in mammarenaviral transmission. Nosocomial infection of LASV in Nigeria represents a significant burden to the healthcare

system, with 8% of total cases occur among healthcare workers in the 2018 Lassa fever outbreak (Dan-Nwafor et al., 2019). Infection cases of LCMV originated from organ donors have also been reported, indicating a newly discovered route in which mammarenaviruses may be transmitted (Fischer et al., 2006; Mrzljak et al., 2020).

Mammarenaviral infections in humans are usually initiated by contacts of broken or abraded skins with contaminated materials such as secretions and excretions from infected rodents, or via inhalation of aerosolised virus particles (Buchmeier et al., 2013; Charrel et al., 2008; Grant et al., 2012). While majority of the patients remain asymptomatic or manifest relatively mild illnesses, some develop severe diseases including fatal HF. LCMV infection is typically mild with symptoms such as fever, headache, myalgias, nausea, vomiting, malaise, and anorexia (Bonthius, 2012). Following a short period of recovery, some patients, however, go on to develop acute central nervous system disease including aseptic meningitis and encephalitis, which may result in temporary or permanent neurological damage (Bonthius, 2012). LCMV is generally not fatal, with a reported case mortality rate up to 1% (CDC, 2014).

LASV, on the other hand, is estimated to cause up to 300,000 infections annually, with high case morbidity (30%) and mortality rates (15–20%) (McCormick and Fisher-Hoch, 2002; McCormick et al., 1987). The onset of Lassa fever, when symptomatic, typically occurs 1–3 weeks after infection, and includes dengue- and influenza-like symptoms (Hallam et al., 2018b). Severe cases of Lassa fever may result in facial and lung oedema, mucosal bleeding, convulsions, disorientation, internal bleeding, coma, and death (Bausch et al., 2001; Frame et al., 1970; Hirabayashi et al., 1988; McCormick et al., 1987; Monath et al., 1974). Diagnosis of LASV infection is, however, difficult due to the highly similar symptoms that are indistinguishable from malaria, yellow fever, and other HF such as Ebola virus disease, which also circulate in LASV-endemic areas (Richmond and Baglolle, 2003).

Like LASV, other pathogenic mammarenaviruses including JUNV, MACV, GTOV cause similar symptoms and diseases in humans but with subtle differences. For instance, while majority of the LASV infected patients recover after the (gradual) initial phase of malaise and fever, rapid disease progression beyond this stage is the norm for JUNV, MACV, and GTOV infection, resulting in AHF, BHF, and Venezuelan haemorrhagic fever (VHF), respectively (Pfau, 1996). Indeed, neurologic illnesses, internal bleeding, and thrombocytopenia (platelet deficiency) are more common in the three Latin American HFs than Lassa fever (Pfau, 1996). This differential pathogenicity between OW LASV and NW JUNV, MACV, GTOV is reflected in their relative case fatality rates, with an estimated 15–20% for LASV but up to 35% for JUNV, MACV, and GTOV (Pfau, 1996; Yaro et al., 2021). Hepatitis (liver inflammation), on the other hand, is more frequent and severe in LASV infection, but unusual and mild in JUNV, MACV, and GTOV infections (Pfau, 1996).

Newly discovered mammarenaviruses harbouring unknown zoonotic and pathogenicity potential represent a major public health concern. Recent identification of WENV and its variant, Cardamones, in China and Cambodia, respectively, signifies a wider distribution of mammarenaviruses in Asia than previously known (Blasdell et al., 2016; Li et al., 2015). While the cross-species transmission potential and pathologic roles of WENV in humans are currently unknown, the Cardamones variant has been isolated from patients presenting symptoms such as fever, headache, nausea, cough, and rhinorrhoea (Blasdell et al., 2016). These patients were tested negative for influenza and dengue infection but positive for anti-arenavirus IgG, although diagnostic tests with higher specificity are required to identify the specific causative agent in the seropositive patients (Blasdell et al., 2016).

1.2.3 Genome Organisation and Gene Expression

Although pathobiologically diverse, mammarenaviruses share a genomic structure comprising a bisegmented, ambisense single-stranded RNA genome. Each of the segments, designated small (S) and large (L), respectively, contains two nonoverlapping ORFs separated by an intergenic region (IGR) (**Fig. 1.4**) (Buchmeier et al., 2013). The 3.4 kb S segment encodes for the viral nucleoprotein (NP) and glycoprotein precursor that is later cleaved into stable signal peptide (SSP), GP1, and GP2 subunits (**Fig. 1.4**). Together, the noncovalent association of SSP–GP1–GP2 forms the GP (Crispin et al., 2016; Nunberg and York, 2012). The 7.2 kb L segment encodes for the RNA-dependent RNA polymerase (L) and matrix protein (Z) (**Fig. 1.4**) (Meyer et al., 2002). While the length of the RNA genomes varies slightly for individual arenaviruses, the general genome organisation is well conserved across the mammarenavirus family. Indeed, the RNA sequences of the S and L segments are often used to derive phylogenetic relationships among known and newly discovered arenaviruses (Blasdell et al., 2016; Meyer et al., 2002; Stenglein et al., 2012).

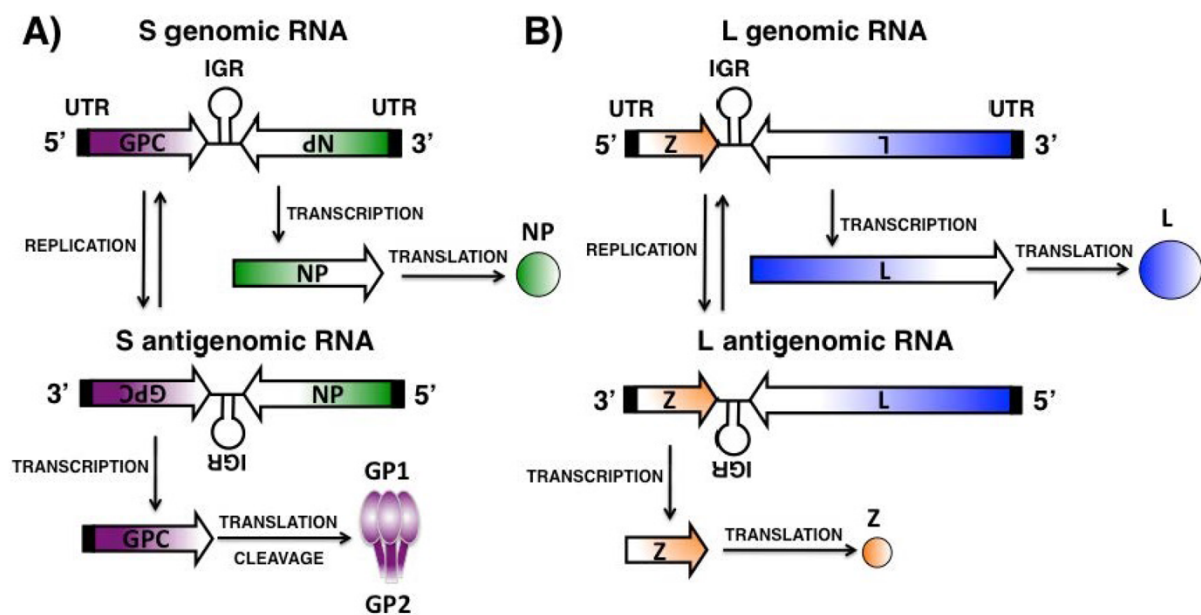


Fig. 1.4. Arenavirus genome replication and gene transcription. The arenavirus replication cycle takes place entirely in the cytoplasm of infected cells. The L polymerase associated with

the viral RNPs initiates transcription from the viral promoter located within the untranslated region (UTR, black boxes) at the 3' termini of the viral RNAs. Primary transcription results in the synthesis of NP (A) and L (B) mRNAs from the S and L segments, respectively. Transcription termination is mediated by a secondary hairpin structure formed by the intergenic region (IGR) found in both viral RNA segments between each of the two encoded genes. Subsequently, L adopts a replicase mode and moves across the IGR to generate a copy of the full-length antigenome S and L viral RNAs. The antigenomic RNA S and L segments serve as templates for the synthesis of GP (panel A; S segment) and Z (panel B; L segment) mRNAs. The antigenomic S and L RNA segments also serve as templates for the amplification of the corresponding viral RNA genome species. This figure is reproduced from *Viruses* (Martínez-Sobrido and de la Torre, 2016), where this work is licensed under a Creative Commons Attribution License (CC BY 4.0).

The coding organisation of mammarenavirus S and L segments does not conform to the typical negative-sense coding arrangements (Meyer et al., 2002). On S segment, the NP gene is encoded in a negative-sense polarity whereas the GP gene is encoded in a pseudo positive-sense polarity (**Fig. 1.4**) (Auperin et al., 1984a; Southern et al., 1987). On L segment, the L and Z genes are encoded in an identical fashion as the way NP and GP genes are encoded on the S segment, respectively (Singh et al., 1987). In brief, the NP (from S segment) and L (from L segment) ORFs are first transcribed into genomic complementary messenger RNA (mRNA), followed by replication of the genomic RNAs by the polymerase to generate the full-length antigenomic RNAs. The GP (from S segment) and Z (from L segment) ORFs are then transcribed using the antigenomic RNAs as templates (Buchmeier et al., 2013; Ferron et al., 2017). Termination of mRNA transcription occurs at sites within the non-coding IGR, which was predicted to adopt a stable hairpin secondary structure as a mean of transcription termination (**Fig. 1.4**) (Auperin et al., 1984a; Auperin et al., 1984b; Pinschewer et al., 2005).

The ambisense coding strategy allows the arenavirus to regulate its early and late gene expression (Auperin et al., 1984b; Southern et al., 1987). Indeed, mRNA for NP can be detected merely 2 h post infection, even in the absence of protein synthesis (Franze-Fernández et al., 1987). On the other hand, mRNA for GP can only be detected 4 h after *in vitro* infection when

protein translation and synthesis are fully active (Franze-Fernández et al., 1987). Such pattern of gene transcription regulation indicates that GP and Z protein productions are highly dependent on the presence of antigenomic RNA templates manufactured by early gene transcription products, NP and L proteins. The viral RNAs are encapsidated by the NP, and together with L, form ribonucleoprotein (RNP) complexes which are active in replication and transcription (Papageorgiou et al., 2020). Once GP mRNA and protein synthesis have started, its production level continues to ramp up, overtaking the protein production level from L segment-derived genomic mRNA (Fuller-Pace and Southern, 1988). Furthermore, qT-PCR data suggests that the level of viral RNA transcription peaks at 48 h post infection, followed by a decline in the abundance of viral RNA (Haist et al., 2015).

Viral genomic and antigenomic RNAs are uncapped, exposing the 3' end of the S and L segments, which contain highly conserved 19-nucleotide (nt) untranslated regions (UTR) (Auperin et al., 1982b). This region has been proposed to serve as a promoter for the viral polymerase (Auperin et al., 1982a). Interestingly, the 3' 19-nt sequence is reverse complement to the 5' end of viral genomic and antigenomic RNAs, allowing the formation of a panhandle structure between 5' and 3' termini of both S and L segments (Auperin et al., 1982a; Auperin et al., 1982b; Meyer et al., 2002). Indeed, circular RNPs within arenavirus virion particles have been observed by electron microscopy, indicating a putative functional role of such double-stranded characteristic of the 5' and 3' UTR in promoter recognition by the viral polymerase for transcription initiation (Hass et al., 2006; Perez and de la Torre, 2003; Radoshitzky et al., 2015). The 5' termini of the arenaviral genomic RNAs also contain an additional nontemplated G nucleotide, which has been proposed to allow L polymerase to initiate replication using a prime-and-realign mechanism (Garcin and Kolakofsky, 1990, 1992; Polyak et al., 1995). The UTR duplex with the additional nontemplated G overhang also allows the RNA to escape

immune detection by retinoic acid-inducible gene I (RIG-I) receptor, an innate immune sensor responsible for type-1 interferon (IFN) response (Marq et al., 2010).

In contrast to viral genomic and antigenomic RNAs, arenaviral mRNAs bear 5'–7-methylguanylate caps, which the virus obtains via cap snatching, a process by which a short segment of host cell RNA is removed through endonucleolytic cleavage and used as a 5' cap for priming viral RNA synthesis (Meyer and Southern, 1993; Polyak et al., 1995; Raju et al., 1990). Many viruses utilise a cap snatching mechanism, including *Arenaviridae*, *Orthomyxoviridae*, *Peribunyaviridae*, and *Phenuiviridae* (Olschewski et al., 2020). As described, viral transcription is terminated by the IGR which contains termination signal. The exact sequence of the IGR appears to be non-crucial as utilising heterologous IGR sequences from distantly related arenaviruses does not interrupt viral transcription termination, indicating that the folding of a stable hairpin structure likely plays a major role in viral transcription termination (Iwasaki et al., 2016). Of note, mammarenaviral mRNAs are not polyadenylated at the 3' termini (Meyer and Southern, 1993; Singh et al., 1987; Southern et al., 1987).

1.2.4 Structure and Components of the Virion

The electron cryomicroscopy single-particle analysis (cryo-EM SPA) method involves grouping protein or virus particles that are similar or identical in size, shape, and conformation together to boost image signal-to-noise ratio (Stass et al., 2019). However, members of the genus *Mammarenavirus* form pleomorphic (variable in shape) virus particles and thus, are not amenable to the SPA method, as the lack of regularity in virion size and shape hinders particle alignment and averaging in a coherent manner (Stass et al., 2019; Tani et al., 2012). Consequently, the virion ultrastructure (typically in ~10–20 Å resolution range) is more usefully studied using electron cryotomography (cryo-ET) with subtomogram averaging (STA), whereas high-resolution structures of the isolated viral components are determined by X-ray

crystallography and cryo-EM SPA. When combined, these data provide a holistic structure-based molecular overview of the architectures and functions of the viruses and their components.

The presence of host ribosomes acquired during virus assembly and budding gives mammarenaviruses a sandy appearance when viewed under transmission electron microscopes (**Fig. 1.5A**). The virion particles are roughly spherical but highly heterogeneous in size, with a diameter of 50–200 nm (Buchmeier, 2002; Charrel and de Lamballerie, 2003; Jay et al., 2005; Meyer et al., 2002). The virions possess a lipid bilayer envelope, flanked by the membrane-bound GP spikes on the extra-viral side, and intraviral domains of GP in contact with a layer of Z protein on the intraviral side (**Fig. 1.5B**). The interior of the virus particles is partially filled with RNP, which is composed by NP, L, and viral RNA genomes (Li et al., 2016).

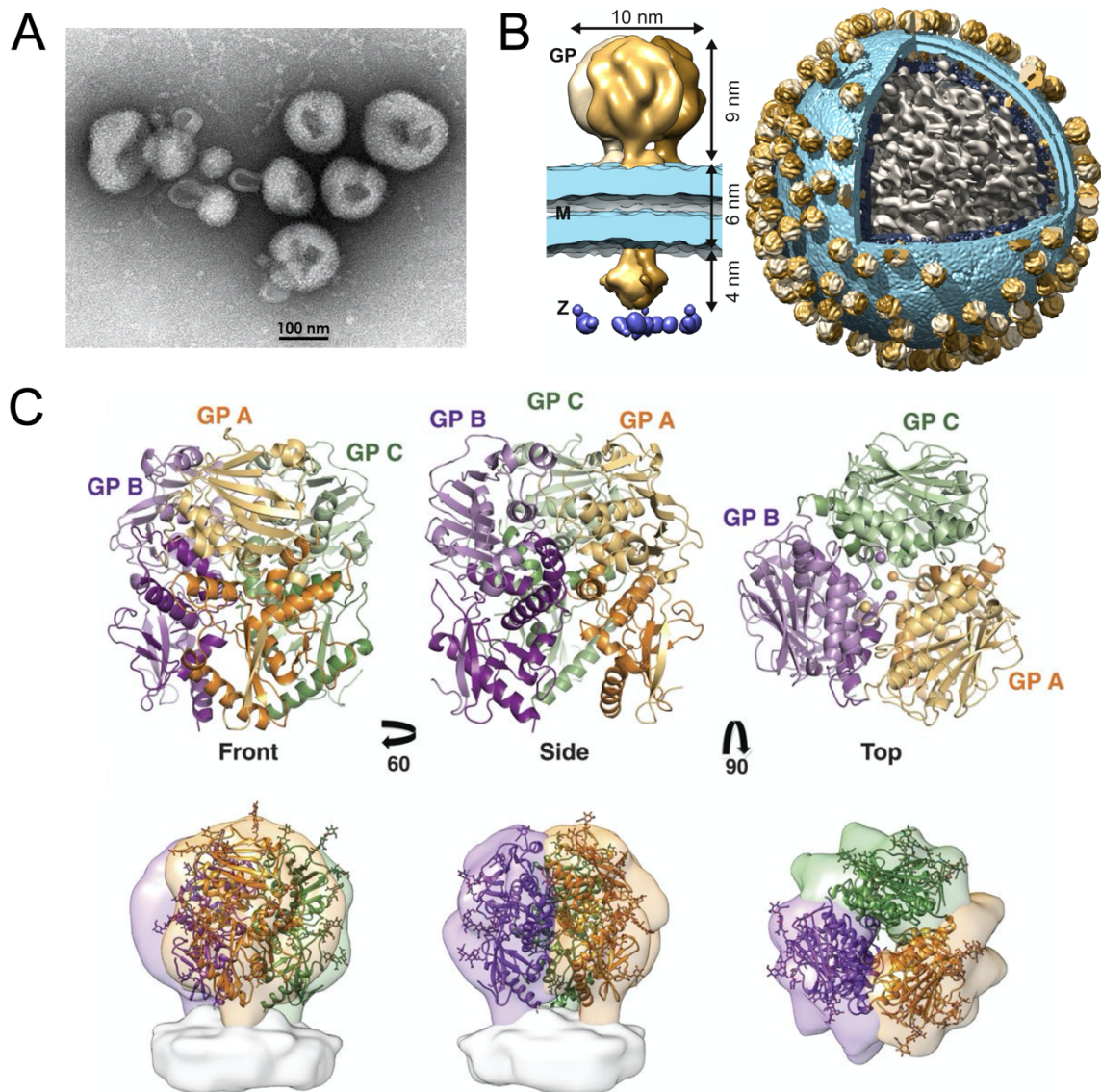


Fig. 1.5. Structure overview of LASV virion and GP spike. (A) Electron microscopy image of arenavirus particles displaying grainy/sandy appearance given by the presence of acquired host ribosomes during virus assembly. *Scale bar* = 100 nm. (B) Structure of LASV spike determined using cryo-ET and STA. (*Left*) LASV GP spike structure reconstructed by STA segmented to three spike protomers and coloured different shades of brown. The membrane (M) and the Z layer are coloured light blue and dark blue, respectively. The GP intraviral tails are located between the M and Z layers. (*Right*) A LASV virion is shown cut open highlighting the irregular organisation of the GP spikes, membrane, Z layer, and RNP density (grey). (C) Organisation of the trimeric GP. (*Top*) The LASV GP1–GP2 trimer (PDB ID: 5VK2) (Hastie et al., 2017) is shown as cartoon and coloured orange, green, and purple for each of the three protomers (GP A–C). The GP1 subunit is in a light shade and the GP2 subunit in a dark shade. (*Bottom*) The crystal structure of the LASV GP trimer docked into the STA reconstructed volume (white) of the LASV GP spike, in the same orientation as shown in the top panel. The N-linked glycans are shown as sticks. Note that panel A is adapted from CDC webpage (www.cdc.gov/vhf/virus-families/arenaviridae.html) and panel B from *PLoS Pathogens* (Li et

al., 2016) under a Creative Commons Attribution License (CC BY 3.0 and 4.0). Panel C is adapted from *Science* (Hastie et al., 2017) with permission from the AAAS.

Chemical cross-linking studies and sucrose density gradient analysis revealed that the GP complexes assemble on the virion surface as trimeric spikes (Schlie et al., 2010). The stoichiometry was further confirmed by STA analysis of 3D volumes of GP spikes extracted from tomograms of inactivated LASV virions, with each spike consisting of three protomers of GP1–GP2 heterodimers organised into a tripartite assembly, measuring approximately 9 nm in height above membrane and 10 nm in width (**Fig. 1.5B**) (Li et al., 2016). The tripodal GP complexes distributed randomly across the surface without higher-order clustering, as shown when the STA sub-volumes were mapped back (template matching) onto the whole virion structure (**Fig. 1.5B**) (Li et al., 2016). Fitting of the high-resolution crystal structure of LASV GP1–GP2 heterodimer (PDB ID: 5VK2) into the STA volume places the GP1 receptor-binding glycoprotein atop the membrane-anchored GP2 class I fusion glycoprotein (**Fig. 1.5C**) (Hastie et al., 2017). Earlier studies have established that the unique 58-amino acid long myristoylated SSP retains with the GP1–GP2 heterodimer in the membrane (Nunberg and York, 2012; York and Nunberg, 2007a). The structural topology and arrangement of SSP, however, have so far not yet been resolved by the 14-Å resolution STA map. Detailed discussion of the mammarenaviral GP complex is further outlined in section 1.3.

Examination of the tomographic slices of inactivated LASV virions revealed the presence of two layers of density immediately underneath the viral membrane (Li et al., 2016) (**Fig. 1.5B**). These layers were assigned to the three 41-residue long intra-viral tails of the GP2 glycoproteins (total molecular mass of 14 kDa), followed by the 99-residue long (molecular mass of 11 kDa) Z protein (**Fig. 1.5B**). The mammarenaviral Z is a multifunctional protein due to its multi-domain architecture (Fehling et al., 2012). The crystal and nuclear magnetic resonance (NMR) structures of LASV Z revealed that the protein consists of three domains: a

N-terminal domain (NTD), which contains a membrane-anchoring myristoylation site, a central zinc-binding ‘Really Interesting New Gene’ (RING) domain, and a C-terminal (CTD) late (L) domain (**Fig. 1.6A**) (Hastie et al., 2016b).

Each Z domain have specific roles. For example, the highly flexible NTD contains a 14-carbon fatty acid chain (myristate), which is covalently attached to a glycine residue at position two (G2) (Perez et al., 2004). The myristyl moiety is important for membrane anchoring, Z self-assembly, and Z–GP association functionalities (Capul et al., 2007; Perez et al., 2004). Z has also been shown to directly interact with SSP but the molecular details of such interaction remain unresolved (Fehling et al., 2012). Furthermore, the NTD of Z exhibits an antagonistic role against host immune response by binding to RIG-I receptors to downregulate cellular IFN response (Xing et al., 2015). The highly conserved central RING domain (~40–60 residues) is well-ordered with a single α -helix and multiple β -strands (Fehling et al., 2012; Hastie et al., 2016b). The RING domain is capable of interacting with NP and L proteins, therefore important for RNP packaging and incorporation at the budding site (Capul et al., 2011; Loureiro et al., 2011; Wilda et al., 2008). This conserved Z–NP association has been observed in LASV, LCMV, Mopeia virus (MOPV), JUNV, and TCRV (Casabona et al., 2009; Groseth et al., 2010; Livingston Macleod et al., 2011; Ortiz-Riaño et al., 2011; Shtanko et al., 2010). Proline rich motifs (e.g. PT/SAP and PPxY, where ‘x’ represents any amino acid) in the CTD-L domain mediate virus budding by interacting with components of the endosomal sorting complex required for transport (ESCRT) pathway (Freed, 2002; Perez et al., 2003). Indeed, expression of Z protein alone is sufficient for the production of lipid-enveloped virus-like particles (VLPs) (Hoenen et al., 2007; Strecker et al., 2003; Ziegler et al., 2016). Similar to the NTD, the NMR structure of the CTD of Z revealed that it is structurally disordered (Volpon et al., 2008; Volpon et al., 2010). Its flexibility is likely important to mediate its interactions with various binding partners (May et al., 2010). Interestingly, previous studies of LCMV and

LASV Z showed that the protein self-assembles into an oligomeric state, indicating a functional role of this higher order assembly (Hastie et al., 2016b; Kentsis et al., 2002). Crystal structure of LASV Z revealed a ring of dodecameric state formed by six dimers (Hastie et al., 2016b) (**Fig. 1.6A**). It has been shown that the RING domain is highly crucial for LCMV Z self-assembly. Such higher order assembly increases local Z concentration and thus enhances its biochemical activities (Hastie et al., 2016b).

As described, Z proteins act as an interaction hub (**Fig. 1.6D**), connecting viral surface GP spikes, viral RNP, and host ESCRT machinery to drive virus assembly and budding (Fehling et al., 2012). While the molecular basis of these interactions remains to be determined, structural and functional investigations of Z binding partners including the viral NP and L proteins provide important information. The crystal structure of LASV NP revealed that the ~570-residue protein consists of two domains connected by a flexible linker: an α -helical NTD and an α/β fold of CTD (Brunotte et al., 2011) (**Fig. 1.6B**). The presence of a deep nucleotide-binding pocket of ~20 Å long in the NTD reflects its function in RNA binding (Brunotte et al., 2011). Additionally, the NTD has been implicated in NP oligomerisation, as shown by the crystal structure and cryo-EM analysis, revealing a ring-shaped trimeric NP assembly (Brunotte et al., 2011) (**Fig. 1.6B**). The precise roles of NP trimers warrant further investigations, but it has been hypothesised that the assembly prevents NP from binding to cellular RNA (Hastie et al., 2011). The structure of NP NTD–single-stranded RNA suggested that the CTD pivoted slightly away from its position to make RNA-binding pocket accessible. This rotation, however, disrupts the trimeric arrangement. It is postulated that the NP may then form a linear, continuous chain along the viral RNA molecule (Hastie et al., 2011). The NP–Z interaction is mediated through the CTD of NP, which also displays anti-IFN activity (Ortiz-Riaño et al., 2011).

L polymerase is the largest structural protein encoded by mammarenaviruses. At ~2200-residue long and ~200 kDa molecular mass, cryo-EM structures of OW LASV and NW MACV L revealed that the polymerase consists of three domains: a N-terminal protease-associated domain, a central RdRp domain, and a C-terminal PB2-like domain (Peng et al., 2020) (**Fig. 1.6C**). The central RdRp domain forms a functional core, sandwiched by the NTD cap-snatching endonuclease region and the NTD core region. The RdRp domain adopts a highly conserved “cupped” right hand-like architecture, consisting of thumb, palm, and fingers subdomains (Choi, 2012; Jacome et al., 2015; Peng et al., 2020). The palm subdomain contains a six-stranded β -sheet flanked by multiple α -helices. The inner surfaces of the NTD core and RdRp thumb subdomain form a highly positive-charged RNA template-binding channel that leads to the catalytic active site within the palm subdomain. Interestingly, the binding of 3'-viral RNA induces dimerisation of L proteins (Peng et al., 2020). A recent study demonstrated that L–L interaction stabilises the CTD, which in turn facilitates transcription and replication activities of arenaviral polymerases (Sánchez and de la Torre, 2005). The RdRp core domain also mediates Z–L interaction (Wilda et al., 2008). It is postulated that Z protein binding blocks the entry of nucleoside triphosphate (NTP) to the catalytic active site to initiate virion assembly and budding (Peng et al., 2020). The CTD of L is composed of two subdomains: a helical subdomain and an apical domain which contains four aromatic residues responsible for cap binding (Peng et al., 2020).

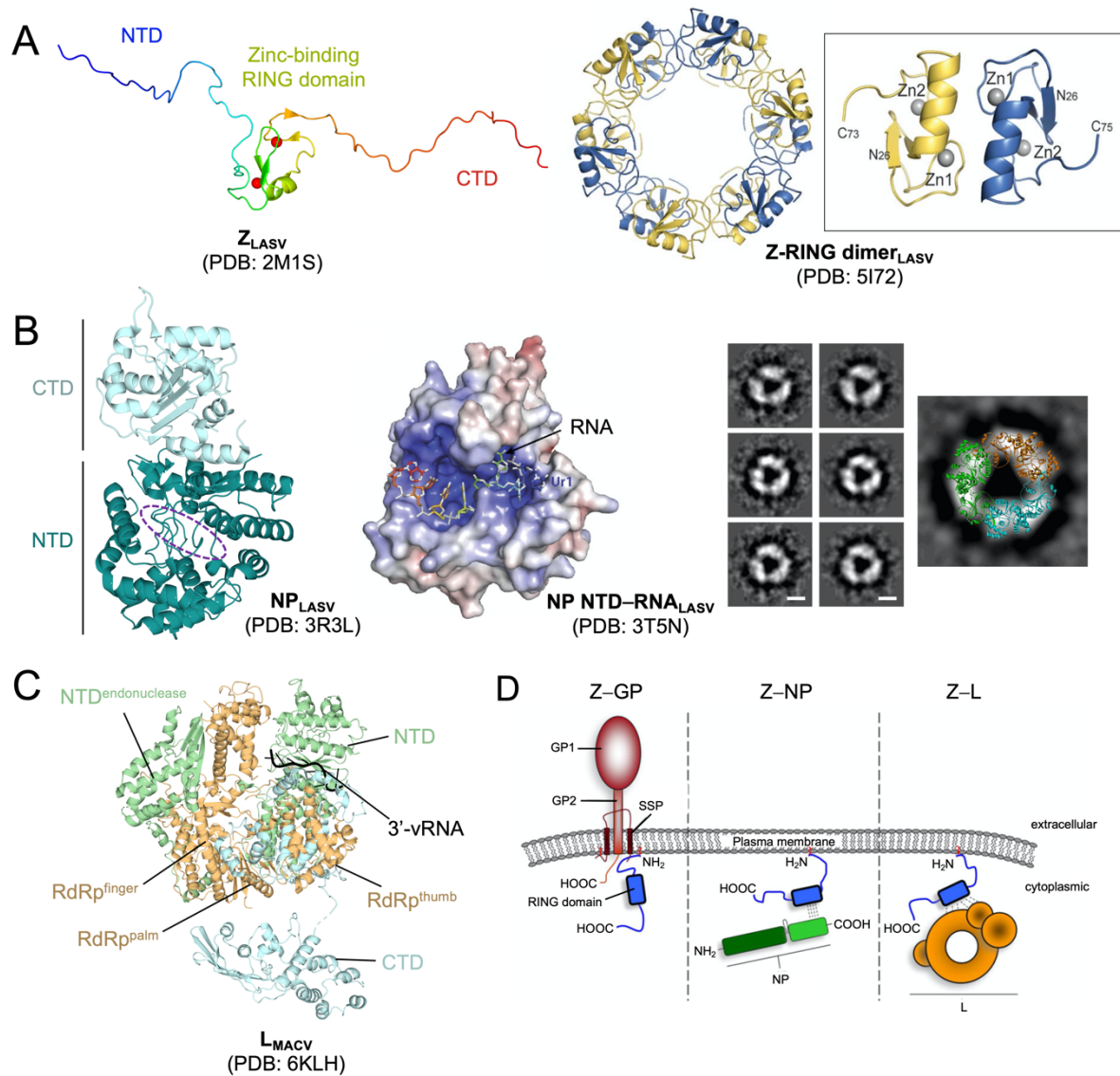


Fig. 1.6. Structure overview of mammarenaviral Z, NP, and L proteins. (A) Structure of LASV Z. (*Left*) NMR structure of LASV Z. The structure (Volpon et al., 2010) is shown in cartoon representation and coloured as a rainbow ramped from N- (blue) to C-terminus (red). Zinc ions are shown as red spheres. The structure shown represents only one snapshot of an NMR solution structure ensemble. The inherent flexibility of NTD and CTD are indicated by the disordered loops. (*Right*) The crystal structure of dimeric Z-RING domain. The structure (Hastie et al., 2016b) is shown as cartoon with the RING monomers coloured as yellow and blue, respectively. A ring of dodecameric state formed by six RING dimers is shown on the left, with a close-up of a single dimer shown on the right. Zinc ions are shown as grey spheres. (B) Structure of LASV NP. (*Left*) The crystal structure (Brunotte et al., 2011) of LASV NP is shown as cartoon. The NTD is coloured in dark green and the CTD is coloured in light cyan. The NTD contains a deep cavity (dashed purple oval) previously shown to bind pyrimidine nucleotides. (*Middle*) The structure of NP-NTD bound to RNA (Hastie et al., 2011) is shown as surface representation. Electrostatic surface potential of NP is calculated using Adaptive Poisson-Boltzmann Solver (APBS), showing a deep basic groove through which the single-stranded RNA channels. (*Right*) Negative stained EM analysis of ring-shaped trimeric NP

assembly (Brunotte et al., 2011). Class averages obtained and overlay of the NP trimer onto the class average are shown. *Scale bar* = 5 nm. (C) The crystal structure of MACV L polymerase. The structure (Peng et al., 2020) was determined in dimeric form but only one of the monomers is shown here for clarity. The NTD, central RdRp domain, CTD, and RNA are coloured green, orange, light cyan, and black, respectively. (D) Schematic drawing showing the interactions of Z protein with GP, NP, and L proteins. Z protein (blue) associates with GP (red) through interaction with the SSP of GP. Z protein directly interacts with the CTD of NP (light green) and the L polymerase (orange) through the Z RING motif. Some parts of this figure are adapted from previous publications under Creative Commons Attribution Licenses (CC BY 3.0 and 4.0): (A, right) is adapted from *Journal of Virology* (Hastie et al., 2016b); (B, middle) from *PNAS* (Hastie et al., 2011); (B, right) from *Journal of Biological Chemistry* (Brunotte et al., 2011); and (D) from *Viruses* (Fehling et al., 2012).

1.2.5 Lifecycle and Replication

1.2.5.1 Host Cell Entry

Human infection occurs when broken skin of an individual comes into direct contact with rodent excrement contaminated with mammarenavirus particles, or through inhalation of aerosolised excreta and secretions from infected rodents (Buchmeier et al., 2013; Grant et al., 2012). The initial cellular targets remain unclear, but it is hypothesised that alveolar macrophages are the first cell types infected (González et al., 1980; Walker et al., 1982). The viruses then spread to other host tissues following migration of infected macrophages to the draining lymph node (González et al., 1980).

As previously described, mammarenaviruses are categorised into two subclasses, termed Old World (OW) and New World (NW) serogroups, based on their genetic and geographical relationships (Charrel et al., 2008; Radoshitzky et al., 2015) (**Fig. 1.2** and **1.3**). Host cell entry pathways differ between OW and NW mammarenaviruses, and are initiated by the specific interaction of the viral GP1 with host cell receptors. For instance, OW viruses such as LASV and LCMV interact with *O*-mannose glycans containing xylose-glucuronic acid sugar repeats presented on α -dystroglycan (α -DG), an extracellular matrix protein, to mediate entry (Cao et al., 1998). LUJV, on the other hand, has recently been shown to use the cell surface

receptor neuropilin-2 (NRP2) as primary receptor (Cohen-Dvashi et al., 2018; Raaben et al., 2017).

Several recent studies have identified alternative receptors for LASV and LCMV, indicative of the utilisation of α -DG-independent host cell entry pathways by OW mammarenaviruses. Indeed, in the absence of α -DG, DC-SIGN (dendritic cell-specific intercellular adhesion molecule-3-grabbing-nonintegrin) and LSECtin (liver and lymph node sinusoidal endothelial calcium-dependent lectin) from the C-type lectin family, and Tyro3 and Axl from the TAM family have been shown to play a role in LASV and LCMV cellular infection (Fedeli et al., 2018; Goncalves et al., 2013; Shimojima and Kawaoka, 2012; Shimojima et al., 2012). It is postulated that in endothelial cells where α -DG is expressed at low level, other receptors are engaged for a productive host cell entry (Fedeli et al., 2018).

Following cellular receptor attachment, virus particles are internalised via endocytosis and delivered to acidified endosomes. In endosomes, LASV switches from using the cell surface α -DG to an intracellular receptor, lysosomal-associated membrane protein 1 (LAMP1), to facilitate viral-host membrane fusion activity of GP2 (Cohen-Dvashi et al., 2015; Cohen-Dvashi et al., 2016; Jae et al., 2014). Interestingly, LAMP1 utilisation is not a strict requirement for LASV entry and infection. However, binding of LAMP1 by LASV has been shown to promote membrane fusion events at a higher pH threshold, thus increasing the efficiency of LASV infection before the virions encounter an increasingly acidic and proteolytic environment in the late endosomes (Hulseberg et al., 2018). Apart from LASV, LAMP1 recognition has so far not been reported in other OW mammarenaviruses. This receptor switching mechanism is, however, not unique, and has been observed in other HF virus such as Ebola virus (Jae and Brummelkamp, 2015), which binds to several different receptors at the cell surface but requires the lysosomal protein, Niemann-Pick C1, for membrane fusion.

Similarly, LUJV has been suggested to switch from NRP2 to endosomal/lysosomal tetraspanin (CD63) that assists pH-dependent GP2-mediated membrane fusion (Raaben et al., 2017).

NW mammarenaviruses are further subdivided into four clades, namely clade A–D (Bowen et al., 1997; Clegg, 2002). Clade B viruses such as JUNV, MACV, SABV, GTOV, CHAV, and clade D (also named clade A/Rec) WWAV recognise transferrin receptor 1 (TfR1) of their respective rodent host (Abraham et al., 2009; Flanagan et al., 2008; Radoshitzky et al., 2007; Radoshitzky et al., 2008; Zong et al., 2014). The ability of these NW mammarenaviruses to also use human TfR1 (hTfR1) constitutes a primary determinant of zoonosis and pathogenicity in humans (Flanagan et al., 2008). In contrast to the carbohydrate-driven interaction between OW mammarenaviral GP and receptors, protein–protein interactions dominate the mode of receptor recognition by NW arenaviral GP (Abraham et al., 2010). However, not all NW viruses use TfR1. Clade C NW mammarenaviruses such as Oliveros virus (OLV) and Latino virus (LATV) utilise α -DG as host receptors (Spiropoulou et al., 2002). To date, the cellular receptor for clade A NW mammarenaviruses has yet to be established.

Like OW viruses, NW mammarenaviruses has been observed to recognise more than one receptor. For example, human T-cell immunoglobulin and mucin-domain containing proteins 1 (hTIM1) promotes host cell entry for clade B NW JUNV, MACV, TCRV, and Amapari virus (AMAV) (Jemielity et al., 2013). Furthermore, C-type lectins such as DC-SIGN and its homologue, L-SIGN, have been shown to substantially augment JUNV entry and infection by acting as a co-receptor to hTfR1 (Martinez et al., 2013). The ability of both OW and NW mammarenaviruses to recognise C-type lectins underscores a functional role of the glycans on the GP. Additionally, siRNA screens uncovered voltage-gated calcium channels (VGCCs) as potential co-receptor for JUNV entry, especially during early infection (Lavanya et al., 2013). Indeed, VGCCs co-downregulate with TfR1 upon virus binding and are required for efficient acidic pH-dependent GP2-mediated membrane fusion (Lavanya et al., 2013).

1.2.5.2 Endocytosis and pH-Dependent Fusion

GP-mediated receptor engagement triggers internalisation of the attached virions via endocytosis. OW mammarenaviruses are endocytosed through an unknown clathrin, caveolin, and dynamin-independent pathway (**Fig. 1.7**) but multiple components of the ESCRT have been identified to be important for cell entry (Pasqual et al., 2011; Rojek et al., 2008b). Interestingly, productive infection of OW viruses does not rely on small GTPases Rab5 and Rab7, which are important for vesicular trafficking to and from early and late endosomal compartments, suggesting an unconventional endosomal delivery of virus particles (Rojek et al., 2008b). Indeed, the optimal pH for LASV GP2-mediated membrane fusion seems to be remarkably low ($< \text{pH } 4.5$), as opposed to pH 5.0–6.0 of fusion observed in other viruses (e.g. influenza virus and JUNV) that utilise Rab5- and Rab7-dependent endosomal trafficking route (Cosset et al., 2009; Martinez et al., 2009; Sieczkarski and Whittaker, 2003).

LASV entry involving a pathway resembling micropinocytosis (uptake of macromolecules by invagination of the plasma membrane) facilitated by heparan sulfate and LAMP1 has been reported for Axl receptor-mediated entry (Fedeli et al., 2018; Goeijenbier et al., 2013). On the other hand, the role of macropinocytosis (non-selective ingestion of extracellular materials by actin-dependent extension of plasma membrane ruffles) in LCMV and LASV viral entry remains elusive. Indeed, LCMV and LASV were shown to enter human epithelial cells via α -DG using macropinocytosis, which is highly dependent on the actin dynamics (Oppliger et al., 2016). The loss or perturbation of the structural integrity and dynamics of actin cytoskeletons in monkey cells, however, does not affect host cell entry, highlighting the differential phenotypes of viral entry in different cell types (Rojek et al., 2008b).

In contrast to OW viruses, entry of NW mammarenaviruses relies on clathrin-dependent endocytosis (Rojek et al., 2008b). Following TfR1 attachment and rapid internalisation by host

cell, virus particles are trafficked to early and late endosomes (Oldenburg et al., 2007; York and Nunberg, 2006). However, depending on cell types, NW viruses have been observed to bypass the early compartments (Martinez et al., 2007; Rojek et al., 2008b; Roldán et al., 2016). Delivery of the virus to late endosomal compartments seems to be a requirement for both TfR1-dependent and TfR1-independent entry pathways, as low pH environment is essential for GP2-mediated membrane fusion activity (e.g. < pH 5.5 for JUNV GP) (Martin et al., 2010; Roldán et al., 2016; York and Nunberg, 2006).

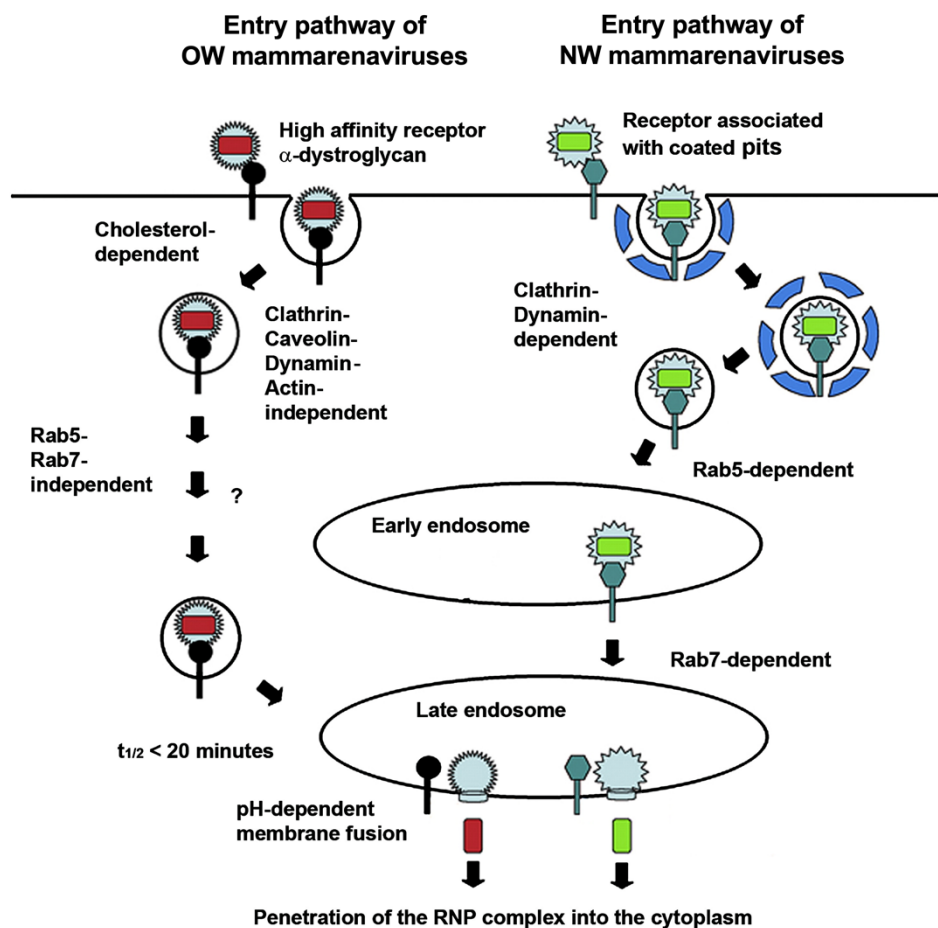


Fig. 1.7. Host cell entry pathways utilised by OW and NW mammarenaviruses. Upon receptor engagement (e.g. α -DG), OW viruses such as LASV and LCMV enter host cells via a clathrin-, caveolin, dynamin-, and actin-independent endocytotic pathway that delivers the virus particles to endosomes without involving Rab5- and Rab7-dependent processes, such as early-to-late endosomal trafficking. In contrast, NW viruses such as JUNV, enter cells by clathrin-mediated endocytosis. Here, viruses and receptors are represented by spiky circles and

ball-and-stick schematics, respectively. This figure is adapted from *Virology* (Kunz, 2009) with the permission from Elsevier.

Upon exposure to acidic pH in late endosomes, mammarenaviral GP undergoes conformational changes leading to membrane fusion. For OW viruses, the GP1 subunit has been observed to adopt a structural state at low pH, where it is no longer α -DG-binding competent (Cohen-Dvashi et al., 2015). For LASV, the utilisation of LAMP1 as intracellular receptor promotes fusion event (Cohen-Dvashi et al., 2016; Hulseberg et al., 2018). Eventually, structural changes in GP1 destabilise the metastable prefusion GP complex and lead to the molecule detaching from GP2 (Branco et al., 2010; Elagoz et al., 2002). The class I GP2 fusion glycoprotein subsequently undergoes a structural reorganisation from the prefusion state to a highly stable, extended six-helix bundle post-fusion conformation, which exposes and inserts the hydrophobic fusion peptide into endosomal membrane (Hastie et al., 2016a; Hastie et al., 2017; Igonet et al., 2011; Parsy et al., 2013). Collapse of the GP2 protein ultimately brings the viral and host membranes into proximity for fusion (White et al., 2008). In contrast to the structural changes observed for OW arenaviral GP1, the conformation of recombinantly expressed NW arenaviral GP1 has not been observed to alter upon pH change (Abraham et al., 2010; Clark et al., 2018; Mahmutovic et al., 2015; Shimon et al., 2017; Zeltina et al., 2017). While it has yet to be experimentally shown, it seems likely that NW arenaviral GP1, like OW GP1, dissociates from GP2 at acidic pH to allow fusogenic rearrangement of GP2 and to activate fusion event (Pasquato et al., 2018).

Further description and discussion of GP2-mediated membrane fusion is detailed in section 1.3.

1.2.5.3 Transcription, Translation, and Genome Replication

Viral–host membrane fusion releases viral RNA genomes into the cytoplasm. As outlined in section 1.2.3, mammarenaviruses contain two ambisense single-stranded RNA segments: S and L, where the former encodes for GP and NP proteins and the latter encodes for L and Z proteins (Buchmeier et al., 2013) (**Fig. 1.4**). To briefly summarise, early gene expression involves the transcription and translation of NP and L proteins from S and L segments, respectively, as they are encoded in the sense orientation (Auperin et al., 1984b; Southern et al., 1987). Together with the viral RNAs, they form RNP, representing the minimal active complex required for initiation of replication (Papageorgiou et al., 2020). NP and L replicate the viral genomes into viral antigenomes (Emonet et al., 2009). GP and Z are encoded in an antisense orientation and cannot be directly transcribed. Late gene transcription of GP and Z, therefore, uses the viral antigenomes as templates. Gene transcription are terminated by the stable hairpin secondary structure adopted by non-coding IGR situated between the coding genes in the S and L segments (Auperin et al., 1984a; Auperin et al., 1984b; Pinschewer et al., 2005). However, the L polymerase reads through the IGR hairpins to make complimentary antigenomes. The antigenomes also serve as intermediate templates for viral genome replication (Emonet et al., 2009; Ferron et al., 2017).

L polymerase has been proven to be error-prone, as shown for OW LCMV, in which the mutation rate was estimated to be on the order of $1\text{--}5 \times 10^{-4}$ (Grande-Pérez et al., 2016). The polymerase also contains endonuclease activity required for cap snatching activity of the 5'–7-methylguanylate cap sequence from host cellular mRNAs (Kranzusch et al., 2010; Morin et al., 2010; Raju et al., 1990). The 5' mRNA cap serves the purposes of translation initiation and stability (Both et al., 1975; Furuichi et al., 1977). Viral mRNAs are translated into proteins by host ribosomes.

GP is translated in the endoplasmic reticulum (ER), where it undergoes signal peptidase (SPase) cleavage of SSP from GP1 and GP2, and asparagine (*N*)-linked glycosylation (Eichler et al., 2003, 2004). SSP retains and associates with the GP2 subunit, an unusual protein processing strategy unique to arenaviruses (York and Nunberg, 2007a). The SSP–GP1/2 glycoprotein is further cleaved in the cis and trans-Golgi network by cellular proprotein convertase subtilisin kexin isozyme-1 (SKI-1)/site-1 protease (S1P) to yield a functional tripartite SSP–GP1–GP2 spike (Beyer et al., 2003; Burri et al., 2013a; da Palma et al., 2014; Kunz et al., 2005; Lenz et al., 2001; Rojek et al., 2008a). Detailed discussion of the mammarenaviral GP is outlined in section 1.3. A schematic representation of mammarenaviral gene transcription, replication, and translation is shown in **Fig. 1.8**.

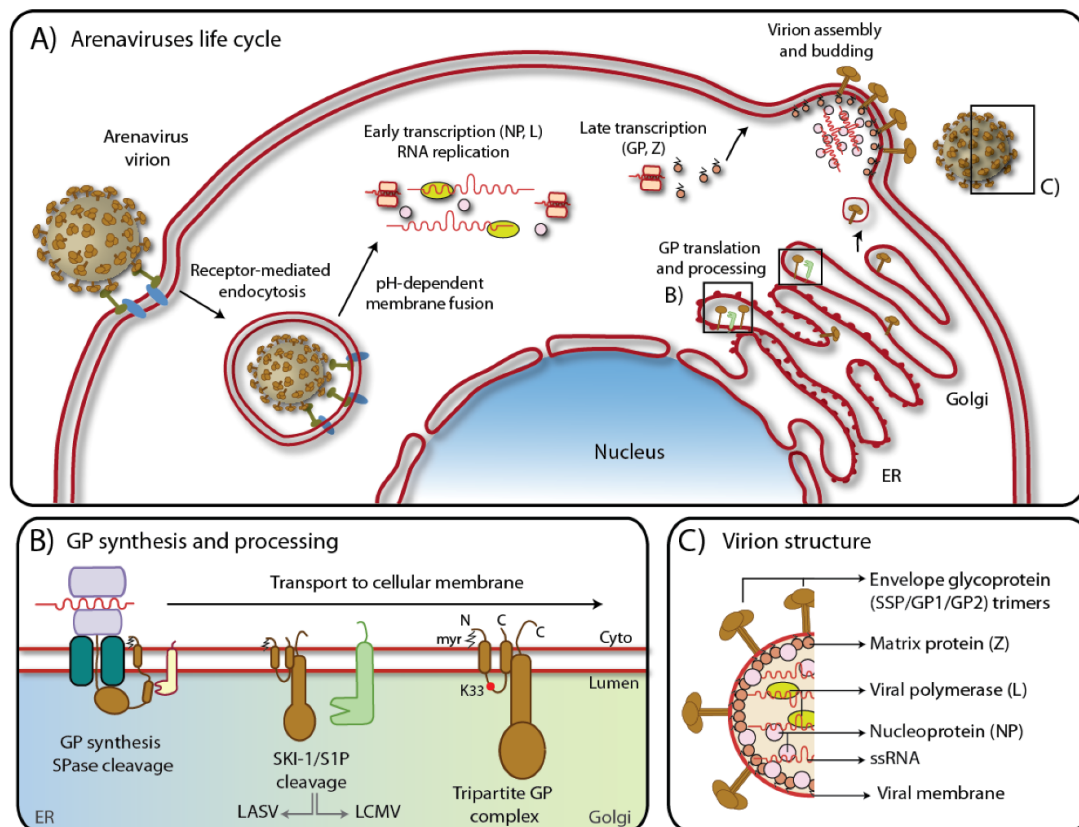


Fig. 1.8. Schematic representation of mammarenavirus gene transcription, replication, and translation. (A) Mammarenavirus life cycle. Arenaviruses enter host cells via receptor-mediated endocytosis. Acidic pH associated with endosomal trafficking promotes merger of the host and viral membranes, releasing the viral RNPs. Early transcription involves transcription and translation of NP and L, which in turn transcribes the viral genomes into

antigenomes. Late gene transcription uses the antigenomes as templates, producing GP and Z. GP are translated and processed in the ER and Golgi networks. Localisation of all viral components at the membrane budding site initiates virus assembly. (B) Biosynthesis of GP spike in the ER with SSP cleavage by SPase and maturation by SKI-1/S1P along the secretory pathway. SSP is myristoylated (myr) and remained associated with GP2. (C) Assembled arenaviral virion structure showing the tripartite GP spike, Z, L, NP, and the ambisense, segmented single-stranded RNA genome. This figure is reproduced from Viruses (Burri et al., 2012) under a Creative Commons Attribution License (CC BY-NC-SA 3.0).

1.2.5.4 Virion Assembly and Budding

Once sufficient amounts of NP and Z are synthesised, capsid assembly begins in the cytoplasm. Each of the structural protein plays a major role in virus assembly and budding. For instance, the packaging of viral genome involves the interactions between L, NP and viral RNAs (Buchmeier et al., 2013). While NP and L are crucial in RNA synthesis, Z is not strictly required for RNA replication or transcription (Cornu et al., 2004). Instead, Z has multiple functions due to its multi-domain structure (Hastie et al., 2016b) (**Fig. 1.6A**). In particular, its NTD contains a myristylation site that allows the protein to anchor itself to the host plasma membrane (Perez et al., 2004). A glycine residue at position two (G2) is critical for the myristylation of Z and its subsequent membrane localisation (Perez et al., 2004). Furthermore, its central RING domain has been shown to bind both NP and L proteins, sequestering the viral RNPs from sites of RNA synthesis in the cytoplasm to the virion membrane budding site (Capul et al., 2011; Loureiro et al., 2011; Wilda et al., 2008). Finally, the C-terminal proline-rich L domain (e.g. PT/SAP, PPPY, and APPY) interacts with ESCRT machinery such as tumour susceptibility gene 101 (Tsg101) protein, which is required for sorting of endocytic cargos into multivesicular bodies (Freed, 2002; Perez et al., 2003; Urata et al., 2006). Indeed, the role of L domain in matrix protein for viral budding has also been observed in other viruses such as Sendai virus and Marburg virus (Irie et al., 2007; Urata et al., 2007). Furthermore, mammarenaviral Z protein has been proven to induce membrane budding even in the absence of other arenaviral structural proteins (Strecker et al., 2003).

The co-localisation of Z with GP protein at the plasma membrane is highly dependent on the myristylation of Z (Capul et al., 2007). While the exact molecular mechanism of Z-GP association remains unknown, it is postulated that Z protein interacts with the SSP and intraviral C-terminal domain of GP2 (Capul et al., 2007). In *toto*, Z acts as a major hub (Fig. 1.9) that recruits and interconnects GP, NP, viral RNAs, membranes, and a myriad of host cell factors (Urata and de la Torre, 2011). The formation of higher-order Z layer at the inner leaflet of the membranes, along with GP spikes and ESCRT complexes, is thought to provide the membrane curvature require for budding (Urata and Yasuda, 2012). Host cellular ribosomes are often packaged together within the virion during assembly and budding (which gives arenaviruses their sandy appearance) but the role, if any, of these ribosomes is unknown (Sarute and Ross, 2017).

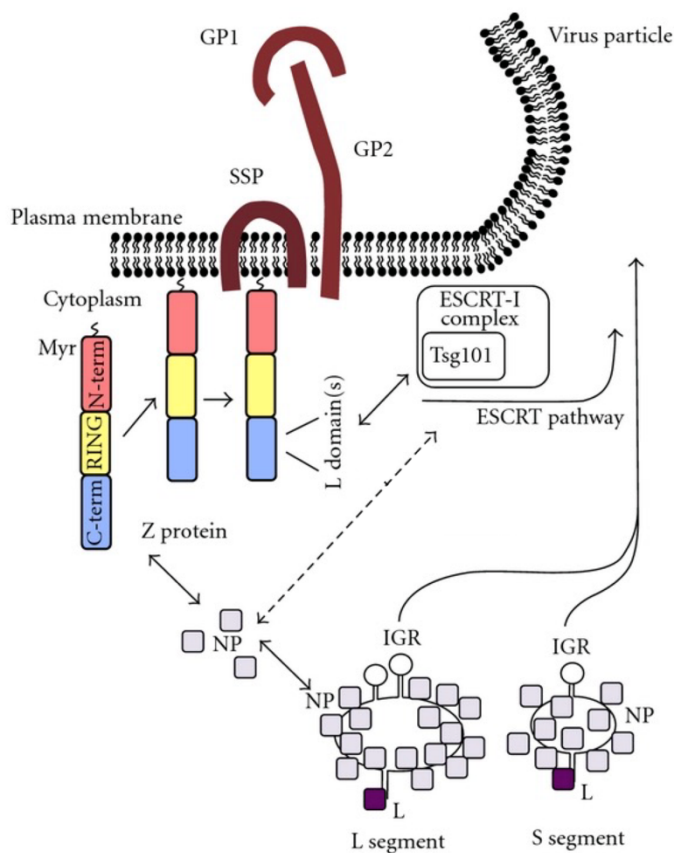


Fig. 1.9. Model of arenavirus budding. Myristoylation (myr) of Z at residue G2 facilitates its interaction with the plasma membrane, where Z likely forms higher-order complexes. The central RING domain of Z is involved in binding and recruiting NP, L, and viral RNA to the membrane budding site. The proline-rich L domain located within the CTD of Z facilitate its interactions with host cellular factors including the ESCRT machinery for virus egress. The NP has also been described as an accessory factor for recruiting additional ESCRT complex components to the budding sites. The figure is adapted from *Advances in Virology* (Urata and de la Torre, 2011), under a Creative Commons Attribution License (CC BY 3.0).

1.2.6 Therapeutics

As outlined in sections 1.2.1 and 1.2.2, pathogenic mammarenaviruses cause a range of diseases from influenza- and dengue-like fever, body aches, malaise, dehydration to cases of human HFs and deaths (Hallam et al., 2018b; Pfau, 1996). Therapeutics and prophylactics are, thus, crucial to minimise the severity and danger of the corresponding illnesses, and more importantly to avoid deaths.

One of the most common treatments involves passive immunisation of patient sera containing high titre of neutralising antibodies (nAbs) (Enria et al., 2008). Such strategy has been particularly effective for JUNV-induced acute diseases (Enria et al., 1984; Maiztegui et al., 1979; Zeitlin et al., 2016). However, treatment of Lassa fever patients with convalescent plasma seems to constitute a limited efficacy (Hallam et al., 2018a; McCormick et al., 1986). Indeed, plasma sera only provided partial protection, with a general consensus that cell-mediated immune responses may play a bigger role in LASV clearance (Hallam et al., 2018a; ter Meulen et al., 2000). This divergence in treatment outcome highlights the differential immunogenicity of OW and NW mammarenaviruses, and the requirement to tailor treatments for individual mammarenavirus-derived disease.

In cases where nAbs are not the dominant player, such as during OW LASV infection, other therapeutics have been proposed, tested, and approved. Ribavirin, a ribonucleoside analogue antiviral medication often used to treat chronic hepatitis C infection (Martin and Jensen, 2008), has been proven to be effective as a prophylaxis to treat NW JUNV and OW LASV (Carrillo-Bustamante et al., 2017; McCormick et al., 1986; Salazar et al., 2012; Veliziotis et al., 2020; Weissenbacher et al., 1986). Indeed, intravenous ribavirin treatment within 6 days of the onset of Lassa fever resulted in a significant reduction of fatality rate to 5% (McCormick et al., 1986; Vela, 2012). Further studies have also implicated the potential antiviral function of ribavirin for LCMV infection (Moreno et al., 2012). While ribavirin has

been used in the treatment of BHF (MACV), its limited usage and results warrant more extensive investigations to assess its role in combating BHF (Kilgore et al., 1997).

Currently, there are no internationally approved vaccines for mammarenaviruses. However, the development of vaccines, with a predominant focus on LASV and JUNV, have shown promising results. For instance, the JUNV vaccine strain, Candid#1, licensed in Argentina, has been proven successful for AHF (Maiztegui et al., 1998). In addition, other vaccine strategies and platforms have also been explored and investigated. For example, heterologous vaccination with LASV and apathogenic MOPV, a closely related OW mammarenavirus species, showed cross-protection in rhesus monkeys (Kiley et al., 1979). Likewise, immunisation using JUNV with apathogenic TCRV led to heterologous protection of guinea pigs against JUNV challenge (Coto et al., 1967). Interestingly, unlike JUNV/TACV vaccine, clearance of LASV/MOPV was achieved without the appearance of nAbs, indicating that other components in cellular immunity (e.g. CD4⁺ and CD8⁺ T-cell) is highly important for protection against mammarenaviral infections (Kiley et al., 1979; Olschläger and Flatz, 2013).

Vaccine development for LASV is especially important due to the persistent outbreaks of Lassa fever in West Africa, with an estimated 180 million people living at risk of the disease (Fichet-Calvet and Rogers, 2009). The absence of a LASV vaccine, despite developments in vaccine research, is largely due to the lack of safety, immunogenicity, and efficacy data due to the high cost of biocontainment requirement, absence of appropriate small animal models, and high genetic diversity of LASV species (Lukashevich and Pushko, 2016). The most advanced candidates, which show promising results in non-human primates (NHPs), include recombinant vesicular stomatitis virus (VSV)–LASV and ML29, a live reassortant of LASV and MOPV, a non-pathogenic OW mammarenavirus closely related to LASV (Huttner et al., 2015; Lukashevich et al., 2008). A recent study utilising a recombinant measles vaccine strain

simultaneously expressing LASV GP and NP has shown protection against LASV challenge in NHPs after a single dose (Mateo et al., 2019). This candidate has been selected by the Coalition for Epidemic Preparedness Innovations (CEPI) to enter human clinical trials (Mateo et al., 2019).

In addition to passive immunisation and ribavirin, a number of antiviral medications and small molecules have been tested for protection against mammarenaviral infection using animal models. For example, favipiravir (T-705), an anti-influenza compound that targets RNA polymerases of RNA viruses (Furuta et al., 2002; Furuta et al., 2005), was found to inhibit NW arenaviral infection including JUNV, MACV, GTOV and Pichinde virus (PICV) (Gowen et al., 2007; Mendenhall et al., 2011a; Mendenhall et al., 2011b). Small molecule ST-336 showed antiviral activity by binding to the GP spikes and disrupting host cell entry by JUNV, MACV, and GOTV (Bolken et al., 2006), whereas ST-193 and ST-294 demonstrated activity for inhibition of pH-dependent GP2 fusion activity (Bolken et al., 2006; York et al., 2008). Genistein, a tyrosine kinase inhibitor, showed promising prophylactic potential when it inhibited PICV cell entry and virus propagation (Vela et al., 2008). Other compounds exhibit differential efficacy against different mammarenavirus species. Trifluoperazine and chlorpromazine, which inhibit viral replication, were effective against JUNV and TACV but not PICV, indicative of the challenge in searching and designing broad-spectrum antivirals against diverse arenaviruses (Candurra et al., 1996).

Shared usage of hTfR1 as cellular receptor by clades B and D NW mammarenaviruses represents a possible route of intervention (Abraham et al., 2009; Choe et al., 2011; Radoshitzky et al., 2007). While engineered protein decoys mimicking viral cellular receptors constitutes a promising strategy, they have to compete with natural host receptors that generally exist in high abundance, thus suffer from low potency at the clinically achievable dosage (Wycoff et al., 2015). Therefore, instead of hTfR1, a recent study has utilised modified rodent

host-derived TfR1 orthologues, termed ‘Arenacept’, that are recognised by NW arenaviruses with high affinity due to their co-adaptation and co-evolution in rodent host (Cohen-Dvashi et al., 2020). Arenacept were found to be effective against pseudotyped and live infectious JUNV and MACV, indicating its potential as a broad spectrum and high potency immunotherapeutic agent against TfR1-utilising arenaviruses (Cohen-Dvashi et al., 2020).

1.3 MAMMARENAVIRAL GLYCOPROTEIN COMPLEX (GP)

1.3.1 Structure and Conformations

Encoded by the S genome segment, the mammarenaviral GP is expressed as a polypeptide precursor, composed of SSP, GP1 receptor binding glycoprotein, and GP2 fusion glycoprotein (Nunberg and York, 2012). Maturation of the GP precursor requires two cleavage events (**Fig. 1.8B**). First, the 58-amino acid SSP is cleaved from the GP1 and GP2 subunits by cellular signal peptidases (SPases) in the ER but remains associated with the GP (Eichler et al., 2003, 2004; York and Nunberg, 2007a). The second event involves cleavage of GP1/2 by cellular proprotein SKI-1/S1P (recognises the motif RXLX, where X is a hydrophobic residue except cysteine) in the Golgi compartments, yielding a non-covalently associated SSP–GP1–GP2 (molecular mass of ~75kDa per protomer) heterotrimer which assembles on the envelope membrane as tripartite GP spikes (Burri et al., 2013a; da Palma et al., 2014; Lenz et al., 2001; Li et al., 2016).

Briefly, SSP is a 58-amino acid peptide that contains a myristoylation site at the G2 position for membrane anchoring, allowing it to remain associated with the GP2 subunit (Briknarová et al., 2011; York and Nunberg, 2007b; York et al., 2004). The retained SSP has been reported to play essential roles in GP biosynthesis, fusion activity, cleavage, and maturation (Agnihothram et al., 2006; Albariño et al., 2011; Bederka et al., 2014). GP1 glycoprotein is responsible for receptor attachment. The specific interaction between GP1 and

target receptors initiates host cell entry, and thus, constitutes an important determinant of cell and tissue tropism and cross-species transmission (Wang et al., 2016). The GP2 glycoprotein is a class I viral fusion protein that mediates membrane fusion after viral particles are internalised into acidic endosomes (Eschli et al., 2006). Fusion of viral and host endosomal membranes releases the viral RNA genomes for transcription and replication (Fehling et al., 2012).

As described in section **1.2.5.1**, the 40–45 kDa mammarenaviral GP1 recognises a number of cellular receptors for host cell entry. OW mammarenaviruses such as LCMV and LASV (Cao et al., 1998) and clade C NW mammarenaviruses (Spiropoulou et al., 2002) have been shown to recognise α -DG. Members of the C-type lectin family such as DC-SIGN and LSECtin, and members of the TAM family such as Axl and Tyro3 have also been shown to be utilised by OW mammarenaviruses (Fedeli et al., 2018; Goncalves et al., 2013; Shimojima and Kawaoka, 2012; Shimojima et al., 2012). On the other hand, clades B and D NW viruses use rodent host TfR1 to mediate entry (Abraham et al., 2009; Choe et al., 2011; Radoshitzky et al., 2007). Similarly, other host receptors such as DC-SIGN, L-SIGN, hTIM1, and VGCCs have been shown to promote NW viral entry (Jemielity et al., 2013; Lavanya et al., 2013; Martinez et al., 2013). Interestingly, LUJV GP1 utilises NRP2 as receptor, indicative of a functional divergence from the other OW and NW viruses (Raaben et al., 2017). Furthermore, LASV and LUJV GP1s have been demonstrated to bind LAMP1 and tetraspanin, respectively, to facilitate membrane fusion activity (Jae et al., 2014; Raaben et al., 2017). The ability of mammarenaviral GP1 to recognise more than one host receptors allows efficient viral entry, especially in cells that exhibit low expression of the primary receptors (α -DG and TfR1) (Fedeli et al., 2018).

The first GP1 structure of NW MACV (PDB ID: 2WFO) was reported in 2009 by Bowden *et al.*, and subsequently in complex with hTfR1 (PDB ID: 3KAS) by Abraham *et al.* in 2010. The monomeric MACV GP1 structure adopts a α/β architecture (**Fig. 1.10A**). A

central seven-stranded β -sheet is sandwiched between three α -helices and connecting loops. The overall fold is stabilised by four disulfide bonds (Abraham et al., 2010; Bowden et al., 2009). Despite the low sequence conservation (estimated ~24–47% sequence identity) between NW mammarenaviral GP1s, the observed MACV GP1 fold is recapitulated by JUNV GP1 (PDB ID: 5EN2, 5NUZ, and 5W1K) and WWAU GP1 (PDB ID: 5NSJ) (**Fig. 1.10A**), reflective of their common TfR1-binding functionality (Clark et al., 2018; Mahmutovic et al., 2015; Shimon et al., 2017; Zeltina et al., 2017). The crystal structure of MACV GP1–hTfR1 complex also revealed that the apical domain of the receptor interacts with a central groove formed by loop 3, loop 7, and the central β -sheet of MACV GP1 (**Fig. 1.10B**). Notably, residue Tyr211_{hTfR1} inserts and occupies this GP1 pocket (Abraham et al., 2010).

Identification of the receptor-binding site (RBS) on the GP1 molecule is critical for antibody design and drug development. Indeed, structural studies of antibody-mediated neutralisation of NW JUNV GP1 (PDB ID: 5NUZ, 5EN2, and 5W1K) revealed that the antibodies engage the GP1 molecule at the RBS, precluding TfR1 recognition (Clark et al., 2018; Mahmutovic et al., 2015; Zeltina et al., 2017) (**Fig. 1.10B**). Furthermore, the insertion of a tyrosine residue by the complementary-determining region (CDR) loop encoded by the antigen-binding fragment (Fab) into the central JUNV GP1 pocket suggests that this signature interaction represents a neutralisation mechanism involving receptor mimicry, and that the RBS constitutes a key site of vulnerability (Zeltina et al., 2017) (**Fig. 1.10B**). However, a recent structural investigation of a cross-reactive MACV/JUNV nAb (PDB ID: 5W1M) did not observe the aforementioned tyrosine-mediated mode of GP1 recognition, indicative of the flexibility and variability of GP1–antibody interaction in accommodating sequence and structural variations among NW mammarenaviral GP1s (Clark et al., 2018).

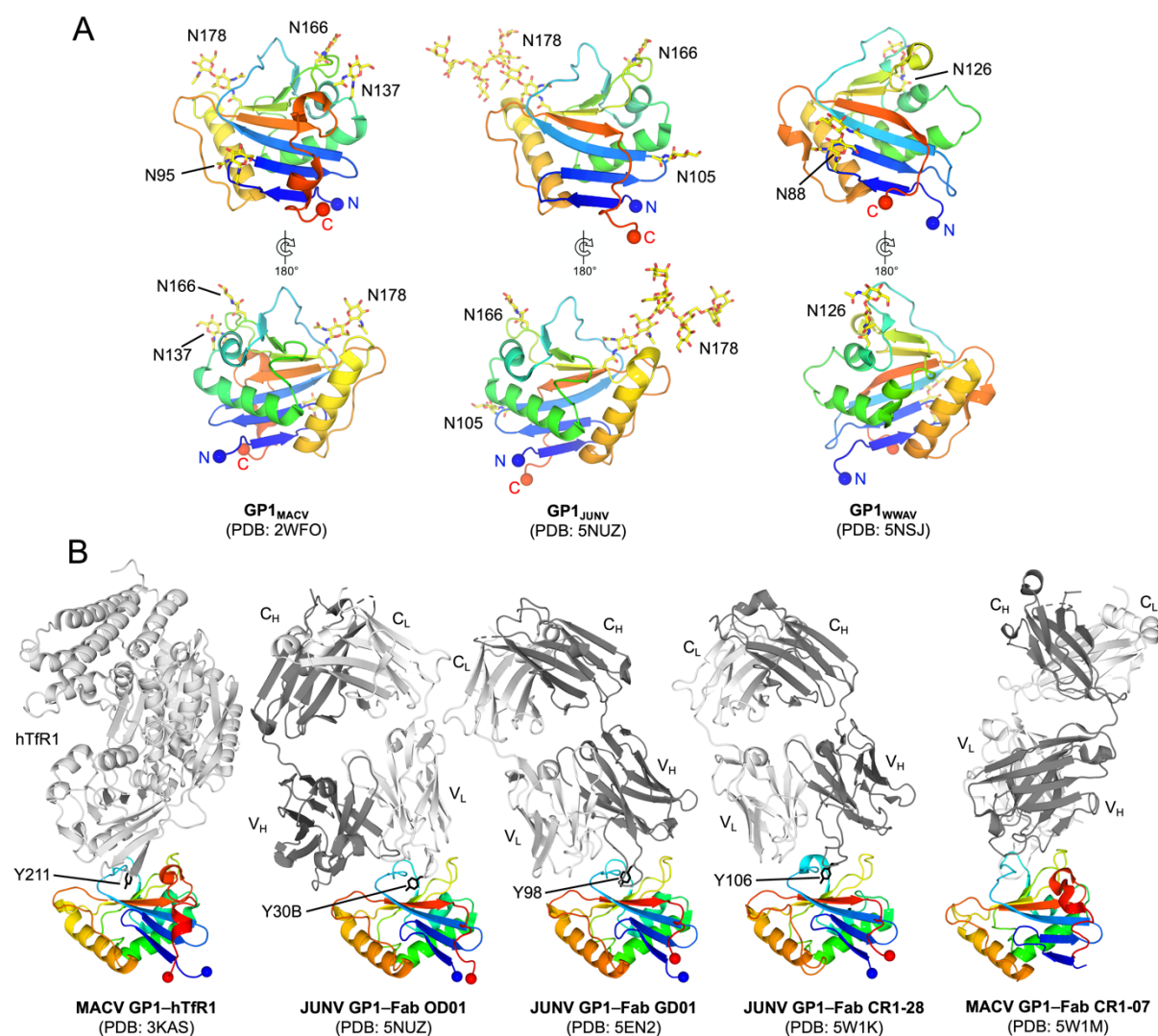


Fig. 1.10. Structure overview of NW mammarenaviral GP1s. (A) The crystal structures of monomeric MACV, JUNV, and WWAV GP1s. The structures are shown in cartoon representation and coloured as a rainbow ramped from N- (blue) to C-terminus (red). The N- and C-termini are shown as blue and red spheres, respectively. Crystallographically observed N-linked glycans are shown as yellow sticks, with constituent nitrogen and oxygen atoms coloured blue and red, respectively. (B) The crystal structures of MACV-hTfR1, MACV-Fab, and JUNV-Fab complexes. The structures are shown as cartoon. GP1 molecules are coloured as a rainbow, hTfR1 in light grey, Fab heavy chain in dark grey, Fab light chain in light grey. V_H, V_L, C_H, and C_L denote the antibody variable heavy, variable light, constant heavy, and constant light-chain domains, respectively. The GP1 pocket-inserting tyrosine residues encoded by the Fab CDR loops are shown as black sticks and labelled accordingly.

The monomeric GP1 structure of OW LASV (PDB ID: 4ZJF) was reported in 2015 by Cohen-Dvashi *et al* (**Fig. 1.11A**). Structural comparison between LASV GP1 and MACV GP1 revealed significant differences outside of the conserved central β -sheet core, reflecting the

evolutionary distance between OW and NW viruses. Indeed, the differential arrangement of α -helices and termini between LASV and MACV GP1s is consistent with their differential receptor usage (**Fig. 1.11B**) (Cohen-Dvashi et al., 2015). The reported structure of isolated LASV GP1 could pull down endogenous LAMP1 at pH 5.0 but not α -DG at pH 5.0 and pH 8.0, indicating that the observed structure adopts a primary receptor-binding incompetent conformation and thus, likely does not resemble the native structure presented on the virion spike (Cohen-Dvashi et al., 2015).

Structural elucidation of OW LCMV and LASV GP1–GP2 heterodimers (PDB ID: 5INE and 5VK2, respectively) revealed that the GP1 molecule sits atop the cognate GP2 subunit where the elongated N-terminus of GP1 extends downwards and interacts with GP2 (Hastie et al., 2016a; Hastie et al., 2017) (**Fig. 1.11A**). This GP1–GP2 arrangement was further confirmed by a 14-Å STA map of LASV GP (**Fig. 1.5B**) (Li et al., 2016). Comparison of a LASV GP1–GP2 structure obtained at pH 8.0 and the monomeric LASV GP1 structure obtained at pH 5.0 revealed major structural differences (Hastie et al., 2017) (**Fig. 1.11B**). The observed conformational differences in the GP1 suggests that the change in pH from neutral to acidic associated with virion internalisation and endosomal trafficking causes the GP1 molecule to change from the neutral-pH conformation (α -DG-binding) to the acidic-pH conformation (LAMP1-binding) (Li et al., 2016) (**Fig. 1.11C**). Indeed, the putative LAMP1-binding histidine triad (His92, His93, and His230) on LASV GP1 has been hypothesised to act as a pH sensor (Cohen-Dvashi et al., 2015). Structure of monomeric Morogoro virus (MORV) GP1 (PDB ID: 5NFF) (**Fig. 1.11A**), an apathogenic OW virus, highlights its high structural similarity with the monomeric LASV GP1 conformation, and the sequence and structural conservation of the histidine triad (Israeli et al., 2017). Although several residues crucial for α -DG binding were identified in a mutagenesis study (Hastie et al., 2016a), the structural basis

of GP1- α -DG recognition has yet been reported, likely due to the flexibility of glycan-glycan interactions, which are typically non-conductive for structural biology studies.

Unlike NW viruses where GP2-free GP1 is sufficient for TfR1 recognition (Abraham et al., 2010), OW viruses such as LASV and LCMV require the tripartite assembly to achieve α -DG engagement (Cohen-Dvashi et al., 2015; Hastie et al., 2016a). Similarly, most nAb against OW LASV recognise epitopes formed by the extended N-terminus of GP1 and parts of GP2 that interact with the N-terminus of GP1, as shown by the crystal structure of LASV GP1-GP2-Fab complex (PDB ID: 5VK2) (Hastie et al., 2017; Robinson et al., 2016), whereas NW GP1 alone can elicit strong antibody-mediated host immune response (Borenstein-Katz et al., 2019).

In contrast to the conformational plasticity displayed by the acidic-pH monomeric GP1 and neutral-pH heterodimeric GP1-GP2 structures, monomeric NW GP1s from JUNV, MACV, and WWAV adopt a similar structural state at both acidic and neutral pH (Bowden et al., 2009; Shimon et al., 2017; Zeltina et al., 2017). Given that the isolated NW mammarenaviral GP1s are capable of binding to TfR1 and vaccine and infection-elicited monoclonal antibodies (mAbs), it has been postulated that the observed NW GP1 conformation likely resembles that presented on the virion spike (Abraham et al., 2010; Clark et al., 2018; Mahmutovic et al., 2015; Zeltina et al., 2017). However, this hypothesis requires further confirmation as the structure of NW GP1-GP2 heterodimer has yet been reported.

A recent structural study of GP2-free LUJV GP1 in complex with NRP2 (PDB ID: 6GH8) revealed a novel mammarenaviral GP1 architecture (**Fig. 1.11A**) that is dissimilar to the previous OW and NW GP1 folds discussed above (Cohen-Dvashi et al., 2018). Interestingly, structural alignment of LUJV GP1 with OW and NW GP1s showed that LUJV GP1 contains features from both the OW and NW GP1 folds (**Fig. 1.11B**), suggesting that LUJV structurally and functionally diverge from the established OW and NW viruses early in the evolution

(Cohen-Dvashi et al., 2018). Of note, the observed LUJV GP1 structure adopts a receptor-binding competent conformation, indicating that the structure likely resembles the native conformation displayed on the virion spike (Cohen-Dvashi et al., 2018). To summarise, an outline of mammarenaviral GP1 structures determined to date, including the new structures reported in this thesis, is listed in **Table 1.1**.

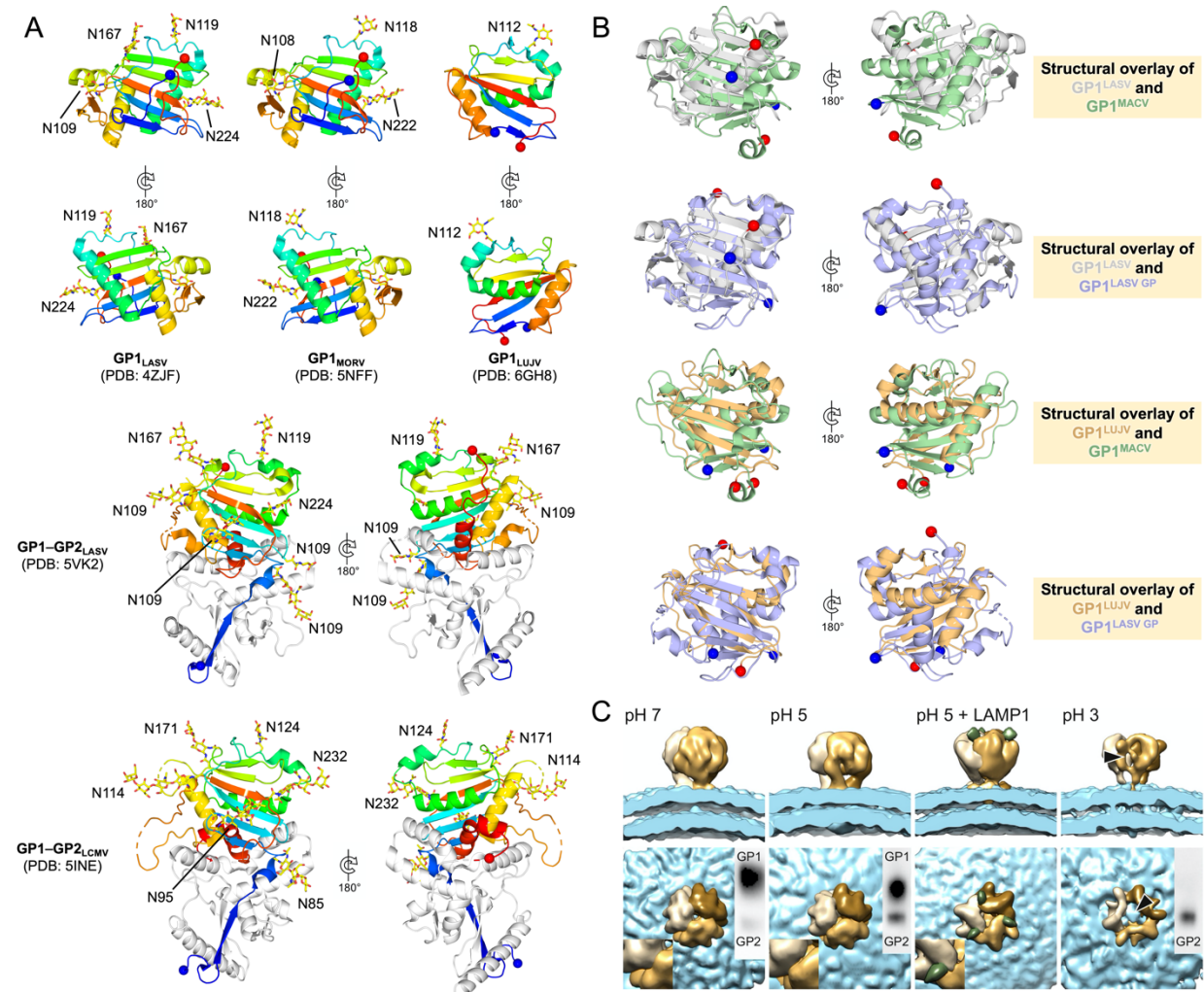


Fig. 1.11. Structure overview of OW mammarenaviral GP1s. (A) The crystal structures of monomeric LASV (*top left*), MORV (*top middle*), LUJV (*top right*) GP1s and LASV (*middle*) and LCMV (*bottom*) GP1–GP2. Water molecules, ligands, and protein chains (e.g. receptor and antibody fragments) not comprising GP1 molecules are not shown. The GP1 structures are shown in cartoon representation and coloured as a rainbow ramped from N- (red) to C-terminus (blue). The N- and C-termini are shown as blue and red spheres, respectively. The GP2 subunits are shown as white cartoons. Unresolved regions are shown as dashed lines. Crystallographically observed N-linked glycans are shown as yellow sticks, with constituent nitrogen and oxygen atoms coloured blue and red, respectively. (B) Structure comparison of OW and NW GP1s. The crystal structures of LASV GP1, MACV GP1, and LUJV GP1 were

superimposed. The GP1 structures are shown in cartoon representation. Monomeric LASV GP1 is coloured light grey, MACV GP1 in green, LASV GP1 (determined in the presence of GP2; denoted as GP1^{LASV GP}) in blue, and LUJV GP1 (determined in the presence of NRP2) in orange. GP2 and NRP2 are not shown for clarity. (C) Structural transitions and LAMP1 binding of the LASV GP spikes upon acidification as determined by STA. GP structures at different pH are shown from side (*top row*) and top (*bottom row*). All volumes were filtered to 17-Å resolution. The GP protomers are coloured in different brown shades, with the residual density corresponding to LAMP1 coloured in green. Insets in the lower left corners show a close-up of the interface between two spike GP1 monomers. Insets in the top right corners show Western blot analysis of GP1 and GP2 subunits. The arrowheads indicate the missing density in the central top part and side of the pH 3.0 structure. Note that the figure in panel C is reproduced from *PLoS Pathogens* (Li et al., 2016) under a Creative Commons Attribution License (CC BY 4.0).

Table 1.1. Summary of mammarenaviral GP1 structures determined to date.

	Mammarenavirus	pH at which the structure was solved	PDB ID	References
OW	LASV GP1 (monomeric, apo)	5.0	4ZJF	(Cohen-Dvashi et al., 2015)
	LASV GP1 (trimeric, associated with GP2, bound to Fab 37.7H)	8.0	5VK2	(Hastie et al., 2017)
	LCMV GP1 (dimeric, associated with GP2, apo)	7.5	5INE	(Hastie et al., 2016a)
	MORV GP1 (monomeric, apo)	4.0	5NFF	(Israeli et al., 2017)
	LUJV GP1 (monomeric, bound to NRP2)	9.5	6GH8	(Cohen-Dvashi et al., 2018)
	LORV GP1 (monomeric, apo)	5.0	6HJ6	(Pryce et al., 2018) <i>Discussed in Chapter 3</i>
	LORV GP1 (monomeric, apo)	8.0	6HJC	(Pryce et al., 2018) <i>Discussed in Chapter 3</i>

NW	MACV GP1 (monomeric, apo)	5.5	2WFO	(Bowden et al., 2009)
	MACV GP1 (monomeric, bound to hTfR1)	5.5	3KAS	(Abraham et al., 2010)
	MACV GP1 (monomeric, bound to Fab CR1-07)	5.5	5W1M	(Clark et al., 2018)
	MACV GP1 (monomeric, bound to Fab MAC1)	6.0	To be reported	<i>Discussed in Chapter 4</i>
	JUNV GP1 (monomeric, bound to Fab GD01)	8.0	5EN2	(Mahmutovic et al., 2015)
	JUNV GP1 (monomeric, bound to Fab OD01)	5.6	5NUZ	(Zeltina et al., 2017)
	JUNV GP1 (monomeric, bound to Fab CR1-28 and Fab CR1-10)	7.7	5W1K	(Clark et al., 2018)
	JUNV GP1 (monomeric, bound to Fab JUN1)	7.5	To be reported	<i>Discussed in Chapter 4</i>
	WWAV GP1 (monomeric, apo)	8.7	5NSJ	(Shimon et al., 2017)
	WWAV GP1 (monomeric, apo)	5.6	6HJ5	(Pryce et al., 2018)
	WWAV GP1 (monomeric, apo)	7.5	6HJ4	(Pryce et al., 2018)

Membrane fusion constitutes an important host cell entry event that merges viral and host membrane bilayers into one to initiate the release of viral genomes into the cytoplasm for transcription and replication (White et al., 2008). The fusion pathway of membrane apposition to hemifusion intermediate (outer contacting lipid monolayers merge but the inner leaflets remain intact) to full fusion pore formation is orchestrated by the viral fusion glycoprotein (Harrison, 2008). The fusion protein is displayed on the mature virion surface in a prefusion conformation that is often, but not always, metastable. Following virion internalisation, acidic

endosomal pH and/or receptor binding trigger structural rearrangement of the fusion protein from pre to post-fusion conformation, through a number of intermediate structures, to expose the hydrophobic fusion peptide/loop which embeds into the target membrane (Halldorsson et al., 2018; Harrison, 2008; Kielian and Rey, 2006). Sequential conformational changes involving the formation of an extended structure and a fold-back compact structure overcome the large energy barrier to viral and target cell membrane merger (Cohen and Melikyan, 2004).

Based on their distinct structural properties, viral membrane fusion proteins are typically classified into class I, II, and III. Class I and II fusion proteins contains mostly α -helices and β -sheets, respectively, whereas class III fusion protein consists of a combination of both α -helices and β -sheets (White et al., 2008). The ~35 kDa mammarenaviral GP2 subunit belongs to the class I viral fusion protein (Eschli et al., 2006). The first two structures of the mammarenaviral GP2 from OW LCMV (PDB ID: 3MKO) and NW GTOV (PDB ID: 4C53) were reported in 2011 and 2013 by Igonet *et al.* and Parsy *et al.*, respectively (**Fig. 1.12A**). Structural overlay analysis revealed that the two GP2 structures adopt a similar overall architecture (a 2.2-Å root mean square deviation [RMSD] over 96 Ca residues) (Igonet et al., 2011; Parsy et al., 2013). The structures present an extended α -helix-rich post-fusion conformation consists of an 85-Å-long ~45-amino acid N-terminal helix and a ~10-amino acid short parallel C-terminal helix, connected by a ~50-residue “T region”, which is composed of a long “T-loop” and a small α -helix (“T-helix”) (**Fig. 1.12A**). While the N-terminal helices from the two structures superimpose well (~1.0-Å RMSD), a greater structural variability is observed in the T region and C-terminal helix (> 3.0-Å RMSD) despite the 50% amino acid sequence identity between OW LCMV and NW GTOV GP2, reflecting the intrinsic differences between OW and NW GP architecture.

Determination of the prefusion structure of OW LCMV GP2 (in association with GP1; PDB ID: 5INE) revealed several drastic differences when compared with the post-fusion

structure (**Fig. 1.12B**) (Hastie et al., 2016a). In the prefusion structure, the elongated N-terminal helix is folded as four distinct segments: three α -helices and a loop. Furthermore, the T-loop forms antiparallel β -strands instead of the short helix observed in the post-fusion structure. The C-terminal helix forms hydrophobic contacts with T-region β -stands and the N-terminus of GP1 in the prefusion state, instead of packing against the long N-terminal helix observed in the post-fusion structure (**Fig 1.12B**).

High GP2 amino acid sequence homology (approximately 60%) is routinely observed amongst OW and NW mammarenaviruses, substantially more so than the GP1 counterparts (~20–40%), due to the conserved membrane fusion functionality of GP2 (Brouillette et al., 2017; Meulen et al., 2004). Indeed, the crystal structure of OW LASV GP1–GP2 trimer (PDB ID: 5VK2) showed that the GP2 subunit adopts a similar prefusion architecture when compared with the GP2 from the LCMV GP1–GP2 heterodimer structure, with a 2.0-Å RMSD over the entire structure and a 1.0-Å RMSD over the α -helices and β -sheets of GP1 and all of GP2 (**Fig. 1.12B**) (Hastie et al., 2017). The trimeric LASV GP1–GP2 structure also reinforces the necessity of SKI-1/S1P enzymatic processing of GP1/2, as the cleavage (RRLL motif) of C-terminus of GP1 from the N-terminus of GP2 is required to pivot the GP1 C-terminus away from the GP2 and allow the formation of the tripartite assembly without steric clashes at the trimerisation interface (Hastie et al., 2017). In *toto*, comparison of various GP2 structures provides a molecular blueprint of the structural rearrangement involved in the conversion from a compact tripod prefusion state to the extended six-helix bundle post-fusion state. To summarise, an outline of mammarenaviral GP2 structures published to date is listed in **Table 1.2**.

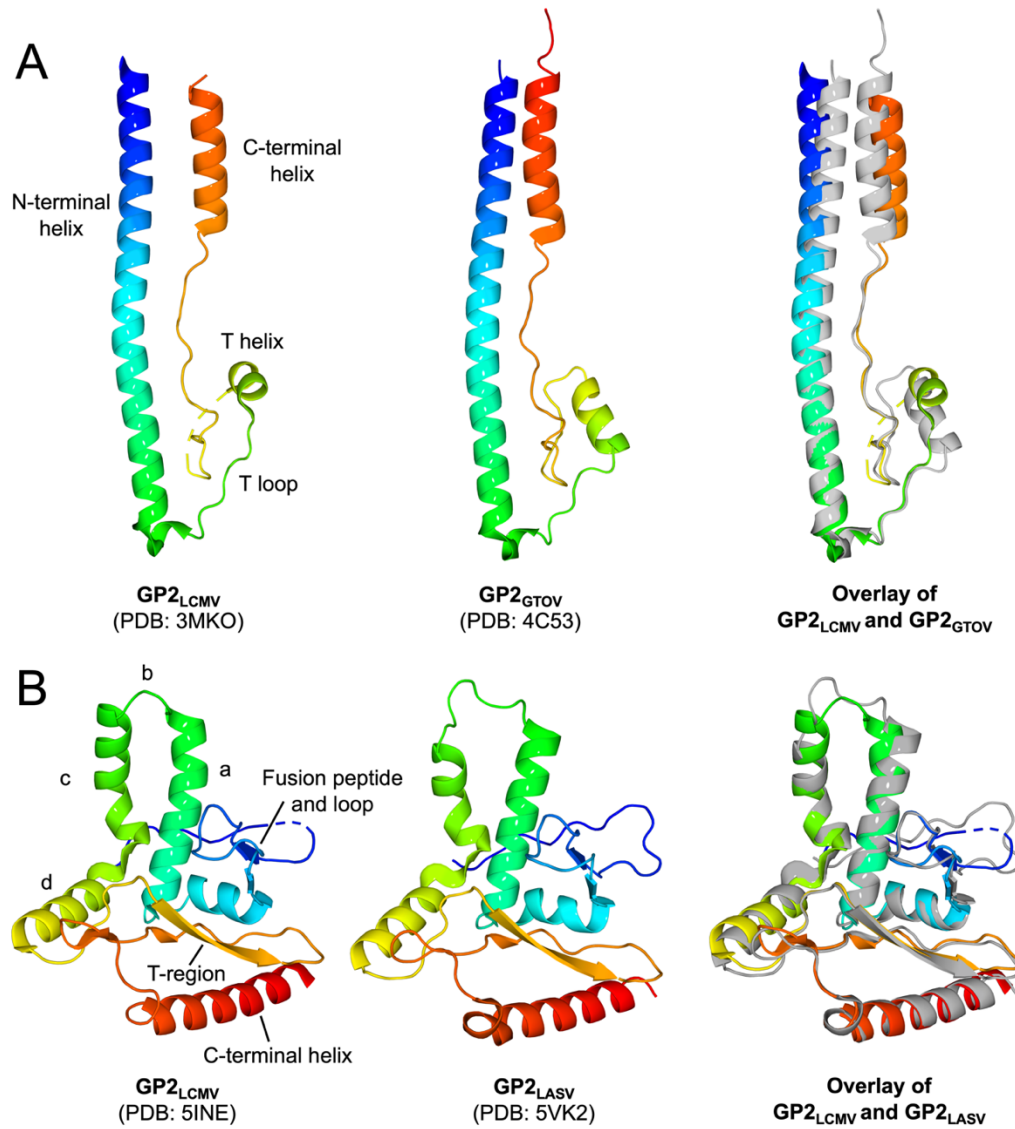


Fig. 1.12. Structure overview of OW and NW mammarenaviral GP2s. (A) Structures of OW LCMV GP2 (*left*) and NW GTOV GP2 (*middle*) in post-fusion conformation. The structures are shown in cartoon representation and coloured as a rainbow from N- (blue) to C-terminus (red). Unresolved regions are shown as dashed lines. (*Right*) Structural overlay of LCMV GP2 (rainbow) and GTOV GP2 (grey). (B) Structures of OW LCMV GP2 (*left*) and LASV GP2 (*middle*) in prefusion conformation. The GP2 structures were determined in the presence of the cognate GP1 subunits, which are not shown here. The structures are presented as in panel A. The long N-terminal helix in post-fusion conformation is folded as four distinct segments: three α -helices (a, c, d) and a loop (b). The T-loop observed in the post-fusion structures forms antiparallel β -strands. (*Right*) Structural overlay of LCMV GP2 (rainbow) and LASV GP2 (grey).

Table 1.2. Summary of mammarenaviral GP2 structures determined to date.

	Mammarenavirus	pH at which the structure was solved	PDB ID	References
OW	LCMV GP2 (post-fusion)	7.0	3MKO	(Igonet et al., 2011)
	LCMV GP2 (prefusion)	7.5	5INE	(Hastie et al., 2016a)
	LASV GP2 (prefusion)	8.0	5VK2	(Hastie et al., 2017)
NW	GTOV GP2 (post-fusion)	3.5	4C53	(Parsy et al., 2013)

The 58-amino acid retained myristoylated SSP is highly conserved amongst mammarenaviruses and contains two hydrophobic transmembrane domains connected by a short extra-viral loop (Eichler et al., 2003, 2004; Saunders et al., 2007). Importantly, SSP plays essential roles in GP cleavage maturation and membrane fusion activity (Albariño et al., 2011; Bederka et al., 2014). It was previously shown that SSP association is crucial in allowing JUNV to exit from the ER and enter the Golgi network where SKI-1/S1P protease is located (Agnihothram et al., 2006). Moreover, the specific interaction between SSP and GP2 stabilises the prefusion GP complex. Multiple residues at the transmembrane domain and intraviral tail of GP2 have been identified as essential for SSP association (Wang et al., 2016). This SSP–GP2 interaction provides a target for drug compounds such as ST-294, ST-193, ST-161, 16G8, and F3406 (Messina et al., 2012; Wang et al., 2016). These inhibitors disrupt the SSP–GP2 interface and thus destabilise the prefusion GP spike (Larson et al., 2008; Lee et al., 2008). Of note, the structural basis of SSP, SSP–GP1–GP2, and SSP–GP2 interactions remains unresolved.

1.3.2 Shielding of the Protein Surface by N-linked Glycosylation

The mammarenaviral GP spike is heavily glycosylated, ranging from eight to fifteen sites (N-X-S/T, where X is not P), in which the N-linked glycosylation is estimated to account for approximately 30–40% of the GP molecular mass (Parsy et al., 2013; Watanabe et al., 2018). The location of N-linked glycosylation sites in the GP1 vary across OW and NW mammarenaviruses, but are more conserved among viruses within the same serogroups (**Table 1.3**) (Sommerstein et al., 2015; Wang et al., 2016). Furthermore, OW GP1s (~6–10 sites) and clade C NW GP1s (9–10 sites) are generally more glycosylated than clade B and D GP1s (~4–7 sites) (Sommerstein et al., 2015; Wang et al., 2016), likely due to the requirement for α -DG and C-type lectin receptor recognition (Cao et al., 1998; Goncalves et al., 2013). Interestingly, clade A NW GP1s are also highly glycosylated (~8–11 sites), though their cellular receptor has not yet been established. On the other hand, N-linked glycosylation sites in the GP2 are less abundant (~3–5 sites) than the cognate GP1 and are highly conserved across OW and NW viruses (Wang et al., 2016).

N-linked glycosylation plays a number of important roles, including mammarenaviral GP folding, processing, maturation, receptor-binding, and immune evasion (Bonhomme et al., 2011; Bonhomme et al., 2013; Braakman and van Anken, 2000; Wright et al., 1989). For instance, N-linked glycans may contribute to constructing the RBS on GP1. Together with the pH-sensing histidine triad, the second N-linked glycosylation site (Asn89) on LASV GP1 forms the LAMP1-binding site (Cohen-Dvashi et al., 2015). On MACV GP1, the pyranose ring of N-acetylglucosamine (GlcNAc) moiety from Asn95 glycosylation site forms a stacking interaction with Phe98, which in turns interacts with hTfR1 (Abraham et al., 2010; Bowden et al., 2009). Furthermore, the first two N-linked glycosylation sites (Asn85 and Asn95) on LCMV GP1s are essential to ensure viral fitness and rescue (Bonhomme et al., 2013). Indeed,

the functional importance of these two sites are reflected by the high degree of conservation across mammarenaviruses (Bonhomme et al., 2013).

Table 1.3. N-linked glycosylation location in mammarenaviral GP sequences. Predicted N-linked glycosylation was determined according to NXT/S motifs, where $X \neq P$, using the NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>). This table is adapted from *Virologica Sinica* (Wang et al., 2016), with permission from Springer Nature.

	Virus	GP1										GP1 ^a GP2					GP2 ^b Length								
OW	Lujo	–	73	–	–	93	104	–	–	112	–	148	–	–	194	6	–	327	335	352	357	4	454		
	Dandenong	–	80	85	95	–	114	–	–	124	–	–	171	–	–	232	7	–	371	–	396	401	3	498	
	LCMV	–	–	85	95	–	114	–	–	124	–	–	171	–	–	232	6	–	371	–	396	401	3	498	
	Lunk	–	77	82	92	–	111	–	–	121	–	–	168	–	–	229	6	–	368	–	393	398	3	495	
	Merino walk	–	73	82	92	–	102	–	–	112	122	164	170	–	–	230	9	–	371	379	396	401	4	497	
	Ippy	66	–	78	88	97	107	–	–	117	–	159	165	197	–	228	10	–	369	377	394	399	4	495	
	LASSA	–	–	79	89	99	109	–	–	119	–	–	167	–	–	224	7	–	365	373	390	395	4	491	
	Loei river	65	77	–	87	97	107	–	–	117	–	159	165	–	–	225	10	332	366	374	391	396	5	492	
	Morogoro	–	–	78	88	98	108	–	–	118	–	–	166	–	–	222	7	–	363	371	388	393	4	489	
	Mopeia	–	–	78	88	98	108	–	–	118	–	–	166	–	–	222	7	–	363	371	388	393	4	489	
NW C	Luna	–	–	78	88	98	108	–	–	118	–	–	166	–	–	224	7	–	365	373	390	395	4	491	
	Mobala	–	–	78	88	98	108	–	–	118	–	–	166	–	–	224	7	–	365	373	390	395	4	491	
	Oliveros	–	75	–	90	101	112	117	–	131	–	–	180	–	234	251	9	–	389	397	414	419	4	518	
	Latino	–	75	–	90	101	112	117	–	131	138	–	180	–	231	248	10	–	386	394	–	416	3	515	
	Junín	–	–	–	95	105	–	–	–	–	–	166	178	–	–	–	4	–	357	365	382	387	4	485	
	Machupo	–	–	83	95	–	–	–	–	–	137	166	178	–	–	–	5	–	368	376	393	398	4	496	
	Tacaribé	–	–	83	95	–	–	–	–	124	129	–	176	188	–	–	6	–	367	375	392	397	4	495	
	NW B	Cupixi	–	–	–	–	88	99	–	–	125	–	–	174	–	–	214	5	315	352	360	377	382	5	480
		Amaparí	–	–	–	–	88	99	–	–	128	–	–	174	–	–	214	5	315	352	360	377	382	5	481
		Guanarito	–	–	–	–	88	–	–	–	125	–	–	174	–	202	214	5	314	351	359	376	381	5	479
Chapare		–	69	–	88	89	–	–	–	125	–	171	178	–	–	218	7	–	356	364	381	386	4	484	
Sabiá		–	69	–	88	89	–	–	–	125	–	171	178	–	–	222	7	–	360	368	385	390	4	488	
NW A	Pichindé	67	74	–	89	100	111	116	121	132	–	–	181	217	–	241	11	–	379	387	404	409	4	503	
	Pirital	–	75	–	90	101	112	117	122	133	–	–	181	218	–	243	10	–	381	389	406	411	4	510	
	Paraná	–	74	–	89	–	111	116	119	130	–	–	170	215	219	240	10	–	378	386	403	408	4	507	
	Flexal	–	74	–	89	–	111	116	–	130	–	–	179	–	223	240	8	–	378	386	403	408	4	507	
	Allpahuayo	–	74	–	89	–	111	–	119	130	–	–	179	–	223	240	8	–	378	386	403	408	4	507	
NW D	Whitewater arroyo	–	73	–	88	–	–	–	123	126	–	165	176	–	–	215	7	315	352	360	377	382	5	480	
	Skinner tank	–	73	–	88	–	–	–	–	129	–	168	178	–	–	214	6	316	353	361	378	383	5	481	
	Tamiami	–	73	–	88	–	–	–	117	128	–	–	179	–	–	218	6	320	357	365	382	387	5	485	
	Bear canyon	–	73	–	88	–	–	–	–	130	–	–	179	–	–	216	5	–	355	363	380	385	4	483	

Note: ^a The number of N-Linked glycosylation sequons on GP1. ^b The number of N-Linked glycosylation sequons on GP2.

A recent mass spectrometry study revealed the composition of LASV GP N-linked glycans, and when mapped onto the structure of LASV GP1–GP2 ectodomain, the glycans were observed to form a dense shield consisting of oligomannose-type glycans patches (Fig.

1.13) (Watanabe et al., 2018). The dense coverage of the GP proteinaceous surface by N-linked glycans is thought to facilitate protection from the host immune system, including detection by nAbs (Sommerstein et al., 2015; Zeltina and Bowden, 2017). Indeed, the coexistence of high LASV viral load and high antibody titres highlights the inefficient antibody-mediated neutralisation response against LASV infection (Fisher-Hoch and McCormick, 2001; Johnson et al., 1987). A negative correlation has been observed between the GP glycan density and the humoral immune response. In contrast to LASV, where a majority of the nAbs target the quaternary epitopes or the GP2 subunit (Robinson et al., 2016), JUNV GP1 with four N-linked glycans, elicits a higher nAb titre than viruses with GP1s consisting of more glycans (Sommerstein et al., 2015). Moreover, deletion of selected glycan sites in JUNV GP1 has demonstrated a more potent nAb response than wild type GP1, suggestive of increased antibody accessibility to neutralising epitopes (Sommerstein et al., 2015).

Glycan-mediated immune evasive properties displayed by the mammarenaviral GP (e.g. LASV) are reminiscent to those observed on the human immunodeficiency virus type-1 (HIV-1) envelope glycoprotein (Env) (Doores, 2015). Indeed, the glycan-mediated immunoevasive properties of LASV (Watanabe et al., 2018) and HIV (Crispin et al., 2018) seem to exceed other highly glycosylated viruses for which glycan coat has been characterised, including the haemagglutinin glycoprotein of influenza virus (Wanzeck et al., 2011) and the glycoprotein spikes of coronavirus (Watanabe et al., 2020) and Ebola virus (Ritchie et al., 2010). Interestingly, anti-HIV broadly nAbs (bnAbs) recognise patches of mannose-, hybrid-, and complex-type glycans presented on the Env (Falkowska et al., 2014; Pancera et al., 2013; Scanlan et al., 2002). Similarly, bnAbs against Dengue virus (DENV) E surface glycoprotein are able to cross-neutralise all four DENV serotypes by binding to conserved epitopes comprised of glycans (Rouvinski et al., 2015). Therefore, isolation of bnAbs targeting conserved mammarenaviral GP glycans may constitute a feasible therapeutic strategy against

“evasion strong” mammarenaviruses (Burton, 2017), such as LASV, with heightened resistance to antibody-mediated neutralisation.

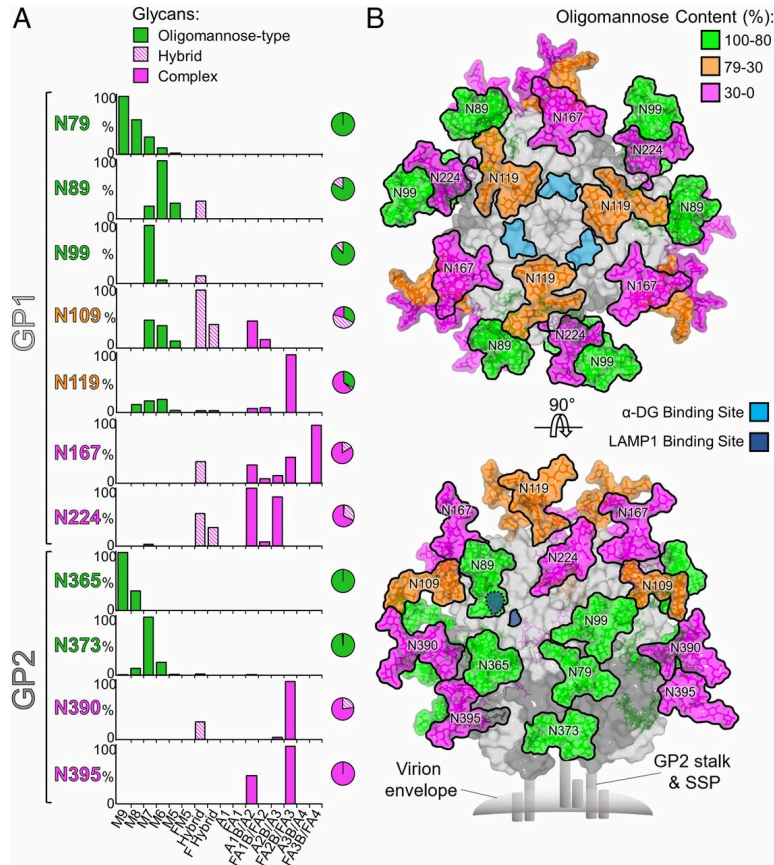


Fig. 1.13. Compositional analysis and structure-based mapping of LASV GP N-linked glycans. (A) Quantitative site-specific N-linked glycan analysis of LASV GP. The bar graphs represent the relative quantities of each glycan group with oligomannose-type glycan series (M9 to M5; $\text{Man}_9\text{GlcNAc}_2$ to $\text{Man}_5\text{GlcNAc}_2$) (green), afucosylated and fucosylated hybrid glycans (Hybrid & F Hybrid) (dashed pink), and complex glycans grouped according to fucosylation and number of antennae (A1 to FA4) (pink). The pie charts summarize the quantification of these glycans. (B) Modelling experimentally

observed glycosylation onto the crystal structure of the trimeric LASV GP1–GP2 (PDB ID: 5VK2) (Hastie et al., 2017). The glycans are coloured according to the oligomannose content, as defined by the upper-right-hand key. The locations of α -DG and LAMP1 binding sites have been previously proposed and are annotated in cyan and dark blue, respectively. GP1 and GP2 are coloured light grey and dark grey, respectively. This figure is reproduced from *PNAS* (Watanabe et al., 2018), under a Creative Commons Attribution License (CC BY 4.0).

1.4 RABIES VIRUS

1.4.1 Rabies Disease

The genus *Lyssavirus* (Greek ‘*lyssa*’, meaning ‘rage’), within the family *Rhabdoviridae* (Greek ‘*rhabdos*’, meaning ‘rod’) of the order *Mononegavirales*, contains 17 members, including rabies virus (RABV), the prototype species of the genus (Walker et al., 2018). Although most

lyssaviruses are carried by bats, RABV has extremely wide host range, encompassing many mammalian species such as foxes, racoons, dogs, and humans (Marston et al., 2018). Indeed, RABV is endemic within domesticated and wild dog populations, particularly in Asia and Africa (Fooks et al., 2017). Today, an estimated of 3 billion people are living at risk of RABV infection, making the virus one of the most globally prevalent zoonotic pathogens (Hellert et al., 2020; Knobel et al., 2005).

Human RABV infection is usually initiated via a bite or scratch, where >99% of human cases are accounted by dog bites (Brunker and Mollentze, 2018). Following initial infection, RABV amplifies in the local muscle tissues before retrogradely transports via peripheral nerves to dorsal root ganglia, and ultimately to the central nervous system (CNS) as a primary site of replication (Guo et al., 2019). Infected patients generally present one to ten-day period of fever, headache, malaise, myalgias, photophobia, and a tingling, prickling, or burning sensation at the site of bite (Giesen et al., 2015). Disease progression is indicated by excessive motor activity, followed by hydrophobia (fear of water) and dysphagia (difficulty in swallowing) (Jackson, 2018). During the latter stages of the disease, muscle spasms, paralysis, and hallucinations are common. Eventually, patients become comatose, and death occurs shortly thereafter (Kauffman and Goldmann, 1986). Following the onset of symptoms, RABV infection almost always results in mortality and therefore, poses a persistent threat to human health and economy.

Although a human vaccine is available against RABV, the virus causes approximately 60,000 deaths each year (Fooks et al., 2017). Administration of post-exposure prophylaxis (PEP) consisting of three to four doses of vaccine over the first week and one dose of rabies immunoglobulin (RIG) without delay is crucial for virus clearance before it enters the nervous system (Cantaert et al., 2019). The expense and complications associated with the production and administration of rabies vaccine and serum-derived RIG, combined with the low

availability of PEP in affected countries, highlights the need for optimised vaccines and prophylactics capable of preventing against infection.

1.4.2 RABV Lifecycle

RABV is an enveloped virus with a non-segmented, negative sense RNA genome of approximately 12 kb in length (Zhang and Fu, 2012). The genome encodes for five structural proteins in a conserved order: nucleoprotein (N; ~57 kDa), phosphoprotein (P; ~38 kDa), matrix protein (M; ~25 kDa), envelope glycoprotein (G; ~65 kDa), and the large RNA polymerase (L; ~244 kDa). When viewed under the electron microscope, RABV virions appear as 130–250 nm long bullet-shaped particles (Guichard et al., 2011). The N protein encapsidates the RNA genome to form a tightly coiled helical RNP, which extends along the longitudinal axis of the virion (Wunner, 2007). The RNP is surrounded by the lipoprotein viral envelope composed of the lipid bilayer and M. The membrane-anchored G assembles as a trimeric spike on the virion surface and is responsible for host cell entry, including receptor attachment and membrane fusion activity. The N, P, and L proteins are involved in polymerase activity. Indeed, P interacts with N to prevent non-specific binding of N to cellular RNA (Mavrakakis et al., 2003), while P acts as a non-catalytic factor in transcription and replication of the viral genome by stabilising the P–L complexes on the RNA template (Chenik et al., 1998; Fu et al., 1994).

Viral entry is initiated by the interaction between RV-G and host receptors. A number of receptors have been implicated to facilitate this host cell recognition event, including the nicotinic acetylcholine receptor (nAChR) (Lentz et al., 1982; Lentz et al., 1983), nerve cell adhesion molecule (NCAM) (Thoulouze et al., 1998), the low-affinity nerve growth factor receptor (p75NTR) (Tuffereau et al., 1998), and metabotropic glutamate receptor 2 (mGluR2) (Wang et al., 2018). nAChR is expressed in neuromuscular junctions and has been suggested to facilitate RABV infection of and replication in muscle cells (Lafon, 2005), NCAM is a cell

adhesion glycoprotein expressed on neurons and lymphocytes (Thoulouze et al., 1998) while p75NTR mediates retrograde axonal transport of RABV (Gluska et al., 2014). Expression of mGluR2 is localised to neuronal cells of the brain (Wang et al., 2018). The ability of RABV to recognise different receptors reflects the efficient spread of the virus through the CNS.

Upon receptor attachment, virus particles are internalised via receptor-mediated endocytosis and delivered to acidified endosomes. Low pH in the late endosomal compartment triggers fusogenic rearrangement in RV-G from prefusion to post-fusion conformation to mediate fusion of virus and endosomal membranes and the release of helical RNP into the cytoplasm (Yang et al., 2020). The five structural genes (N, P, M, G, and L) are transcribed into five separate positive sense 5'-capped and 3'-polyadenylated mRNAs, a process driven by the viral L and P proteins. In addition, transcription of the full-length negative sense viral genome produces a positive sense antigenome cRNA, which serves as a template for the replication of negative sense progeny viral RNA. The N, P, M and L proteins are translated from their respective mRNAs by free cytoplasmic ribosomes, whereas the RV-G is synthesised by the rough ER-bound ribosomes (Wunner and Conzelmann, 2013). The polypeptide transits through the lumens of ER and Golgi network where protein folding, quality control, and glycosylation take place.

Virus assembly and budding is initiated by the formation of RNP when N, P, and L interacts with the progeny viral RNA genome (Liu et al., 2004). Subsequently, the association of M protein with the RNP inhibits transcription and promotes replication (Finke et al., 2003). Furthermore, M–RNP interaction condenses the RNP into a tightly coiled helical state in which the polymerase activity is halted, and is proposed to give the RABV virion its bullet-shape morphology (Wunner and Conzelmann, 2013). Localisation of M to the cellular membrane where glycosylated trimeric RV-G are located also allows M proteins to interact with the intracellular tails of RV-G (Mebatsion et al., 1999). The PPEY motif (L domain) in M is

associated with virus budding (Harty et al., 1999). Oligomerisation of M at the membrane increases the binding affinity of M for the RV-G and lipid bilayer, and may play a major role in membrane curvature and virus egress (Graham et al., 2008). The RABV lifecycle is summarised in **Fig. 1.14**.

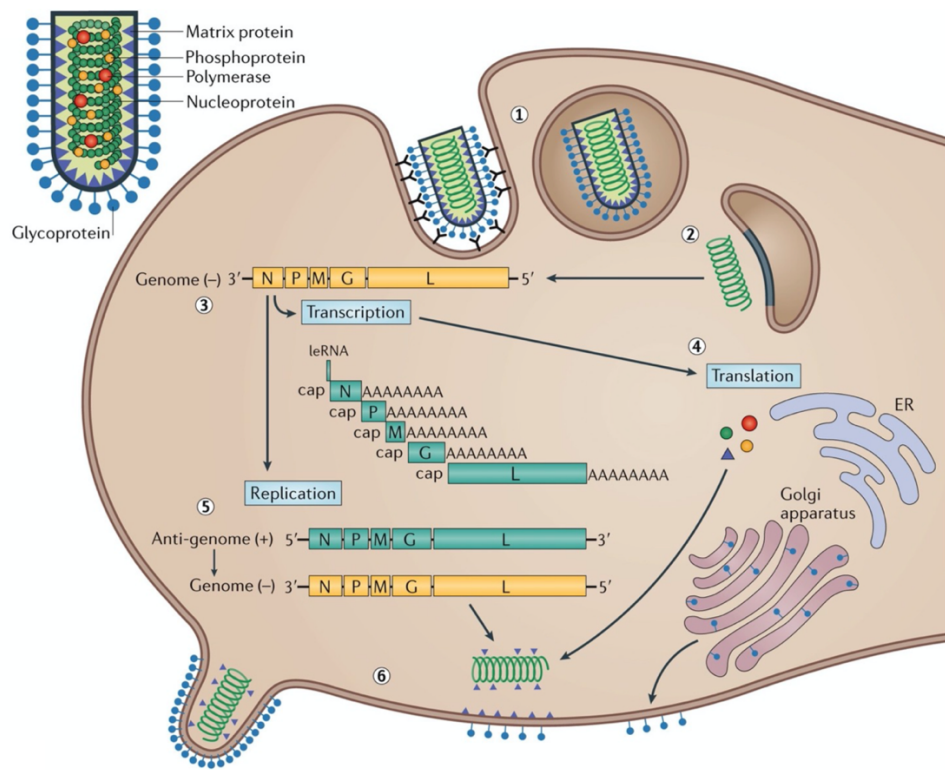


Fig. 1.14. RABV lifecycle. Host cell entry of RABV is mediated by receptor-mediated endocytosis, a process initiated by the interaction between viral surface RV-G and cell surface receptors (step 1). Subsequently, acidification of the endosomal compartments associated with the vesicle trafficking promotes the fusion of the viral membrane with the endosomal membrane to release the viral genome (uncoating, step 2). The encapsidated negative-sense single stranded RNA genome is transcribed by the polymerase complex, starting with a short, uncapped leader RNA (leRNA), followed by the transcription of 5' end-capped (cap) and polyadenylated (A) mRNAs and their translation into the viral proteins nucleoprotein (N), phosphoprotein (P), matrix protein (M), glycoprotein (G) and polymerase (L). N, P, M, and L are synthesised in the cytoplasm, while G is translated and processed in the ER and Golgi networks (steps 3 and 4). In replication mode, the full-length anti-genomic RNA is produced. The synthesised anti-genome functions as a template for the synthesis of additional copies of genomic viral RNA (step 5). The viral components are assembled, and the RABV virions bud and are released, starting a new round of infection (step 6). This figure is adapted from *Nature Reviews Microbiology* (Fisher et al., 2018), with the permission from Springer Nature.

1.4.3 Rabies Virus Glycoprotein (RV-G)

The envelope surface of RABV is decorated by a spike glycoprotein, RV-G, which facilitates receptor recognition and entry of the virus into the host cell (Guo et al., 2019). Endocytic uptake of RABV is initiated following the interaction between RV-G and host cell receptors. Although a number of receptors have been implicated to facilitate this host cell recognition event, including nAChR (Lentz et al., 1982; Lentz et al., 1983), NCAM (Thoulouze et al., 1998), p75NTR (Tuffereau et al., 1998), and mGluR2 (Wang et al., 2018), none of these molecules have been shown to be absolutely necessary for RABV infection, and little is known about the molecular determinants of this process. Following host cell attachment, pH changes associated with endosomal trafficking of RABV trigger conformational rearrangement of RV-G from the prefusion form to the low-energy post-fusion state, using released energy to promote merger of the viral and host endosomal membranes (Yang et al., 2020).

As described in section 1.3.1, viral fusion glycoproteins are divided into three structural classes: I, II, and III (Harrison, 2008). While the native trimeric architecture of RV-G remains unknown, RV-G has been shown to adopt a class III architecture similar to the fusion glycoproteins of vesicular stomatitis virus (VSV), Mokola virus (MOKV), Chandipura virus (CHAV) and herpesviruses (e.g. Epstein-Barr virus (EBV), herpes simplex virus (HSV), and human cytomegalovirus (HCMV)) (Hellert et al., 2020; Yang et al., 2020). Unlike Class I fusion protein such as arenaviral GP2 glycoprotein where pH-dependent fusogenic rearrangement is irreversible (Hastie et al., 2016a; Hastie et al., 2017; Igonet et al., 2011; Parsy et al., 2013), Class III fusion proteins undergo reversible pH-triggered conformational changes between pre- and post-fusion states (Backovic and Jardetzky, 2011). To date, post-fusion structures are available for VSV-G (Roche et al., 2006), MOKV-G (Belot et al., 2020), EBV glycoprotein B (gB) (Backovic et al., 2009), HSV gB (Heldwein et al., 2006), HCMV gB (Chandramouli et al., 2015), and CHAV (Baquero et al., 2015) fusion glycoproteins (for CHAV

a structure of a putative intermediate is also known) (**Fig. 1.15A**). However, the trimeric prefusion structures of a class III fusogen are only known for VSV-G (Roche et al., 2007) and HCMV gB (Liu et al., 2021) (**Fig. 1.15B**), which exhibit a number of features that set it apart from class I and II fusion proteins, including a tripod-like assembly where the membrane-linked C-termini are separated by ~60 Å. The structure of RV-G in the trimeric pre- or post-fusion states remains unknown. Indeed, the unstable nature of RV-G trimer upon extraction from membranes using detergents poses an obstacle to recombinant protein production for structural studies (Gaudin et al., 1991; Whitt et al., 1991).

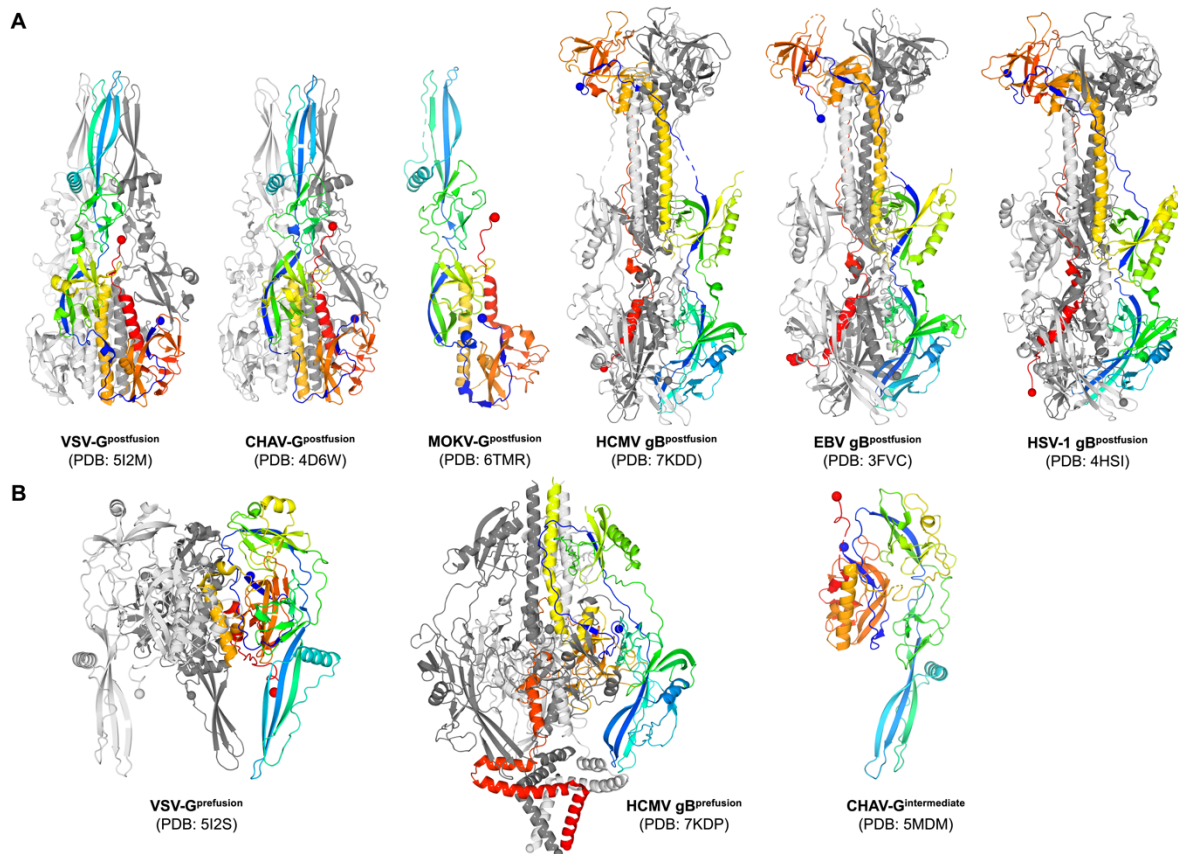


Fig. 1.15. An overview of the structures of class III viral membrane fusion proteins. (A) Post-fusion structures of VSV-G (PDB ID: 5I2M), CHAV-G (PDB ID: 4D6W), MOKV-G (PDB ID: 6TMR), HCMV gB (PDB ID: 7KDD), EBV gB (PDB ID: 3FVC), and HSV-1 gB (PDB ID: 4HSI). The structures are shown in cartoon representation. One of the protomers is coloured as a rainbow from N- (blue) to C-terminus (red), with the other two protomers coloured light grey and dark grey for clarity. The structures are presented such that the fusion domains or fusion loops are oriented upright (towards the host cell membrane). The N- and C-termini are shown as blue and red spheres, respectively. Unresolved regions are shown as dashed lines. (B) Prefusion structures of VSV-G (PDB ID: 5I2S) and HCMV gB (PDB ID:

7KDP) and intermediate structure of CHAV-G (PDB ID: 5MDM). The structures are shown in cartoon representation and coloured as in panel *A*. The structures are presented such that the fusion domains or fusion loops are oriented downwards (towards the viral membrane).

Structural elucidation of monomeric RV-G ectodomain at both pH 8.0 (PDB ID: 6LGX) and pH 6.5 (PDB ID: 6LGW) conformations provides a molecular blueprint for the pre-to-post structural transition of RV-G (Yang et al., 2020) (**Fig. 1.16A**). The ~80 Å tall basic pH-structure of RV-G ectodomain is composed of three domains: a membrane-distal pleckstrin homology domain (PHD), a membrane-proximal fusion domain (FD) and a central domain (CD) that is involved in the trimerisation of the molecule. The PHD sits atop the FD while the CD is positioned at the side of the PHD–CD interface. This domain arrangement is similar to that observed in the prefusion structure of VSV-G (PDB ID: 5I2S) (Roche et al., 2007). Structural superimposition of the basic- and acidic-pH RV-G structures further revealed that the pH-dependent structural rearrangement involves dramatic inter-domain reorientations and concomitant domain–linker reconstructions of the structure (**Fig. 1.16B**). Indeed, while the individual domains remain largely unaltered, the CD was rotated by approximately 180° and the FD is flipped by 60°. The antigenic deconfiguration associated with the pH-dependent conformational changes is reflected by the elongation of the RV-G structure to ~120 Å tall where the PHD is now sandwiched between the CD and FD (**Fig. 1.16B**).

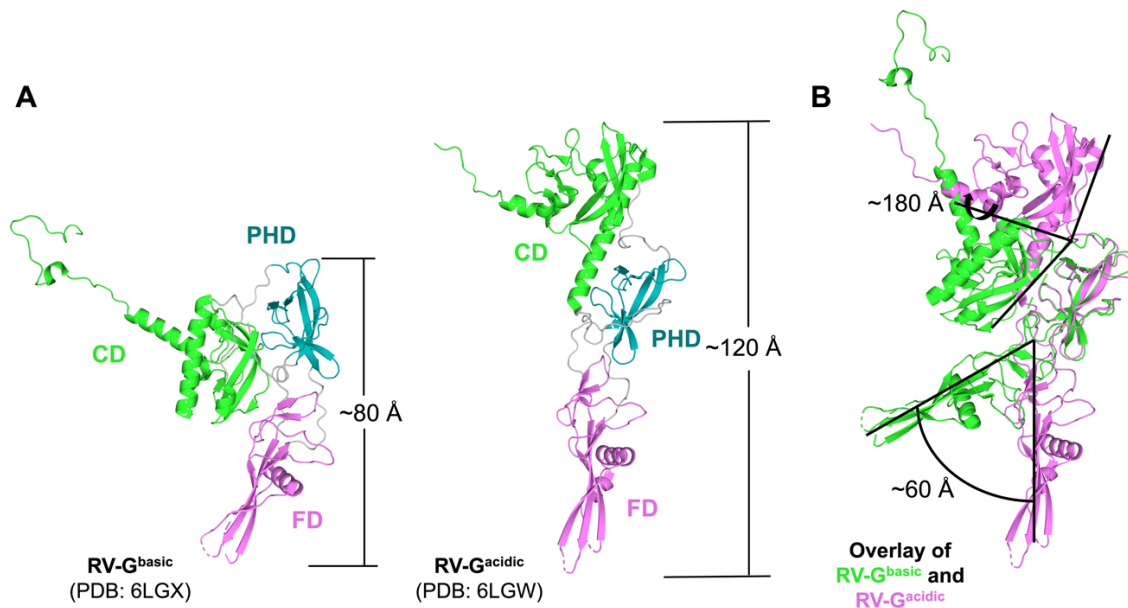


Fig. 1.16. An overview of reported RV-G ectodomain structures at pH 8.0 and pH 6.5. (A) The crystal structures of RV-G ectodomain at basic pH 8.0 (*left*; PDB ID: 6LGX) and acidic pH 6.5 (*right*; PDB ID: 6LGW). The structures are shown in cartoon representation with domains CD, PHD, and FD coloured green, cyan, and magenta, respectively, and domain-linkers in grey. The heights of the molecules are labelled. The regions that are not visible in the structure are indicated with dashed lines. (B) Basis of low-pH-induced structural transition in RV-G. Structural superimposition of the basic- and acidic-pH RV-G ectodomain structures based on PHD. The structures are shown as in panel A, with the RV-G^{basic} and RV-G^{acidic} coloured green and magenta, respectively. The domain orientation observed in the CD and FD are marked and labelled.

RV-G spikes protrude from the virus envelope and constitute the primary target of the nAb response arising during infection and immunisation. Extensive epitope mapping studies have demonstrated that nAbs predominantly target five regions on RV-G, which are defined as major sites I, II, III, IV, and a minor site termed ‘a’ (Benmansour et al., 1991; De Benedictis et al., 2016; Lafon et al., 1983; Nagarajan et al., 2008) (**Fig. 1.17A**). Furthermore, recent structural studies of monomeric RVG ectodomain have provided some structure-based models for how some nAbs may interrupt p75NTR recognition (Yang et al., 2020) and block fusogenic rearrangements (Hellert et al., 2020). However, there is limited information for how the mature and trimeric RV-G spike complex is recognised by nAbs, including those currently being used clinically for PEP in India (e.g. nAb Rab1 or ‘Rabishield’; produced by Mass Biologics and

Serum Institute of India Pvt. Ltd.). Additionally, there is a paucity of knowledge with regards to how RV-G interacts with receptors during host-cell entry.

Current licenced nAbs for rabies treatment do not offer full coverage against all circulating RABV strains, posing a risk for lack of efficacy and viral escape (De Benedictis et al., 2016). Indeed, the clinical trial of a nAb CL184 (produced by Crucell) had been halted due to a lack of broad RABV coverage. On the other hand, bnAbs RVC20 and RVC58 exhibit high neutralisation potency and were able to neutralise 35 tested RABV isolates and 25 non-RABV lyssaviruses (De Benedictis et al., 2016), indicative of the importance to identify bnAbs that are able to recognise conserved RV-G sequences from all lineages. Indeed, based on an analysis of the sequence conservation of the known RV-G antigenic sites (**Fig. 1.17A**) across the genus *Lyssavirus*, seven sequentially conserved and solvent accessible regions were proposed as possible cross-neutralising epitopes (Buthelezi et al., 2016). The crystal structure of PHD of RV-G in complex with nAb RVC20 suggests that the antibody prevents fusogenic rearrangement of RV-G by locking the glycoprotein in its prefusion state (Hellert et al., 2020) (**Fig. 1.17B**). Additionally, in corroboration with the reversible nature of pH-dependent conformational changes of RV-G, the ability of nAb 523-11 to recognise the CD of RV-G in both pre- and post-fusion structures underscores the functional significance of both prefusion and postfusion-targeting nAbs (Yang et al., 2020) (**Fig. 1.17C**). Therefore, the combination of nAbs consisting of potent and large breadth of reactivity against all RABV strains and multiple RV-G structural states will significantly boost PEP efficiency.

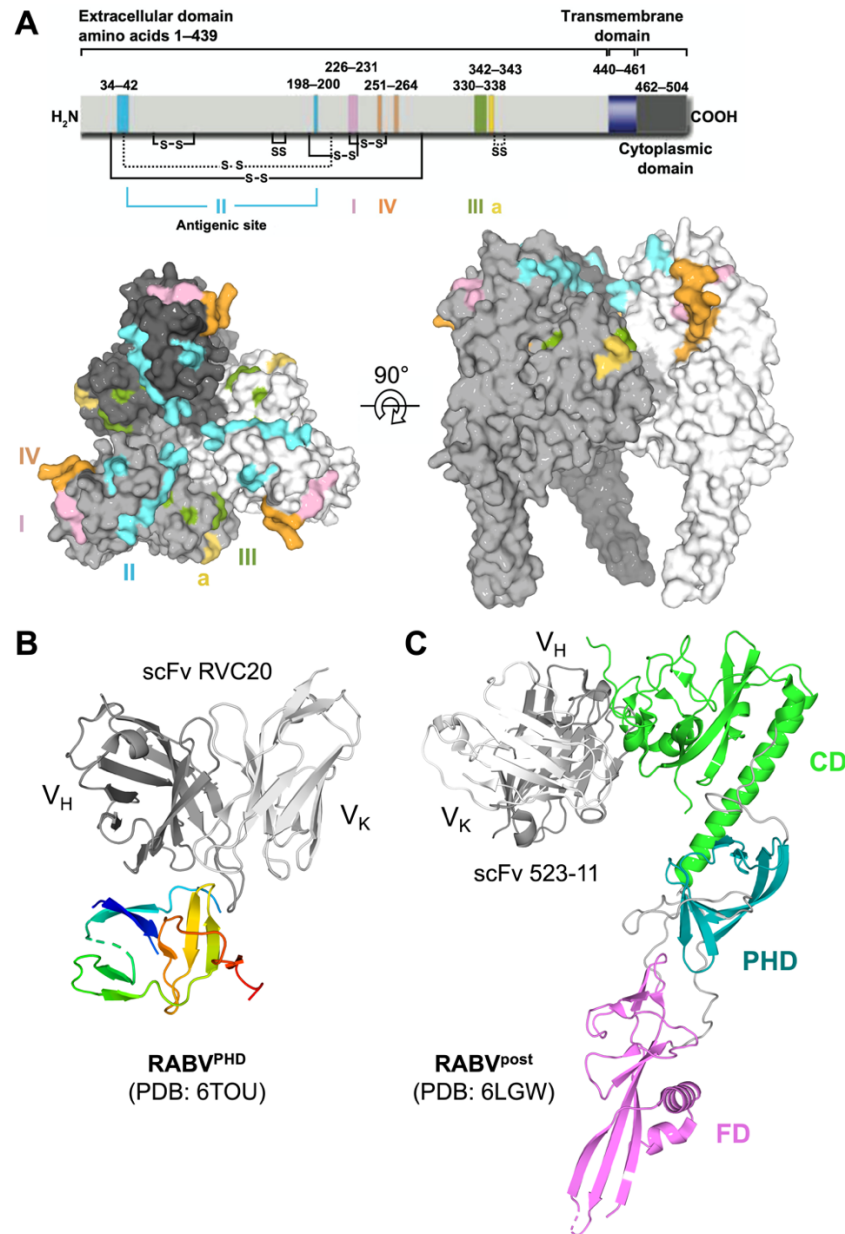


Fig. 1.17. Representation of antigenic sites on the RV-G. (A) Schematic showing the relative position and amino acid numbering of the antigenic sites (I, II, III, IV, and a) within the RV-G ectodomain. (Top) Linear representation of the antigenic sites. The numbering relates to the mature glycoprotein (after removal of the 19-mer signal peptide). Disulfide bonds are shown as solid lines (based on VSV-G) and dashed lines as defined by (Walker and Kongsuwan, 1999). (Bottom) Surface representation of the antigenic sites. Trimeric VSV-G ectodomain (PDB ID: 5I2S) were shown as surface representation, with the three G protomers coloured in white, light grey, and dark grey, respectively. The antigenic sites as shown in the top row are coloured accordingly. (B) The crystal structure of the PHD of RV-G in complex with single-chain variable fragment (scFv) RVC20 (PDB ID: 6TOU). The structure is shown in cartoon representation. The PHD is coloured as a rainbow ramped from N- (blue) to C-terminus (red). V_H and V_K denote variable heavy chain (grey) and variable kappa light chain (white). Unresolved regions are shown as dashed lines. (C) The crystal structure of RV-G in post-fusion conformation bound to scFv 523-11 (PDB ID: 6LGW). The structure is shown as in panel B

with the central domain (CD), pleckstrin homology domain (PHD), and fusion domain (FD) coloured as green, blue, and purple, respectively. V_H and V_K denote variable heavy chain (grey) and variable kappa light chain (white). Unresolved regions are shown as dashed lines. Note that figure in panel A, top row is reproduced from *Antibody Technology Journal* (Nagarajan et al., 2014), under a Creative Commons Attribution License (CC BY-NC 3.0).

1.5 THESIS AIMS

Emergence and re-emergence of zoonotic viruses pose a prominent threat to the public health and social economy. Therefore, understanding of the pathobiological features of the viruses including the molecular determinants of virus-host interactions, allows us to predict and determine their capability to spread into new host reservoirs and formulate efficient prophylaxis and treatments against these pathogens.

The diversity of mammarenaviruses is reflected by the continual discovery of novel species with unknown zoonotic potential in new territories. Indeed, the discovery of Asiatic WENV, LORV, and RYKV, which are closely related to OW LASV, highlights the wider than previously appreciated geographic distribution of mammarenaviruses (Blasdell et al., 2016; Li et al., 2015; Wu et al., 2018). Here, I focus on the mammarenaviral receptor-binding GP1, a key molecule responsible for host cell entry and thus, a primary determinant of cell tropism and cross-species transmission. To better understand the receptor tropism and pathogenic potential of Asiatic mammarenaviruses, I report the crystal structures of monomeric LORV GP1 obtained at both basic and acidic pH. Furthermore, comparison of the two LORV GP1 structures with previously reported OW monomeric GP1s provides structure-based template for studying the RBS, which rationalises the differential receptor tropism observed in these viruses. This work expands our appreciation of the structural landscape sampled by the mammarenaviral GP1s and provides a structure-guided rationale for designing a vaccine that resembles the GP spikes presented on the mature virion.

NAbs have been shown to be effective for the prevention and treatment of HF mammarenaviral infection (Enria et al., 1984; Maiztegui et al., 1979). Characterisation of the key sites of vulnerability on the mammarenaviral GP is thus important for the rational design of nAb-based prophylactics that target the mammarenavirus surface. In collaboration with Dr. Katie Doores (King's College London, U.K.) and Prof. Dr. Daniel Pinschewer (University of Basel, Switzerland), we utilised a mouse model of MACV and JUNV infection combined with antigen-specific B cell sorting to isolate nAbs specific to MACV and JUNV GP1s. Crystallographic investigation of the Fab fragments of nAbs in complex with their cognate GP1 antigens reveals the mechanisms of nAb-mediated neutralisation of NW arenaviruses and a structure-based model for how cross-reactive nAbs may recognise multiple mammarenaviral species.

Structure elucidation of the full-length trimeric GP spike provides a template for understanding the antigenic surface of the GP and identifying the higher-order quaternary nAb epitopes. While the structural studies of mammarenaviral GP or GP1–GP2 were hampered by the current COVID-19 pandemic, our ongoing collaboration with Dr. Alexander Douglas (Jenner Institute, University of Oxford, U.K.) on the structural characterisation of RV-G yields promising preliminary results. In this project, the structure of the trimeric, full-length RV-G spike in complex with nAbs was characterised using cryo-EM and SPA. The initial cryo-EM data illuminate the trimeric arrangement of the RV-G spike and the distinct nAb-targeting epitopes, providing structural framework for immunogen design and molecular basis for nAb-mediated neutralisation of RABV.

2 MATERIALS AND METHODS

2.1 MOLECULAR BIOLOGY AND MICROBIOLOGY

2.1.1 Complementary DNA (cDNA) Templates

Genes encoding nucleotide sequences of target protein constructs were codon optimised for mammalian cell expression system and synthetically manufactured by GeneArt™ (Life Technologies) as lyophilised DNA fragments or within a delivery vector. Constructs were designed with restriction sites (e.g. AgeI, KpnI, EcoRI, XhoI sites) at both 5' and 3' ends to facilitate cloning into the desired expression vectors such as pHLsec (Aricescu et al., 2006).

Construct design was guided by factors including: number of cysteine residues, number of N-linked glycosylation site (lesser number of glycans facilitates crystallogenes), removal of transmembrane regions for soluble protein expression (using TMHMM server: <http://www.cbs.dtu.dk/services/TMHMM/>) (Möller et al., 2001), and sequence alignment (using Clustal Omega server: <https://www.ebi.ac.uk/Tools/msa/clustalo/> and MultAlin server: <http://multalin.toulouse.inra.fr/multalin/>) (Corpet, 1988; Sievers and Higgins, 2014; Sievers et al., 2011) with structurally characterised homologous and orthologous proteins. The studied protein constructs with their corresponding Genbank accession numbers are listed in **Table 2.1**.

Table 2.1. List of studied protein constructs.

Protein Name	Protein construct boundaries	Genbank accession number
LORV GP1	Met80–Thr238	AHE76159.1
JUNV GP1	Asp87–Asn232	ACO52428.1
MACV GP1	Glu87–Phe257	AAS77647.1
RABV G	Met1–Leu524	AAA47209.1

Anti-JUNV Fab JUN1	See Appendix 8.1 for nucleotide and amino acid sequences
Anti-MACV Fab MAC1	See Appendix 8.1 for nucleotide and amino acid sequences
Anti-RABV Fab 17C7	See Appendix 8.1 for nucleotide and amino acid sequences
Anti-RABV Fab 1112-1	Derived from murine hybridoma cells. The sequence is currently being determined.

2.1.2 Cloning of cDNA into Expression Vectors

Lyophilised DNA (strings or within delivery vector) and primers were resuspended in sterile deionised water or 10 mM Tris (pH 8.0) buffer solution. Target genes were amplified using polymerase chain reaction (PCR) with gene-specific primers (section **2.1.2.1**). PCR-amplified cDNAs were then treated with restriction enzymes and ligated into appropriate expression vectors (section **2.1.2.2**).

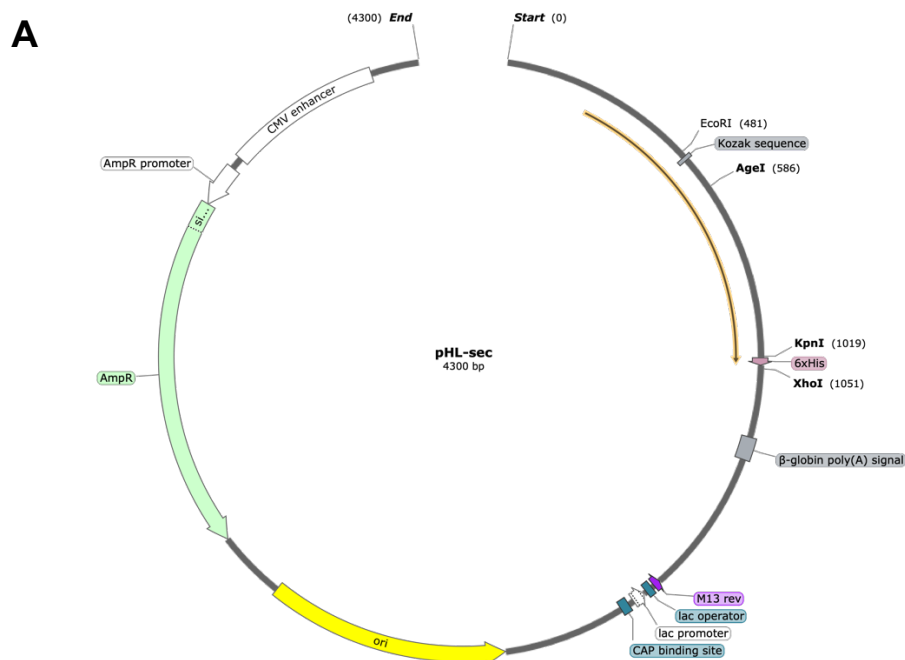
2.1.2.1 Primer Specifications

DNA primers (synthetically manufactured by Thermo Fisher Scientific) were designed to contain restriction endonuclease sites with short CG-rich overhangs to facilitate restriction digestion and overlapping sequence complementary to the start (5') or end (3') of the desired protein construct. The lengths of the primers were adjusted based on the melting temperature (T_m) of the primer–construct complementary pair. The T_m was set to between 64°C and 68°C, calculated using the online NEB T_m calculator (<https://tmcaculator.neb.com/>). Lyophilised DNA primers were resuspended in sterile deionised water or 10 mM Tris (pH 8.0) buffer solution to a final concentration of 100 µM.

2.1.2.2 Expression and Transfer Vectors

PCR-amplified cDNAs were inserted into the pHLsec expression vector (Aricescu et al., 2006) (**Fig. 2.1A**) using restriction site ligation. The pHLsec vector contains a secretion signal sequence (MGILPSPGMPALLSLVSLLSVLLMGCVA) and a C-terminal hexa-histidine (His₆) tag, facilitating the production of secreted mammalian soluble constructs and protein purification by immobilised metal ion affinity chromatography, respectively. Typically, cDNAs encoding the protein of interest were inserted into the vector between AgeI and KpnI restriction sites. The EcoRI site was used for constructs that included their signal peptide sequences. XhoI was used to create constructs lacking a C-terminal His₆ tag.

For protein production using an established lentivirus system (Elegheert et al., 2018), the pHR-CMV-TetO₂ transfer plasmid (**Fig. 2.1B**) was used. Protein constructs were cloned into the plasmid using the EcoRI and XhoI sites. Transfection of Lenti-X human embryonic kidney (HEK) 293T cell line with this plasmid allowed the production of lentivirus carrying target gene that was capable of infecting and transducing a target cell line (see section 2.2.2 for stable cell line production using lentivirus).



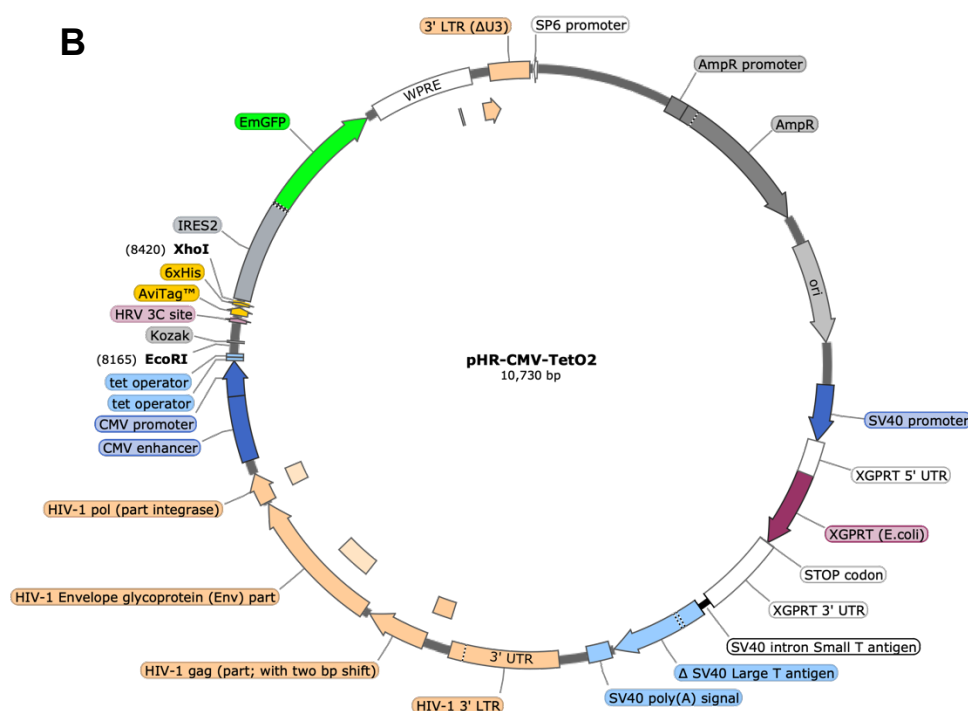


Fig 2.1. Plasmid maps. (A) pHLsec (Aricescu et al., 2006) and (B) pHR-CMV-TetO₂ (Elegheert et al., 2018) are annotated with features and relevant restriction enzyme sites. All construct inserts were cloned into the plasmids at the multiple cloning site (MCS), where restriction sites such as EcoRI, AgeI, KpnI, and XhoI are encoded. Plasmid maps are generated in SnapGene software.

2.1.3 Polymerase Chain Reaction (PCR)

Target genes (**Table 2.1** and **Appendix 8.1**) were amplified using the manufacturer's recommended PCR protocol for the Phusion® High-Fidelity PCR Mastermix with HF buffer (New England Biolabs, NEB). Each PCR reaction was assembled in a 50 μ L volume (22.5 μ L deionised water, 1.5 μ L dimethyl sulfoxide (DMSO), 25 μ L 2x Phusion Master Mix, 0.5 μ L template DNA, and 0.5 μ L primers) with 15–20 ng of template DNA and 0.5 μ M of each forward and reverse primer. A representative thermocycling protocol is shown in **Table 2.2**. The resulting PCR products were analysed using agarose gel electrophoresis, followed by DNA band excision from the agarose gel and purification using QIAquick® Gel Extraction Kit (Qiagen), as per the manufacturer's standard protocol.

Table 2.2. PCR thermocycling conditions.

Step	Temperature (°C)	Duration
Initial Denaturation*	98	30 s
Annealing*	98	10 s
Elongation*	72	30 s
Final Elongation	72	10 min
Hold	4	∞

*These steps are repeated for 35 cycles.

2.1.4 Agarose Gel Electrophoresis

Agarose gel electrophoresis was used to monitor the success of molecular cloning and identify amplified PCR products and restriction digested gene fragments. Agarose gels were prepared by adding 1% (w/v) of agarose powder to 1x TAE buffer (Tris, acetic acid, ethylenediaminetetraacetic acid (EDTA)) and heating the mixture until the powder fully dissolved. SYBRTM Safe stain (Thermo Fisher Scientific) was added at a 1:10,000 v/v ratio before the gel had cooled and set. 6x gel loading dye was added to the samples at a 1:6 dilution before the samples were loaded onto the gel. Analyte DNA was separated by electrophoresis at 100 V for 20 minutes, followed by visualisation using a blue light transilluminator for DNA band excision, or imaged with UV light using a ChemiDocTM System (BioRad).

2.1.5 Restriction Enzyme Digestion

Purified PCR products, gene fragments, and vector DNA were prepared for ligation by digestion using restriction enzymes. Each restriction reaction was assembled as per the manufacturer's protocol with DNA samples (< 1 µg), restriction enzymes, and CutSmart® buffer (NEB). The most commonly used enzymes include AgeI-HF® (NEB), KpnI-HF®

(NEB), EcoRI-HF® (NEB), and XhoI® (NEB). The reaction mixture was incubated at 37°C for 1 h and the reaction was stopped by adding gel loading dye. The digested fragments were separated by agarose gel electrophoresis as described in section 2.1.4. The target DNA bands at correct molecular weights were excised and the DNA was purified using the QIAquick® Gel Extraction Kit (Qiagen).

2.1.6 DNA Ligation

Restriction digested DNA fragments/inserts and vectors were ligated using Quick Ligation™ Kit (NEB) following the manufacturer's Quick Ligation protocol. DNA inserts and vectors were mixed at a 3:1 molar ratio with Quick Ligase and Quick Ligase Reaction Buffer from the kit. Reactions were incubated for 15 min at room temperature (25°C) prior to transformation of competent *Escherichia coli* (*E. coli*) or storage at -20°C.

2.1.7 Transformation of Chemically-Competent *E. coli*

Following the ligation of target constructs into expression vectors, plasmids were transformed into Library Efficiency® DH5α™ competent *E. coli* (Invitrogen). 30 µL aliquots of competent cells (stored at -80°C before use) were gently thawed on ice and pipetted into sterile 1.5 mL Eppendorf tubes prior to supplementation with 3 µL of ligated plasmid DNA. When purified DNA of higher concentration (e.g. from miniprep) was used for transformation, the volume used was greatly reduced or the DNA was diluted to a maximum concentration of 10 ng. The cell-DNA mixture was incubated on ice for 30 mins before subjected to heat shock for 45 s at 42°C using a water bath. The cells were immediately cooled on ice for 2 min prior to the addition of 300 µL S.O.C. medium (super optimal broth; Invitrogen). The cells were subsequently transferred to a 50 mL Falcon tube and incubated at 37°C for 1 h in a shaking incubator (180 rpm) for recovery. 130 µL of cell cultures were plated onto Luria broth (LB)-

agar plates containing ampicillin at 100 µg/mL. The plates were incubated at 37°C for 18 h to enable selection of colonies that were successfully transformed with circularised vector, as the pHLsec and pHR-CMV-TetO₂ plasmids contain ampicillin resistance (Amp^R) gene.

2.1.8 Plasmid Amplification and Purification

A single colony from the LB-agar plate was used to inoculate 5 mL of LB medium in a Falcon tube supplemented with ampicillin at 100 µg/mL. The culture was incubated at 37°C for 18 h in a shaking incubator (180 rpm). The bacterial cells were subsequently harvested by centrifugation at 4000 x g for 8 min, supernatant discarded, and plasmid DNA was extracted using the QIAprep® Spin Miniprep Kit (Qiagen), as per the manufacturer's protocol. To confirm the integrity and quality of all ligated DNA inserts, 5 µL of purified plasmids at 80–100 ng/µL and 5 µL of primers at a concentration of 5 µM were mixed in Eppendorf tubes and sent for sequencing analysis (Eurofins genomics, previously GATC Biotech). Sequencing primers complementary to the regions immediately adjacent to the MCS in pHLsec and pHR-CMV-TetO₂ were used:

forward pHLsec sequencing primer: TTCCTACAGCTCCTGGGCAAC;

reverse pHLsec sequencing primer: CACCAGCCACCACCTTCTGATAG;

forward pHR sequencing primer: CGTCGTCGTGCTCGTTTAGTGAAC;

reverse pHR sequencing primer: CCAGAGGTTGATTATCGATACGTACGATATCAAG

Once the target DNA construct was confirmed by sequencing analysis and tested by small-scale transient protein expression (see section 2.3.1), *E. coli* transformation was repeated as previously described using the miniprep-purified plasmids. The colonies were subsequently inoculated into four 2 L flasks, each containing 500 mL LB medium supplemented with 100 µg/mL ampicillin. The cell cultures were incubated at 37°C for 18 h in a shaking incubator (180 rpm). Bacterial cells were pelleted by centrifugation at 4000 x g for 15 min. Plasmids

were purified from cell pellets using HiSpeed® Plasmid Giga EF Kits (Qiagen) following manufacturer's protocol. Purified plasmids (2–3 mg/mL) were stored at -20°C.

2.1.9 Gene Mutagenesis Using Splicing by Overlap Extension PCR

To introduce a site-directed mutation into the gene construct, overlap extension PCR (OE-PCR) was performed as described in (Heckman and Pease, 2007). Three separate PCR experiments (**Fig. 2.2**) involving four primers (two flanking primers and two mutagenic primers) were performed. PCR #1 using 5' flanking primer and the 3' mutagenic primer and PCR #2 using 5' mutagenic primer and 3' flanking primer produced intermediate PCR products that contained overlapping regions of the entire gene construct. PCR #3 used 5' and 3' flanking primers and PCR products from reactions #1 and #2 as DNA templates, producing a complete DNA containing the desired mutation. Subsequent cloning of the amplified gene proceeded as outlined in sections 2.1.4–2.1.8.

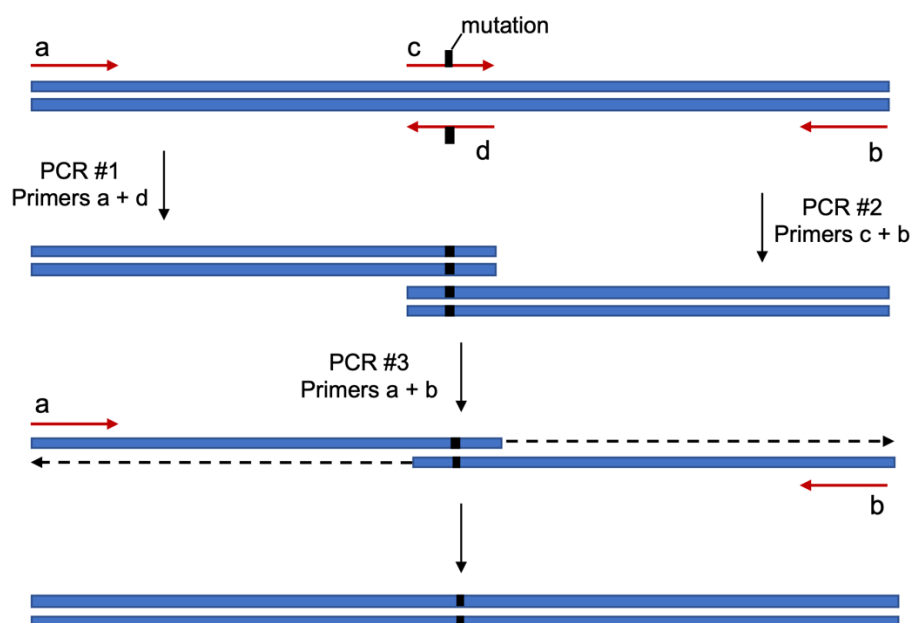


Fig 2.2. A schematic outline of site-directed mutagenesis accomplished by OE-PCR. The blue DNA template represents the gene of interest. The red arrows represent the flanking primers (a and d) and the mutagenic primers (b and c) containing the intended mutation (black bar).

2.2 MAMMALIAN TISSUE CULTURE

2.2.1 Cultivation of Mammalian Cell Lines

Cell culture was performed in a laminar flow cabinet using good microbiological practices. Cells were cultured and maintained in T175 flasks (Corning) at 37°C with 5% CO₂ until ~80–90% cell confluency was reached. To passage, medium was removed from the flasks and the cell monolayer were washed with 10 mL of phosphate buffered saline pH 7.4 (PBS, Gibco). Cell detachment from the flask was done through 5 min incubation of 5 mL of trypsin-EDTA (Gibco). To stop the trypsin reaction, 20 mL of cell culture medium was added to resuspend the cells. The resuspended cells (2–3 mL) were added to a new T175 flask and diluted with medium to a final volume of 25 mL. The passage steps were repeated again upon reaching ~80–90% cell confluency. The mammalian cell lines used are listed in **Table 2.3**.

Table 2.3. Summary of cell types used and their corresponding growth media.

Cells	Cell culture media
Human embryonic kidney (HEK) 293T	Dulbecco's Modified Eagle's Medium (DMEM, Sigma) 10% fetal calf serum (FCS, Sigma) 1% non-essential amino acids (NEAA, Gibco) 1% L-glutamine (Gibco)
Madin-Darby Canine Kidney II (MDCK-II)	DMEM (Sigma) 10% FCS (Sigma) 1% NEAA (Gibco) 1% L-glutamine (Gibco)
Baby hamster kidney (BHK) 21	Glasgow's Minimum Essential Medium (GMEM, Gibco) 10% FCS (Sigma) 1% NEAA (Gibco)

2.2.2 Production of Stable Cell Lines

Recombinant protein and virus-like particle (VLP) expression using stable cell lines is a cost and time effective method that has proven to be useful for increasing expression yield in a scalable and reliable way. All stable cell lines (**Table 2.4**) were produced using an established protocol involving lentiviral transduction, as described in Elegheert *et al.* 2018.

Table 2.4. Summary of the production and usage of stable cell lines.

Cells	Method	Use
HEK293T	Lentiviral transduction	Protein and VLP expression
MDCK-II		VLP expression

2.2.3 Storage and Revival of Cell Lines

When cells were not in use for protein and VLP production, they were stored at -80°C to ensure sustained viability. To prepare the cells for long-term storage, medium was removed from the flasks and cells were washed with PBS (Gibco) followed by 3 mL of trypsin-EDTA (Gibco) for detachment, as described in section 2.2.1. Cells were resuspended with 5 mL of appropriate medium supplemented with FCS. The cells were aliquoted into cryo-vials and DMSO was added as cryo-protectant to a final concentration of 10%. Vials were inserted into a CoolCell® container (Biocision) for a controlled freezing of the cells within the container.

To thaw and revive the cells, the cryo-vials were placed into a 37°C water bath up to maximum 1 min, and immediately followed by centrifugation at 1000 x g for 3 mins to separate the cell pellet from the medium containing trypsin-EDTA and DMSO. The cell pellet was resuspended in 20 mL of medium, transferred to T175 flasks, and allowed to grow at 37°C and 5% CO₂.

2.3 RECOMBINANT PROTEIN EXPRESSION AND PURIFICATION

2.3.1 Small-Scale Protein Expression

Prior to large-scale protein production, small-scale expression tests were performed to ensure that the designed protein constructs were optimal for expression. HEK293T cells were typically used for this step due to their rapid growth rate and the ability for enzymatically controlled asparagine-linked glycosylation processing (e.g. with the use of alpha-mannosidase inhibitor) (Aricescu et al., 2006; Elbein et al., 1990). Following cell passaging (section 2.2.1), 2mL of resuspended cells were diluted with 10mL of DMEM supplemented with 10% FCS, 1% NEAA, and 1% L-glutamine. The cells were subsequently inoculated into 24-well plates (0.5 mL per well) and allowed to grow to ~80% confluency overnight at 37°C and 5% CO₂.

Per well, 0.8 µg of purified DNA (e.g. from miniprep) was diluted in 50 µL of serum-free DMEM (1% NEAA, 1% L-glutamine) and 2 µL of lipofectamine® 2000 (Invitrogen) transfection agent was also diluted in separate 50 µL of serum-free DMEM in parallel. The DNA and lipofectamine were allowed to sit for 5 min before mixing to enable DNA-lipofectamine complex formation. The mixture was incubated for 20 min at room temperature. The medium was exchanged with 400 µL fresh DMEM (10% FCS, 1% NEAA, 1% L-glutamine) per well, and topped up with 100 µL of the DNA–lipofectamine mixture after 20 min. The steps above were repeated for two wells as positive (a construct known to express protein) and negative (no DNA added, only DMEM) controls. Cells were incubated at 37°C and 5% CO₂ for 72 h. Subsequently, the cell culture supernatants and dislodged cells were harvested and analysed using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and Western blots.

2.3.2 SDS-PAGE

The separation of protein samples according to molecular mass using SDS-PAGE allows the identification of target protein and monitoring the purity of protein samples. NuPAGE 4–12% Bis-Tris gels (Life Technologies) were used in a mini gel tank (Life Technologies) filled with 1x NuPAGE MES SDS Running Buffer. The samples were mixed with reducing loading dye (250 mM Tris-HCl pH 6.8, 10% SDS, 50% glycerol, 25% β -mercaptoethanol, 0.01% bromophenol-blue) or non-reducing loading dye (125 mM Tris-HCl pH 6.8, 20% glycerol, 4% SDS, 0.1% bromophenol-blue) before loading into the gel columns. BenchMark™ protein ladder (Invitrogen) was added to allow protein molecular weight estimation. If the gel was to be used for western blot, BenchMark™ pre-stained protein ladder (Invitrogen) was also added to allow visualisation of successful transfer. SDS-PAGE was performed by applying a 180 V current for 45 min (when the dye running front reached the end of the gel). The gel was removed from its plastic cast and stained with SimplyBlue™ SafeStain (Thermo Fisher Scientific) or InstantBlue™ Protein Stain (Expedeon), or used for western blot. Stained gels were visualised using the ChemiDoc™ System (BioRad).

2.3.3 Western Blot

The SDS-PAGE gel was assembled into a commercial iBlot™ 2 Transfer Stacks (Life Technologies) and blotted onto a nitrocellulose membrane using an iBlot™ 2 Gel Transfer Device (Life Technologies) per manufacturer's instructions. The preset manufacturer's protocol P0 (20 V for 1 min, 23 V for 4 min, and 25 V for 2 min) was used to transfer protein bands from SDS-PAGE to the nitrocellulose membrane.

The nitrocellulose membrane was incubated with 5% (w/v) milk powder in PBS for 45 min at room temperature. The membrane was then further incubated with primary antibody diluted in 5% (w/v) milk-PBS for 1 h at room temperature. The membrane was washed three

times with PBS for 15 min prior to incubation with a secondary antibody diluted in 5% (w/v) milk-PBS for 1 h at room temperature. Most commonly used primary antibodies (diluted 1:1000 v/v) included mouse-derived monoclonal anti-penta-histidine antibody (Qiagen, ID: 34660) and glycoprotein-specific neutralising antibodies. Most commonly used secondary antibody (diluted 1:2000 v/v) included goat-derived horseradish peroxidase (HRP)-conjugated anti-mouse IgG antibody (Sigma). The membrane was washed six times with PBS containing 0.05% (v/v) Tween20 (PBS-T) for 30 min. Clarity™ Western ECL substrate kit (Bio-Rad) was used for detection (horseradish peroxidase-luminol system) and the signals were visualised and captured using the ChemiDoc™ system (Bio-Rad).

2.3.4 Large-Scale Protein Expression

Once the expression of the target protein constructs had been confirmed by small-scale expression tests and analysed with SDS-PAGE and Western blot, large-scale expression of the target protein was performed. HEK293T cells were first cultured in T175 flasks as detailed in section 2.2.1 and subsequently transferred to 2,125 cm² area roller bottles (Cellmaster™, Greiner BioOne). Roller bottles were placed in roller cabinet at 37°C for 3–4 days for the cells to reach ~80–90% confluency.

Each roller bottle of cells was transfected with 0.5 mg of purified DNA (e.g. from Gigaprep). Prior to mixing the DNA with polyethylenimine (PEI, Polyscience Inc.) transfection agent, 10% (v/v) chloroform was added to the DNA and mixed vigorously to sterilise the DNA, followed by centrifugation for 10 min at 13,400 x g. The sterilised DNA was carefully pipetted off the upper layer of the mixture and added to 50 mL FCS-free DMEM (1% NEAA, 1% L-glutamine) and 1 mg of PEI. The mixture was incubated at room temperature for at least 10 min to allow DNA–PEI complex formation. The medium in the roller bottles was replaced with 200 mL of fresh DMEM (2% FCS). The DNA–PEI mixture was added into the roller bottles

in the presence of 5 μ M of kifunensine, a class 1 α -mannosidase inhibitor. Cells were incubated in a rolling cabinet at 37°C for 72–96 h for recombinant protein expression.

2.3.5 Cell Supernatant Harvesting and Diafiltration

Cell culture supernatants were harvested and centrifuged for 20 min at 4,000 x g to remove the cell debris. Supernatants were filtered through 0.22 μ m Stericup® filter units (Millipore) followed by diafiltration using the ÄKTA flux tangential flow filtration system (GE Healthcare), fitted with three 10 kDa molecular weight cut-off membranes. Supernatants were first concentrated to a smaller volume (~100–150 mL) and then buffer exchanged into 10 mM Tris pH 8.0, 150 mM NaCl to remove components in the FCS-supplemented DMEM medium that can strip nickel ions from nickel affinity columns.

2.3.6 Immobilised Metal Ion Affinity Chromatography (IMAC)

Target proteins containing C-terminal His₆ tag were purified using 5 mL HisTrap™ FF (GE Healthcare) nickel affinity columns that were prepacked with pre-charged Ni Sepharose 6 Fast Flow resin. ÄKTA liquid chromatography systems (GE Healthcare) were used throughout the purification steps. The columns were first washed with filtered deionised water to remove the 20% ethanol used for nickel resin storage, followed by equilibration with buffer containing 10 mM Tris pH 8.0, 150 mM NaCl for at least five column volumes. Concentrated and diafiltrated supernatants were then loaded onto the equilibrated column at a flow rate of 5 mL/min to allow the His₆-tagged target proteins to bind to the immobilised nickel ions. The flowthrough solution was collected for SDS-PAGE analysis to confirm that all target proteins were bound to the resin. To remove non-specific contaminants that were weakly bound to the resin, imidazole-containing wash buffer was passed through the columns in the manner of a sequential increase in imidazole concentration (20 mM imidazole for 10–15 min, 30 mM imidazole for 10–15 min,

40 mM imidazole for 10 min, and 50 mM imidazole for 10 min). Buffer containing 250 mM imidazole was used to elute target proteins from the resin.

2.3.7 Enzymatic Deglycosylation

Endoglycosidase F₁ (endoF₁) was added at a mass ratio of 1:100 (endoF₁:protein) to the IMAC-purified protein samples, and incubated for 18 h at room temperature. Glycoproteins expressed in the presence of kifunensine contain N-linked high-mannose glycan chains (GlcNAc₂Man₅₋₉). EndoF₁ cleaves these glycans, resulting in a single *N*-acetylglucosamine (GlcNAc) moiety linked to the asparagine residue. Deglycosylated protein samples were subsequently subjected to size exclusion chromatography (SEC) prior to crystallisation experiment.

2.3.8 Size Exclusion Chromatography (SEC)

SEC was performed to ensure high purity of protein samples, estimate molecular weights and oligomerisation state of the proteins, and for buffer exchange of samples. Superdex™ 200 Increase 10/300 GL column (Sigma) was connected to the ÄKTA liquid chromatography systems (GE Healthcare). The column was first washed with filtered and degassed deionised water to remove the 20% ethanol used for column storage, and then equilibrated with buffer containing 10 mM Tris pH 8.0, 150 mM NaCl. Protein samples were concentrated to 500 µL and filtered using Millipore™ Amicon™ centrifugal filter units (10 kDa or 30 kDa molecular weight cut-off) and Millipore™ Ultrafree 0.22 µm centrifugal filter units, respectively. The samples were then injected into the column via a 500 µL sample delivery loop. SEC was performed at a flow rate of 0.75 mL/min with a column pressure limit of 2.2 MPa. Sample solutions passing through the column were tracked by the ÄKTA UV absorbance unit (mAU) device, fractionated, and collected. Fractions that gave high mAU corresponding to the presence of proteins were assessed using SDS-PAGE and western blot analysis. Fractions

containing the target proteins were pooled and concentrated. Protein concentration was measured using Nanodrop 1000 Spectrophotometer (Thermo Fischer Scientific).

2.3.9 Enzyme-Linked Immunosorbent Assay (ELISA)

Antibodies JUN1–7 and MAC1 binding: High binding ELISA 96 half-well microplates (Corning) were coated with purified JUNV GP1 or MACV GP1 (25 μ L, 3 μ g/mL in PBS) overnight at 4°C. Plates were washed five times with PBS containing 0.05% Tween-20 (PBS-T) and blocked with blocking solution (5% non-fat milk in PBS-T) for 1 h at room temperature. The blocking solution was removed and serially diluted antibodies (starting at 10 μ g/mL, 1:5 dilution in blocking solution) was added for 2 h at room temperature. Plates were washed five times with PBS-T. Secondary antibody (goat anti-mouse IgG Fc, biotin-conjugate, Thermo Fisher Scientific, 1:1,000) was added for 30 min and plates were washed as described above before alkaline phosphatase (AP)-conjugated streptavidin (1:10,000, Invitrogen) was added for 30 min. Following a final wash, p-nitrophenyl phosphate substrate (Sigma) was added to detect binding and the optical densities (ODs) were measured at 405 nm.

Antibodies JUN1–7 binding competition: 96-half well ELISA microplates (Corning) were coated with purified JUNV GP1, washed with PBS-T, and blocked with blocking solution, as described above. The blocking solution was removed and serial diluted Fab JUN1 (at 50 μ g/mL, 1:5 dilution in blocking solution) was added for 30 min at room temperature. IgG JUN1–7 and OD01 were added at the EC₈₀ concentration and incubated for a further 1.5 h. Plates were washed five times with PBS-T. Secondary antibody (goat anti-mouse IgG Fc biotin-conjugate, Thermo Fisher Scientific, 1:1,000) was added for 30 min. Plates were washed as described above and streptavidin-AP (1:10,000, Invitrogen) was added for 30 min. Following a final wash, p-nitrophenyl phosphate substrate (Sigma) was added to detect binding and the ODs were

measured at 405 nm. The competition is reported for JUN1 Fab at 50 µg/mL and is reported relative to the % competition measured for JUN1 IgG.

Antibodies 17C7, RVC20, and 1112-1 binding competition: 96-half well ELISA microplate was coated with purified RV-G (50 µL, 1 µg/mL in PBS), washed with PBS-T, and blocked with blocking solution, as described above. The blocking solution was removed and immobilised RV-G was saturated with Fab 17C7 (50 µL, 20 µg/mL in PBS) for 1 h at room temperature. Fab 17C7 was removed and serial diluted IgG 1112-1 (at 0.6 µg/mL, 1:6 dilution in blocking solution) was added for 1 h at room temperature. Plate was washed five times with PBS-T and secondary antibody (goat anti-mouse IgG Fc, AP-conjugate, Thermo Fisher Scientific, 1:1000) was added for 1 h at room temperature. Following a final wash, p-nitrophenyl phosphate substrate (Sigma) was added to detect binding and the ODs were measured at 405 nm. This experiment was repeated with IgG 1112-1 alone and with Fab RVC20 replacing Fab 17C7.

2.4 VIRUS-LIKE PARTICLE (VLP) PRODUCTION AND PURIFICATION

2.4.1 VLP Stable Cell Line Culture

MDCK-II cells (kindly provided by Dr. Thomas Strecker, Philipps University of Marburg) stably expressing LASV glycoprotein (GP) complex were cultured in 25 mL of DMEM supplemented with 10% FCS, 1% NEAA, and 1% L-glutamine at 37°C and 5% CO₂ until ~80–90% cell confluency was reached (~72 h) (as described in section 2.2.1). To start VLP expression, the medium was discarded, and cells were washed with 10 mL of phosphate buffered saline pH 7.4 (PBS, Gibco). 20–25 mL of fresh DMEM supplemented with 2% FCS was then added to the cells and incubated at 37°C and 5% CO₂ for 96 h for VLP expression.

2.4.2 VLP Purification

After 96 h, the cell culture supernatant was cleared by centrifugation at 3,000 x g, 4°C for 30 min. The supernatant was aliquoted into 25 x 89 mm Open-Top Polyclear™ Centrifuge Tubes (Seton Scientific). Prechilled buffer (6 mM succinate, 22 mM sodium dihydrogen phosphate, 22 mM glycine, 100 mM NaCl, pH 7.5; SPG50N100) containing 20% sucrose was added to the bottom of the supernatant slowly using a blunt pipetting needle (Sigma) to avoid mixing of the two solutions. VLPs were pelleted through 20% sucrose cushion by ultracentrifugation at 27,000 x g for 3 h at 4°C. The supernatant and sucrose buffer were discarded by gentle pipetting, taking extra care to not disturb the VLP pellets. Pellets were resuspended in 100 µL SPG50N100 buffer for at least 24 h at 4°C.

2.5 STRUCTURAL BIOLOGY METHODS

2.5.1 Protein Crystallisation

Crystallisation trials were performed using the full range of crystallisation screens available at the Division of Structural Biology (STRUBI), University of Oxford, obtained from Hampton Research, Emerald BioSystems, and Molecular Dimensions Limited. Typically, trials were performed in a nanolitre scale sitting drop vapour diffusion format by mixing 100 nL reservoir solution with 100 nL protein solution in the central shelves of 96-well Greiner CrystalQuick® plates using a crystallisation robot (Cartesian Technologies). Reservoirs contained 95 µL of mother liquor. Plates were immediately sealed with plastic films and stored at 20°C (Formulatrix® system) to allow crystal formation.

Crystals formed from initial crystallisation conditions were optimised by varying precipitant concentration, protein concentration, and through the addition of additives (Hampton Research) to the reservoir or protein–precipitant solutions. The specific crystallisation experiments and conditions are outlined in the relevant result chapters.

2.5.2 Cryo-Protection and Flash Cooling of Protein Crystals

Plastic films directly above the wells containing protein crystals were carefully removed to allow access. To make crystal harvesting easier from a larger volume, 1 μ L of mother liquor solutions was carefully pipetted onto the crystallisation drops. Mother liquor solutions supplemented with cryo-protectant such as 25% (v/v) glycerol or ethylene glycol (prepared in advance) was pipetted into the adjacent shelf. Protein crystals were subsequently harvested from the drop using LitholoopsTM and dipped into the cryo-protectant solution. Crystals were rapidly flash cooled by plunging into liquid nitrogen.

2.5.3 X-Ray Data Collection and Processing

All crystal diffraction data were collected at Diamond Light Source, Didcot, UK. Data were indexed, integrated, and scaled with XIA2 (Winter et al., 2013). Crystal structures were phased by molecular replacement with PHASER (McCoy et al., 2007). For all structures, iterative rounds of model building and refinement were performed using COOT (Emsley and Cowtan, 2004) and PHENIX (phenix.refine with translation-libration-screw (TLS) parameterisation) (Adams et al., 2002), respectively, to improve electron density maps. R_{work} and R_{free} values which describe the agreement between the model and the data were used to assess the refinement progress. MolProbity was used to validate final model quality (Chen et al., 2010). The PyMOL Molecular Graphics System (<https://www.schrodinger.com/pymol>) was used to analyse and generate the structural models presented in the figures. Specific information about data collection, processing strategy, and refinement statistics for each crystal structure are detailed in the relevant result chapters and crystallographic tables.

2.5.4 Cryogenic Electron Microscopy

A standard cryogenic electron microscopy (cryo-EM) workflow starts with grid preparation (section **2.5.4.1**), followed by sample screening, leading to high resolution cryo-EM data collection and structure analysis (section **2.5.4.2**) (Passmore and Russo, 2016). Cryo-EM grid preparation was performed using Vitrobot (FEI/Thermo Fisher Scientific). Cryo-EM imaging was performed using the 200 kV Glacios and 300 kV Titan Krios microscopes (FEI/Thermo Fisher Scientific) equipped with Falcon III (FEI/Thermo Fisher Scientific) and K2 Summit® (Gatan) direct electron counting detection cameras, respectively. Aforementioned infrastructures are located at the Oxford Particle Imaging Centre (OPIC), STRUBI.

2.5.4.1 Grid Preparation

Cryo-EM grids used included Quantifoil® Holey Carbon Film R 1.2/1.3 Copper 200 mesh, Quantifoil® Holey Carbon Film R 2/1 Copper 200 mesh, and C-Flat™ Holey Carbon Film R 2/2 Copper 200 mesh. Grids were glow-discharged for 30–35 s at the highest intensity in a plasma glow discharger (Harrick Plasma) and used within 1 h. The climate chamber of the Vitrobot (FEI/Thermo Fisher Scientific) was pre-cooled to 4°C and maintained at 100% humidity. Typically, 3.5–4.0 µL of samples were deposited onto the grids held firmly by a tweezer within the chamber. Excess sample was blotted from the grids with filter papers (Agar Scientific) fixed on rotating foam pads for 3.5–4.0 s at -15 N blotting pressure to create a thin layer of protein suspension on the grid. The grids were rapidly frozen (plunge freezing) in liquid ethane in a copper cup that was insulated and chilled by a reservoir of liquid nitrogen in an outer container. The grids were transferred and stored in cryo grid boxes (Euro-Techniek) in liquid nitrogen dewar. The cryo-EM grids were clipped into AutoGrids (Thermo Fisher Scientific) prior to sample screening and data collection on the Glacios and Titan Krios microscopes.

2.5.4.2 Data Collection and Processing

Samples on cryo grids were first screened and evaluated using the in-house 200 kV Glacios microscope (FEI/Thermo Fisher Scientific) equipped with a Falcon III direct electron counting detection camera and EPU, a SPA software for cryo-TEM sample screening and data acquisition. Typically, an initial data set was collected for 18–20 h for preliminary data analysis. Cryo-EM images were recorded using linear mode, at defocus values ranging from -2.0 to -3.5 μm at a calibrated magnification of 92,000x, corresponding to a pixel (px) size of 1.6 $\text{\AA}/\text{px}$. The total electron dose over 40 frames was $60\text{ e}^-/\text{\AA}^2$.

Once suitable cryo-EM grids had been identified, data acquisition was performed for ~66 h using the in-house 300 kV Titan Krios microscope (FEI/Thermo Fisher Scientific) equipped with a K2 direct electron detector and a 20 kV Gatan Imaging Filter (GIF) energy filter to obtain a large high-resolution data set. SerialEM software was used to record the cryo-EM images. Cryo-EM images were recorded using electron counting mode, at defocus values ranging from -0.8 to -2.6 μm at a calibrated magnification of 130,000x, corresponding to a pixel size of 1.05 $\text{\AA}/\text{px}$. The total electron dose was $45\text{ e}^-/\text{\AA}^2$ over 40 frames.

Cryo-EM SPA data processing was performed using cryoSPARC v2 and v3 software (Punjani et al., 2017). Averaged micrographs were generated using global movie alignment. Contrast transfer function (CTF) parameters were estimated using global CTF estimation in cryoSPARC. For reference-free particle picking, blobs with a diameter of 80–120 px were used. The picked particles were extracted from the micrographs with box sizes of 260–300 px and subsequently subjected to 2D classification. Typically, 100–200 2D classes were generated. 2D classes that did not contain clearly defined protein features were discarded while 2D classes displaying distinctive secondary structure features and varied particle orientations were selected as templates for reference-based automatic particle picking. Picked particles were extracted and subjected to 2D classification. Particles from the best 2D classes that displayed

strong signal-to-noise ratio, clear protein features such as α -helices, and varied orientations were used for *ab initio* 3D model reconstruction. 3–5 initial models were generated, which served as templates for 3D classification (heterogeneous refinement) and non-uniform refinement, during which suitable symmetry was applied. Map resolution was estimated by gold-standard Fourier shell correlation (FSC) plot at the 0.143 threshold. To improve map resolution and quality, local (per-particle) CTF refinement and motion correction were performed. The model figures were generated using UCSF Chimera software.

2.6 STRUCTURE-BASED PHYLOGENETIC ANALYSIS (SBPA)

To quantitatively investigate the structural relationships among mammarenaviral GP1 receptor-binding glycoproteins, structures of available arenaviral GP1 were used to construct a structure-based phylogenetic tree. Structures used were as follows: LORV, Loei River virus (PDB ID: 6HJ6) (Pryce et al., 2018); LASV, Lassa virus (4ZJF and 5VK2) (Cohen-Dvashi et al., 2015; Hastie et al., 2017); MORV, Morogoro virus (5NFF) (Israeli et al., 2017); LCMV, Lymphocytic choriomeningitis virus (5INE) (Hastie et al., 2016a); LUJV, Lujo virus (6GH8) (Cohen-Dvashi et al., 2018); WWAV, Whitewater Arroyo virus (6HJ5) (Shimon et al., 2017); MACV, Machupo virus (2WFO) (Bowden et al., 2009); and JUNV, Junín virus (5NUZ) (Zeltina et al., 2017). Water molecules, ligands, and protein chains (e.g. GP2, receptor, and antibody fragments) not comprising GP1 molecules were removed prior to structural alignment. GP1 structures were applied to the Structural Homology Program (SHP) (Stuart et al., 1979) to create a pairwise evolutionary distance matrix which was subsequently displayed as an unrooted phylogenetic tree using PHYLogeny Inference Package (PHYLIP) (Felsenstein, 1989). SBPA results are discussed in section 3.4.3.

3 DIFFERENTIAL CONFORMATIONS OF OLD AND NEW WORLD ARENAVIRAL GP1 GLYCOPROTEINS

3.1 INTRODUCTION

Reported in 2016, LORV is a novel mammarenavirus species isolated from bandicoot rats (*Bandicota indica* and *B. savilei*) and nivistenter (*Nivistenter fulvescens*) in Loei, Thailand (Blasdell et al., 2016). With the exception of LCMV that is primarily transmitted by house mouse (*Mus musculus*) worldwide, the known distribution of mammarenaviruses in Asia was previously limited to WENV (Li et al., 2015) and RYKV (Wu et al., 2018) discovered in Wenzhou and Yunnan Province, China, respectively. Together with WENV and RYKV, LORV is one of the only three known mammarenaviruses with Asiatic origin reported to date.

Metagenomic studies and sequence phylogenetic analysis of LORV revealed that the virus belongs to the OW serogroup and is closely related to WENV and Cardamones (a WENV variant identified in Veal Renh, Cambodia) (**Fig. 3.1**) (Blasdell et al., 2016). However, unlike WENV and Cardamones that have been shown to infect humans and cause febrile illnesses, zoonotic potential of LORV has not yet been confirmed (Blasdell et al., 2016) and therefore, is a cause of concern as it may constitute an emerging human pathogen.

Structural analysis of the OW mammarenavirus GP has, thus far, been limited to those displayed by arenaviruses circulating in the African countries (Cohen-Dvashi et al., 2015; Hastie et al., 2016a; Hastie et al., 2017; Israeli et al., 2017). The absence of structural information on the Asiatic mammarenaviral GP hinders our understanding of structural conservation, receptor tropism, and zoonotic potential of WENV, RYKV, and LORV, when compared to the characterised and established OW mammarenaviruses such as LASV and LCMV. Therefore, this chapter aims to elucidate the first structure of the receptor-binding GP1 from an Asiatic mammarenavirus, LORV. The data allow a detailed comparison with existing

OW GP1 structures and provide a structure-based prediction of LORV receptor tropism. On a broader level, this work defines the discrete conformational classes assumed by arenaviral GP1 glycoproteins.

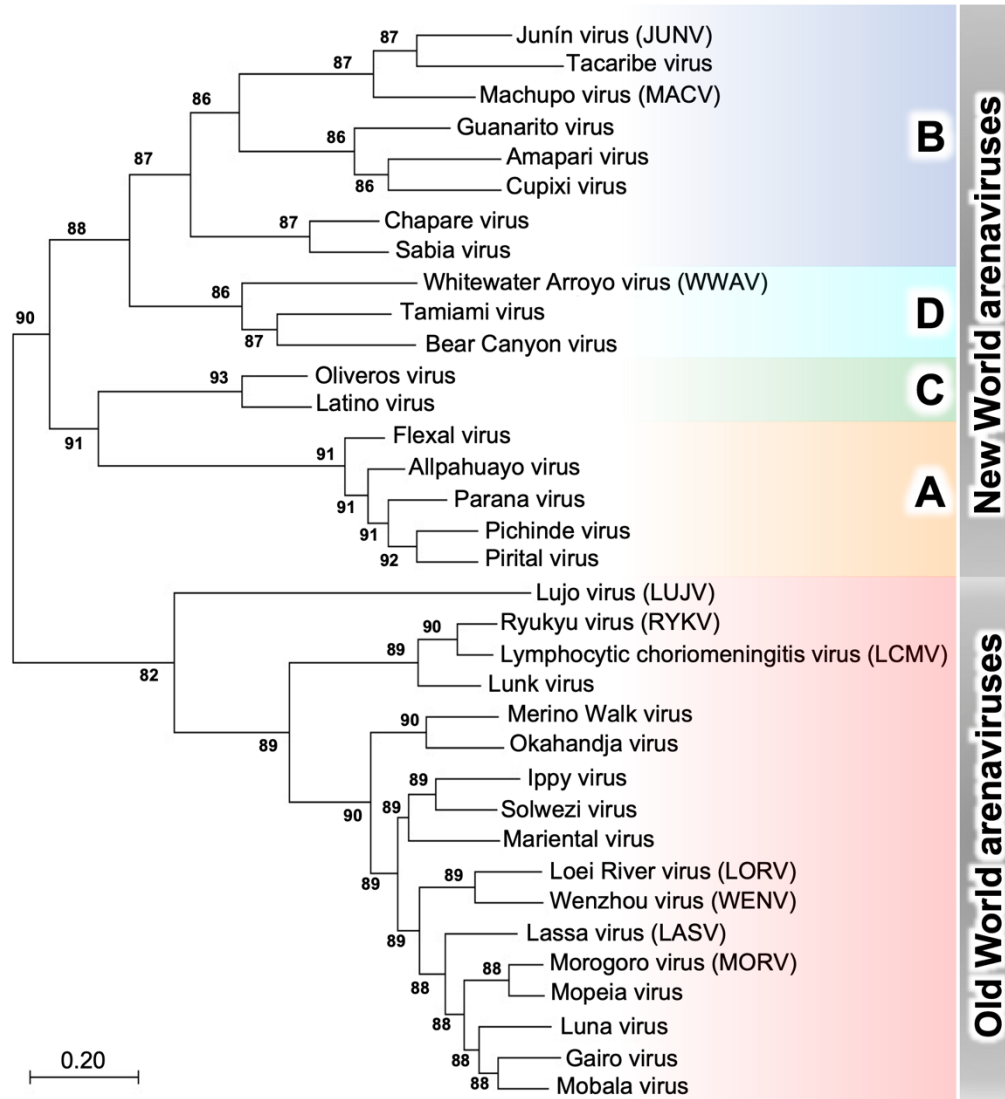


Fig 3.1. A maximum-likelihood phylogeny of 35 arenaviral GP amino acid sequences dividing the genus into OW and NW (clades A to D) serogroups. An evolutionary history was inferred using the maximum-likelihood method based on the model of Le and Gascuel (Le and Gascuel, 2008). OW serogroup is highlighted in red and NW clades A to D are highlighted in orange, green, cyan, and dark blue, respectively. The tree is drawn to scale, with branch length measured in the number of amino acid substitutions per site. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Evolutionary analyses were conducted in MEGA X (Kumar et al., 2018). Abbreviations for mammarenaviruses frequently mentioned in this chapter are shown to allow convenient reference. [Note that this figure has also been reported in Pryce *et al.* 2018 (listed in **Appendix 8.2**) containing work carried out as part of this thesis.]

3.2 PROTEIN PRODUCTION AND STRUCTURE DETERMINATION OF OLD WORLD LOEI RIVER VIRUS (LORV) GP1

3.2.1 Construct Design and Small-Scale Transient Expression

LORV GP1 constructs were generated from a human codon-optimised gene of the full length LORV GP (GenBank accession number AHE76159.1). The LORV GP1 subunit boundaries were estimated by sequence alignment (**Fig. 3.2A**) with the amino acid sequences of structurally characterised OW arenaviral GP1s including LASV GP1 (PDB ID: 4ZJF and 5VK2) (Cohen-Dvashi et al., 2015; Hastie et al., 2017), LCMV GP1 (PDB ID: 5INE) (Hastie et al., 2016a), and MORV GP1 (PDB ID: 5NFF) (Israeli et al., 2017). Six constructs of slightly differing lengths were designed (**Fig. 3.2B**), cloned into pHLsec expression vector (between AgeI and KpnI restriction sites) (Aricescu et al., 2006) using construct-specific primers, and tested for small-scale transient expressions in HEK293T cells.

Five out of the six constructs expressed and secreted LORV GP1, as detected by Western blot (**Fig. 3.2C**) of the supernatant of the cell growth media. However, only three constructs displayed thick and intense bands on the Western blot, a visual judgement that often indicates a satisfactory protein expression yield. Furthermore, the N-terminus (~20 residues) of the GP1 is observed to interact with the GP2, as observed in the crystal structures of LCMV and LASV GP1–GP2 (Hastie et al., 2016a; Hastie et al., 2017). Here, in the absence of GP2, the N-terminus of GP1 is likely flexible and therefore, removed from the expression construct. Finally, the construct (Met80–Thr238) with the least number of putative N-linked glycosylation sites was chosen from amongst the three for large-scale transient protein expression, as the presence of more flexible N-linked glycans could impede crystallogenesis.

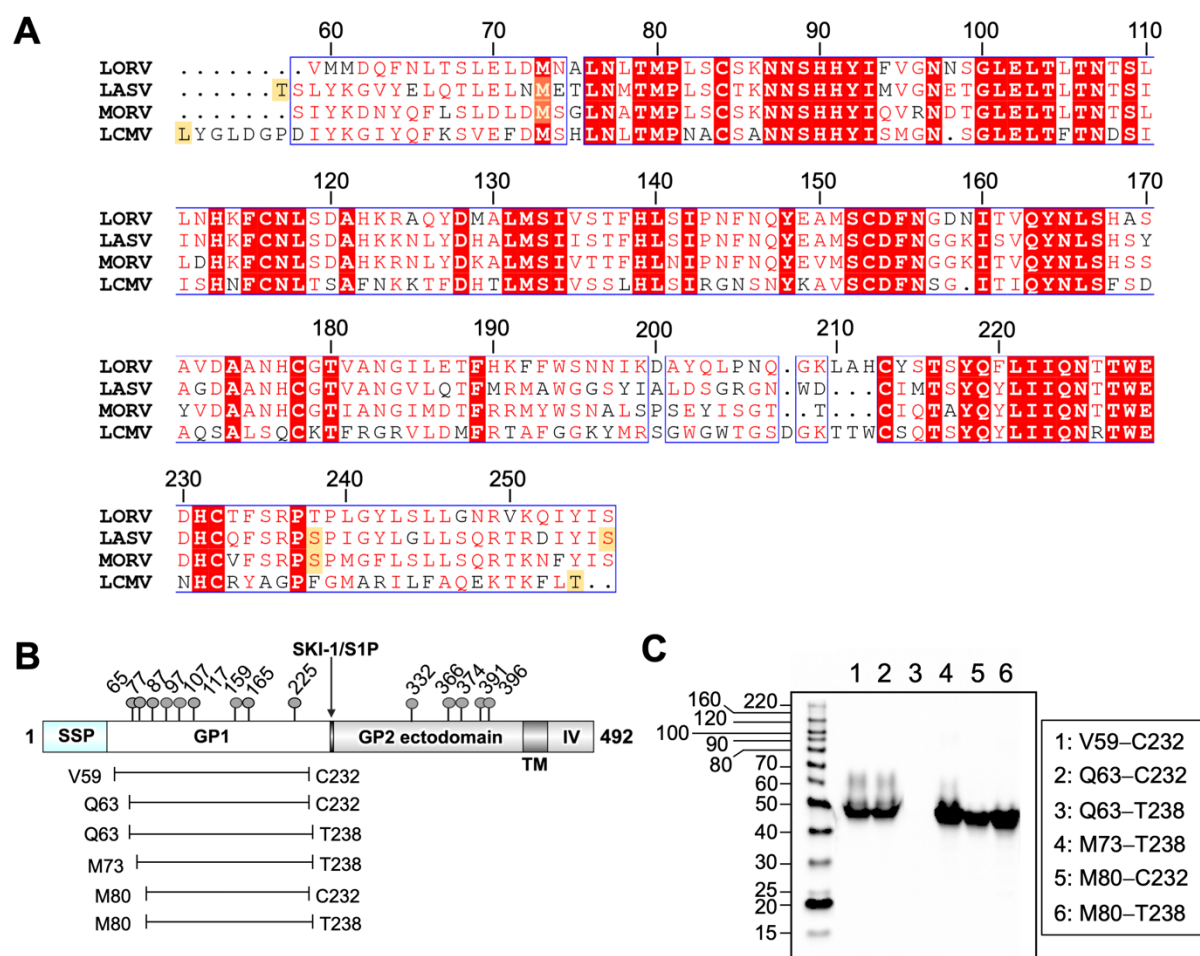


Fig 3.2. Protein construct design and expression test of LORV GP1. (A) Sequence alignment of LORV GP1 with other OW arenaviral GP1s that have been structurally characterised (LASV, MORV, and LCMV). Construct boundaries for LASV, MORV, and LCMV are highlighted with yellow boxes. Note that the structures of LASV GP1 have been characterised both with and without its cognate GP2 (Cohen-Dvashi et al., 2015; Hastie et al., 2017) while the structure of LCMV GP1 was characterised in complex with its GP2 subunit (Hastie et al., 2016a). Sequence numbering of LORV GP1 is used for annotation. Sequence alignment was determined with MultAlin (Corpet, 1988) and plotted with ESPript (Robert and Gouet, 2014). (B) Schematic organisation of LORV GP (generated with DOG (Ren et al., 2009)). The stable signal peptide (SSP), GP1 glycoprotein, subtilisin-like kexin protease 1-site 1 protease (SKI-1/S1P) cleavage site, GP2 glycoprotein, transmembrane region (TM), and intravirion domain (IV) are annotated. Putative N-linked glycosylation sequons (NXT/S, where X $\neq$ P) are labelled as pins above the schematic. The designed constructs are illustrated below the schematic. (C) Western blot of small-scale expression test of constructs tagged with a C-terminal His₆ tag, where an SDS-PAGE gel of cell supernatants was performed under reducing conditions.

3.2.2 Large-Scale Transient Expression and Purification

Large-scale transient expression of LORV GP1 (Met80–Thr238) was performed in the presence of kifunensine, a class 1 α -mannosidase inhibitor that prevents processing to complex- and hybrid-type N-linked glycans (Elbein et al., 1990). Cell supernatant was harvested and clarified 96 h post-transfection. Supernatant was concentrated and dialysed into purification buffer (10 mM Tris pH 8.0, 150 mM NaCl). Diafiltrated supernatant was purified with IMAC using HisTrapTM FF (GE Healthcare) nickel affinity columns (**Fig. 3.3A**). The sample was further purified using SEC (**Fig. 3.3B**). Purified LORV GP1 was added with endoF₁ (18 h at room temperature or 37°C for 2 h), truncating the mannose-rich N-glycans (GlcNAc₂Man₅₋₉) to yield a single GlcNAc moiety remaining on the asparagine residue (Elbein et al., 1990). SEC was performed on the deglycosylated sample to remove cleaved sugar molecules and endoF₁ enzymes, obtaining pure deglycosylated LORV GP1 for the protein crystallisation experiment. Protein purity and deglycosylation were monitored using SDS-PAGE throughout protein purification (**Fig. 3.3C**).

3.2.3 LORV GP1 Crystallisation Experiment

Purified deglycosylated LORV GP1 was concentrated to 4.5 mg/mL and 5.3 mg/mL, and used to set up a number of crystallisation trial screens (Hampton Research). Crystallisation experiments were performed at 20°C using the sitting-drop vapour diffusion method. Crystals of LORV GP1 grew under two conditions (**Fig. 3.3D**): (i) 4.5 mg/ml protein, 30% (w/v) polyethylene glycol (PEG) 6000, and 0.1 M citrate, pH 5.0, and (ii) 5.3 mg/ml protein, 10% (w/v) ethanol, and 1.5 M NaCl (buffered with 10 mM Tris, pH 8.0, from the protein solution). In both instances, crystals were cryoprotected by transfer into a solution of the respective precipitant supplemented with 25% (v/v) glycerol prior to flash cooling in liquid nitrogen. Crystal freezing was performed by Dr. Karl Harlos at STRUBI.

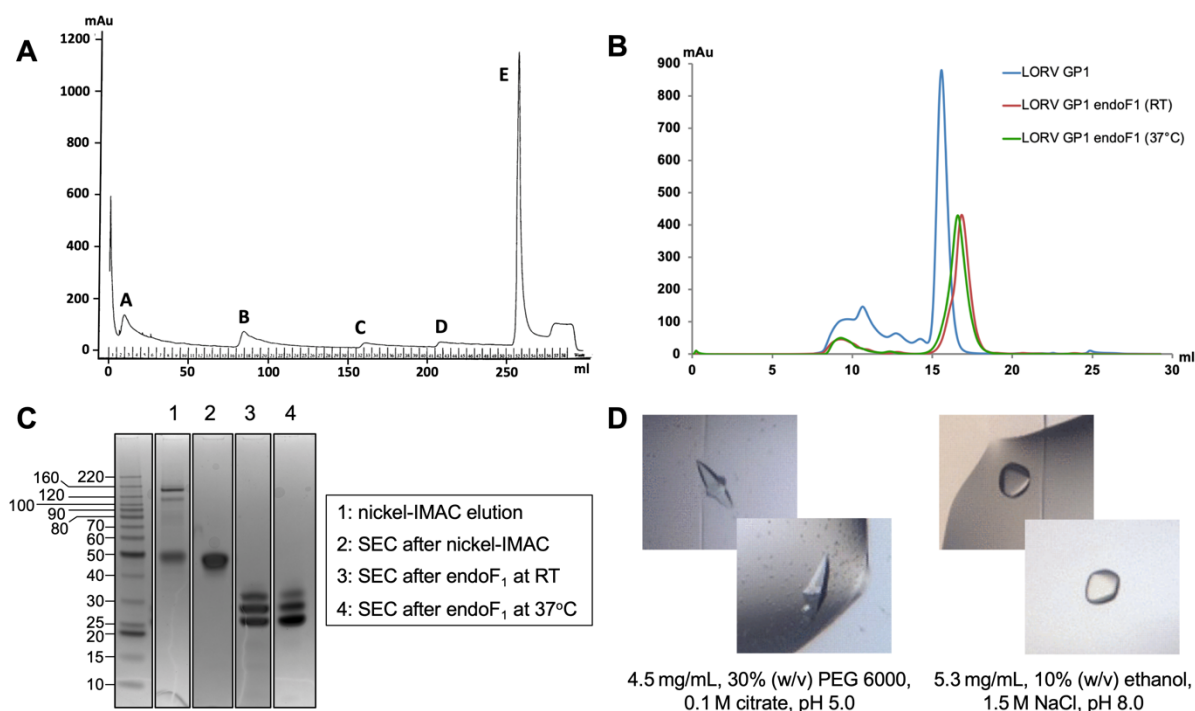


Fig 3.3. Purification and crystallisation of LORV GP1. (A) Nickel-affinity chromatogram of LORV GP1. HisTrapTM column was loaded with dialysed and filtered supernatant containing LORV GP1, washed with buffer containing low imidazole concentration (peak A–D corresponding to 20-, 30-, 40-, and 50-mM imidazole) prior to elution with 250 mM of imidazole (peak E). (B) Size exclusion chromatograms of LORV GP1 following nickel-affinity chromatography. Following SEC (SuperdexTM 200 Increase 10/300 GL column (Sigma)), GP1 was deglycosylated with endoF₁ at room temperature or at 37°C, after which the samples were subjected to SEC again. The reduction of molecular mass after GP1 deglycosylation was tracked by following the chromatogram peak shift (rightward). (C) Reducing SDS-PAGE analysis of LORV GP1 purification, indicating glycosylated GP1 at ~50 kDa. Successful deglycosylation of the GP1 was shown by the downward shift of the protein bands to 25–30 kDa. The presence of more than one band indicates that only partial deglycosylation was achieved. (D) Images of protein crystals. Distinct crystal forms were obtained from two crystallisation conditions at basic and acidic pH, respectively, as annotated.

3.2.4 X-Ray Diffraction Data Collection and Processing

X-ray diffraction data were recorded at Diamond Light Source (DLS), United Kingdom. Crystal data for LORV GP1 obtained at pH 8.0 and pH 5.0 were indexed, integrated, and scaled with XIA2 (3dii) (Winter et al., 2013) to 2.5- and 2.0-Å resolution, respectively. Both structures of LORV GP1 were solved by molecular replacement with PHASER (McCoy et al., 2007), using LASV GP1 (PDB ID: 4ZJF (Cohen-Dvashi et al., 2015)) as a search model. For both

structures, iterative rounds of model building and refinement were performed using COOT (Emsley and Cowtan, 2004) and PHENIX (Adams et al., 2002) with TLS parameterisation, respectively. MolProbity was used to validate model quality (Adams et al., 2002). Data collection and refinement statistics are presented in **Table 3.1**. Coordinates and structure factors of LORV GP1, crystallised at basic and acidic pHs, have been deposited in the PDB with the accession codes 6HJC and 6HJ6, respectively. The PyMOL Molecular Graphics System (<https://www.schrodinger.com/pymol>) was used to generate the structural models presented in the figures.

3.3 STRUCTURE ANALYSIS

3.3.1 Structure Overview of LORV GP1

LORV GP1 crystals were generated under both basic (pH 8.0) and acidic (pH 5.0) conditions, and X-ray diffraction data were collected to 2.5- and 2.0-Å resolution, respectively (**Table 3.1**). Basic- and acidic-pH-derived LORV GP1 structures were solved by molecular replacement, using the crystal structure of monomeric (GP2-free) LASV GP1 (PDB ID: 4ZJF) (Cohen-Dvashi et al., 2015) as a search model. Structural overlay analysis revealed that the two LORV GP1 structures are highly similar (0.7-Å RMSD), supportive that isolated OW arenaviral GP1s adopt the same conformation at basic and acidic pH (**Fig. 3.4A** and **3.4B**).

Table 3.1. Crystallographic data collection and refinement statistics for LORV GP1.

Data Collection Statistics	LORV GP1 (pH 5.0)	LORV GP1 (pH 8.0)
Beamline	DLS I03	DLS I03
Wavelength (Å)	0.9763	0.9763
Space Group	$P4_32_12$	$P4_12_12$
Cell dimensions <i>a</i> , <i>b</i> , <i>c</i> (Å)	57.3, 57.3, 113.2	60.5, 60.5, 96.4
α , β , γ (°)	90, 90, 90	90, 90, 90
Resolution range (Å)	40–1.98 (2.01–1.98) ^a	39–2.51 (2.55–2.51)
Rmerge	0.062 (>1)	0.112 (>1)
<i>I</i> /σ (<i>I</i>)	22.3 (1.5)	15.4 (1.4)
CC _{1/2}	1.000 (0.585)	0.999 (0.522)
Completeness (%)	100 (97.9)	99.2 (94.9)
Multiplicity	15.2 (5.5)	11.8 (3.7)
Refinement Statistics		
Resolution (Å)	33–1.98	39–2.51
No. reflections	13,746	6,480
<i>R</i> _{work} / <i>R</i> _{free}	0.189/0.237	0.208/0.238
No. atoms		
Protein	1,268	1,260
Ligand/ion	6	6
Water	70	14
Glycan	70	56
Average B-factors		
Protein	56.2	71.9
Ligand	87.2	103.1
Water	61.9	69.0
Glycan	87.1	103.2
Ramachandran (%)		
Favoured	98.7	96.8
Allowed	1.3	3.2
Outlier	0	0
Root mean square deviations (RMSD)		
Bond lengths (Å)	0.010	0.002
Bond angles (°)	0.96	0.42

^a The value for the highest-resolution shell is shown in parentheses.

LORV GP1 presents a compact α/β fold comprising seven β -strands, five of which (designated $\beta 1$, $\beta 2$, $\beta 3$, $\beta 4$, and $\beta 7$) comprise the central β -sheet that is sandwiched between three α -helices ($\alpha 1$ to $\alpha 3$) and connecting loops (**Fig. 3.4A**). The remaining two antiparallel β -strands ($\beta 5$ and $\beta 6$) slots in between helices $\alpha 2$ and $\alpha 3$ on one side of the central β -sheet while the N and C termini extend in the same direction on the other side. The secondary structures are stabilised by three disulfide bonds, one of which bridges the loops leading to the N and C termini, with another stabilising loop 2 and strand $\beta 3$, and the third linking $\alpha 2$ and $\beta 6$ (**Fig. 3.4A**). All three disulfide bonds are fully conserved in other OW mammarenaviruses, including the structurally characterised LASV, LCMV, and MORV GP1s (see **section 3.3.2**).

Five N-linked glycosylation sites (Asn87, Asn107, Asn159, Asn165, and Asn225) observed in the crystal structures lie on the periphery of the central β -sheet (**Fig. 3.4A**). Although multiple protein bands were observed in the reducing SDS-PAGE gel (**Fig. 3.3C**) following GP1 deglycosylation with endoF₁, electron density corresponding to only one ordered GlcNAc moiety was observed at each site (**Fig. 3.4A**), suggesting that the remaining untruncated glycan residues were protected by the proteinaceous environment from enzymatic deglycosylation and may be disordered in the crystal. LORV GP1–glycan interactions were observed for all five sites. For example, GlcNAc residues at Asn87, Asn107, Asn165, and Asn225 form hydrogen bonds with LORV GP1 (**Fig. 3.4C**).

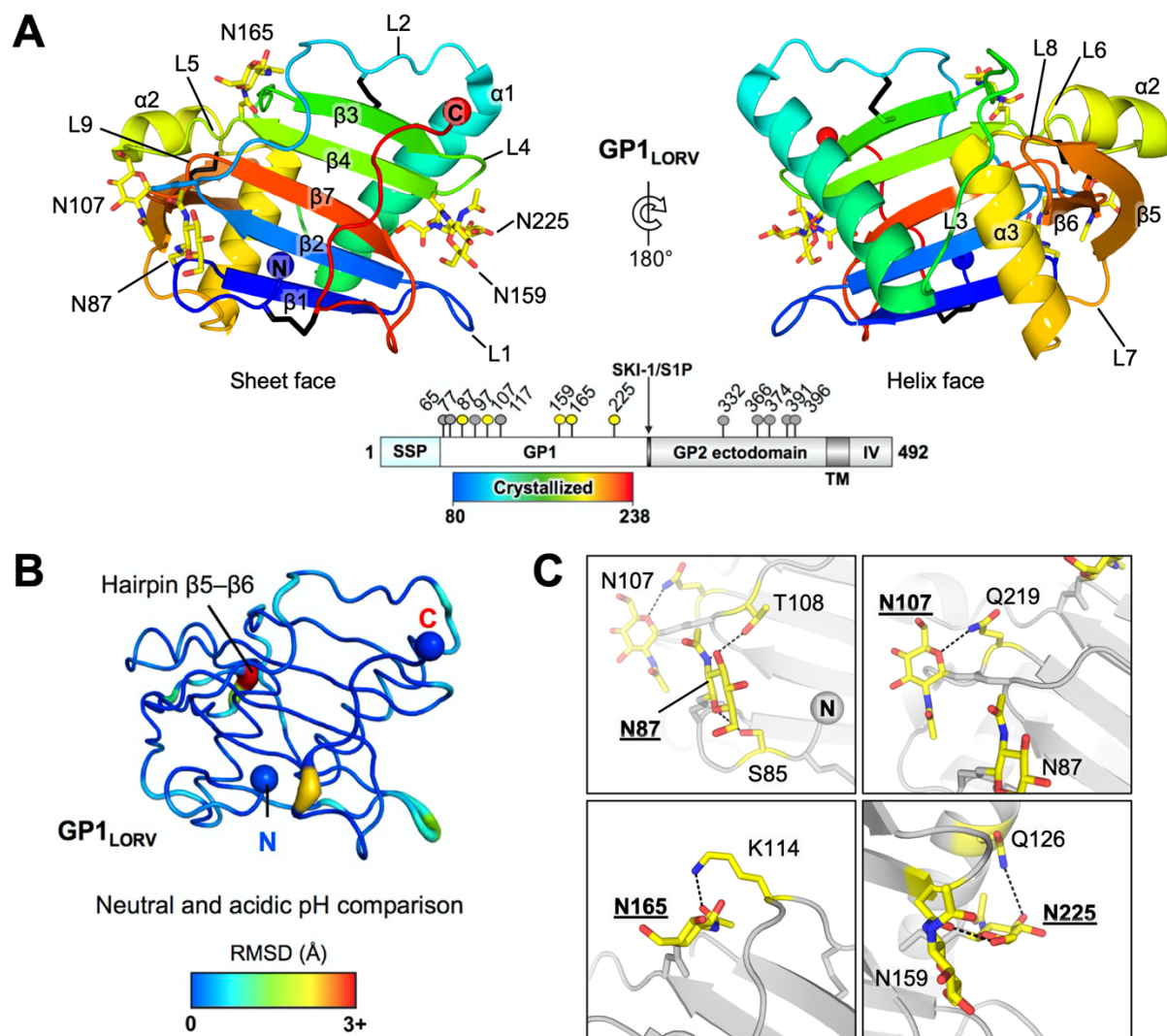


Fig. 3.4. Structure and organisation of LORV GP1. (A) Structure of LORV GP1. (*Top*) LORV GP1 (pH 5.0) is shown as a cartoon and coloured as a rainbow ramped from blue (N-terminus) to red (C-terminus). The N- and C-termini are shown as blue and red spheres, respectively. The β -strands are labelled as $\beta 1$ – $\beta 7$, α -helices as $\alpha 1$ – $\alpha 3$, and loops as L1–L9. N-linked glycosylation sites are shown as yellow sticks, with constituent nitrogen, and oxygen atoms coloured blue, and red, respectively. Disulfide bridges are shown as black sticks. (*Bottom*) Schematic organisation of LORV GP1 (generated with DOG (Ren et al., 2009)). The SSP, GP1, SKI-1/S1P cleavage site, GP2, transmembrane region (TM), and intravirion domain (IV) are annotated. Putative N-linked glycosylation sites are labelled as pins above the schematic, with sites observed to be occupied in either pH 8.0 or pH 5.0 crystal structure coloured yellow. Crystallised construct is indicated by a rainbow-coloured bar. (B) Structural comparison of LORV GP1 at pH 8.0 and pH 5.0. RMSDs between equivalent C α positions are represented by both colour (ramped from blue to red) and tube width (thin to thick). (C) Close-up views of the crystallographically observed interactions between GlcNAc moieties with LORV GP1 residues. The GP1 molecule is coloured light grey, with the GlcNAc moieties and the interacting GP1 residues shown as sticks and coloured according to their elements (carbon, yellow; nitrogen, blue; oxygen, red). Intermolecular hydrogen bonds are shown as dashed lines.

3.3.2 Structure Comparison of LORV GP1 with Characterised Arenaviral GP1

LORV GP1 exhibits a high level of structural conservation with other OW mammarenaviral GP1 glycoproteins in monomeric (GP2-free) states (**Fig. 3.5A**), where superimposition of LORV GP1 with monomeric LASV GP1 (PDB ID: 4ZJF) (Cohen-Dvashi et al., 2015) and MORV GP1 (PDB ID: 5NFF) (Israeli et al., 2017) resulted in a remarkably low overall RMSD (approximately 0.8 Å and 0.7 Å, respectively). On the other hand, overlay of LORV GP1 with LASV GP1 (PDB ID: 5VK2) (Hastie et al., 2017) and LCMV GP1 (PDB ID: 5INE) (Hastie et al., 2016a) in the GP2-associated state revealed stark differences between the glycoprotein structures (**Fig. 3.5B**).

Consistent with previous comparisons of LASV GP1 structures (Hastie and Sapphire, 2018; Hastie et al., 2017), more than 50% of C α atoms failed to align upon structural superimposition of GP2-free LORV GP1 and GP2-associated LASV and LCMV GP1s (**Fig. 3.5B**), suggesting that GP2 plays a role in stabilising GP1 in the GP2-associated conformation likely to exist on the mature virion. Indeed, the conformational switch from GP2-associated to GP2-free state displayed by GP1 is reflected by the propensity of the GP1–GP2 protomer to dissociate from one another and for GP2 to rearrange from pre- to post-fusion conformation, an elongated six-helix bundle structure (Burri et al., 2013b; Igonet et al., 2011).

Upon closer inspection, both N- and C-termini of the GP1, which are positioned at the same side of the β -sheet in LORV, appeared on opposite faces in LASV and LCMV (in GP2-associated state), where the C-terminus is pivoted downwards before crosses the central β -sheet to the helix-loop face and points towards the apex of the molecule (**Fig. 3.5B** and **3.6A**). While the GP1 central β -sheet aligned with considerable agreement between the GP2-free and GP2-associated conformations, the orientations of the α -helices in LORV GP1 are nearly perpendicular to those in LASV and LCMV GP1s (**Fig. 3.5B**). Indeed, α 1 and α 2 appear

parallel to the central β -sheet in LASV and LCMV but are joined to form a single long helix that spans the entire vertical length of the β -sheet in LORV (**Fig. 3.5B**).

Structural overlay of LORV GP1 with GP2-free NW JUNV (PDB ID: 5NUZ) (Zeltina et al., 2017), MACV (PDB ID: 2WFO) (Bowden et al., 2009), and WWAV (PDB ID: 5NSJ) (Shimon et al., 2017) GP1s further revealed major differences between the OW and NW arenaviral GP1 (**Fig. 3.5C**). Interestingly, the structural differences between monomeric LORV GP1 and the NW arenaviral GP1s are fairly consistent to those previously observed and described between the GP2-free and GP2-associated states of OW arenaviral GP1s. Indeed, while the core β -sheet remains as the most structurally conserved region of the GP1s, significant dissimilarities such as the differential arrangement of termini and perpendicular orientations of the α -helices in LORV GP1 are observed, when compared to the JUNV, MACV, and WWAV GP1s (**Fig. 3.5C**). Combined, these data highlight that OW mammarenaviral GP1 in GP2-associated state is structurally more similar to NW mammarenaviral GP1 in GP2-free state than OW mammarenaviral GP1 in GP2-free state, indicating that the LASV and LCMV GP1–GP2 structures likely constitute the native conformation presented on the mature virus.

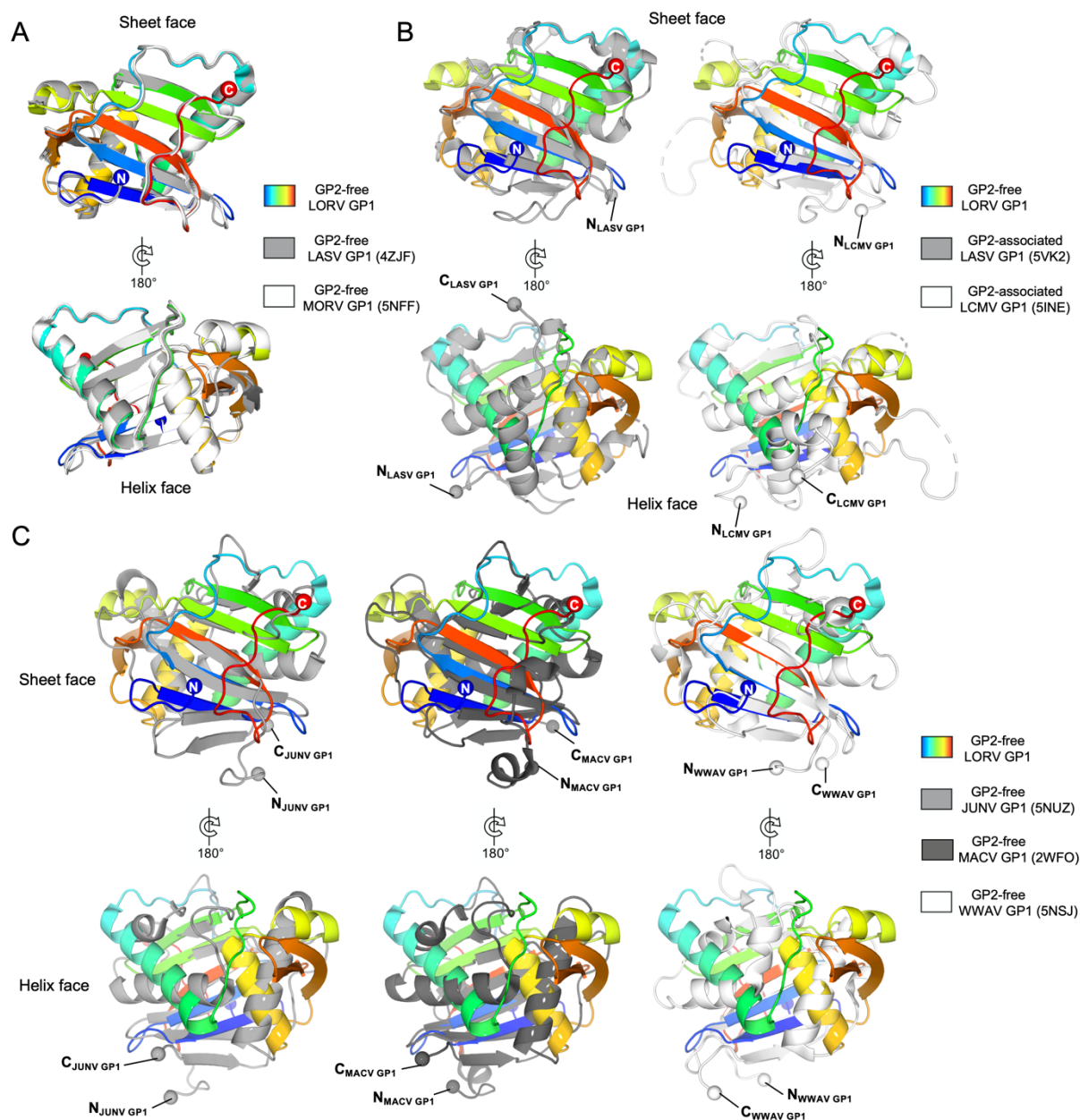


Fig. 3.5. Structure comparison of LORV GP1 with characterised mammarenaviral GP1s.

(A) Structure overlay analysis of monomeric LORV, LASV, and MORV GP1s. The crystal structures of LORV GP1 (pH 5.0), LASV GP1 (PDB ID: 4ZJF) (Cohen-Dvashi et al., 2015), and MORV GP1 (PDB ID: 5NFF) (Israeli et al., 2017) are shown in cartoon representation and coloured as annotated. Both the sheet face (left) and the helix face (right) are shown for clarity. The N- and C-termini of LORV GP1 are shown as blue and red spheres, respectively. The structures of LASV and MORV GP1s were truncated to display only residues Leu84–Ser237_{LASV} and Leu83–Ser235_{MORV}, respectively, to aid comparison with LORV GP1.

(B) Structure superimposition of monomeric LORV GP1 with LASV and LCMV GP1s (determined with cognate GP2). The crystal structures of LORV GP1 (pH 5.0), LASV GP1 (PDB ID: 5VK2) (Hastie et al., 2017), and LCMV GP1 (PDB ID: 5INE) (Hastie et al., 2016a) are shown in cartoon representation and coloured as annotated. The N- and C-termini of the GP1s are labelled accordingly. Unresolved regions in the structure are shown as dotted lines.

The GP2 subunits from the LASV and LCMV GP1–GP2 structures are not displayed. The structures of LASV and LCMV GP1s were truncated to display only residues Thr81–S255_{LASV} and Asn85–Thr261_{LCMV}, respectively, to aid comparison with LORV GP1. (C) Structure overlay analysis of monomeric LORV, JUNV, MACV, and WWAV GP1s. The crystal structures of LORV GP1 (pH 5.0), JUNV GP1 (PDB ID: 5NUZ) (Zeltina et al., 2017), and MACV GP1 (PDB ID: 2WFO) (Bowden et al., 2009), and WWAV GP1 (PDB ID: 5NSJ) (Shimon et al., 2017) are shown in cartoon representation and coloured as annotated. The N- and C-termini of the GP1s are labelled accordingly.

3.3.3 Structural Analysis of Characterised Receptor-Binding Sites

Many OW arenaviruses utilise α -DG (Cao et al., 1998), a cellular receptor protein expressed on the surface of the extracellular matrix, for host cell entry via endocytic pathways that are independent of clathrin, dynamin, caveolin, and actin (Pasqual et al., 2011; Rojek et al., 2008b). While there is a paucity of structural information about the interaction between OW mammarenaviral GP and α -DG, mutagenesis analysis of LCMV GP has highlighted regions on the GP1 that are required for α -DG utilisation (Hastie et al., 2016a).

Of note, residues known to interact with α -DG in LASV GP1 (His141_{LASV}, Asn146_{LASV}, Phe147_{LASV}, Tyr150_{LASV}) (Hastie et al., 2016a; Hastie et al., 2017) are fully conserved in LORV GP1 (His139_{LORV}, Asn144_{LORV}, Phe145_{LORV}, Tyr148_{LORV}), suggestive of shared receptor usage. These residues are located and clustered at the apex of the GP1 molecules of the trimeric LASV GP (**Fig. 3.6A**). Interestingly, structure-based mapping revealed that these residues are spatially dispersed on GP2-free LASV, MORV, and LORV GP1 (**Fig. 3.6B**). For example, in contrast to the close spatial association of critical α -DG binding residues in the trimeric LASV GP1 (GP2-associated), His139 from monomeric LORV GP1 (GP2-free) is displaced by more than 15 Å from the nearest other putative binding site residue (Asn144) (**Fig. 3.6B**). Similar to previous structural and experimental receptor-binding analyses of LASV, MORV, and LCMV GP1 (Cohen-Dvashi et al., 2015; Hastie et al., 2016a; Israeli et al., 2017), the spatial delocalisation of these receptor-binding residues in LORV GP1 is consistent with

the structure constituting an α -DG binding-incompetent conformation formed following detachment from GP2 during host cell entry.

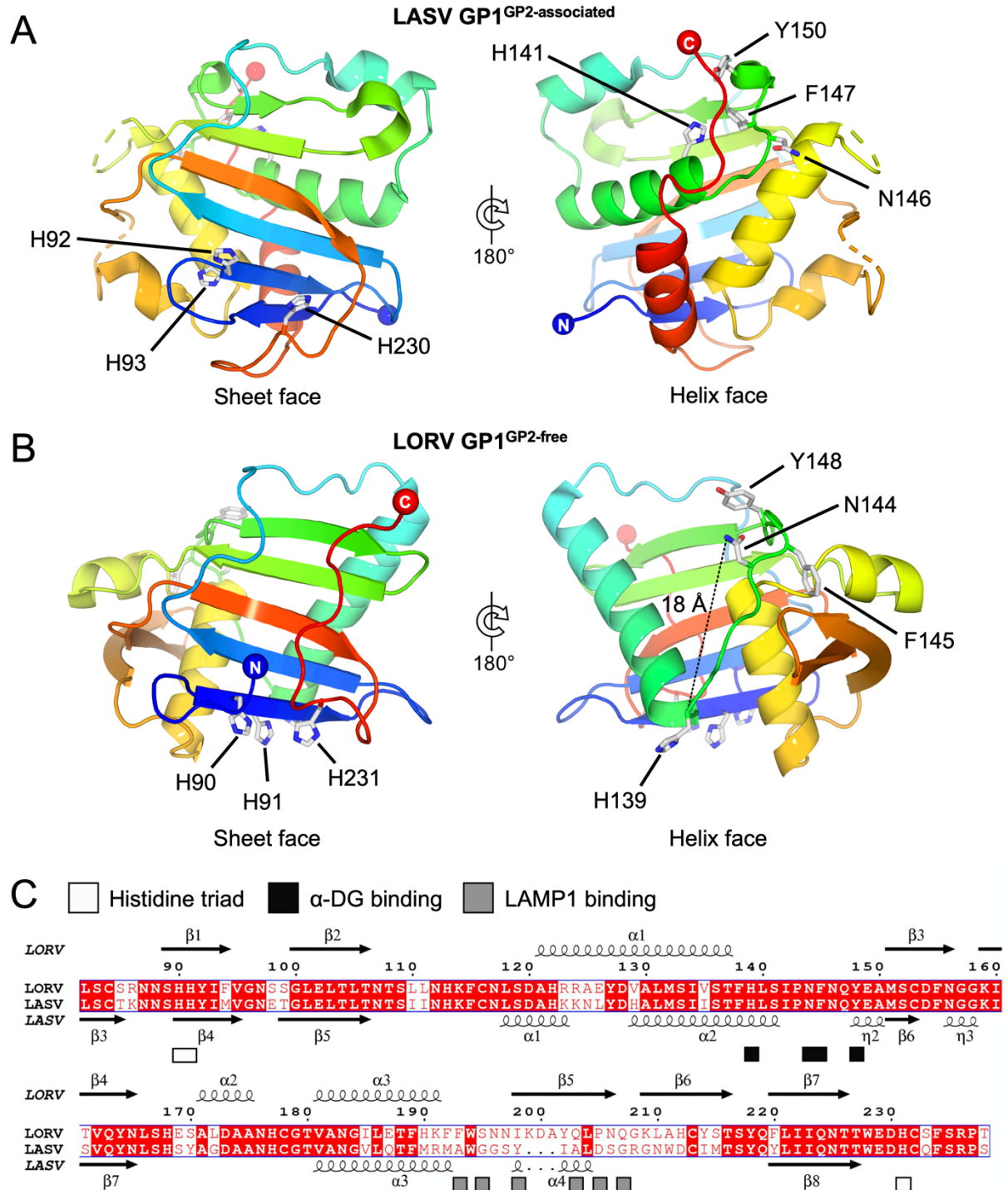


Fig. 3.6. Comparison of LAMP1 and α -DG binding residues on the GP2-associated LASV GP1 and the GP2-free LORV GP1. (A) The crystal structure of LASV GP1 in the GP2-associated conformation (PDB ID: 5VK2) (Hastie et al., 2017) is shown as cartoon and

coloured as a rainbow from blue (N-terminus) to red (C-terminus). The GP2 subunit is not displayed. The GP1 structure was truncated to display only Thr81–S255_{LASV} to aid comparison with LORV GP1. Unresolved regions are displayed as dashed lines. Residues implicated in LAMP1 and α -DG binding are labelled according to LASV GP1 numbering and shown as white sticks, with constituent nitrogen and oxygen atoms coloured blue and red, respectively. Spatial delocalisation of the α -DG binding residues is implicated by a dashed line. (B) The crystal structure of LORV GP1 (pH 5.0) in the GP2-free conformation is shown as cartoon and coloured as in panel A. Residues highlighted in panel A are shown here in identical manner and labelled according to LORV GP1 numbering. (C) Sequence alignment of the LORV and LASV GP1s. Identical residues are shaded red, while non-identical residues are coloured red. Residues constituting the putative LAMP1 binding site including the histidine triad and the putative α -DG binding site are annotated as indicated. Secondary structure elements of the LORV and LASV GP1s are shown above and below the alignment, respectively. Helices are shown as coils and β -strands as arrows. Sequences are labelled according to LORV GP1 numbering. Sequence alignment was performed with MultAlin (Corpet, 1988) and plotted with ESPript (Robert and Gouet, 2014).

Once inside the endosomal compartments, LASV binds to the intracellular receptor, LAMP1 (Cohen-Dvashi et al., 2015; Cohen-Dvashi et al., 2016; Jae et al., 2014). A cluster of three histidine residues (His92_{LASV}, His93_{LASV}, and His230_{LASV}) on LASV GP1 has been previously identified as a pH-sensor and forms part of the LAMP1 binding site (Cohen-Dvashi et al., 2015). Notably, the presence and location of this histidine triad (His90_{LORV}, His91_{LORV} and His231_{LORV}) are also conserved in the LORV GP1 structures, positioned at the β -sheet face, as observed in GP2-free LASV GP1 and MORV GP1 crystal structures (Cohen-Dvashi et al., 2015; Israeli et al., 2017) (**Fig. 3.6B**). However, when compared to the GP2-free state, the pH-sensing histidine triad on LASV GP1 in the GP2-associated state undergoes C α displacements of up to 7.0 Å, with the aromatic side chains reoriented 90° to 180° (**Fig. 3.6A** and **3.6B**). In contrast to the spatial closeness of α -DG binding residues in the GP2-associated LASV GP1 molecule required for initial host cell recognition, reorientation of the histidine triad likely renders the trimeric GP incapable of LAMP1 binding. Indeed, LAMP1 binding has only been experimentally shown for GP2-free LASV GP1 (Cohen-Dvashi et al., 2015). Interestingly, a recent cryo-ET and STA study of VLP displaying trimeric LASV GP (GP2-associated) by S.

Li *et al.* revealed the presence of non-GP densities near the cervixes of the trimer at pH 5.0 (**Fig. 1.11C**). Although the densities were attributed to LAMP1 by the authors (Li *et al.*, 2016), the utilisation of deglycosylated LAMP1 in the experiment highlights the questionable reliability of the proposed LASV GP–LAMP1 interaction, as it has been previously shown that mature sialylated glycans on LAMP1 is crucial for LASV GP1 binding (Jae *et al.*, 2014).

Despite evidence that the interaction between LASV GP1 and LAMP1 is crucial for efficient infection, previous studies have shown that the intracellular receptor is not strictly required for host cell entry (Hulseberg *et al.*, 2018). Furthermore, in parallel with the lack of structural information for OW mammarenaviral GP1–LAMP1 complex, to date there is no data that supports LAMP1 recognition by other OW mammarenaviruses apart from LASV. For instance, MORV GP1 has been experimentally shown to have no interaction with LAMP1 (Israeli *et al.*, 2017). The following structural comparison of LORV, LASV, and MORV GP1 in the GP2-free conformation reveals several key compositional and conformational differences of the putative LAMP1 binding residues on strand β 5 and loop L7.

Residues Ala195_{LASV}, Gly197_{LASV}, and Tyr200_{LASV} are located on LASV GP1 loop L7. These residues have been experimentally shown to be crucial for LAMP1 binding (Israeli *et al.*, 2017). However, they are not present in LORV GP1, which instead has Phe193_{LORV}, Ser195_{LORV}, and Ile198_{LORV} at these equivalent positions on the structure (**Fig. 3.6C**). Indeed, Ser195_{LORV} and Ile198_{LORV} are not able to form hydrogen bond like that observed between the side-chain hydroxyl of Tyr200_{LASV} and the main-chain carbonyl of Gly197_{LASV}, causing the loop L7_{LORV} to adopt a different main-chain conformation to L7_{LASV} (**Fig. 3.7A** and **3.7B**). In addition, aforementioned residues (Ala195_{LASV}, Gly197_{LASV}, and Tyr200_{LASV}) involved in LAMP1 binding are also not conserved in loop L7 of MORV GP1 (Tyr194_{MORV}, Ser196_{MORV}, and Asn197_{MORV}, respectively), which has been experimentally shown to not utilise LAMP1 (Israeli *et al.*, 2017). Structural alignment of GP2-free LORV and MORV GP1s further reveals

that loop L7_{LORV} adopts a similar conformation to L7_{MORV} (**Fig. 3.7A and 3.7B**), suggesting that LORV GP1, like MORV GP1, does not bind LAMP1.

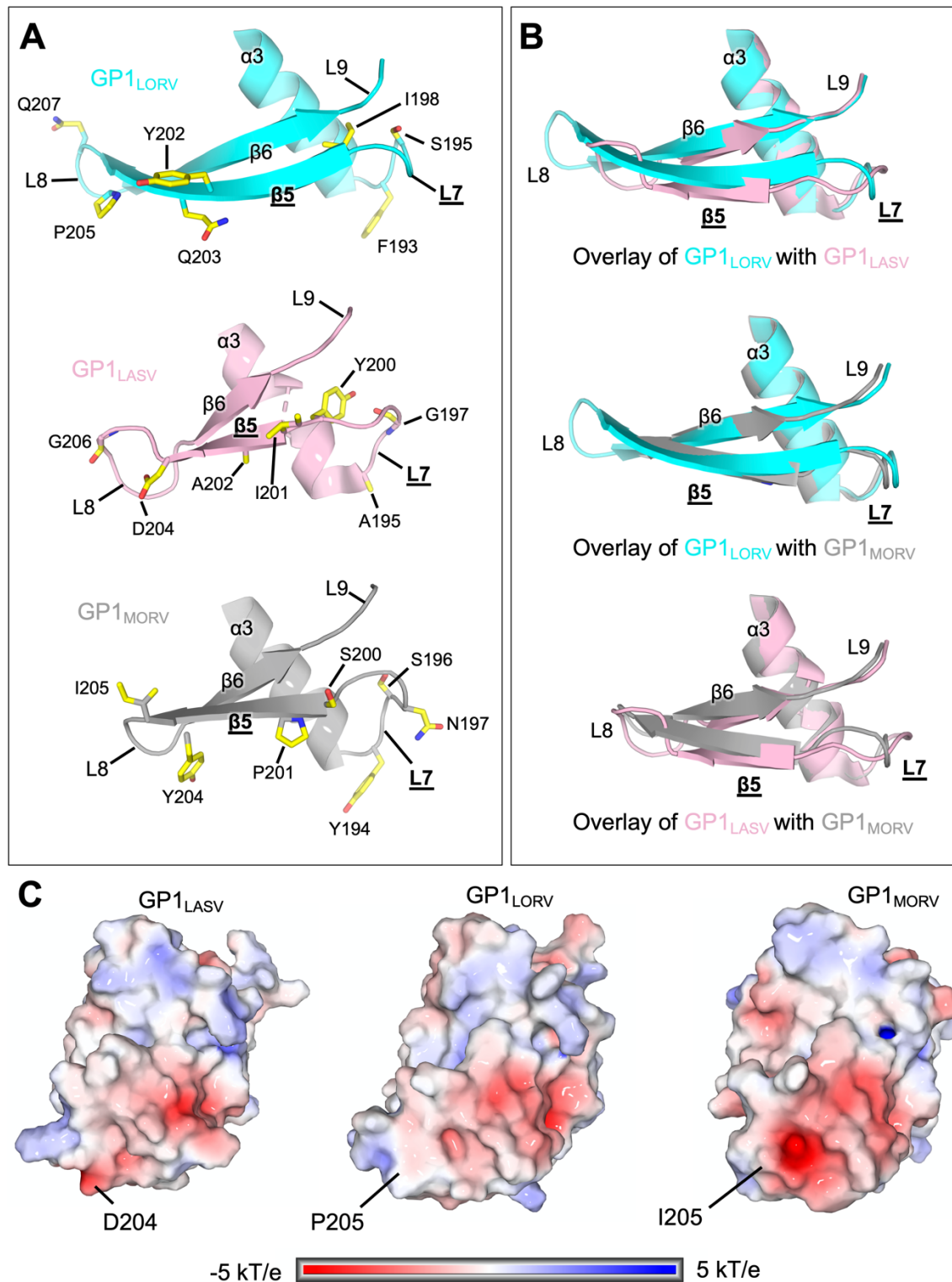


Fig. 3.7. Structural comparison of strand $\beta 5$ and loop L7 of GP2-free LORV, LASV, and MORV GP1s. (A) Residue composition and conformation of strand $\beta 5$ and loop L7 of monomeric LORV (*top*), LASV (*middle*) (Cohen-Dvashi et al., 2015), and MORV (*bottom*)

(Israeli et al., 2017) GP1s. Only the regions of $\beta 5$, L7, and their immediate vicinities are shown for clarity. The structures are shown as cartoons and coloured blue, pink, and grey for LORV GP1 (pH 5.0), LASV GP1, and MORV GP1, respectively. Residues previously implicated in LAMP1 binding on LASV GP1 are labelled and shown as yellow sticks, with constituent nitrogen and oxygen atoms coloured blue and red, respectively. The equivalent residues in LORV and MORV GP1s are shown in an identical manner. (B) Structural overlay of strand $\beta 5$ and loop L7 of monomeric LORV, LASV, and MORV GP1s. Only the regions of $\beta 5$, L7, and their immediate vicinities are shown for clarity. (*Top*) Overlay of LORV and LASV GP1 $\beta 5$ and L7. (*Middle*) Overlay of LORV and MORV GP1 $\beta 5$ and L7. (*Bottom*) Overlay of LASV and MORV GP1 $\beta 5$ and L7. (C) Presentation of the solvent accessible surfaces of monomeric LASV (*left*), LORV (*middle*), and MORV (*right*) GP1s. The surfaces are coloured according to the electrostatic potential in the range of ± 5 kT/e, as calculated by APBS in PyMol. Residue located at the end of the strand $\beta 5$ is labelled for each molecule to reveal the differences in electrostatic surface potential at the putative LAMP1 binding site.

Immediately adjacent to loop L7, the strand $\beta 5$ is oriented such that it is solvent exposed. In addition to the residues described above, Ile201_{LASV}, Ala202_{LASV}, Asp204_{LASV}, and Gly206_{LASV} on LASV GP1 strand $\beta 5$ have also been identified to play a role in LAMP1 binding (Israeli et al., 2017), but these residues are not conserved in LORV GP1 (**Fig. 3.6C** and **3.7A**). While substitutions at these sites do not abolish LAMP1 recognition, they greatly diminish the binding affinity (Israeli et al., 2017). Comparison of the strand $\beta 5$ from LASV GP1 and LORV GP1 in the GP2-free state reveals differences in the main-chain conformation and divergence in the residue composition (**Fig. 3.7B**). Furthermore, the corresponding residues on the strand $\beta 5$ of LORV GP1 (Tyr202_{LORV}, Gln203_{LORV}, Pro204_{LORV}, and Gln207_{LORV}, respectively) present a different calculated electrostatic surface potential. Indeed, in contrast to the extra negatively charged patch on LASV GP1 (Asp204_{LASV}) at the end of the strand $\beta 5$, a charge-neutral patch is observed at this location on LORV GP1 (Pro205_{LORV}) instead (**Fig. 3.7C**).

Based on the above observations, the lack of sequence and structural conservation of the putative LAMP1-binding residues on strand $\beta 5$ and loop L7 in LORV GP1 supports the hypothesis that LORV does not bind LAMP1. Owing to time constraint, this hypothesis had

not been tested in this thesis. Therefore, future investigation of the interaction between LORV GP1 and LAMP1 will be required to confirm this hypothesis.

3.3.4 LORV GP1 Is Highly Glycosylated

Glycosylation on the arenaviral GP has been shown to promote evasion of the humoral immune response (Sommerstein et al., 2015; Watanabe et al., 2018). This is particularly true for OW arenaviruses, as a mass spectrometric characterisation of site-specific N-glycan composition on LASV GP provided a structural model that demonstrates extensive coverage of the GP proteinaceous surface by underprocessed, high-mannose N-linked glycans (**Fig. 1.13**) (Watanabe et al., 2018). While the role of high-mannose glycans in DC-SIGN-mediated entry of LASV into monocyte-derived dendritic cells has been established (Goncalves et al., 2013), the abundance of glycans, which account for nearly 30% of the total mass of LASV GP, may also augment infection through immunological resistance by sterically impeding antibody accessibility (Watanabe et al., 2018).

LORV GP1 encodes nine putative N-linked glycosylation sequons (NXT/S, where X $\neq$ P), seven of which are present in the crystallised construct (**Fig. 3.4A**). Electron density corresponding to well-ordered N-linked GlcNAc moieties was observed at five of the seven sequons in LORV GP1 (Asn87, Asn107, Asn159, Asn165, and Asn225). However, no clear density was observed at the remaining sites (Asn97 and Asn117), indicating that these sites being either disordered in the crystal or not glycosylated during protein folding. Additional glycosylation sites, Asn65 and Asn77, are located outside the studied boundaries of the LORV GP1 construct. Mapping of these residues onto the crystal structure of the trimeric LASV GP (PDB ID: 5VK2) (Hastie et al., 2017) reveals that they are likely located in a membrane-proximal region of the glycoprotein spike (**Fig. 3.8**).

Seven out of the nine N-linked glycosylation sequons encoded by LORV GP1 are highly conserved with those by other OW mammarenaviruses, including LASV, LCMV, and MORV (**Fig. 3.8**), indicative of the importance of these sites in protein folding and processing, receptor binding, and immune evasion (Bonhomme et al., 2011; Bonhomme et al., 2013; Braakman and van Anken, 2000; Wright et al., 1989). While no glycan density was observed at Asn117_{LORV GP1}, the presence of glycans at this site on LASV GP1 and the high conservation of this asparagine residue across other OW arenaviruses may indicate an important role of the N-linked glycans at this site. Indeed, when mapped onto the trimeric LASV GP structure, the proximity of Asn117_{LORV GP1} to the conserved residues proposed to interact with α -DG (His139, Asn144, Phe145, Tyr148) suggests that the glycan, if occupied on the virus, protects the receptor binding site from antibody-mediated neutralisation.

Interestingly, LORV GP1 contains two putative N-linked glycosylation sites (Asn65 and Asn159) that are not observed on most OW arenaviruses. When mapped, these sites are proximal to areas that have been observed to present underprocessed, high mannose-type glycans on LASV GP (Watanabe et al., 2018), suggesting that glycosylation at Asn65 and Asn159 may contribute to an expanded glycan patch (**Fig. 3.8**). The prominence of N-linked glycosylation on the LORV GP, which exceeds that of LASV, suggests that LORV may also be an “evasion strong” virus (Burton, 2017; Watanabe et al., 2018) with heightened resistance to antibody-mediated neutralisation. Notably, mapping of the location of these sites reveals the close spatial association of Asn65 to an epitope recognised by a LASV neutralising antibody, 37.7H (Hastie et al., 2017). If N-linked glycans are present at this site, it seems possible that the glycans may help to shield this region of the GP surface. Additionally, like LASV, the high glycan density presented on the LORV GP may also facilitate a C-type lectin-mediated host cell entry pathway (e.g. DC-SIGN) (Goncalves et al., 2013).

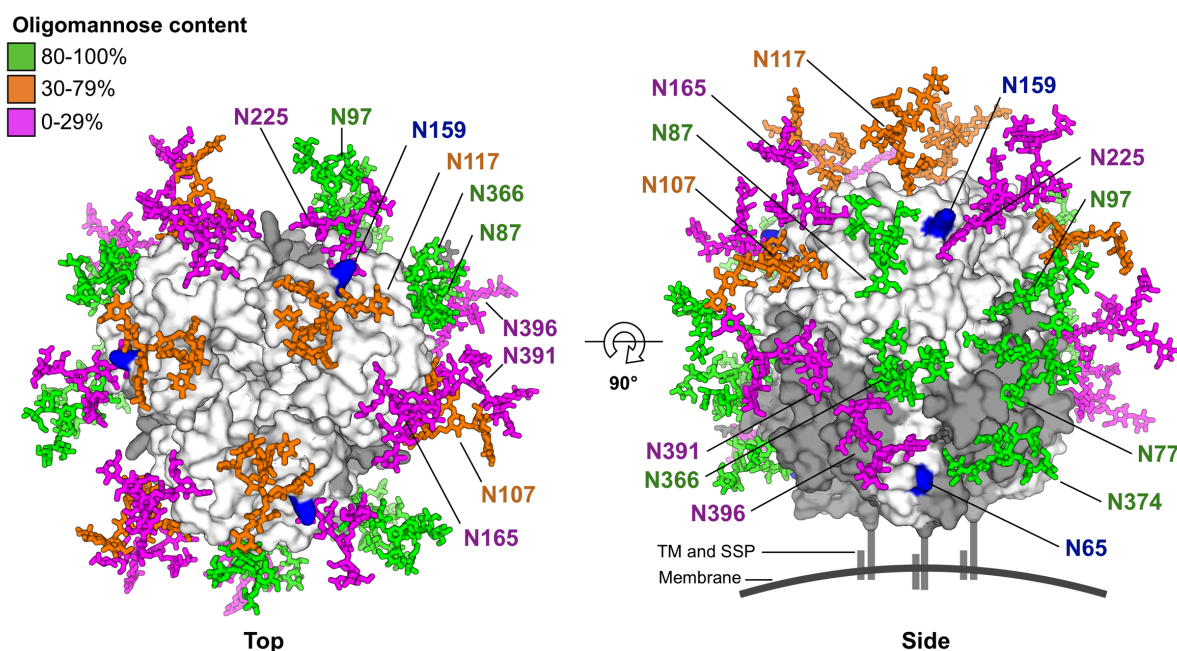


Fig 3.8. Mapping of LORV N-linked glycosylation sites onto the structure of trimeric LASV GP1–GP2. LASV GP1 and GP2 are shown as white and grey surfaces (PDB ID: 5VK2), respectively. Glycans are modelled as sticks and coloured according to the oligomannose content defined for the LASV GP (Hastie et al., 2017). For clarity, only glycans from a single GP1–GP2 protomer are labelled (LORV GP numbering). LORV GP1 contains two additional N-linked glycosylation sites, Asn65 and Asn159, which are not found in LASV GP1 (coloured blue). A schematic representation of the viral membrane, the TM region of GP2, and the SSP is shown. [Note that this figure has also been reported in Pryce *et al.* 2018 (listed in **Appendix 8.2**), which was adapted from (Watanabe et al., 2018).]

3.4 STRUCTURE-BASED PHYLOGENETIC ANALYSIS (SBPA) OF THE ARENAVIRAL GP1

3.4.1 Introduction

The conventional method of comparing and classifying viruses relies heavily on their genomic sequences. While this works well for viruses within the same family that contain high sequence similarity, the feasibility and reliability of this approach diminish rapidly when evolutionary distant viruses are considered. However, it has been observed that viral protein structures are often conserved despite low sequence conservation, a phenomenon that may, in part, be attributed to geometric, stereochemical, and functional constraints on protein architectures. Indeed, protein structure is restricted to approximately 2,000 unique folds predicted to exist

among naturally occurring proteins (Abrescia et al., 2012; Bamford, 2003; Oksanen et al., 2012). Therefore, when combined with modern sequence-based approaches (Simmonds et al., 2017; Wolf et al., 2018) and advanced machine-learning neural network (e.g. AlphaFold) (AlQuraishi, 2019; Senior et al., 2020), comparison of viral protein structures constitutes an important consideration when establishing commonalities of protein function and evolutionary relatedness from genetically distinct viruses (Bamford et al., 2005). Indeed, I have successfully used SBPA to relate commonalities in origin and function for several well-studied protein folds, including paramyxoviral attachment glycoproteins, influenza virus neuraminidases and haemagglutinin, and coronavirus S1-CTD glycoproteins. I have also reported these analyses in a review manuscript (Ng et al., 2020) (listed in **Appendix 8.2**).

Here, SBPA is utilised to classify the conformational states of OW and NW arenaviral GP1 glycoproteins. The result informs the potential physiological roles of these discrete structural classes and highlights their differential antigenicity during viral infection. Note that the following SBPA sections have also been discussed in the published review (Ng et al., 2020).

3.4.2 Background of SBPA

The human eye is remarkably good at qualitatively visualising and comparing the topological organisation and orientation of protein structures. However, quantitative details such as residue charges, sizes, angles, and distances are not always apparent, and therefore often overlooked when attempting to establish structural equivalence. Over the past decades, numerous structural alignment programmes have been developed to determine the relatedness of three-dimensional protein structures. These tools, which often rely on the availability of medium-to-high resolution ($<3 \text{ \AA}$) structures, includes but are not limited to HOMLOGY (Rao and Rossmann, 1973), SHP (Stuart et al., 1979), COMPARER (Sali and Blundell, 1990), DALI (Holm and Sander, 1993), CE (Shindyalov and Bourne, 1998), MAMMOTH (Ortiz et al.,

2002), FATCAT (Ye and Godzik, 2004), SSM (Krissinel and Henrick, 2004), Espresso/3D Coffee (Armougom et al., 2006; Poirot et al., 2004), MUSTANG (Konagurthu et al., 2006), SALIGN (Braberg et al., 2012), CLICK (Nguyen et al., 2011), and HSF (Ravantti et al., 2013). In this section, three underlying steps for deriving structural relationship information from SBPA are discussed: (i) Alignment, (ii) Classification, and (iii) Phylogeny.

(i) Alignment. A pairwise structural alignment is performed to establish the degree of equivalence between two superimposed structures. Conventionally, a measure of the RMSD between the corresponding C α atom pairs of superimposed protein structures has been used to define structural similarity (Rossmann and Argos, 1976). While such C α alignment calculations work well for architecturally conserved proteins, they can be less adequate when comparing protein structures that are only remotely related. Therefore, in addition to atom pair distances, pairwise structural alignments may be improved by added analysis of an array of parameters, such as solvent accessibility, dihedral angles, and local secondary structure. For example, MUSTANG (Konagurthu et al., 2006) aligns residues on the basis of similarity in residue-residue contact patterns and local structural topology, CLICK (Nguyen et al., 2011) analyses residue depth and solvent accessibility to optimise alignments of flexible or multi-domain structures, SHP (Stuart et al., 1979) utilises the shape of the local polypeptide chain and a rotation function that infers the relative orientation of structures, and HSF (Ravantti et al., 2013) analyses physiochemical and geometrical properties of the residues as well as local secondary structure.

Such parameters can be given different weightage depending on how the programmes treat them during the determination of structural equivalence. For example, COMPARER (Sali and Blundell, 1990) defines a protein structure as a multi-level hierarchy of features: a ‘property’ such as the residue type (e.g. polar, hydrophobic, aromatic), engaged in a number of

‘relationships’ (e.g. hydrogen bonding, hydrophobic clustering, π - π stacking). A residue-by-residue matrix consisting of the weighted sum of the relative contributions by ‘properties’ and ‘relationships’ is used to align two proteins under comparison. In contrast, HSF (Ravantti et al., 2013) implements a uniform representation of all ‘properties’ and ‘relationships’, allowing these parameters to independently boost the agreement of residues that exhibit localised regions of similarity.

(ii) *Classification*. Following pairwise structural alignments, the analysis is expanded to include all alignments into a hierarchical classification, a step that is performed by programmes such as HSF and MUSTANG. In HSF (Ravantti et al., 2013), the initial set of equivalent residues identified by pairwise comparison forms a common structural core between the two structures. The alignment establishes invariant information (e.g. residue size and charge) that remain constant, and variant information (e.g. C α -C α pair distances) that vary throughout the alignment process. The invariant information determine the initial set of equivalent residues (common core) while the invariant information refine the pairwise alignment. Comparison of the similarities between structural cores established from all of the pairwise alignments generates new combined cores (i.e. new sets of equivalent residues), which will subsequently be used to repeat this recursive alignment process until the pool of input structures under comparison is diminished. As each class grows to encompass more closely related structures, a full hierarchical classification is generated. Similarly, MUSTANG (Konagurthu et al., 2006) applies analogous steps in its multiple structural alignment and classification: residue-residue similarity scores are generated by alignments of equivalent contiguous structural cores, followed by comparison of all pairwise alignments in the context of multiple structures to recalculate similarity scores for each newly aligned class.

(iii) *Phylogeny*. Subsequently, the relatedness of protein structures can be calculated quantitatively and plotted as an unrooted tree (of unknown origin), reflecting the distances (branch length) between protein structures. For instance, SHP, MUSTANG, and HSF calculate so-called ‘evolutionary distances’ from similarity scores or probabilities of equivalence, as detailed in Bamford, Grimes, and Stuart (2005; SHP), Konagurthu et al. (2006; MUSTANG), and Ravantti, Bamford, and Stuart (2013; HSF). Finally, structure-based phylogenetic trees can be plotted from distance matrices and visualised using programmes such as PHYLIP (Felsenstein, 1989) and Dendroscope (Huson et al., 2007).

3.4.3 SBPA of the Arenaviral GP1

Structures of available mammarenaviral GP1 glycoproteins from LORV, LASV, LCMV, MORV, LUJV, JUNV, MACV, and WWAV were aligned using SHP (Stuart et al., 1979) to create pairwise evolutionary distance matrices. The distances, as iterated above, represent the structural similarities between the protein structures, and were plotted as an unrooted phylogenetic tree using PHYLIP (Felsenstein, 1989) (**Fig. 3.9**). The smaller the distance (branch length) between the two GP1 molecules, the more similar they are in their structures.

Consistent with the sequence-based phylogenetic analysis of mammarenaviral glycoproteins (**Fig. 3.1**), SBPA of arenaviral GP1s revealed that the GP1s broadly divide into two branches (**Fig. 3.9**) according to their OW and NW origins (Emonet et al., 2009). Upon finer inspection, the structure-based phylogenetic tree reveals that OW arenaviral GP1 glycoproteins further bifurcate into two distinct structural classes: one class comprises structures of OW GP1s determined in association with their respective cognate GP2 (GP2-associated), and the other contains GP1 structures determined in isolation (GP2-free). Indeed, the utility of SBPA highlights the higher structural similarity between LASV GP1 and LCMV

GP1 in their GP2-associated conformation than that observed between the two known conformations of LASV GP1, which possess identical primary sequences.

The large structural distance of GP2-free state from the GP2-associated structures reflects the pronounced antigenic deconfiguration and variation of the GP1 molecules once shed from the GP during host cell entry (Branco et al., 2010; Hastie et al., 2017; Pryce et al., 2018). Such observation provides a structure-based rationale to the functional importance of shed OW GP1. Indeed, the GP2-free GP1 may resemble the shed OW arenaviral GP1 detected in patient sera during acute LASV infection (Branco et al., 2010), suggesting that the presentation of dramatically distinctive epitopes not present on the mature virion may serve as an immunological decoy (Fisher-Hoch and McCormick, 2001; Hastie and Saphire, 2018; Robinson et al., 2016). This hypothesis is further supported by the absence of an effective humoral immune response upon immunisation with recombinant OW GP1 (Borenstein-Katz et al., 2019).

Examination of the SBPA reveals that the GP1 structures from NW JUNV, MACV, and WWAV separate into two branches corresponding to their respective classification into clades B and D (Ng et al., 2020; Pryce et al., 2018) (**Fig. 3.9**). Although JUNV, MACV, and WWAV GP1s all utilise TfR1 as a receptor (Abraham et al., 2009; Choe et al., 2011; Radoshitzky et al., 2007), WWAV is approximately equidistant between JUNV/MACV and GP2-associated LASV/LCMV GP1s. This highlights the GP1 structural variation within TfR1-tropic NW arenaviruses and is likely a result of coevolution with individual rodent host TfR1 orthologues and selection pressure from host immune system. Due to the current paucity of GP2-bound NW GP1 structures, it is unknown if NW GP1s exhibit similar conformational plasticity as observed in OW GP1s. However, this seems to be unlikely as GP2-free NW GP1s have demonstrated the ability to recognise both host TfR1 and vaccine-elicited neutralising monoclonal antibodies, and elicit a nAb immune response following immunisation in animal

models (Abraham et al., 2010; Borenstein-Katz et al., 2019; Clark et al., 2018; Mahmutovic et al., 2015; Zeltina et al., 2017). Further discussion of NW arenaviral GP1 structures is detailed in Chapter 4.

Another striking feature of the structure-based phylogeny is the unique placement of GP2-free LUJV GP1 outside of the established OW and NW branches (**Fig. 3.9**). The near-equal structural distance of LUJV GP1 from both GP2-associated OW and GP2-free NW GP1 classes supports the unique receptor tropism characteristics and structure of the glycoprotein (Cohen-Dvashi et al., 2018). Indeed, the structure of GP2-free LUJV GP1 solved in complex with NRP2 receptor revealed a different protein architecture to those of OW and NW arenaviral GP1s reported to date. Interestingly, the crystal structure of GP2-free LUJV GP1 exhibits features observed in both NW mammarenaviral GP1s and OW mammarenaviral GP2-associated GP1s (Cohen-Dvashi et al., 2018). Furthermore, the receptor-bound structure suggests that the observed conformation likely represents that of the virion-displayed GP. Combined, the structural and SBPA data suggest that LUJV GP1 structurally and functionally diverged early in the evolutionary bifurcation of OW and NW arenaviruses (Ng et al., 2020).

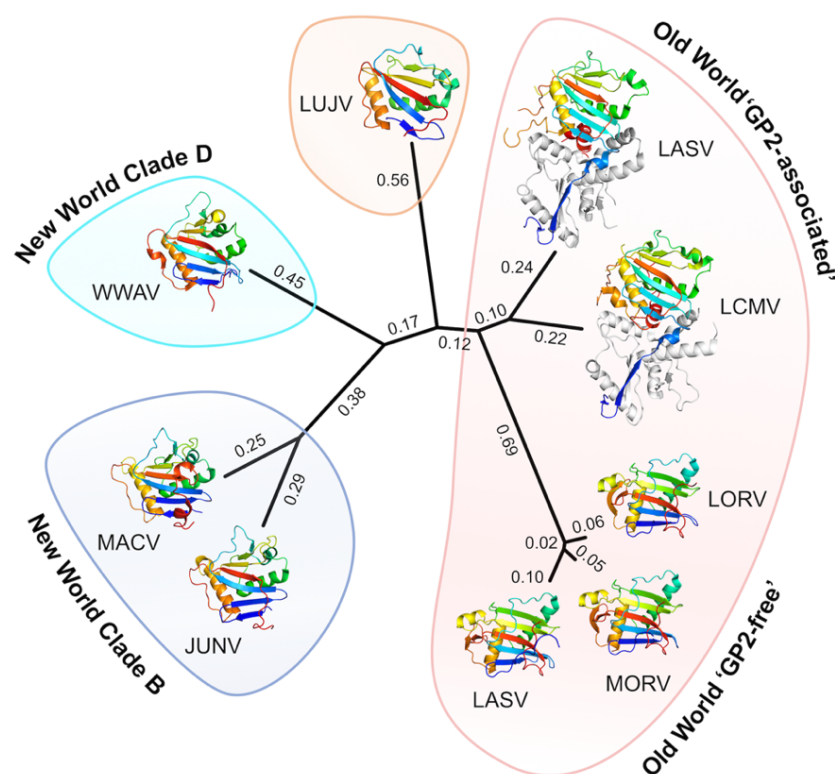


Fig. 3.9. Structure-based phylogenetic tree showing the conformations adopted by arenaviral GP1 attachment glycoproteins. Structures of available arenavirus GP1 glycoproteins used were as follows: LORV, Loei River virus (PDB ID: 6HJ6); LASV, Lassa virus (4ZJF and 5VK2); MORV, Morogoro virus (5NFF); LCMV, Lymphocytic choriomeningitis virus (5INE); LUJV, Lujo virus (6GH8); WWAV, Whitewater Arroyo virus (6HJ5); MACV, Machupo virus (2WFO); and JUNV, Junín virus (5NUZ). For 5VK2, 5INE, 6GH8, and 5NUZ, all chains not comprising GP1 molecules (e.g. GP2, receptor, and antibody fragments) were removed prior to structural alignment. A pairwise evolutionary distance matrix was created using SHP (Stuart et al., 1979) and displayed as an unrooted phylogenetic tree using PHYLIP (Felsenstein, 1989). GP1 structures are shown in cartoon representation and coloured as a rainbow from the N- (blue) to C-terminus (red). Although the GP2 component of the GP was not included in the structure comparison, it is shown and coloured as a white cartoon for clarity. Calculated evolutionary distances derived from this analysis are indicated next to each branch. [Note that this figure has also been reported in Ng *et al.* 2020 (listed in **Appendix 8.2**).]

3.5 SUMMARY AND DISCUSSION

Over the past decades, major efforts have been spent on investigating cellular tropism characteristics and protein structures of established OW and NW arenaviruses that are known to persistently circulate in Africa and the Americas, respectively. However, the recent

discovery of novel Asiatic OW arenaviruses, such as WENV (Li et al., 2015), RYKV (Wu et al., 2018), and LORV (Blasdell et al., 2016) is an indication of the breadth in genetic diversity and geographical distribution of these viruses. As more emerging arenaviruses are being identified using modern high-throughput sequencing techniques, the urgency in understanding their evolutionary relationships, zoonotic potentials, and host range becomes even greater. With a focus on LORV, I sought to characterise the first structure of an OW Asiatic mammarenaviral GP1 and delineate the roles of detachment and acidification in determining the GP1 conformation. Here, I report the structures of LORV GP1 obtained at basic (pH 8.0) and acidic pH (pH 5.0) and refined to 2.5- and 2.0-Å resolution, respectively.

Both the basic and acidic structures of LORV GP1 are highly conserved with LASV (Cohen-Dvashi et al., 2015) and MORV (Israeli et al., 2017) GP1s that had been structurally characterised in isolation without their cognate GP2 (GP2-free) (**Fig. 3.4A**). However, major reorganisation of the GP1 molecules can be observed when compared to LASV (Hastie et al., 2017) and LCMV (Hastie et al., 2016a) GP1s solved in association with their respective GP2 subunits (GP2-associated). Indeed, apart from the central core β -sheet that remains structurally conserved, GP1 undergoes a major structural rearrangement involving a reorientation of its C-terminus and α -helices when it dissociates from the GP2 (**Fig. 3.4B**). Furthermore, the high structural similarity between the basic and acidic pH structures of LORV GP1 suggests that pH-derived alteration of the electrostatic surface potential does not lead to any appreciable local conformational changes. Therefore, the observed two structural states adopted by OW arenaviral GP1 are more likely due to the association and dissociation of GP1 from the GP2 subunit than pH change alone. This investigation expands our appreciation of the structural landscape covered by mammarenaviral glycoproteins and provides evidence that pH change does not directly modulate the conformation of isolated GP1.

In addition, SBPA of all available GP1 structures highlights the conformational duality of OW arenaviral GP1 and their biological importance (Ng et al., 2020; Pryce et al., 2018) (**Fig. 3.9**). Indeed, the large structural phylogenetic distance of OW mammarenaviral GP1s in the GP2-free state from the GP2-associated GP1s corresponds to the antigenic deconfiguration of GP1 molecules and their inability to bind the primary receptor, α -DG (Cohen-Dvashi et al., 2015; Israeli et al., 2017). The presentation of dramatically different epitopes by shed OW arenaviral GP1, with respect to virion-displayed GP2-associated GP1, may thus potentially serve as an immunological decoy and contribute to an ineffective humoral immune response during a viral infection (Ng et al., 2020; Pryce et al., 2018). The detection of shed OW arenaviral GP1 in patient sera during LASV infection further supports this hypothesis (Branco et al., 2010).

The dense coverage of the mammarenaviral GP surface by the abundant N-linked glycosylation further rationalises the difficulty in mounting an effective antibody-mediated immune response to OW arenaviruses (Sommerstein et al., 2015; Watanabe et al., 2018; Zeltina and Bowden, 2017). Interestingly, mapping analysis of LORV GP1 (**Fig. 3.8**) revealed that the glycan-mediated masking of LORV GP is likely to be even more pronounced than that of the LASV GP, suggestive of the existence of arenaviruses with even greater glycan-mediated immune evasive properties. Indeed, the presence of glycan sites near nAb-binding epitopes on LASV GP may indicate glycan shielding on this region of LORV GP1, hindering antibody recognition.

The often inadequate antibody immune response arising from infection combines with the lack of approved vaccines and efficacious treatments represent a continuous threat in combating highly pathogenic OW mammarenaviruses. Vaccine development focusing on the envelope GP is, therefore, of paramount importance and urgency. Structural data detailed in this chapter provide a molecular blueprint for understanding how conformational plasticity and

glycan shielding render the OW mammarenaviral GP spikes immunologically resistant targets. Therefore, the data presented here will be integral for the creation of an immunogen capable of mimicking the mature OW mammarenaviral GP surface. For instance, the design of a protective OW mammarenaviral immunogen may likely require GP1–GP2 as the minimal component, which presents a conformation capable of eliciting nAbs against the GP (Hastie and Saphire, 2018; Zeltina and Bowden, 2017). Furthermore, the abundance of underprocessed, high-mannose type glycans on LASV GP (Watanabe et al., 2018) has important implications in the development and production of vaccine candidates. Given the functional importance of the glycan processing states, it is likely that the current recombinant expression, which utilises varied cell lines (e.g. insect cells), mannosidase inhibitor, and endoglycosidase-mediated deglycosylation, for monomeric OW mammarenaviral GP1 (Cohen-Dvashi et al., 2015; Israeli et al., 2017; Pryce et al., 2018) and GP ectodomains (Hastie et al., 2016a; Hastie et al., 2017) may not reproduce native-like glycosylation, and may be presenting epitopes that are natively inaccessible as a result of glycan removal and/or truncation. Therefore, future mammarenaviral GP immunogen design efforts must suppress the exposure of these epitopes through the presentation of more-native-like glycosylation.

4 ANTIBODY-MEDIATED NEUTRALISATION OF NEW WORLD JUNÍN VIRUS AND MACHUPO VIRUS

4.1 INTRODUCTION

NW HF mammarenaviruses constitute a group of rodent-borne zoonotic pathogens that cause severe disease with high case fatality rates upon transmission into human populations. These agents belong to the clade B and D viruses and include JUNV, MACV, GTOV, CHAV, SABV, and WWAV (Paessler and Walker, 2013; Pryce et al., 2018). Although there are currently no internationally approved therapeutic countermeasures available for preventing or treating NW arenaviral HFs, a live attenuated vaccine, Candid#1, has been successfully utilised throughout endemic areas of Argentina to protect against JUNV and shows evidence of partial cross-protection against MACV (Koma et al., 2016; Maiztegui et al., 1998). Similarly, ribavirin administered intravenously in high dosage has shown promise in treating the disease (Enria et al., 2008), and passive immunisation with immune plasma containing high titres of nAbs from convalescent patients has been shown to be successful in treating infected individuals (Enria et al., 1984; Maiztegui et al., 1979).

Entry of NW clade B and D mammarenaviruses into a host cell is initiated by the specific interaction of the GP with host cell receptors, including C-type lectins (Martinez et al., 2013), TIMS (Jemielity et al., 2013), and TfR1 of their respective rodent host (Flanagan et al., 2008; Radoshitzky et al., 2007). The ability of NW viruses to also utilise hTfR1 is a primary determinant of zoonosis and pathogenicity in humans (Abraham et al., 2010; Abraham et al., 2009; Choe et al., 2011; Radoshitzky et al., 2007). TfR1 is a homodimeric type II transmembrane protein, where each monomer is composed of a cytoplasmic subunit, a single-pass transmembrane region, and an extracellular domain (ECD). The ECD is further subdivided into three subdomains: a membrane-proximal protease-like domain, a central helical domain

that facilitates dimerisation, and a membrane-distal apical domain (Lawrence et al., 1999). The GP1 subunit contacts the apical domain of TfR1, away from transferrin and hereditary hemochromatosis protein recognition sites (Abraham et al., 2010; Bowden et al., 2011). The GP1–TfR1 interaction involves the insertion of a tyrosine residue (Tyr211 in hTfR1) at a position in the central pocket of GP1 (Abraham et al., 2010). Structural determination of Fab fragments from several nAbs (OD01, GD01, and CR1-28) of different germline origins bound to JUNV GP1 revealed a similar mode of recognition, establishing receptor mimicry as a common means of antibody-mediated neutralisation (Clark et al., 2018; Mahmutovic et al., 2015; Zeltina et al., 2017). However, this mode of antibody recognition was not observed in the structure of the Fab fragment from cross-reactive MACV/JUNV nAb (CR1-07) in complex with MACV GP1 (Clark et al., 2018) and there is currently no reported structure of a MACV GP1-specific antibody.

The generation of nAbs has been shown to be effective for the prevention and treatment of HF mammarenaviral infection (Enria et al., 1984; Maiztegui et al., 1979). Indeed, passive immunisation of patient sera containing nAbs reduces the case mortality rate of JUNV infection from 15–30% to 1% (Enria et al., 1984; Maiztegui et al., 1979). However, the limited quantity and variable quality of immune plasma represent an obstacle in the treatment of AHF and BHF. Therefore, mAb-based therapy has been developed as an alternative to plasma, offering a highly specific, potent, and generally safe platform for HF treatment (Enria et al., 2008). When tested *in vivo* using an animal model (Kenyon et al., 1986), mAbs shows high protection efficacy up to 100%, indicating that mAb-based interventions may be superior to the broad-spectrum polyclonal sera (Zeitlin et al., 2016).

There is anecdotal evidence that Candid#1 vaccination confers partial protection against MACV disease in humans. Furthermore, NHPs vaccinated with an attenuated JUNV strain (XJ clone 3) survived MACV challenge through an accelerated anti-MACV nAb

response (Peters et al., 1987). Despite the diverse sequence and structural variability across NW mammarenaviral GP (Demogines et al., 2013; Kerr et al., 2015), these data show that GP-directed antibody-mediated cross neutralisation of JUNV and MACV is possible, indicative of the presence of conserved epitopes within the GP that can elicit bnAbs. Given the reliable efficacy of mAb-based treatment, identification and characterisation of these key sites of vulnerability have important implications to the development of bnAbs-based therapeutics against JUNV, MACV, and other NW mammarenaviruses.

Here, a mouse model of MACV and JUNV infection combined with antigen-specific B cell sorting was utilised to isolate a nAb specific to MACV GP1 (termed MAC1) and seven nAbs specific to JUNV GP1 (termed JUN1-7), which recognise an overlapping epitope. Crystallographic investigation of the Fab fragments of nAb MAC1 and a representative JUNV-specific nAb, JUN1, in complex with their cognate GP1 antigens reveals that both nAbs target the TfR1 binding site. Close examination of the antigen–antibody interfaces reveals that whilst nAb JUN1 originates from a unique germline, it utilises a mode of antigen recognition involving TfR1 receptor mimicry common to all characterised infection and vaccine-elicited JUNV-specific nAbs. In contrast, nAb MAC1 undergoes a mode of antigen recognition that is distinct from that utilised by reported anti-JUNV nAbs and the cross-reactive MACV/JUNV nAb. Combined, these molecular-level findings expand our appreciation of the antibody-mediated immune response to NW arenavirus infection and the varied means by which the GP1 of NW HF mammarenaviruses can be targeted by the antibody-mediated immune response.

4.2 PROTEIN PRODUCTION AND STRUCTURE DETERMINATION OF NEW WORLD JUNIN VIRUS (JUNV) AND MACHUPO VIRUS (MACV) GP1s IN COMPLEX WITH FAB FRAGMENTS

4.2.1 Large-Scale Transient Protein Expression and Purification

Crystal structures of JUNV GP1 (PDB ID: 5NUZ (Zeltina et al., 2017), 5EN2 (Mahmutovic et al., 2015), and 5W1K (Clark et al., 2018)) and MACV GP1 (PDB ID: 2WFO (Bowden et al., 2009), 3KAS (Abraham et al., 2010), and 5W1M (Clark et al., 2018)) have been previously determined, therefore, the published constructs were replicated exactly here. Mammalian expression vectors (pHLsec) containing constructs encoding JUNV GP1 (Asp87–Asn232) and MACV GP1 (Glu87–Phe257) were kindly provided by Dr. Antra Zeltina from the Bowden group in STRUBI.

Protein expression and purification were performed in the similar fashion described for LORV GP1. Briefly, HEK293T cells were transiently transfected with the desired protein constructs using PEI in the presence of kifunensine (Elbein et al., 1990). Cell supernatants were harvested 96 h post-transfection and diafiltrated against the purification buffer (10 mM Tris pH 8.0, 150 mM NaCl). Glycoproteins were purified by IMAC using HisTrap™ FF followed by SEC.

The heavy- and light-chain genes of the JUN1 and MAC1 Fab fragments were PCR-amplified from cDNA (provided by research collaborator, Dr. Katie Doores) and cloned into the pHLsec expression vector (Aricescu et al., 2006). A C-terminal His₆-tag was included in the heavy chain construct and both chains were co-expressed [1:1 (wt/wt) ratio of Fab heavy to light chain plasmids] in HEK293T cells and purified, as previously described. Fab JUN1 and MAC1 were subsequently mixed with purified JUNV GP1 and MACV GP1, respectively, in a 1:1.1 molar ratio with excess GP1 to allow clear separation of SEC chromatogram peaks corresponding to the ~75 kDa GP1–Fab complex and ~25 kDa GP1 alone. To aid

crystallogenesis, JUNV GP1 and MACV GP1 were partially deglycosylated with endoF₁ (25°C, 18 h). Following deglycosylation, the GP1–Fab complexes were repurified by SEC, as described above (**Fig. 4.1**).

4.2.2 Crystallisation of GP1–Fab Complexes

Crystallisation experiments were performed at room temperature using the sitting drop vapour diffusion method (Walter et al., 2005). Crystals of JUNV GP1–Fab JUN1 complex were obtained by mixing 100 nL of 5.9 mg/ml protein sample and 100 nL of precipitant containing 20% w/v PEG8000 and 0.1 M HEPES pH 7.5 (**Fig. 4.1**). Crystals of MACV GP1–Fab MAC1 complex were obtained by mixing 100 nL of 8.8 mg/ml protein sample and 100 nL of precipitant containing 1.4 M sodium malonate pH 6.0 (**Fig. 4.1**). Crystallisation drops were equilibrated against 95 µL of a precipitant-containing reservoir. In all instances, crystals were cryo-protected by transfer into a solution of the respective precipitant supplemented with 25% (v/v) glycerol, prior to flash cooling in liquid nitrogen. Crystal freezing was performed by Dr. Karl Harlos at STRUBI.

4.2.3 X-Ray Diffraction Data Collection and Processing

X-ray diffraction data were recorded at Diamond Light Source, United Kingdom. Crystal data were indexed, integrated, and scaled with XIA2 (3dii or DIALS programmes) (Winter et al., 2013). The structure of JUNV GP1–Fab JUN1 complex was solved by molecular replacement with PHASER (McCoy et al., 2007), using the crystal structures of JUNV GP1 and Fab OD01 (PDB ID: 5NUZ (Zeltina et al., 2017)) as search models. The structure of MACV GP1–Fab MAC1 complex was phased by molecular replacement with PHASER using MACV GP1 (PDB ID: 2WFO (Bowden et al., 2009)) and Fab OD01 (PDB ID: 5NUZ (Zeltina et al., 2017)) as search models. For both structures, iterative rounds of model building and refinement were

performed using COOT (Emsley and Cowtan, 2004) and PHENIX with TLS parameterisation (Adams et al., 2002), respectively. MolProbity (Chen et al., 2010) was used to validate model quality. Data collection and refinement statistics are presented in **Table 4.1**. The PyMOL Molecular Graphics System (<https://www.schrodinger.com/pymol>) was used to generate the structural models presented in the figures.

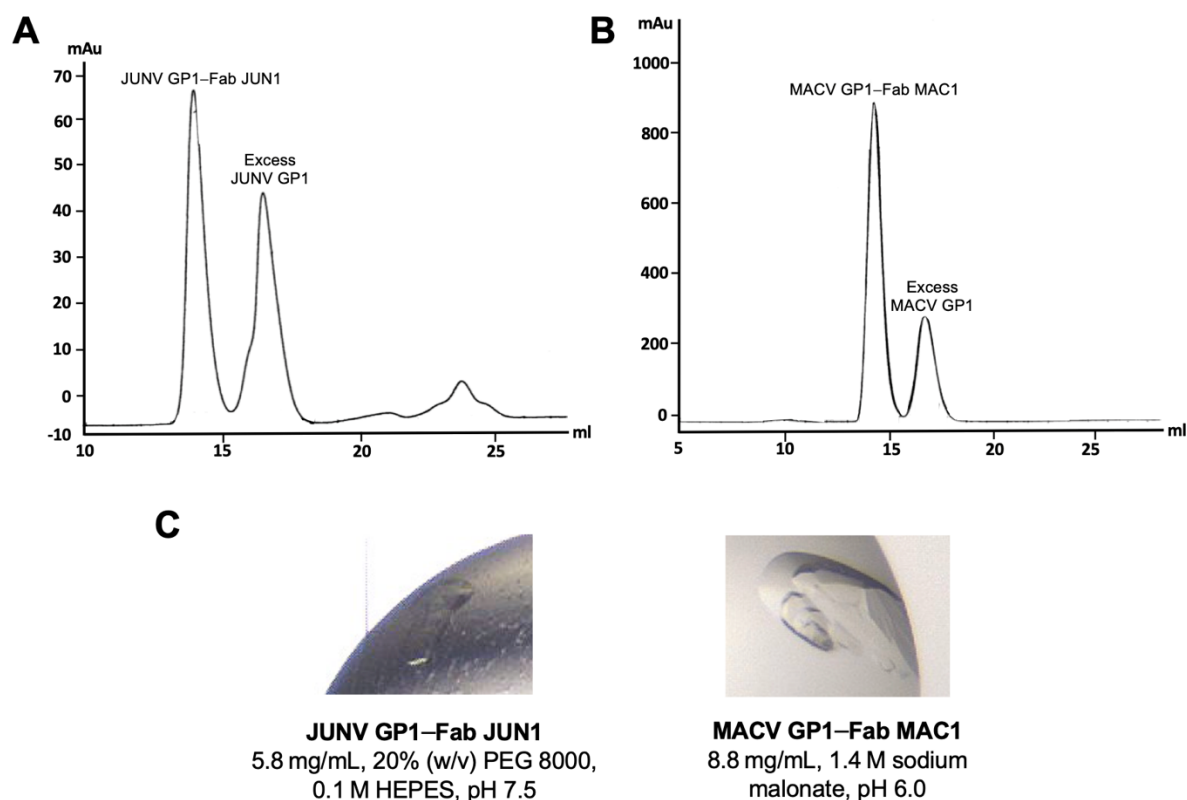


Fig. 4.1. Purification and crystallisation of JUNV GP1–Fab JUN1 and MACV GP1–Fab MAC1 complexes. (A) Size exclusion chromatograms (Superdex™ 200 Increase 10/300 GL column) of deglycosylated JUNV GP1–Fab JUN1. Deglycosylated JUNV GP1 and Fab JUN1 were mixed at 1:1:1 molar ratio and allowed to incubate for 30 min. The mixture was concentrated and then subjected to SEC. Peaks corresponding to the deglycosylated JUNV GP1–Fab JUN1 complex and excess deglycosylated JUNV GP1 are labelled accordingly. Sample fractions corresponding to the GP1–Fab complex peak were pooled and concentrated for crystallisation experiment. (B) Size exclusion chromatograms (Superdex™ 200 Increase 10/300 GL column) of deglycosylated MACV GP1–Fab MAC1. Deglycosylated MACV GP1 and Fab MAC1 were mixed at 1:1:1 molar ratio and allowed to incubate for 30 min. The mixture was then subjected to SEC. Peaks corresponding to the deglycosylated MACV GP1–Fab MAC1 complex and excess deglycosylated MACV GP1 are labelled accordingly. Sample fractions corresponding to the GP1–Fab complex peak were pooled and concentrated for crystallisation experiment. (C) Images of protein crystals. The crystallisation conditions for JUNV GP1–Fab JUN1 and MACV GP1–Fab MAC1 are annotated accordingly.

Table 4.1. Crystallographic data collection and refinement statistics for JUNV GP1–Fab JUN1 and MACV GP1–Fab MAC1 complexes.

Data Collection Statistics	JUNV GP1–Fab JUN1	MACV GP1–Fab MAC1
Beamline	DLS I03	DLS I04
Wavelength (Å)	0.96859	0.97951
Space Group	<i>P</i> 2 ₁	<i>I</i> 222
Cell dimensions <i>a</i> , <i>b</i> , <i>c</i> (Å)	69.8, 66.8, 73.1	91.7, 100.7, 164.8
α , β , γ (°)	90, 103.5, 90	90, 90, 90
Resolution range (Å)	67.92–2.50 (2.54–2.50) ^a	50.37–1.91 (1.94–1.91)
Rmerge	0.131 (0.794)	0.118 (>1)
I/ σ (I)	7.1 (1.4)	7.3 (1.0)
CC _{1/2}	0.995 (0.742)	0.998 (0.571)
Completeness (%)	99.8 (96.2)	100 (99.4)
Multiplicity	6.5 (5.9)	6.8 (6.6)
Refinement Statistics		
Resolution (Å)	67.92–2.50	50.37–1.91
No. reflections	22,798	59,314
<i>R</i> _{work} / <i>R</i> _{free}	0.223/0.254	0.190/0.225
No. atoms		
Protein	4,575	4,584
Ligand	54	207
Water	37	358
Average B-factors		
Protein	74.9	45.2
Ligand	93.7	68.7
Water	69.7	46.8
Ramachandran (%)		
Favored	96.9	97.6
Allowed	3.1	2.4
Outlier	0	0
Root mean square deviations (RMSD)		
Bond lengths (Å)	0.002	0.007
Bond angles (°)	0.499	0.875

^a The value for the highest-resolution shell is shown in parentheses.

4.3 STRUCTURE ANALYSIS

4.3.1 Structural Basis of JUN1-Mediated Neutralisation of JUNV GP1

Isolation and *in vitro* characterisation of JUNV and MACV GP1-specific monoclonal antibodies were performed by research collaborators, Dr. Katie Doores (King's College London) and Prof. Dr. Daniel Pinschewer (University of Basel). To briefly summarise, antigen-specific B cell sorting (CD4/CD8-B220⁺IgM⁻IgD⁻GP1⁺) were performed with splenocytes isolated from mice infected with BSL2-arenaviruses engineered to display the surface GP of JUNV and MACV. The sequences of the mAbs were recovered and recombinantly expressed. Due to the efficiency of the sort, seven anti-JUNV GP1 mAbs (JUN1–7) and one anti-MACV mAb (MAC1) were isolated. JUN1–7 and MAC1 exhibited strong neutralising potency against JUNV and MACV GP-pseudotyped viruses, respectively (**Fig. 4.2**). ELISA-based competition assay revealed that nAb JUN1–7 competes for the same binding site, suggesting that these nAbs recognise identical or overlapping epitopes on JUNV GP1 (**Fig. 4.2**). As nAbs JUN1–JUN7 neutralises JUNV GP1/GP2-pseudotyped viruses at a similar level, nAb JUN1 was selected as a representative to determine the structural basis for neutralisation. X-ray diffraction data of the JUNV GP1–Fab JUN1 complex were collected and refined to 2.5-Å resolution (**Table 4.1**) and the structure was solved by molecular replacement with PHASER using structures of JUNV GP1 and Fab OD01 (PDB ID: 5NUZ (Zeltina et al., 2017)) as search models. One JUNV GP1–Fab JUN1 complex was present in the asymmetric unit.

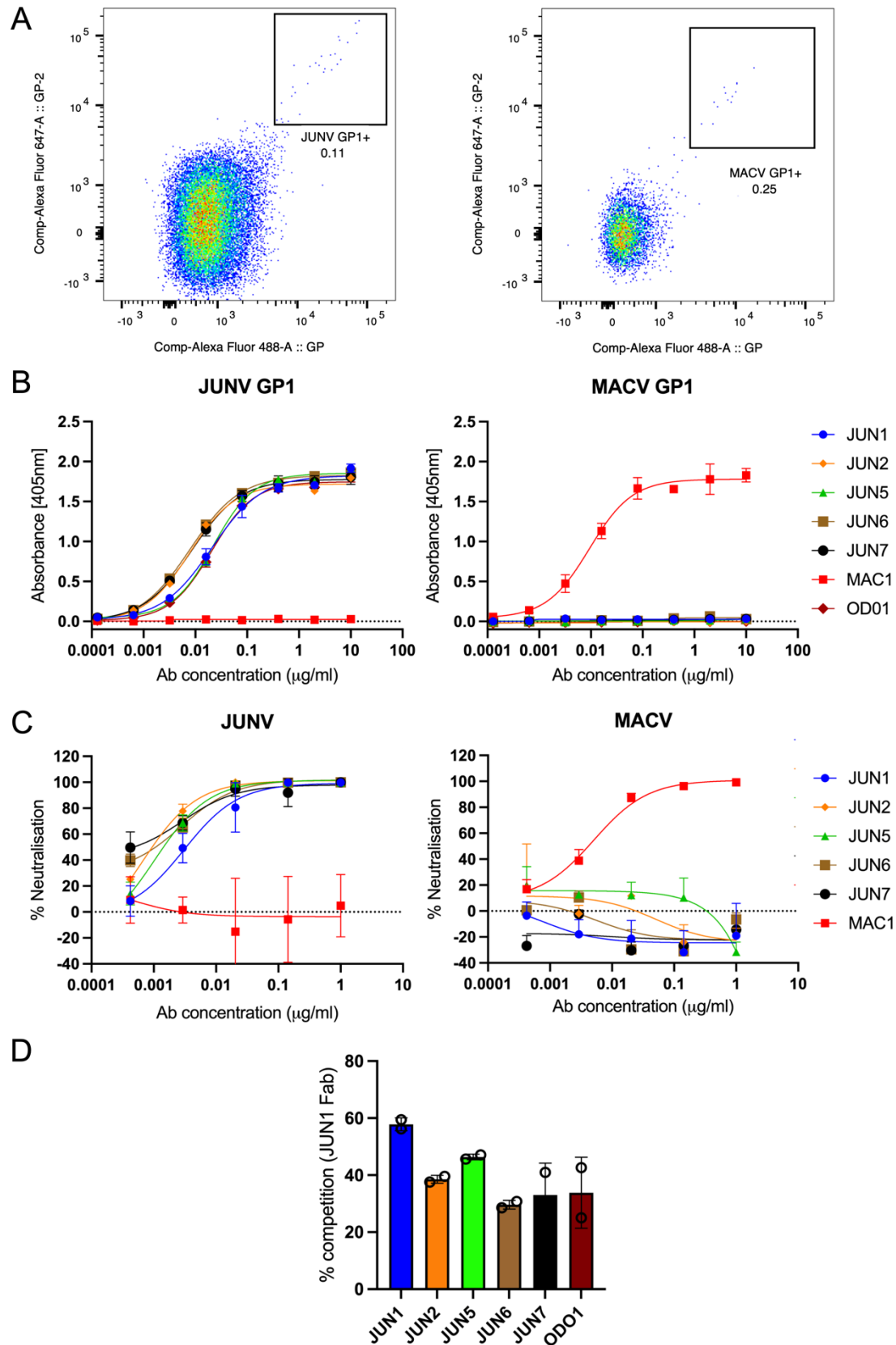


Fig. 4.2. Isolation and characterisation of JUNV and MACV GP1-reactive mAbs. (A) Antigen-specific B cell sorting was used to isolate JUNV and MACV GP1 specific nAbs. Figure shows CD4/CD8-B220⁺IgM-IgD⁻ cells that are either JUNV GP1 or MACV GP1

specific. (B) Binding of mAbs JUN1, 2, 5, 6, 7, MAC1, and OD01 to JUNV and MACV GP1 determined by ELISA. MAbs JUN3 and JUN4 were omitted as sequence analysis showed that their CDR sequences are identical to JUN2 and JUN5. Each data point is an average calculated from the results of $n=2$ biologically independent experiments. The standard deviation of each point is shown by the error bar. (C) Neutralisation by mAbs JUN1, 2, 5, 6, 7, and MAC1 against JUNV and MACV GP1/GP2-pseudotyped viruses. Each data point is an average calculated from the results of $n=2$ biologically independent experiments. The standard deviation of each point is shown by the error bar. (D) Competition between nAbs JUN1, 2, 5, 6, 7, OD01, and Fab JUN1 for binding to JUNV GP1 determined by ELISA. The results of $n=2$ biologically independent experiments are shown. Note that the work included in this figure were performed by Dr. Katie Doores, Jeffrey Seow, and Stefanie Krumm (King's College London) and Prof. Dr. Daniel Pinschewer and Mehmet Sahin (University of Basel).

Consistent with the previously reported structures (Clark et al., 2018; Mahmutovic et al., 2015; Zeltina et al., 2017), JUNV GP1 (Asp87–Asn232) forms a compact α/β fold consisting of a central seven-stranded β -sheet that constitutes the receptor-binding site (RBS) on one face and three α -helices on the opposing face (**Fig. 4.3A** and **4.3B**). Structural overlay analysis reveals that the JUNV GP1 exhibits a high level of structural conservation when superimposed with JUNV GP1 in previously reported Fab-bound structures, with an average of 0.6-Å RMSD over 136 equivalent C α residues.

Fab JUN1 binds to JUNV GP1 at a site previously observed for Tfr1 recognition through the use of all six complementarity-determining regions (CDRs) (**Fig. 4.3B** and **4.3C**), in an extensive interaction that occludes approximately 1,550 Å² of solvent exposed surface area. Although the CDRs of both the heavy and light chains of Fab JUN1 make a similar contribution to the overall amount of surface area occluded in the GP1–JUN1 interface, the CDR3 of the heavy chain (CDR H3) dominates the interaction by extending a 16-amino acid loop into a central pocket formed by the β -sheet and loop 3 of JUNV GP1 (**Fig. 4.3B** and **4.3C**). Within this loop, Tyr100A [Chothia numbering scheme (Martin and Thornton, 1996)] inserts its side chain hydroxyl group into the pocket and forms hydrogen bonds with the Ser111 and Asp113 located at the tip of the third β -strand of the JUNV GP1 (**Fig. 4.3C**).

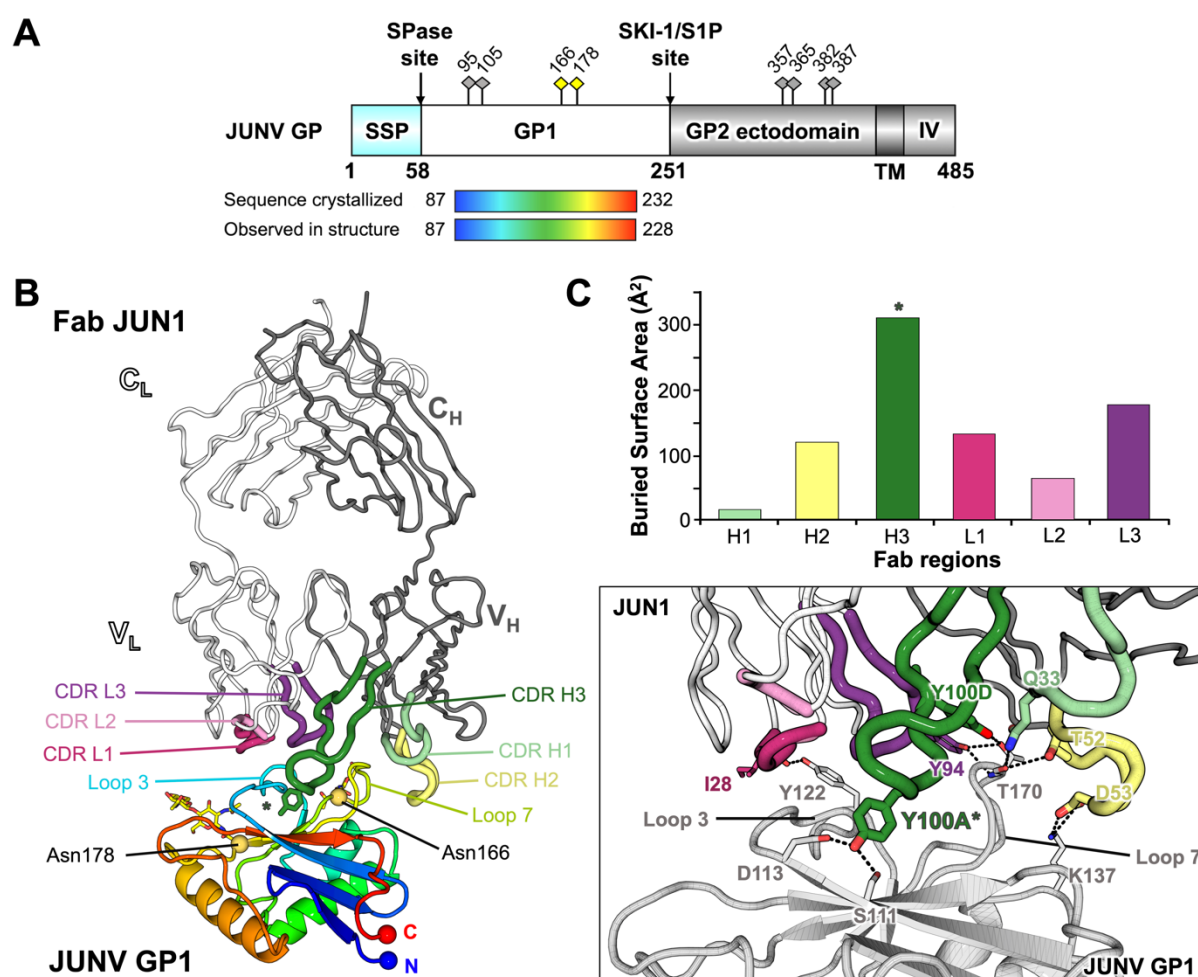


Fig. 4.3. Structural basis for Fab JUN1 recognition of JUNV GP1. (A) Schematic domain organisation of the JUNV GP (generated with DOG software (Ren et al., 2009)). The JUNV GP1 construct used for crystallisation is highlighted as a rainbow. The stable signal peptide (SSP), signal peptidase (SPase) site, GP1 glycoprotein, subtilisin-like kexin protease 1–site 1 protease (SKI-1/S1P) cleavage site, GP2 glycoprotein, transmembrane region (TM), and intravirion domain (IV) are annotated. Putative N-linked glycosylation sites are labelled as diamond-shaped pins above the schematic, with sites observed in the crystal structure coloured yellow. (B) Structure of JUNV GP1 in complex with Fab JUN1. JUNV GP1 is shown as in cartoon representation and coloured as a rainbow ramped from blue (N-terminus) to red (C-terminus). Crystallographically observed N-linked glycosylation sites are labelled and indicated as orange sphere with the glycans shown as sticks. Fab JUN1 is shown in ribbon representation, with heavy chain coloured dark grey and light chain coloured white. CDR loops contributing to JUNV GP1 recognition are labelled and coloured light green (CDR H1), yellow (CDR H2), dark green (CDR H3), pink (CDR L1), light pink (CDR L2), and purple (CDR L3). The side chain of Tyr100A is shown as a stick and denoted with an asterisk. V_H , V_L , C_H , and C_L denote the antibody variable heavy, variable light, constant heavy, and constant light-chain domains, respectively. (C) Interaction between JUNV GP1 and Fab JUN1. (*Upper*) A bar chart representing the contributions of each JUN1 CDR towards JUNV GP1 binding; calculated using PDBePISA server (Krissinel and Henrick, 2007)) and measured in buried surface area ($\text{\AA}^2$). The bar corresponding to the CDR loop carrying Tyr100A is denoted with an asterisk.

(Lower) Close-up view of the JUNV GP1–Fab JUN1 interface. JUNV GP1 is coloured light grey and Fab JUN1 is coloured as in panel *B*. Intermolecular hydrogen bonds are highlighted with dashes and participating residues are shown as sticks.

Despite originating from a unique germline (**Table 4.2** and **Fig. 4.4**), Fab JUN1 employs a tyrosine-mediated mode of antigen recognition shared by the previously characterised nAbs, namely OD01 (Zeltina et al., 2017), GD01 (Mahmutovic et al., 2015), and CR1-28 (Clark et al., 2018). Indeed, the insertion of Tyr100A from nAb JUN1 CDR H3 into the central pocket of JUNV GP1 is mirrored by Tyr30B of the nAb OD01 light chain, and Tyr98 and Tyr106 of nAb GD01 and nAb CR1-28 heavy chains, respectively (**Fig. 4.5**). Furthermore, this mode of tyrosine-mediated GP1 engagement is highly reminiscent to the insertion of Tyr211 from hTfR1 in the MACV GP1–hTfR1 complex (Abraham et al., 2010) (**Fig. 4.5**). In line with previous comparisons of JUNV GP1-bound nAb structures, the identification of this antigen–antibody recognition mode by an infection-elicited nAb confirms TfR1 receptor mimicry as a potent and common means of antibody-mediated neutralisation of JUNV.

Table 4.2. Summary of structurally characterised anti-JUNV and anti-MACV nAbs.

NAb	Specificity	Method	Origin	Germline	Tyr211 ^{hTfR1} - Mimicking Tyrosine
GD01 (5EN2)	JUNV	Infection	Mouse	VK6-13*01 VH1-87*01	CDR H3–Tyr98
OD01 (5NUZ)		Infection	Mouse	VK3-2*01 VH1-5*01	CDR L1–Tyr30B
JUN1		Infection	Mouse	VK6-23*01 VH13-2*01	CDR H3–Tyr113
CR1-28 (5W1K)	JUNV (reacts weakly with MACV)	Vaccination	Human	VK1-5*01 VH3-33*01	CDR H3–Tyr106
MAC1	MACV	Infection	Mouse	VK10-96*01 VH1-80*01	–
CR1-07 (5W1M)	MACV/JUNV	Vaccination	Human	VK4-1*01 VH3-30*15	–

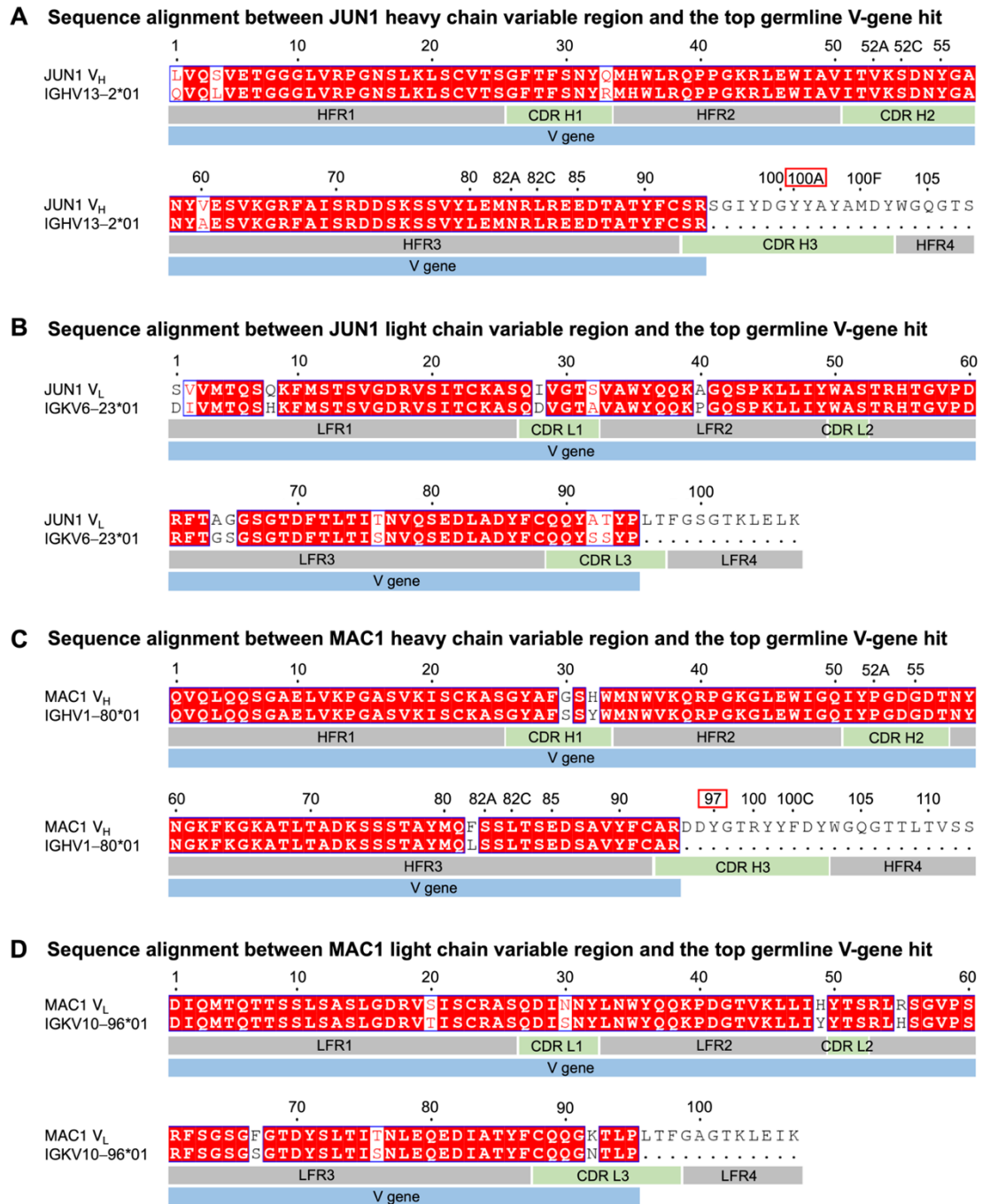


Fig. 4.4. Amino acid sequence alignments of Fab JUN1 and Fab MAC1 with respect to their closest corresponding germline sequences. Alignments of the variable domains from the (A) Fab JUN1 heavy chain, (B) Fab JUN1 light chain, (C) Fab MAC1 heavy chain, and (D) Fab MAC1 light chain are shown. Top germline variable (V)-gene hits were determined using the IgBLAST tool (<https://www.ncbi.nlm.nih.gov/igblast/>, (Ye et al., 2013)), with the heavy and light chain variable domain protein sequences of JUN1 and MAC1 used as queries. Sequence alignments were prepared using MultAlin (Corpet, 1988), plotted with ESPript (Robert and Gouet, 2014), and adjusted by hand. Amino acid numbers corresponding to JUN1 CDR H3 Tyr100A and MAC1 CDR H3 Tyr97 are highlighted with red boxes. HFR and LFR denote heavy and light chain framework, respectively.

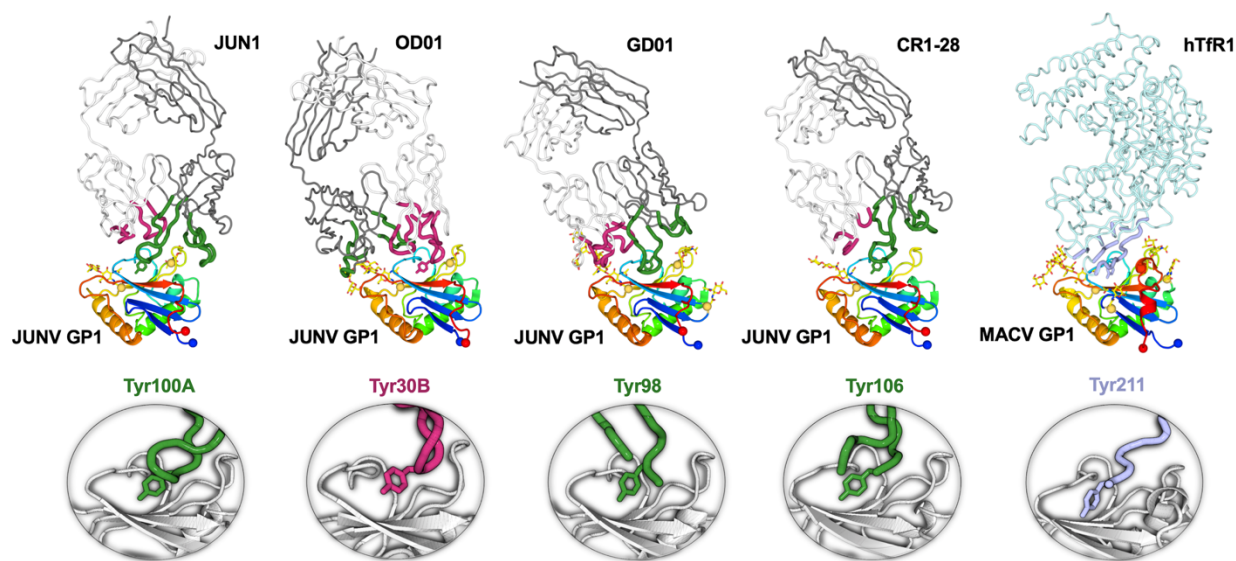


Fig 4.5. Tyrosine-mediated JUNV GP1 and MACV GP1 recognition: a common mode of mAb neutralisation. (*Upper Row*) JUNV GP1 and MACV GP1 in complex with Fab fragments of nAbs (JUN1, OD01 (Zeltina et al., 2017), GD01 (Mahmutovic et al., 2015), and CR1-28 (Clark et al., 2018)) and receptor (hTfR1 (Abraham et al., 2010)), respectively. GP1 is shown in cartoon representation and coloured as a rainbow from blue (N-terminus) to red (C-terminus). N-linked glycosylation sites are indicated as orange sphere with glycans shown as sticks. Heavy and light chains of Fab fragments are shown as dark grey and white ribbons, respectively. CDRs contributing to JUNV recognition are coloured dark green (heavy chain) and pink (light chain). HTfR1 is shown as a light blue ribbon with regions that interact with MACV shown as blue, thicker ribbons. (*Lower Row*) Close-up view of the key tyrosine residues (sticks) that are inserted into the central pocket of the GP1 (light grey cartoon). Only parts of CDR loops containing the tyrosine residues are shown. CDR residues that hinder the view of the tyrosine have been removed for clarity.

Although the insertion of Tyr100A from Fab JUN1 into the central pocket of JUNV GP1 represents the signature feature of the interaction, the antigen–antibody binding interface encompasses a larger footprint contributed by secondary interactions between loops 3 and 7 of JUNV GP1 and CDR H2, H3, L1, and L3 of Fab JUN1 (**Fig. 4.6**). Indeed, an extensive network of hydrogen bonds between Thr170^{JUNV GP1} and Glu171^{JUNV GP1} from loop 7 and Gln33^{CDR H1}, Asn58^{CDR H2}, Tyr100D^{CDR H3}, and Tyr94^{CDR L3} of Fab JUN1 can be observed. Residues Gln27^{CDR L1}–Ser32^{CDR L1} further provide substantial contacts with loop 3 of JUNV GP1, including hydrogen bonding between Ile28^{CDR L1} and Tyr122^{JUNV GP1}. Comparison of JUNV

GP1–Fab JUN1 complex with other Fab-bound structures reveals that Fab JUN1 engages the GP1 molecule at a relative angle of 10° deviation to that of Fab CR1-28, and increases to approximately 20° and 25°, with respect to Fabs GD01 and OD01, respectively (**Fig. 4.5**). Future structure elucidation of JUNV GP1 RBS-targeting nAbs will likely observe a plethora of flexibility displayed by the antigen–antibody interaction while preserving an antigen recognition mode involving the tyrosine-mediated receptor mimicry characteristic.

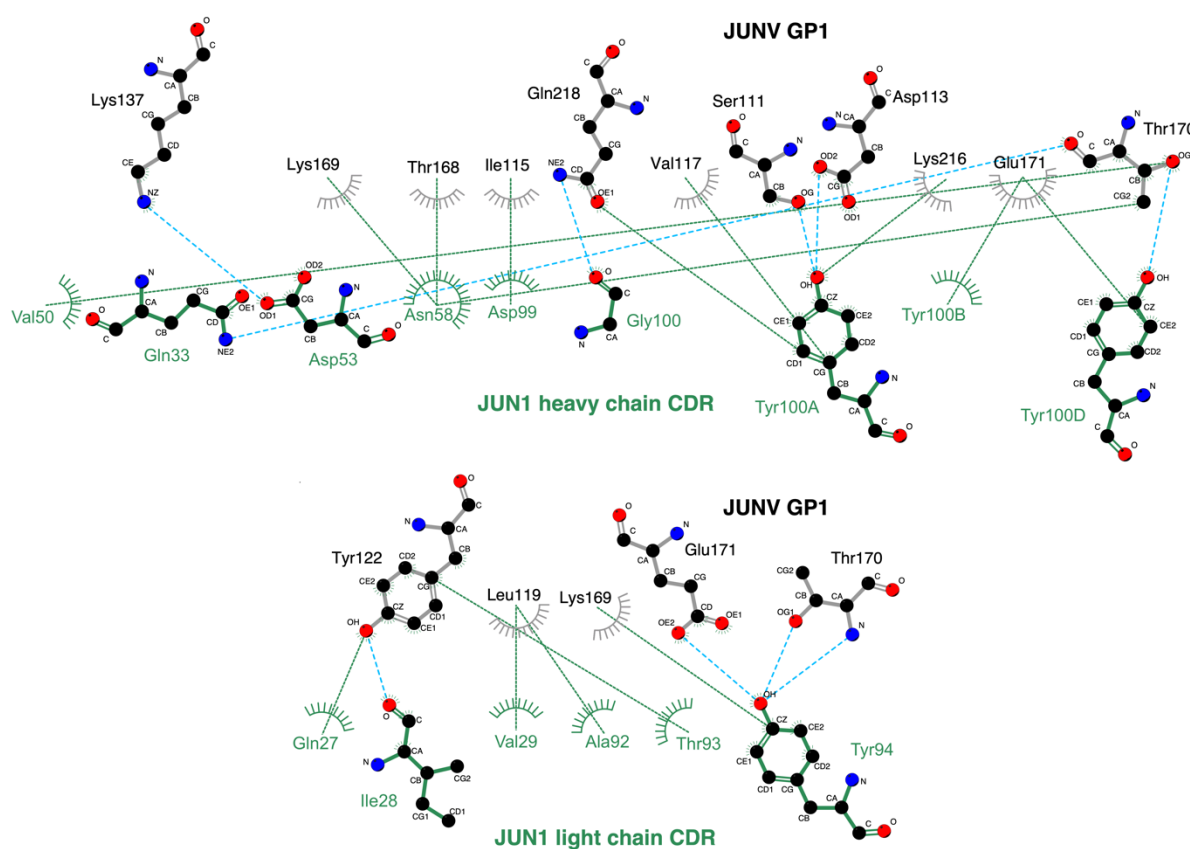


Fig. 4.6. Schematic diagram of contacts formed in the JUNV GP1–Fab JUN1 interface. (*Upper*) GP1–JUN1 heavy chain CDR interactions. (*Lower*) GP1–JUN1 light chain CDR interactions. JUNV GP1 residues are coloured grey and Fab JUN1 residues are coloured green. Atoms corresponding to carbon, nitrogen, and oxygen are shown as black, blue, and red balls, respectively. Residues involved in hydrogen bonding and salt bridges are shown as sticks; residues involved in hydrophobic interactions are shown as spoked arcs. Hydrogen bonds/salt bridges (hydrogen-acceptor and donor-acceptor distance range of 2.70–3.35 Å) and hydrophobic interactions (distance range of 2.9–3.9 Å) are shown as cyan and green dotted lines, respectively. Plots were generated with LigPlot+ (Laskowski and Swindells, 2011).

JUNV GP1 encodes four N-linked glycosylation sequons (Asn95, Asn105, Asn166, and Asn178) that encompass the periphery of the RBS. Electron density corresponding to at least one well-ordered N-linked GlcNAc moiety was observed at two of the four sequons: Asn166 and Asn178 but not at Asn95 and Asn105 (**Fig. 4.7**). Interestingly, electron density corresponding to two GlcNAc moieties was observed at Asn178, suggesting that the two GlcNAc residues were protected by the surrounding proteinaceous environment from endoF₁ digestion. Given that N-linked glycans was present at Asn105 in a previously crystallised JUNV GP1 produced under similar conditions (PDB ID: 5NUZ (Zeltina et al., 2017)), it is likely that glycan extending from this site is disordered in the crystal and hence unresolved in the electron density map. However, the absence of structural evidence for the glycan at Asn95 in this or any other reported JUNV GP1 structures suggests that this site may not be glycosylated to sufficient occupancy during protein folding, at least under the conditions used for recombinant protein expression. Close examination of the JUNV GP1–Fab JUN1 binding interface observes no detectable glycan–antibody interaction, suggesting that glycosylation is unlikely to play a role in the interaction (**Fig. 4.3C**). However, the proximity of these glycans to the JUN1 epitope at the RBS is consistent with their established role in shielding the GP surface and reducing neutralising potential, as has been observed for other anti-JUNV GP1 RBS nAbs, including GD01 and OD01 (Sommerstein et al., 2015).

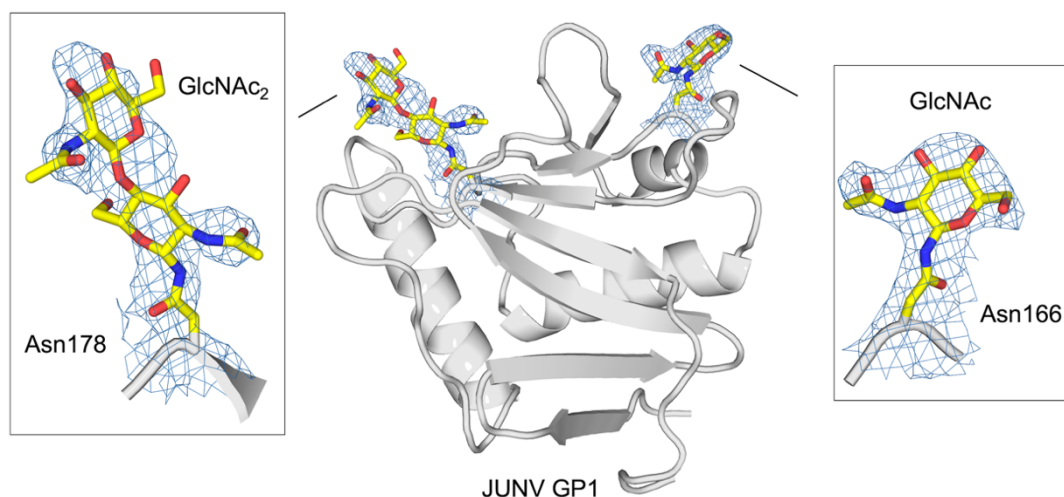


Fig. 4.7. Visualisation of crystallographically ordered N-linked glycans on JUNV GP1. JUNV GP1 is shown in cartoon representation and coloured light grey. N-linked glycans are shown as sticks with atoms corresponding to carbon, nitrogen, and oxygen coloured as yellow, blue, and red, respectively. Electron density (composite omit map contoured at 1.0 σ , blue) is shown for ordered N-linked glycans observed in the crystal structure of JUNV GP1–Fab JUN1. Close-up view of each N-linked glycan and its density are shown and annotated.

4.3.2 Structural Basis for MACV GP1 Recognition by nAb MAC1

MAC1 is currently the only nAb with a known sequence that is solely specific to and neutralises MACV (Table 4.2). To characterise the mode of MAC1 recognition, MACV GP1 was crystallised in complex with the Fab fragment of nAb MAC1 using a similar strategy described for the JUNV GP1–Fab JUN1 complex. X-ray diffraction data were collected and refined to 1.9-Å resolution (Table 4.1) and the structure was solved by molecular replacement using the structures of MACV GP1 (PDB ID: 2WFO (Bowden et al., 2009)) and Fab OD01 (PDB ID: 5NUZ (Zeltina et al., 2017)) as search models. One copy of the MACV GP1–Fab MAC1 complex was present in the asymmetric unit.

MACV GP1 (Glu87–Phe257) adopts the expected compact α/β fold (Fig. 4.8A and 4.8B) and exhibits a high level of structural conservation when superimposed with the unliganded form of MACV GP1 (0.7-Å RMSD; PDB ID: 2WFO (Bowden et al., 2009)), as well as MACV GP1 bound to hTfR1 (0.7-Å RMSD; PDB ID: 3KAS (Abraham et al., 2010))

and the MACV/JUNV cross-reactive Fab, CR1-07 (0.9-Å RMSD; PDB ID: 5W1M (Clark et al., 2018)) (**Fig. 4.8C**). Fab MAC1 interacts with the cognate GP1 molecule at the structurally observed hTfR1 binding site, occluding approximately 1,300 Å² of solvent accessible surface area at the interaction interface. Five of the six Fab MAC1 CDRs are involved in MACV GP1 recognition. Interestingly, the heavy chain contributes to the bulk of the interaction (**Fig. 4.9A**) with ~930 Å² of contact area, nearly triples that from the light chain CDRs (~350 Å²).

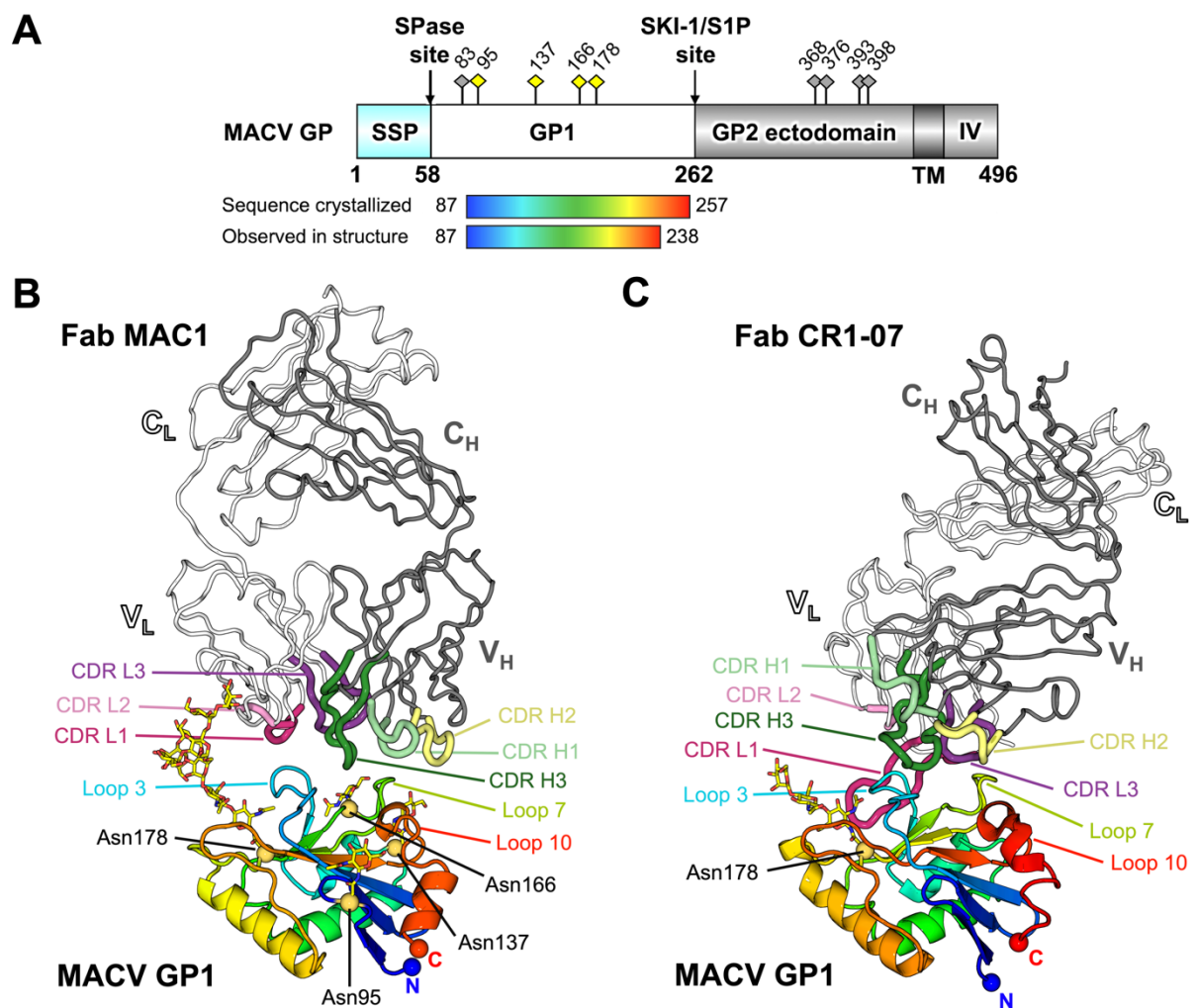


Fig. 4.8. Structural basis for Fab MAC1 recognition of MACV GP1. (A) Schematic domain organisation of the MACV GP (generated with DOG software (Ren et al., 2009)). The MACV GP1 construct used for crystallisation is highlighted as a rainbow. The stable signal peptide (SSP), signal peptidase (SPase) site, GP1 glycoprotein, subtilisin-like kexin protease 1–site 1 protease (SKI-1/S1P) cleavage site, GP2 glycoprotein, transmembrane region (TM), and intravirion domain (IV) are annotated. Putative N-linked glycosylation sites are labelled as

diamond-shaped pins above the schematic, with sites observed in the crystal structure coloured yellow. (B) Structure of MACV GP1 in complex with Fab MAC1. MACV GP1 is shown in cartoon representation and coloured as a rainbow ramped from blue (N-terminus) to red (C-terminus). Crystallographically observed N-linked glycosylation sites are labelled and indicated as orange sphere with the glycans shown as sticks. MAC1 is shown in ribbon representation, with heavy chain coloured dark grey and light chain coloured white. CDR loops are labelled and coloured light green (CDR H1), yellow (CDR H2), dark green (CDR H3), pink (CDR L1), light pink (CDR L2), and purple (CDR L3). V_H, V_L, C_H, and C_L denote the antibody variable heavy, variable light, constant heavy, and constant light-chain domains, respectively. (C) Structure of MACV GP1 in complex with Fab CR1-07 (PDB ID: 5W1M) (Clark et al., 2018). The structure is presented, coloured, and annotated as in panel B.

In contrast to receptor mimicry mode of antigen recognition utilised by the characterised JUNV-specific nAbs, namely the insertion of tyrosine residue into the central pocket of the GP1 glycoprotein (Abraham et al., 2010) (**Fig. 4.5**), this mode of interaction was not recapitulated in the MACV GP1–Fab MAC1 structure. Indeed, the 13-amino acid MAC1 CDR H3 sits shallowly and approximately equidistant between loops 3 and 7 of MACV GP1, while the central TfR1 binding pocket remains unoccupied (**Fig. 4.8B** and **4.9A**). Close examination of the Tyr97 [Chothia numbering scheme (Martin and Thornton, 1996)] from CDR H3 of Fab MAC1 revealed that the residue points towards the RBS pocket. However, unlike Tyr211^{hTfR1} which engages Ser113^{MACV GP1} within the pocket (Abraham et al., 2010), Tyr97^{CDR H3} from Fab MAC1 interacts with Val117^{MACV GP1}, Arg165^{MACV GP1}, and Phe226^{MACV GP1} at the rim of the pocket (**Fig. 4.9A**). Structural comparison with Fab CR1-07, a MACV/JUNV cross-reactive nAb (Clark et al., 2018), revealed that a tyrosine residue (Tyr106^{CDR H3}) of CR1-07 also points towards the pocket in a similar direction as Tyr97^{CDR H3} of MAC1 and interacts with Val117^{MACV GP1} (**Fig. 4.9B**), a residue that is also conserved in JUNV GP1 (Val117^{JUNV GP1}). This conserved interaction indicates that Val117^{MACV GP1} is an important target for both singly and cross-reactive nAbs.

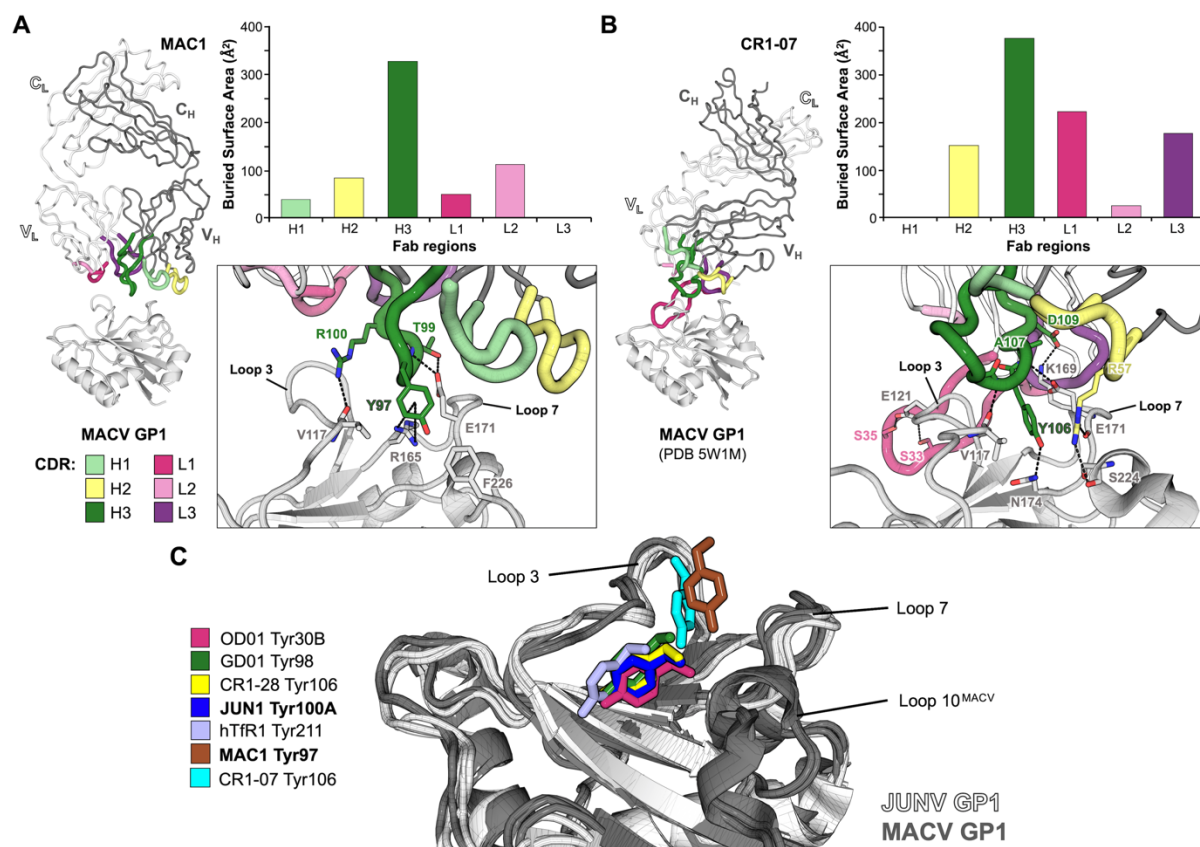


Fig. 4.9. Comparison of MACV GP1–Fab MAC1 and MACV GP1–Fab CR1-07 complex interfaces. (A) Interaction between MACV GP1 and Fab MAC1. (*Left*) MACV GP1 is shown as a light grey cartoon. Fab MAC1 is shown as dark grey (heavy chain) and white (light chain) ribbons, with CDR loops coloured as indicated. (*Upper Right*) A bar chart representing the contributions of each MAC1 CDR towards MACV GP1 binding; calculated using PDBePISA server (Krissinel and Henrick, 2007) and measured in buried surface area (Å²). (*Lower Right*) Close-up view of the MACV GP1–Fab MAC1 interface. Intermolecular hydrogen bonds are highlighted with dashes and participating residues are shown as sticks. (B) Interaction between MACV GP1 and Fab CR1-07. (*Left*) Structure of MACV GP1–Fab CR1-07 (PDB ID: 5W1M) (28), as presented in panel A. (*Upper Right*) A bar chart representing the contributions of each CR1-07 CDR towards MACV GP1 binding, calculated as in panel A. (*Lower Right*) Close-up view of the MACV GP1–Fab CR1-07 interface, as presented in panel A. (C) Overlay of signature CDR tyrosine residues (stick representation) from OD01 (Tyr30B, pink), GD01 (Tyr98, dark green), CR1-28 (Tyr106, yellow), JUN1 (Tyr100A, dark blue), hTfR1 (Tyr211, blue), MAC1 (Tyr97, brown), and CR1-07 (Tyr106, cyan) with respect to JUNV GP1 (white cartoon) and MACV GP1 (grey cartoon).

Superimposition of the crystal structure of MACV GP1 with that of JUNV GP1 revealed key protein architectural difference. Indeed, the presence of a disulfide-linked insert (loop 10) unique to MACV GP1 hinders the accessibility of Fab MAC1 and Fab CR1-07 to the

Tyr211^{hTfR1} pocket at the RBS (**Fig. 4.9C**), providing rationale for the absence of a receptor-mimicking mode of MACV GP1 engagement. Furthermore, structural overlay analysis of MACV GP1 with JUNV GP1s in nAbs-bound states revealed steric clashes between MACV GP1 loop 10 and Fab CDR regions of JUNV-specific nAbs (**Fig. 4.10**), supportive of the lack of MACV/JUNV cross-reactivity by nAbs GD01, OD01, and JUN1. While anti-JUNV nAb CR1-28 cross-reacts weakly with MACV GP1, it neutralises MACV poorly (Clark et al., 2018). Similarly, a previous study has shown enhanced neutralisation of MACV by mouse anti-JUNV GP antisera when MACV loop 10 was removed (Brouillette et al., 2017).

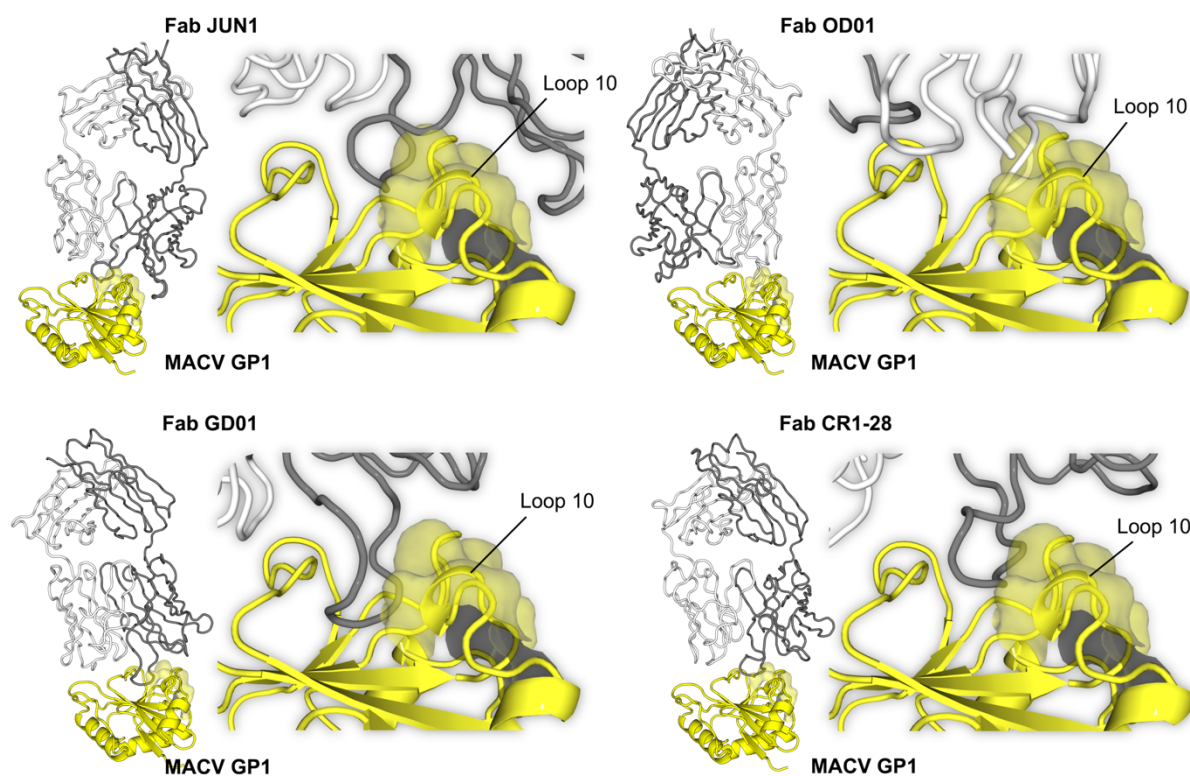


Fig. 4.10. Structure overlay analysis of MACV GP1 with Fabs JUN1, OD01, GD01, and CR1-28. MACV GP1–Fab MAC1 structure was superimposed with the structures of JUNV GP1 in complex with (A) Fab JUN1, (B) OD01, (C) GD01, and (D) CR1-28 by aligning the GP1 subunits. MACV GP1 is shown in cartoon representation and coloured yellow. Loop 10^{MACV GP1} is annotated and shown as a partially transparent surface. Fab fragments of nAbs JUN1, OD01, GD01, and CR1-28 are shown in ribbon representation, with heavy chain coloured dark grey and light chain coloured white. JUNV GP1 and Fab MAC1 are not shown. A close-up view of the steric clashes between loop 10^{MACV GP1} and Fab CDR is presented to the right of each MACV GP1–Fab overlay model.

The Fab MAC1 epitope is further stabilised through hydrogen bonding of paratope residues Thr99^{CDR H3} and Arg100^{CDR H3} with Val117^{MACV GP1}, Lys169^{MACV GP1}, and Glu171^{MACV GP1} (**Fig. 4.9A** and **4.11**). Trp33^{CDR H1} of Fab MAC1 also contacts residue Lys170^{MACV GP1}, with Asp54^{CDR H2} and Asp56^{CDR H2} forming salt bridges with Lys170^{MACV GP1} (**Fig. 4.11**). In contrast to these extensive heavy chain interactions, the Fab MAC1 light chain has a limited role in antigen recognition, with Tyr32^{CDR L1}, His49^{CDR L2}, Tyr50^{CDR L2}, and Arg53^{CDR L2} forming minor contacts with loop 3 residues Glu121^{MACV GP1} and Tyr122^{MACV GP1} (**Fig. 4.11**).

Like JUNV GP1, the crystallised MACV GP1 construct encodes four N-linked glycosylation sequons, which surround the periphery of the MACV GP1 RBS (**Fig. 4.8A** and **4.8B**): Asn95, Asn137, Asn166, and Asn178. Similar to previous structural studies of MACV GP1 in both unbound and hTfR1-bound states (Abraham et al., 2010; Bowden et al., 2009), electron density corresponding to at least a single GlcNAc moiety was observed at each of these sites (**Fig. 4.12**). Interestingly, the N-linked site extending from Asn178 constitutes an ordered chain of nine glycan moieties (Man₇GlcNAc₂), indicating that the di-N-acetylchitobiose core is protected by the surrounding proteinaceous environment from endoF₁ digestion. In contrast to the glycan-independent mode of interaction observed in the JUNV GP1–Fab JUN1 complex, the terminal mannose residues of the Asn178^{MACV GP1}-linked glycan chain contribute a footprint of ~120 Å² through interaction with residues at the C-terminal segment of the Fab MAC1 CDR L2, including Leu54^{MAC1}, Arg55^{MAC1}, Ser56^{MAC1}, Gly57^{MAC1}, Val58^{MAC1}, Pro59^{MAC1}, and Ser60^{MAC1} (**Fig. 4.12**). While the exact glycan composition and occupancy at Asn178^{MACV GP1} is currently unknown and warrants further investigation, the observed epitope is reminiscent to those observed on other highly glycosylated viruses such as HIV-1 Env (e.g. nAbs PGT128 and PGT135) (Doores, 2015), and is consistent with the

requirement for nAbs to penetrate the relatively dense glycan shield presented on the mammarenavirus surface (Sommerstein et al., 2015; Watanabe et al., 2018).

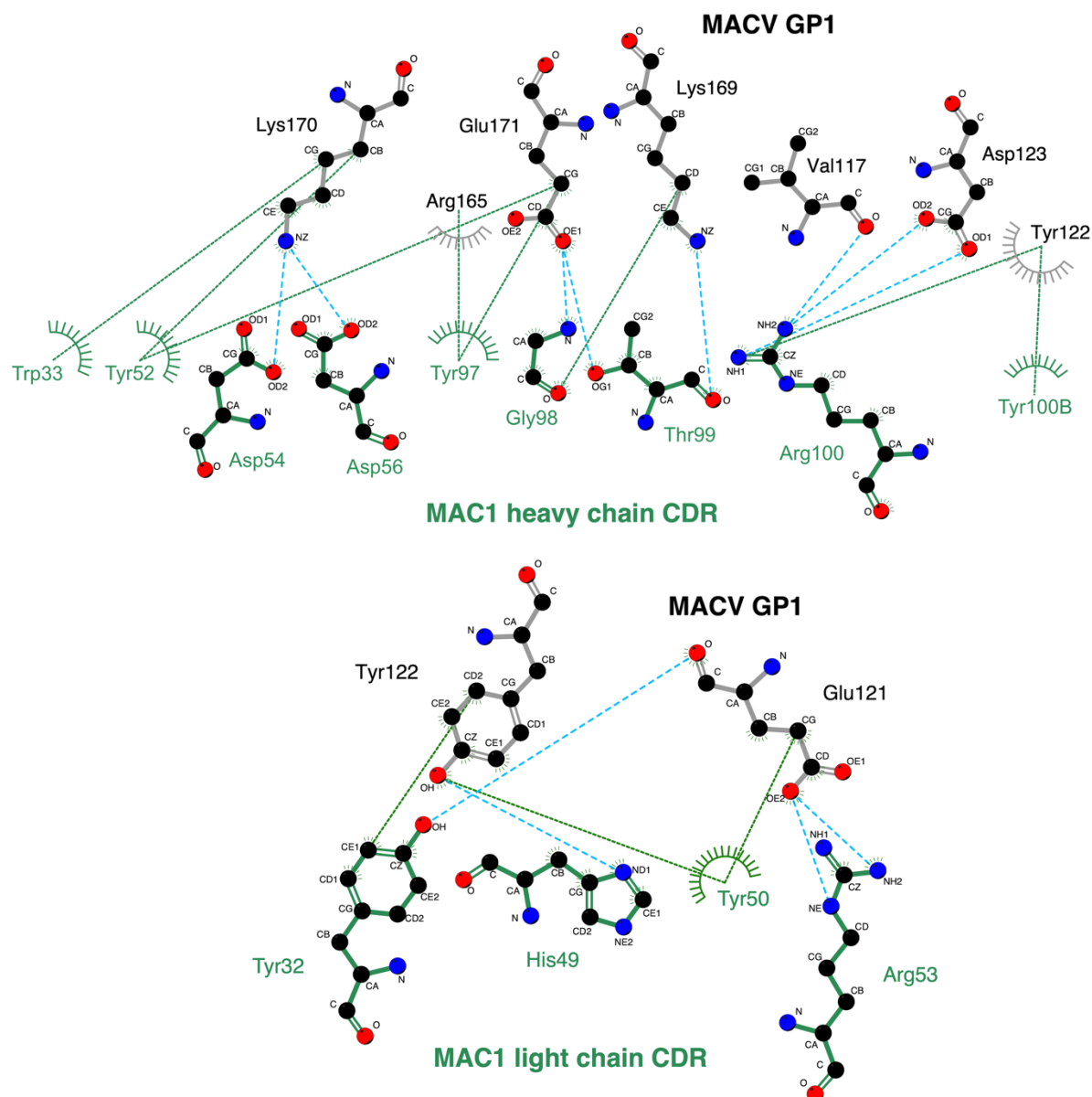


Fig. 4.11. Schematic diagram of contacts formed in the MACV GP1–Fab MAC1 interface. (Upper) GP1–MAC1 heavy chain CDR interactions. (Lower) GP1–MAC1 light chain CDR interactions. MACV GP1 residues are coloured grey and Fab MAC1 residues are coloured green. Atoms corresponding to carbon, nitrogen, and oxygen are shown as black, blue, and red balls, respectively. Residues involved in hydrogen bonding and salt bridges are shown as sticks; residues involved in hydrophobic interactions are shown as spoked arcs. Hydrogen bonds/salt bridges (hydrogen-acceptor and donor-acceptor distance range of 2.70–3.35 Å) and hydrophobic interactions (distance range of 2.9–3.9 Å) are shown as cyan and green dotted lines, respectively. Plots were generated with LigPlot+ (Laskowski and Swindells, 2011).

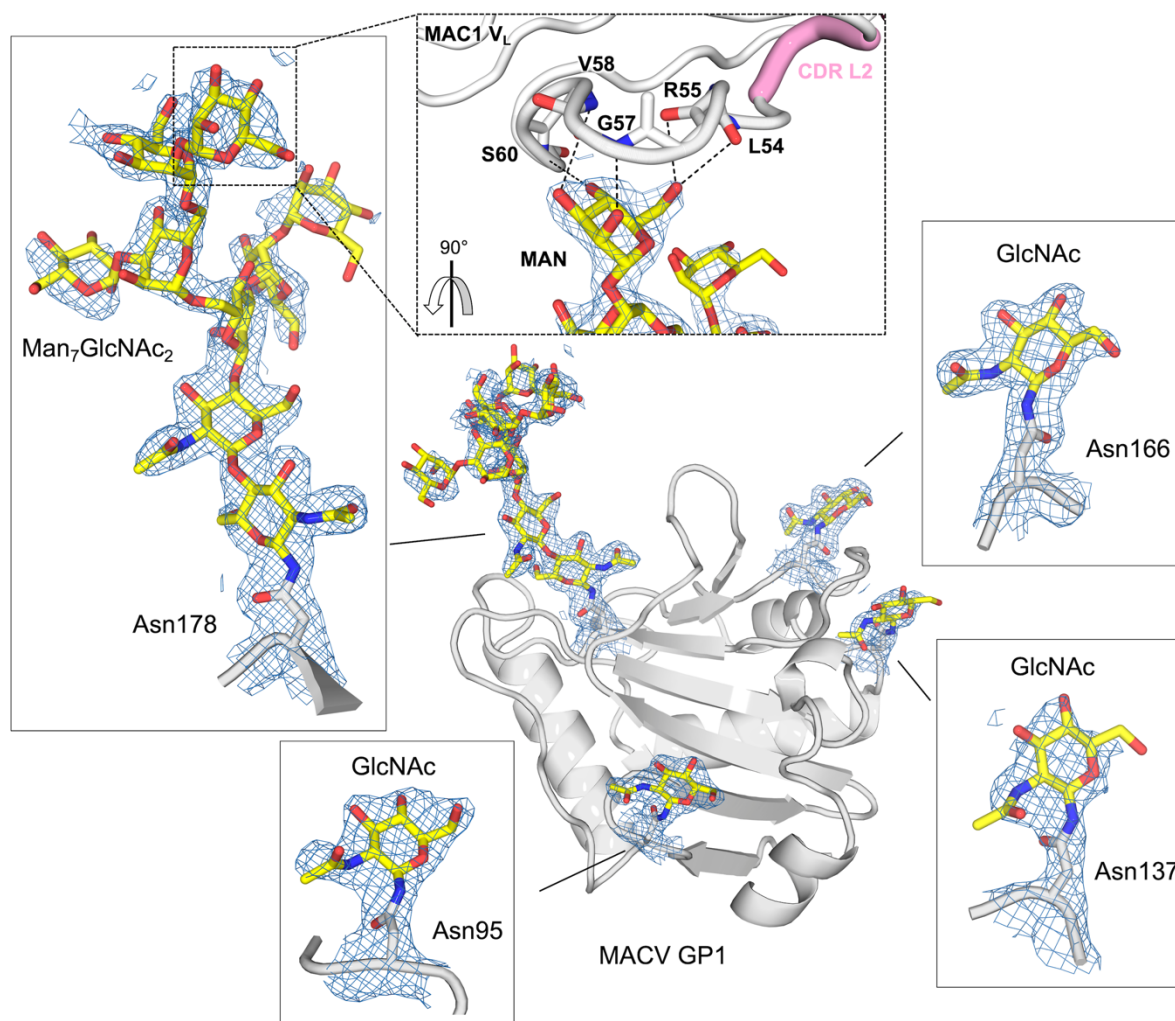


Fig. 4.12. Visualisation of crystallographically ordered N-linked glycans on MACV GP1. MACV GP1 is shown in cartoon representation and coloured light grey. N-linked glycans are shown as sticks with atoms corresponding to carbon, nitrogen, and oxygen coloured as yellow, blue, and red, respectively. Electron density (composite omit map contoured at 1.0 σ , blue) is shown for ordered N-linked glycans observed in the crystal structure of MACV GP1–Fab MAC1. Close-up view of each N-linked glycan and its density are shown and annotated, including a chain of $\text{Man}_7\text{GlcNAc}_2$ on $\text{Asn178}^{\text{MACV GP1}}$ that interacts with residues at the C-terminal region of the Fab MAC1 CDR L2. Intermolecular hydrogen bonds were identified by the PDBePISA server (Krissinel and Henrick, 2007) and are highlighted with dashes and participating residues are shown as sticks. Fab MAC1 is shown as white ribbon with CDR L2 in pink.

4.3.3 Comparison of Anti-MACV GP1 nAbs MAC1 and CR1-07

The MACV GP1–Fab MAC1 structure complements the MACV/JUNV cross-reactive Fab CR1-07 (Clark et al., 2018) as the only other reported MACV GP1-bound nAb structure. To glean insight of the molecular determinants that dictate nAb cross-reactivity at the NW GP1 RBS, comparison of the modes of antigen recognition employed by these two genetically distinct nAbs was performed (**Fig. 4.9**). Both Fab MAC1 and Fab CR1-07 recognise an overlapping binding site at the rim of the central TfR1-binding pocket that includes MACV GP1 loops 3, 7, and 10, and hence precludes TfR1 recognition. Moreover, both of the anti-MACV Fabs engage MACV GP1 with angles of approach that enable them to avoid steric clashes with loop 10, which is unique to MACV GP1 (**Fig. 4.8** and **4.9**).

This mode of GP1 recognition by MAC1 and CR1-07, however, are distinct and differentiated by their respective paratopes, where Fab CR1-07 contacts a significantly larger surface area of MACV GP1 that is conserved between MACV and JUNV GP1s (**Fig. 4.13**). Although sequence and structural variation between MACV and JUNV GP1s at the RBS exists, $\sim 590 \text{ \AA}^2$ of the total Fab CR1-07 footprint is conserved between MACV and JUNV, compared to $\sim 380 \text{ \AA}^2$ for Fab MAC1 (**Fig. 4.13**). This differential degree of GP1 engagement displayed by Fab MAC1 and Fab CR1-07 is largely facilitated by the relative contribution of the Fab heavy and light chain CDRs. Indeed, while Fab MAC1 relies heavily on its heavy chain for GP1 recognition, Fab CR1-07 utilises both the heavy and light chains equally. Thus, while structure overlay analysis does not reveal any apparent steric hindrance between Fab MAC1 and JUNV GP1, the relatively small amount of surface area that is conserved between MACV and JUNV at the Fab MAC1 epitope, likely contributes to the absence of observed cross-reactivity.

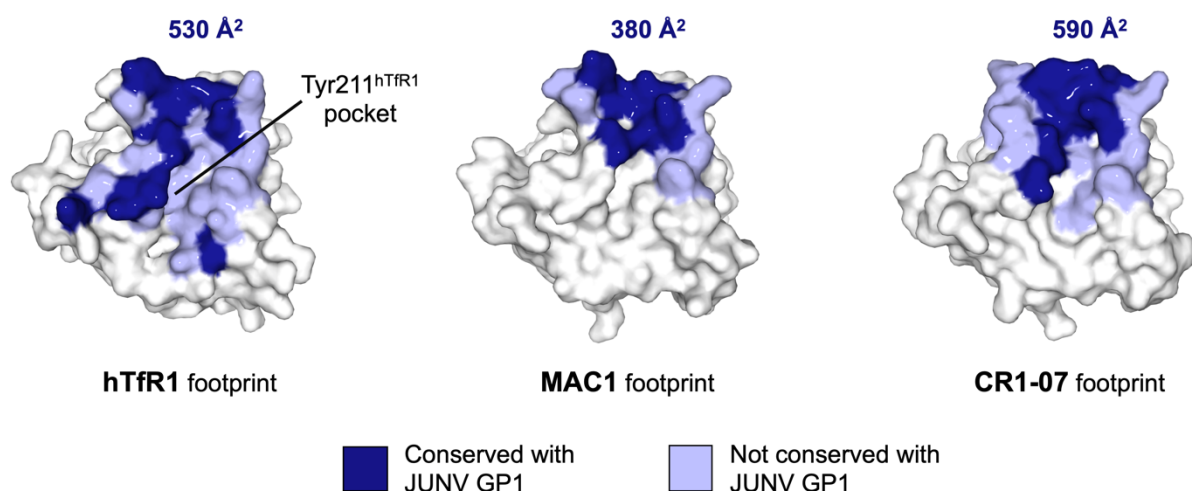


Fig. 4.13. Footprints of hTfR1, Fab MAC1, and Fab CR1-07 on MACV GP1. MACV GP1 is shown in surface representation and coloured white. The footprints of hTfR1 (PDB ID: 3KAS) (Abraham et al., 2010), MAC1, and CR1-07 (PDB ID: 5W1M) (Clark et al., 2018) on MACV GP1 are shown in blue: dark blue represents amino acid residues conserved (identical) between MACV GP1 and JUNV GP1; non-conserved residues are coloured light blue. The footprints exhibited on the solvent accessible surface area of MACV GP1 by hTfR1, MAC1, and CR1-07 that are conserved with JUNV GP1 were calculated with PDBePISA server (Krissinel and Henrick, 2007), measured in buried surface area ($\text{\AA}^2$), and are labelled accordingly.

4.4 SUMMARY AND DISCUSSION

JUNV and MACV, the etiological agents of AHF and BHF, respectively, put more than five million people at risk combined (Mills et al., 1994). Successful treatment of AHF and BHF by passive immunisation has proven that the generation of a nAb response is crucial for controlling JUNV and MACV infection (Enria et al., 1984; Maiztegui et al., 1979). Characterisation of the epitopes targeted by nAbs are therefore essential for the design of prophylactics and immunogens that protect against infection, and empower reverse vaccinology approaches (Burton, 2017) focused on the identification of the minimal subcomponents of the GP spike necessary for eliciting a protective immune response.

Here, eight infection-elicited nAbs JUN1-7 and MAC1 that potently neutralise JUNV and MACV, respectively, were isolated via antigen-specific B cell sorting (**Fig. 4.2**). Sequence analysis reveals that the CDRs of these nAbs are distinct from previously reported anti-NW

GP1 mAb CDR sequences and likely originate from different germ lines (**Table 4.2** and **Fig. 4.4**). Furthermore, structural investigation of nAbs JUN1 and MAC1 in this study demonstrates that both nAbs target the GP1 RBS, precluding TfR1 recognition, albeit through differential modes of antigen recognition (**Fig. 4.3** and **4.8**).

JUN1 engages the JUNV GP1 RBS by insertion of a tyrosine residue into the central pocket of the GP1, an interaction that mimics TfR1–NW GP1 recognition (Abraham et al., 2010) (**Fig. 4.3** and **4.5**). This mode of antigen recognition has been previously observed in the structures of other characterised anti-JUNV nAbs, namely OD01 (Zeltina et al., 2017), GD01 (Mahmutovic et al., 2015), and CR1-28 (Clark et al., 2018) (**Fig. 4.5**). While the aforementioned anti-JUNV nAbs are derived from different germline origins and approaches (**Table 4.2**), the varied means by which these genetically distinct antibodies converge upon a single immunological solution is consistent with previous proposals that the JUNV GP1 RBS constitutes a key site of vulnerability on the NW mammarenavirus surface (Clark et al., 2018; Mahmutovic et al., 2015; Zeltina et al., 2017).

Unlike JUNV-specific nAbs, MAC1 engages the MACV GP1 RBS without occupying the central tyrosine pocket but through the interaction with peripheral residues, including those encoded in loop 10, which are unique to MACV (**Fig. 4.8B** and **4.9A**). Consistent with the hypothesis by Clark *et. al.* 2018, structural analysis indicates that this elongated region of the molecule contributes to the antigenic distinctiveness of MACV GP1 from other NW GP1s and sterically hinders JUNV GP1-specific nAbs from interacting with MACV GP1. Indeed, the inability of anti-JUNV GP1 nAbs JUN1, OD01, and GD01 to cross-react with MACV GP1 may be rationalised by structure overlay analysis, which reveals major clashes between Fab-encoded CDR regions and loop 10 of MACV GP1 (**Fig. 4.10**). The differential modes of antigen recognition utilised by anti-JUNV and anti-MACV nAbs, therefore, reflects the extensive sequence and structural variation between JUNV and MACV GP1, which may have

arisen following co-evolution of the viruses with their individual rodent TfR1 orthologues (Choe et al., 2011).

Interestingly, despite not exhibiting any major clashes with JUNV GP1, MAC1 does not cross-react with JUNV GP1 (**Fig. 4.2**). Comparison of MAC1 with the MACV/JUNV cross-reactive nAb, CR1-07, reveals differences between the two overlapping epitopes that may explicate this different functionality, where the latter exhibits a more extensive footprint with a greater overall surface area that is conserved between the two NW mammarenaviruses (**Fig. 4.13**). This observation indicates that in addition to avoiding clashes with MACV GP1 loop 10, the ability to engage a sufficiently large and conserved surface is likely also an important determinant for MACV/JUNV cross-reactivity.

As discussed in Chapter 3, OW mammarenaviral GP1 has been reported to adopt two different conformational states, namely ‘GP2-free’ and ‘GP2-associated’ (Cohen-Dvashi et al., 2015; Hastie et al., 2017; Ng et al., 2020; Pryce et al., 2018). The presentation of antigenically distinct epitopes by OW GP1 in ‘GP2-free’ structural state, with respect to the virion-displayed GP1 in ‘GP2-associated’ state, renders the free subunit incompatible with cell surface primary receptor and nAb recognition. Given that JUN1, MAC1, and other immunisation-derived nAbs both recognise recombinantly produced NW GP1 in a ‘GP2-free’ state and neutralise NW mammarenaviruses *in vitro* (Clark et al., 2018; Mahmutovic et al., 2015; Zeltina et al., 2017), the data supports a previous study that demonstrates that recombinant NW GP1 is immunogenic and capable of eliciting nAbs (Borenstein-Katz et al., 2019) and the hypothesis that the conformations of monomeric JUNV and MACV GP1 are likely to resemble those existing on the mature trimeric NW mammarenaviral GP (Pryce et al., 2018).

The recent elucidation of the OW LASV GP ectodomain structure (Hastie et al., 2017) (‘GP2-associated’) presents an opportunity to predict how nAbs JUN1 and MAC1 recognise their respective NW GP1 in the context of the higher-order GP. Similar to previous hypotheses

(Clark et al., 2018; Hastie et al., 2017), structure overlay-derived models of trimeric JUNV-GP and MACV-GP demonstrate that the RBS is directed outwards and is compatible with TfR1 recognition (Fig. 4.14A and 4.14B). Similarly, the Fab fragments from the nAbs MAC1 and JUN1 may be concurrently accommodated into each of the three GP1 components of the trimeric GP. This modelling suggests that JUN1 and MAC1 target natively accessible surfaces presented on the mature arenaviral GP (Fig. 4.14C and 4.14D).

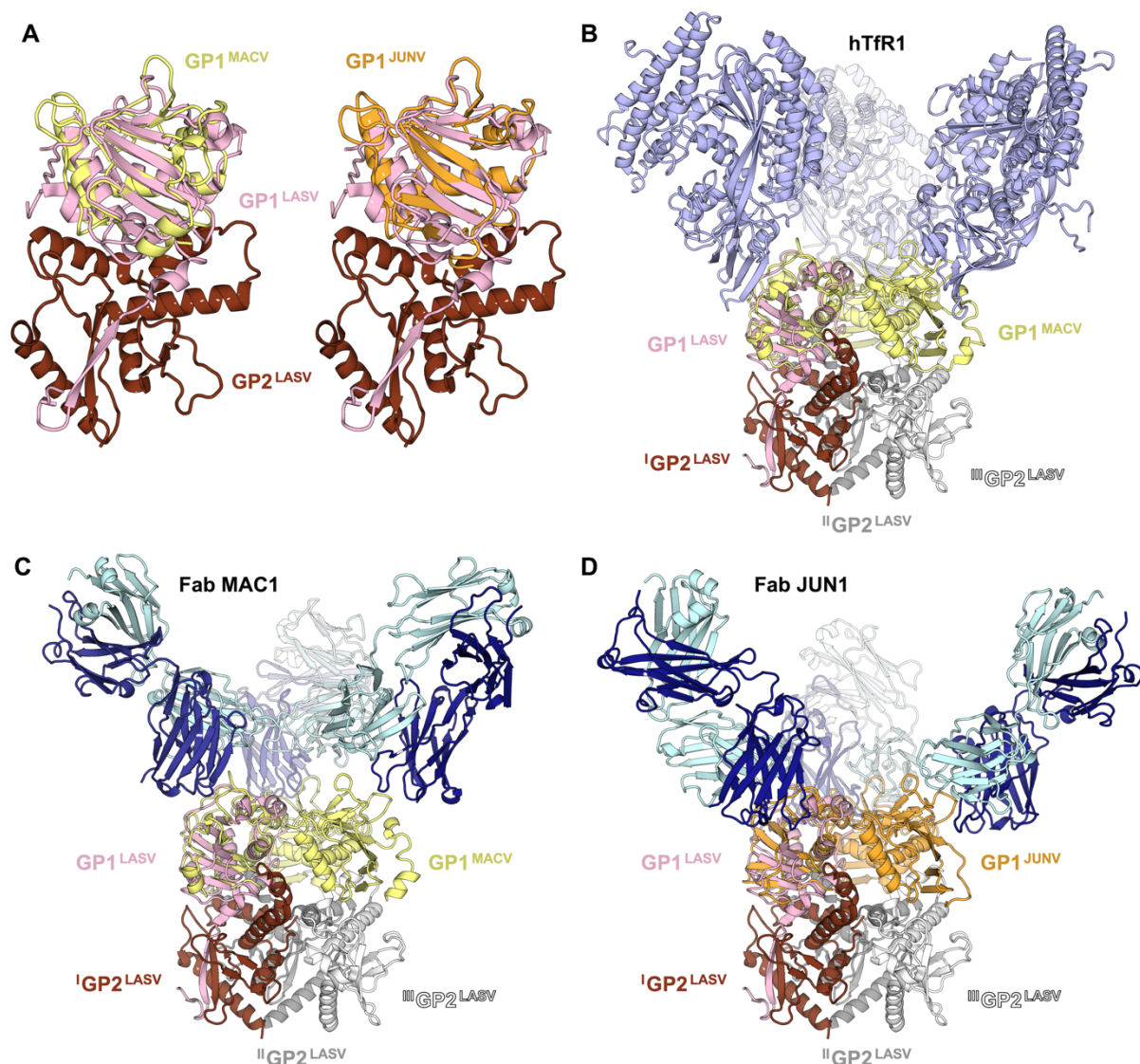


Fig 4.14. Models of JUNV GP1–Fab JUN1, MACV GP1–Fab MAC1, and MACV GP1–hTfR1 complexes in the context of the expected higher-order trimeric GP assembly. (A) The crystal structure of the LASV GP1–GP2 trimer constitutes the only known structure of a higher order arenaviral GP assembly and was used as a template to create models of trimeric NW GPs, as previously described (Clark et al., 2018; Hastie et al., 2017). JUNV

GP1–Fab JUN1 and MACV GP1–Fab MAC1 were overlaid onto the structure of LASV GP1–GP2 ectodomain (PDB ID: 5VK2) (Hastie et al., 2017) by aligning the GP1 subunits. All molecules are shown in cartoon representation. JUNV GP1 is coloured in orange, MACV GP1 in yellow, LASV GP1 in pink, and LASV GP2 in brown. Fab JUN1 and MAC1 are not shown. (B) MACV GP1–hTfR1 (PDB ID: 3KAS) (Abraham et al., 2010), (C) JUNV GP1–Fab JUN1, and (D) MACV GP1–Fab MAC1 were overlaid onto the structure of the trimeric LASV GP1–GP2 ectodomain by aligning the GP1 subunits as in panel A. A LASV GP1 subunit (pink) is shown to provide clarity of the MACV/LASV and JUNV/LASV GP1 superimposition. Only one LASV GP2 subunit is coloured brown, with the remaining two GP2 coloured light grey for better visualisation. LASV GP1, MACV GP1, and JUNV GP1 are coloured as in panel A; hTfR1 is coloured blue. Fab MAC1 and JUN1 are shown in cartoon representation, with heavy chain coloured dark blue and light chain coloured light blue.

The presence of N-linked glycosylation on the surface of viral glycoproteins has been shown to play an important role in shielding the virus from the antibody-mediated immune response arising to infection (Sommerstein et al., 2015; Watanabe et al., 2019). Although not to the same level as observed in OW mammarenaviruses (Pryce et al., 2018; Watanabe et al., 2018), NW mammarenaviruses are extensively glycosylated with between 21–48 N-linked glycosylation sequons per GP trimer (Sommerstein et al., 2015). While the composition and occupancy of these glycans on the mature NW GP remain to be characterised, previous studies have shown that they likely occlude much of the JUNV and MACV GP surface and encircle the GP1 RBS (Abraham et al., 2010; Bowden et al., 2009; Sommerstein et al., 2015; Zeltina et al., 2017). Consistent with N-linked glycosylation providing a limited contribution to the interaction between MACV GP1 and hTfR1 (Abraham et al., 2010), however, much of the GP1 RBS for both JUNV and MACV GP1 remains largely unshielded. It is likely that the exposure of this functionally important region of the NW GP surface may, in part, rationalise why all of the six structurally characterised nAbs target the GP1 RBS. When combined with analysis of glycan occupancy and composition, future studies based upon identifying and characterising GP2 or higher-order GP1–GP2 quaternary epitopes, such as those identified on LASV GP (Hastie et al., 2017; Robinson et al., 2016; Zeltina and Bowden, 2017), will likely aid in the

identification of other sites of vulnerability. Indeed, in-line with this aim, several GP2-specific mAbs have been recently identified (Amanat et al., 2018; Robinson et al., 2016; Ruo et al., 1991). Interestingly, these mAbs are cross-reactive against both OW and NW mammarenaviruses, likely owing to the relatively high level of sequence conservation across the GP2, as compared to the GP1, which is genetically and structurally more diverse (Crispin et al., 2016).

In summary, this study reveals the molecular basis for mAb-mediated neutralisation of JUNV and MACV by nAbs JUN1 and MAC1, respectively. Combined, this work provides a structural template for understanding the antigenic surface of the NW GP1 and augments efforts to develop GP-specific vaccines and synergetic combinations of nAb-based prophylactics that target the NW mammarenavirus surface.

5 STRUCTURE ELUCIDATION OF RABIES VIRUS GLYCOPROTEIN

5.1 INTRODUCTION

RABV is the prototypic member of the genus *Lyssavirus* of the family *Rhabdoviridae* and is the causative agent of human rabies (Walker et al., 2018). The virus is commonly transmitted through dog bites, infecting and replicating in the central nervous system, leading to severe neurological diseases and eventual death (Brunker and Mollentze, 2018). Although a RABV vaccine and RIG are available as PEP, an estimated 60,000 deaths occur each year, indicative of the disease as a major public health concern (Fooks et al., 2017). Furthermore, the high expense and complications associated with vaccine and RIG treatments highlight the need for exploring new therapeutic agents (Cantaert et al., 2019).

RV-G spike presented on the RABV virion is a class III fusion glycoprotein that orchestrates receptor binding and membrane fusion for host cell entry (Hellert et al., 2020; Yang et al., 2020). It has been shown that the protein, when induced by low pH, reversibly samples a series of conformational states from pre-to-postfusion structures (Backovic and Jardetzky, 2011). Despite sharing low sequence identities with VSV-G (~18%) and CHAV-G (~20%), similar reversible structural transitions have also been observed for these two glycoproteins (Baquero et al., 2017; Baquero et al., 2015; Roche et al., 2006; Roche et al., 2007) (**Fig. 1.15**). The crystal structures of monomeric RV-G ectodomain obtained at both basic and acidic pH revealed the structural rearrangement associated with the pre-to-postfusion transition (Hellert et al., 2020; Yang et al., 2020) (**Fig. 1.16**). However, the structural basis of how RV-G undergoes these rearrangements, particularly in the context of the native full-length trimeric glycoprotein, remains unknown and is key for understanding the conformational lability of the molecule.

As RV-G is the only protein exposed on the surface of the virion, it constitutes the main target of the host humoral immune response. The epitopes of hundreds of nAbs have been mapped onto five discrete regions of RV-G, termed major antigenic site I, II, III, IV, and a minor site ‘a’ (Benmansour et al., 1991; De Benedictis et al., 2016; Lafon et al., 1983; Nagarajan et al., 2008) (**Fig. 1.17A**). Two of these epitopes have been structurally characterised in studies of RV-G fragments and monomeric ectodomains (Hellert et al., 2020; Yang et al., 2020) (**Fig. 1.17B** and **1.17C**). The lack of structural information on the nAb quarternary epitopes, however, hinders the understanding for nAb-based RABV neutralisation and the design of nAb-based PEP cocktails.

Here, in collaboration with Dr. Alexander Douglas at Jenner Institute, we aim to illuminate the structure and functionality of RV-G, a key therapeutic target, as follows: (i) by revealing the higher-order architecture of RV-G in prefusion state to facilitate the production of stabilised prefusion RV-G construct for future investigation of RV-G–receptor interaction and as a vaccine candidate; (ii) by defining the exact epitopes targeted by nAbs and providing structure-guided rationales for the nAb-based PEP cocktails; and (iii) by clarifying the molecular basis of fusogenic rearrangements required to mediate fusion of host and viral membranes. In *toto*, this work will not only deliver a coherent structural understanding of RV-G–host interactions at the cell surface, but also provide the molecular-level blueprints necessary to augment ongoing initiatives focused on the development of next generation RABV vaccines and antibody-based antivirals.

5.2 SINGLE-PARTICLE CRYO-EM ANALYSIS OF RV-G IN COMPLEX WITH THE FAB FRAGMENTS OF NABS 17C7 AND 1112-1

5.2.1 Protein Production and Purification

Protein expression and purification were performed by Dr. Sofiya Fedosyuk and Luke Thorley from Douglas group at the Jenner Institute, University of Oxford. Briefly, Lenti651 cells stably expressing untagged full-length RV-G (Met1–Leu524) were pelleted from the growth media through centrifugation (1,000 rpm, 15 min). The RV-G was extracted from the cell membrane using 1% octyl-beta-glucoside (β -OG) detergent in an extraction buffer consisting of 50 mM HEPES pH 7.4, 150 mM NaCl, and protease-inhibitor. Purified Fab 17C7 (1 mg) containing a C-terminal twin-strep tag (WSHPQFEK)₂ was added, and the mixture was incubated at 4°C for 2 h. Following detergent extraction, the mixture was subjected to centrifugation (4,800 rpm, 30 min) to remove cell debris. Supernatant containing extracted RV-G–Fab 17C7 complexes was filtered and used for affinity purification.

All protein purification steps were performed at 4°C. Strep-Tactin®XT Superflow® column (IBA Lifesciences) was used for strep-tag affinity purification to capture RV-G–Fab 17C7 complexes from the supernatant. Next, amphipol exchange was performed to substitute the β -OG detergent by adding Amphipol A8-35 (Anatrace) at three times the amount of RV-G–17C7 complexes. In addition, a second untagged Fab, 1112-1 (200 μ g), was added to the RV-G–17C7–amphipol mixture and incubated at 4°C for 1 h. Note that the combination of Fab 17C7 and Fab 1112-1 enables the best distribution of particle orientation on cryo-EM grids, as determined and detailed in section 5.2.3.1. Subsequently, SEC was performed to obtain homogeneous sample of RV-G–Fab 17C7–Fab 1112-1 complexes. SDS-PAGE was used to monitor protein purity throughout protein purification (**Fig. 5.1**).

In addition to nAbs 17C7 and 1112-1, anti-RABV nAbs RVC20 and E559 were also tested in this study. All Fab fragments were prepared and purified prior to RV-G extraction and

purification. The twin-strep tagged Fab 17C7, Fab RVC20, and Fab E559 were expressed in Expi293F cells and purified using Strep-Tactin®XT column. On the other hand, the untagged Fab 1112-1 was obtained and purified by enzymatic digestion (Pierce™ Fab Preparation Kit, Thermo Fisher Scientific) of murine hybridoma-derived IgG.

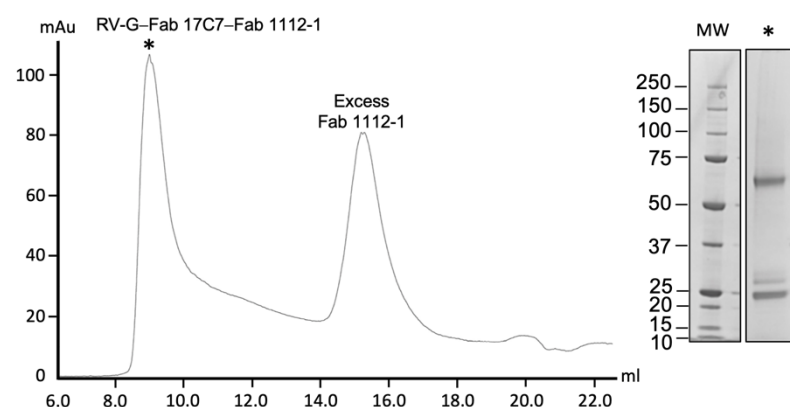


Fig. 5.1. Purification of full-length trimeric RV-G in complex with Fabs 17C7 and 1112-1. (Left) Size exclusion chromatogram (Superose 6 Increase 10/300 GL column (Sigma)) of RV-G in complex with Fabs 17C7 and 1112-1. Prior to SEC (buffer: 50 mM HEPES

pH 7.4, 100 mM NaCl), amphipol-solubilised RV-G–Fab 17C7 was incubated with 200 µg of Fab 1112-1 for 1 h. The first elution peak (marked with asterisk) corresponds to the RV-G–Fab 17C7–Fab 1112-1 complex. The second elution peak corresponds to excess Fab 1112-1. (Right) Reducing SDS-PAGE analysis of RV-G–Fab 17C7–Fab 1112-1 following SEC purification. The protein band at approximately 65 kDa corresponds to the monomeric RV-G while the Fab heavy and light chains are indicated by protein bands at approximately 25 kDa. MW denotes molecular weight (kDa) of the protein ladder and the asterisk (*) corresponds to the first peak in the SEC chromatogram.

5.2.2 Cryo-EM Grid Preparation and Data Collection

Cryo-EM grid preparation was performed using Vitrobot operated at 4°C and 100% chamber humidity. Purified RV-G–Fab 17C7–Fab 1112-1 sample was concentrated to 1.4 mg/mL. 4 µL of the protein solution was then carefully pipetted onto the glow-discharged grid (Quantifoil R 1.2/1.3 200 mesh) in the Vitrobot, followed by blotting (-15 N blot force, and 3 s blot time) with filter paper to remove excess sample solution. The grid was rapidly frozen by plunging into the liquid ethane, which was insulated and chilled by a reservoir of liquid nitrogen. The cryo-EM grids were clipped into AutoGrids and stored in cryo grid boxes in liquid nitrogen dewar prior to sample screening and data collection on the Glacios and Titan Krios microscopes.

For sample screening, cryo-EM grids were imaged using the in-house 200 kV Glacios TEM microscope equipped with Falcon III direct electron detector. In addition to grid quality including suitable ice thickness and minimal ice contamination, a number of criteria such as protein size, shape, purity, stability, concentration, distribution, and orientation were also monitored. For instance, initial screening and analysis indicated that RV-G–Fab complexes requires thicker ice for even particle distribution and orientation. Typically, an initial data set was collected for 18 h to enable preliminary data analysis. Cryo-EM images were recorded using linear mode, at defocus values ranging from -2.0 to -3.5 μm at a calibrated magnification of 92,000x, corresponding to a pixel size of 1.6 $\text{\AA}/\text{px}$. The total electron dose over 40 frames was 60 $\text{e}^-/\text{\AA}^2$.

Several important factors were evaluated during the Glacios preliminary data processing (see section 5.2.3). For instance, reference-free 2D classes were monitored to ensure several distinctive views of the protein particles were visible. Furthermore, the initial 3D reconstruction was assessed to determine if the distributions of particle orientation were sufficient to calculate a 3D structure with isotropic resolution. Identification of issues such as protein aggregation, too low/high concentration, and preferential orientation allowed modification and optimisation of the sample preparation conditions. Therefore, sample screening using the Glacios microscope was an iterative process until a satisfactory sample and/or grid preparation protocol was achieved (Passmore and Russo, 2016).

Image acquisition using the in-house high-end 300 kV Titan Krios microscope (equipped with K2 direct electron detector and a 20 kV GIF energy filter) was performed to obtain a larger data set with the aim of obtaining a high-resolution structure. For the RV-G–Fab 17C7–Fab 1112-1 sample, cryo-EM images were recorded for ~66 h using electron counting mode, at defocus values ranging from -0.8 to -2.6 μm at a calibrated magnification of 130,000x,

corresponding to a pixel size of 1.05 Å/px. The exposure time was 4.8 s, and the total electron dose was 45 e⁻ per Å² over 40 frames, with a dose rate of 9.39 e⁻/px/s.

5.2.3 Cryo-EM Single-Particle Data Processing

Cryo-EM SPA data processing was performed using cryoSPARC v2 and v3 software (Punjani et al., 2017). Global movie alignment and motion correction were executed to align the movie frames and produce averaged micrographs, which were then used for CTF estimation. Subsequently, the micrographs were examined to exclude those containing astigmatism and/or with CTF fitted to less than 10 Å. The selected micrographs were then used for reference-free particle picking using blobs with a diameter of 80–120 px. Picked particles were extracted with a box size (300 px) of at least 1.5 times the size of the particle. Extracted particles were subjected to reference-free 2D classification. Typically, 100–200 2D classes were generated. Particles from good 2D classes such as those with visible secondary structure features and distinctive views/orientations were chosen while the rest were discarded as junk particles. The representative 2D class averaged images were used as templates for automatic particle picking, and the picked particles were subjected to a further round of 2D classification. Here, the best 2D classes that displayed minimal background noise and varied orientation distribution were selected. Using the *ab initio* reconstruction function, particles from the selected 2D classes were used to generate 3–5 initial models, which then served as 3D templates for 3D classification, a processing step called ‘heterogeneous refinement’ in cryoSPARC. Particles were classified into each of the generated initial 3D models and subsequently refined using non-uniform refinement with both C1 and C3 symmetry to achieve a high-resolution map for each 3D class. Non-uniform refinement iteratively accounts for disordered and flexible densities in the map to improve the local resolution and overall map quality.

The maps were assessed for isotropic resolution, as particle preferential orientation problem typically causes a “smearing effect” in the calculated map due to the lack of information in one or more directions in the map (Tan et al., 2017). Furthermore, cryo-EM map resolution and quality were evaluated using the gold-standard Fourier shell correlation (FSC) curve. The FSC curve was calculated from the correlation between two independent maps obtained by splitting the data into half. By refining the two half maps independently, the calculated FSC curve will be free from spurious correlations, and therefore avoid or reduce the risk of overfitting noise (Scheres and Chen, 2012). To further improve the map resolution and quality, particles from the best refined 3D volume were used for per-particle motion correction and CTF refinement, followed by non-uniform refinement. Titan Krios cryo-EM data collection and refinement statistics are presented in **Table 5.1**.

Table 5.1. Cryo-EM data collection and refinement statistics for the RV-G–(Fab 17C7)₃–(Fab 1112-1)₂ complex.

Microscope	FEI Titan Krios
Voltage (kV)	300
Detector	Gatan K2
Magnification (nominal)	130,000
Pixel size (Å/px)	1.05
Dose rate (e ⁻ /px/s)	9.39
Frames per exposure	40
Exposure (e ⁻ /Å ²)	45
Defocus range (μm)	0.8–2.6 (0.3 step size)
Micrographs collected	10,887
Particles extracted / final	2,474,831* / 258,945

Symmetry imposed	C1
Map sharpening B-factor	172.9
Unmasked resolution at 0.143 FSC (Å)	4.50
Masked resolution at 0.143 FSC (Å)	3.88

*This number corresponds to the total number of particles extracted after automatic particle picking. The number of particles selected from 2D classes (after 2D classification) was 696,440.

5.2.3.1 Preliminary Glacios Data Set Processing

The crystal structures of monomeric prefusion RV-G ectodomain (Yang et al., 2020) and trimeric prefusion VSV-G ectodomain (Roche et al., 2007) revealed that G spike is approximately 80–90 Å tall and 80 Å wide (**Fig. 1.15** and **1.16**). To increase the protein size and shape of the globular molecule, several Fab fragments of nAbs were complexed with RV-G, making the protein contains more features and higher contrast for cryo-EM experiment. Amongst the tested nAbs included RVC20, 17C7, 1112-1, and E559.

(i) *RVC20*. RVC20 is a human bnAb that targets antigenic site I on RV-G, as indicated by the crystal structure (PDB ID: 6TOU) of scFv of RVC20 in complex with the PHD of RV-G (**Fig. 1.17B**), which is the most exposed domain of the prefusion spike (Hellert et al., 2020). The structure revealed that the amino acids included in the nAb RVC20 epitope are highly conserved (76%–100%), when compared to 1,412 unique full-length RV-G sequences in GenBank. Furthermore, nAb RVC20 has been shown to lock the RV-G in the prefusion or early intermediate conformation, as the RV-G–Fab RVC20 complex are extremely stable even at pH 5.5 (Hellert et al., 2020). Therefore, RVC20 constitutes an interesting and important target for structural studies.

Here, RVC20 was expressed as Fab fragments and complexed with full-length trimeric RV-G spikes in prefusion state. Cryo-EM images recorded using the Glacios microscope

revealed only one orientation view of the complex, indicating that the protein particles displayed a severe orientation bias. 2D classes further confirmed the presence of preferential orientation, as only the top view of the particle was detected (**Fig. 5.2**). Consequently, the anisotropic distribution of angular projection directions leads to the under-sampling of Fourier components in the 3D reconstruction. Indeed, the calculated 3D volume of RV-G–Fab RVC20 complex (**Fig. 5.2**) exhibits a “smearing effect” artifact due to the loss of structural information parallel to the axis of preferred orientation, which affects the interpretability of the cryo-EM map (Chen et al., 2019; Naydenova and Russo, 2017; Tan et al., 2017). The map reveals a 1:1 RV-G:Fab stoichiometry, exhibiting a C3 symmetry relationship. Fab RVC20 engages the PHD in a horizontal/flat fashion that is parallel to the membrane. This recognition mode likely contributes to the observed orientation bias, as the thin vitreous ice layer in the cryo-EM grids is unlikely to be able to accommodate the width of the side view of the molecule, which will require a relatively substantial ice thickness. Due to the lack of side view projection, the FD of a RV-G protomer and the amphipol regions are not visible or poorly resolved.

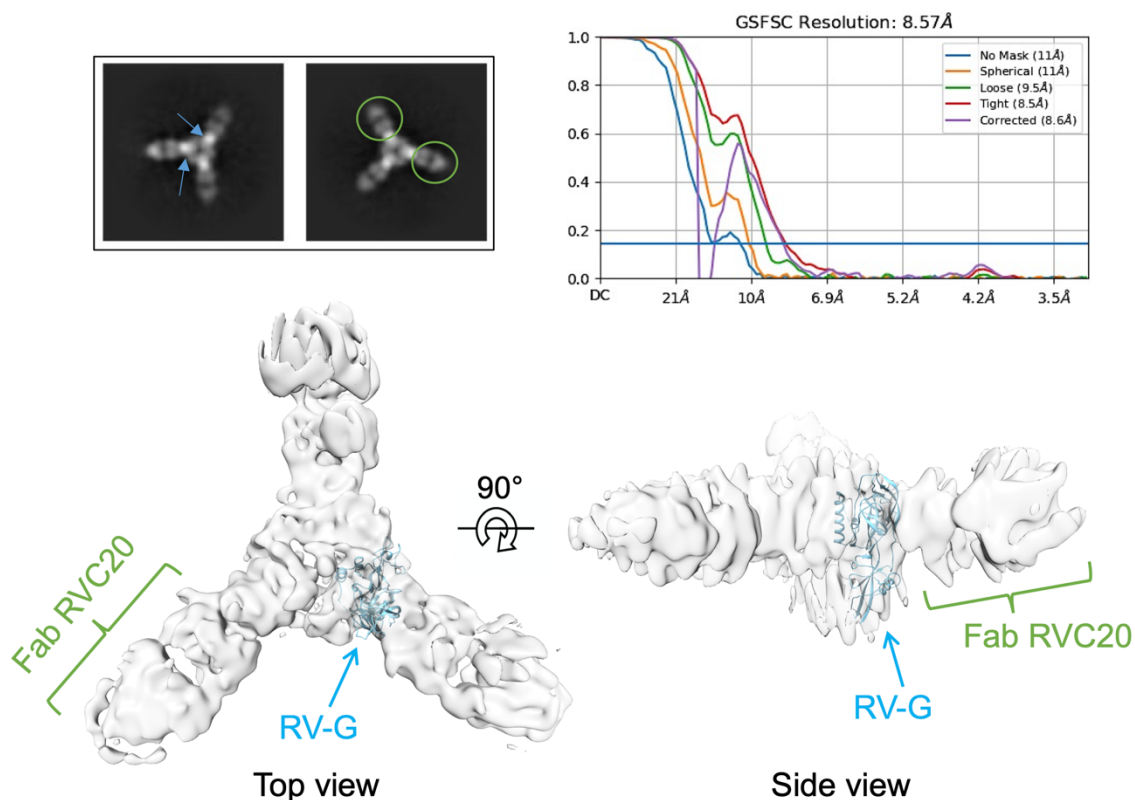


Fig. 5.2. Initial cryo-EM analysis of full-length RV-G in complex with the Fab of a neutralising mAb, termed RVC20. (Top left) 2D class averages derived following collection of an initial dataset on a Glacios 200 kV electron microscope. The positions of the Fab fragment and RV-G are indicated with green circles and blue arrows, respectively. (Top right) A gold-standard Fourier shell correlation (GSFSC) plot indicating an estimated resolution of approximately 8.6-Å for the initial dataset of approximately 80,669 particles using a gold-standard refinement procedure with a FSC = 0.143 criterion (blue horizontal line). Different FSC curves with no mask, spherical mask, loose mask, and tight mask applied to the half maps are also shown. The corrected FSC curve was calculated using the tight mask with correction by noise substitution. This accounts for correlation effects induced by masking. Data was processed with cryoSPARC. (Bottom) The resulting 3D reconstruction (no symmetry applied: C1), calculated with non-uniform refinement, with the crystal structure of the basic-pH RV-G ectodomain (PDB ID: 6LGX, truncated to Lys1–Ile373) fitted into the volume for illustrative purposes (fit using Chimera). The Fab fragments and RV-G are labelled and indicated with green and blue arrows, respectively. This initial map reveals that the epitope is located at the membrane-distal PHD of the RV-G spike. The map also exhibits “smearing” artifact, an indication of particle preferential orientation issue. Similarly, the resolution range at which the corrected and tight FSC curves diverge (a substantial dip in the corrected FSC curve) is suggestive of poor data quality at these resolutions, likely due to particle preferential orientation.

(ii) 17C7. The human nAb, 17C7, exhibited a potent and broad inhibition activity against 23 tested RABV isolates, and protected hamsters from a lethal RABV challenge in a well-established *in vivo* model of PEP (Sloan et al., 2007). Furthermore, nAb 17C7 has been shown to recognise a discontinuous, conformational epitope on the RV-G, which includes antigenic site III (Sloan et al., 2007). To better understand the molecular basis of nAb 17C7-mediated neutralisation of RABV, cryo-EM images of full-length prefusion trimeric RV-G in complex with Fab 17C7 were recorded using the Glacios microscope.

2D classes of the RV-G–Fab 17C7 complexes revealed both the top and side view projections of the molecule (**Fig. 5.3**), and the 3D reconstruction yielded a 7.14-Å cryo-EM map (**Fig. 5.3**). Fab 17C7 exhibits a 1:1 RV-G:Fab stoichiometry and engages the RV-G spike in an upward tilted fashion, slotting in between the RV-G protomers. Although the 2D classes and 3D map of RV-G–Fab 17C7 indicated a substantial improvement in particle orientation distribution, the top view of the molecule continued to dominate the reconstruction. This persistent orientation bias is reflected in the “smeary” cryo-EM map, in which the FD and the amphipol regions remain poorly resolved. Surfactant such as detergent has been proven to be a useful tool in ameliorating preferred particle orientation by coating the air-water interface, thereby reducing the adsorption of the protein particles at the interface and allowing particles to randomly orient within the vitreous ice layer. A number of detergents including β -OG, Tween20, and CHAPS were tested separately for the RV-G–Fab 17C7 sample. However, no improvement in the particle orientation and map quality was observed.

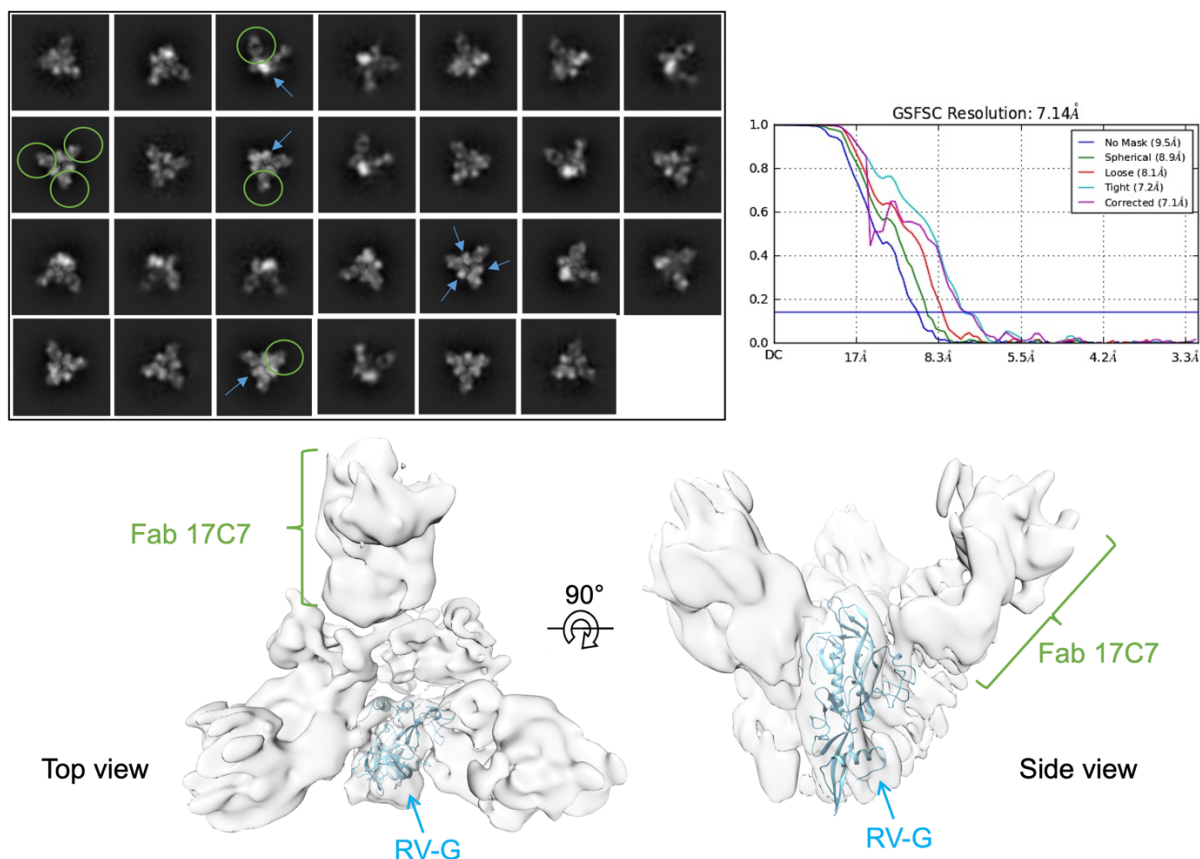


Fig. 5.3. Initial cryo-EM analysis of full-length RV-G in complex with the Fab of a neutralising mAb, termed 17C7. (Top left) 2D class averages derived following collection of an initial dataset on a Glacios 200 kV electron microscope. The positions of the Fab fragment and RV-G are indicated with green circles and blue arrows, respectively. (Top right) A gold-standard Fourier shell correlation (GSFSC) plot indicating an estimated resolution of approximately 7.1-Å for the initial dataset of approximately 164,821 particles using a gold-standard refinement procedure with a FSC = 0.143 criterion (blue horizontal line). Different FSC curves with no mask, spherical mask, loose mask, and tight mask applied to the half maps are also shown. Data was processed with cryoSPARC. (Bottom) The resulting 3D reconstruction (no symmetry applied: C1), calculated with non-uniform refinement, with the crystal structure of the basic-pH RV-G ectodomain (PDB ID: 6LGX, truncated to Lys1–Ile373) fitted into the volume for illustrative purposes (fit using Chimera). The Fab fragments and RV-G are labelled and indicated with green and blue arrows, respectively. This initial map reveals that the epitope is located at the membrane-distal region of the RV-G spike, in between the RV-G protomers. Despite an improvement as compared to the RV-G–Fab RVC20 sample, the map exhibits “smearing” artifact, an indication of particle preferential orientation issue.

(iii) *RVC20 and 17C7*. Superimposition of cryo-EM maps of RV-G–Fab RVC20 and RV-G–Fab 17C7 indicates that the two Fabs can bind to the RV-G spike simultaneously. Indeed, cryo-EM images and 2D classes of the sample containing RV-G spikes bound to Fab RVC20

and Fab 17C7 revealed a snowflake-like-shaped molecule, in which six Fabs concurrently bind to a trimeric RV-G spike with 1:1:1 RV-G:RVC20:17C7 stoichiometry (**Fig. 5.4**). However, unlike the RV-G–Fab 17C7 data set, side view projections are not visible in the RV-G–Fab RVC20–Fab 17C7 data, likely due to the flat and wide engagement mode of RVC20. Similarly, the FD and amphipol regions are poorly resolved in the reconstructed map (**Fig. 5.4**).

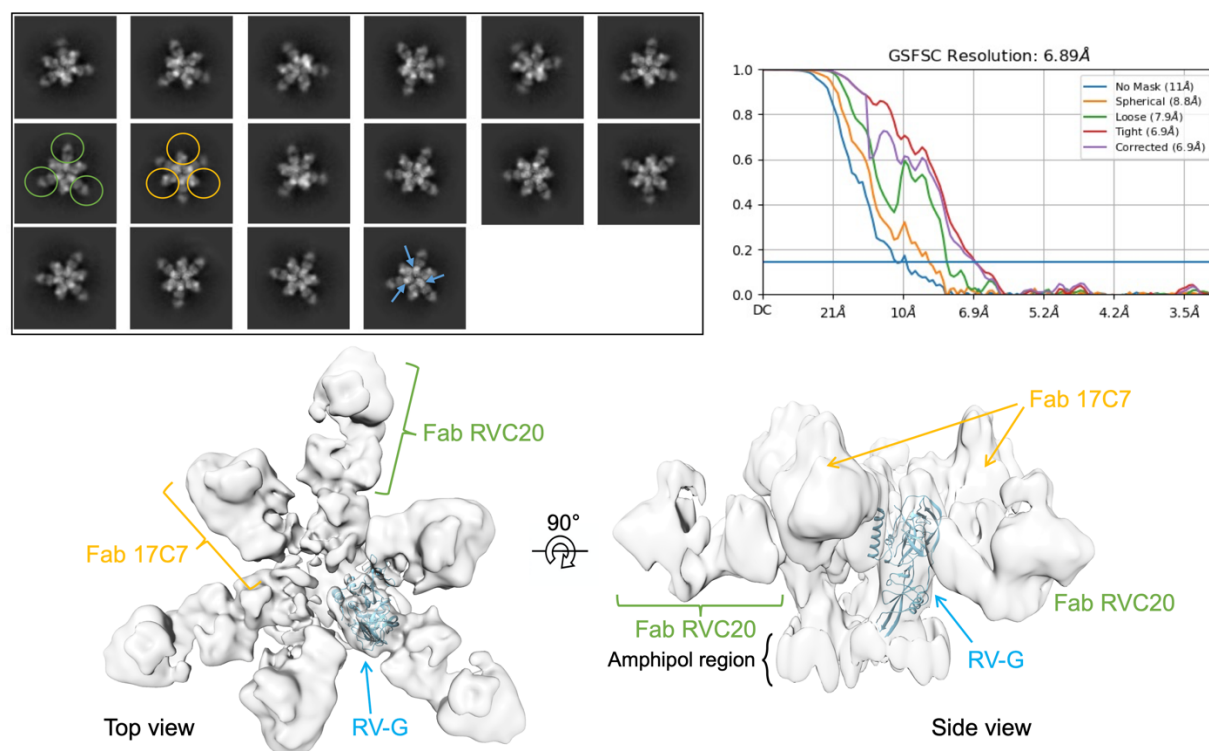


Fig. 5.4. Initial cryo-EM analysis of full-length RV-G in complex with Fabs RVC20 and 17C7. (*Top left*) 2D class averages derived following collection of an initial dataset on a Glacios 200 kV electron microscope. The positions of the Fab RVC20, Fab 17C7, and RV-G are indicated with green circles, yellow circles, and blue arrows, respectively. (*Top right*) A gold-standard Fourier shell correlation (GSFSC) plot indicating an estimated resolution of approximately 6.9-Å for the initial dataset of approximately 179,667 particles using a gold-standard refinement procedure with a FSC = 0.143 criterion (blue horizontal line). Different FSC curves with no mask, spherical mask, loose mask, and tight mask applied to the half maps are also shown. Data was processed with cryoSPARC. (*Bottom*) The resulting 3D reconstruction (C3 symmetry applied), calculated with non-uniform refinement, with the crystal structure of the basic-pH RV-G ectodomain (PDB ID: 6LGX, truncated to Lys1–Ile373) fitted into the volume for illustrative purposes (fit using Chimera). Fabs RVC20, 17C7, RV-G, and amphipol regions are labelled and indicated with green, yellow, blue, and black arrows, respectively. The addition of Fab RVC20 “flattens” the map, in which the volumes

corresponding to Fab 17C7 lack the height observed in *Fig. 5.3*. The map also exhibits “smearing” artifact, an indication of particle preferential orientation issue.

(iv) *1112-1 and 17C7*. RIG constitutes an important component of PEP. Typically, RIGs are derived from human and equine sera. However, the limited supply and health hazard associated with these resources have failed to provide a reliable supply in countries where canine rabies is endemic. Therefore, hybridoma-derived mouse mAbs that display potent and wide breadth of anti-RABV neutralisation activity are often assessed with the ultimate goal to develop a cheaper and efficacious alternative for the treatment of rabies in humans (Müller et al., 2009).

An antigenic site II-targeting mouse IgG nAb, 1112-1, has demonstrated potent inhibition of cell-to-cell virus spread *in vitro* and high protective activity *in vivo*, including 100% survival rate of animals (rats) that were subjected to RABV challenge, with no viral RNA detected in the brains of surviving animals (Dietzschold et al., 1992). Thus, combination of different nAbs in a PEP cocktail should include IgG 1112-1 (Müller et al., 2009). ELISA-based competition assay using nAb 1112-1 and Fab 17C7 showed that the two antibodies recognise non-overlapping epitopes on RV-G (**Fig. 5.5**). In order to delineate the molecular basis of 1112-1 and 17C7-mediated RABV neutralisation, the structure of Fab 1112-1 in complex with RV-G and Fab 17C7 was determined using cryo-EM.

2D class averages of RV-G–Fab 17C7–Fab 1112-1 revealed that the complex displayed multiple particle orientations (**Fig. 5.6**). The 3D reconstruction of these particles generated a 7.91-Å cryo-EM map. Notably, the map does not exhibit any apparent “smearing effect” artifact, and the map regions corresponding to the membrane-proximal FD and amphipol belt surrounding the transmembrane region are visible and better resolved than previous samples (**Fig. 5.6**). The map also shows a 1:1:1 binding stoichiometry of RV-G:17C7:1112-1. The site III-targeting Fab 17C7 slots in between the RV-G protomers while the site II-targeting Fab 1112-1 binds to the PHD of RV-G, suggesting that the 1112-1 epitope partially overlaps with

that of RVC20. Indeed, ELISA-based competitive assay using nAb 1112-1 and Fab RVC20 revealed that the two antibodies cannot bind to RV-G simultaneously (**Fig. 5.5**). Both Fab 17C7 and 1112-1 demonstrate an outward and upward directed binding mode. The combination of these two Fabs likely alters the shape and size of the molecule, enabling the presentation of most particle orientations within the thin vitreous ice layer. As the initial Glacios data set demonstrated promising sample and map quality, the combination of Fab 17c7 and Fab 1112-1 with RV-G was selected for high-resolution data collection using the Titan Krios microscope (see section 5.2.3.2).

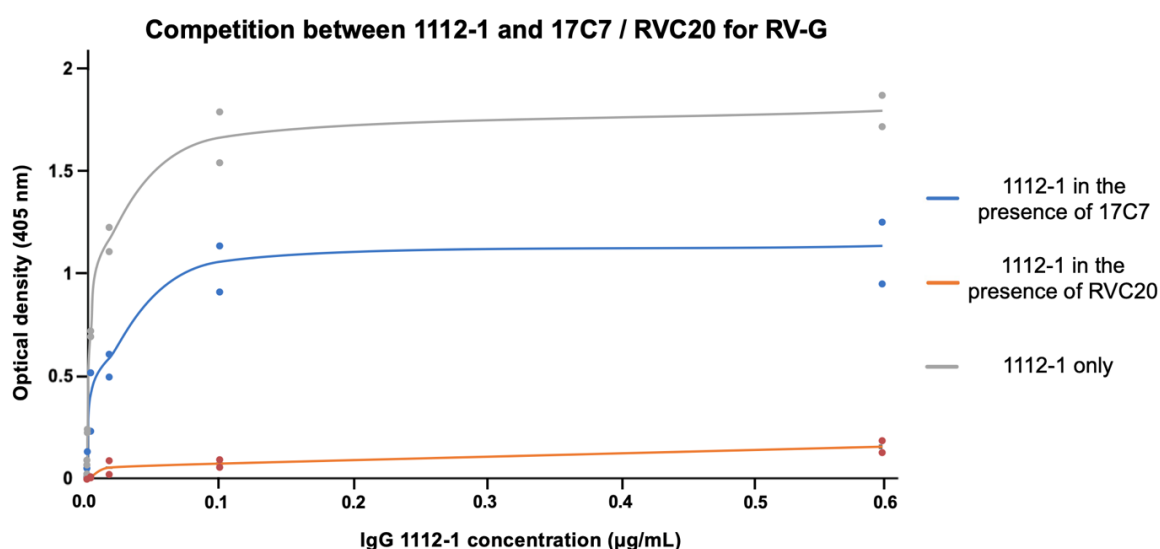


Fig. 5.5. Competition between IgG 1112-1 and Fab 17C7 or Fab RVC20 for binding to RV-G spike as determined by ELISA. RV-G (50µL at 1.0 µg/mL per well) was first immobilised on an ELISA plate overnight at 4°C. After blocking with casein, immobilised RV-G was saturated with Fab 17C7 (50 µL at 20.0 µg/mL per well) for an hour prior to the addition of purified hybridoma-derived IgG 1112-1 (at a six-fold dilution series starting from 0.6 µg/mL). Binding of IgG 1112-1 to RV-G was detected by an anti-mouse Fc-specific secondary antibody. The optical density readings at 405 nm (OD₄₀₅) for 1112-1 binding to the RV-G spikes in the presence of 17C7 were deducted from the OD₄₀₅ signals for RV-G–17C7 interaction alone. The results indicate that 1112-1 and 17C7 can bind to RV-G simultaneously. The results of $n=2$ biologically independent experiments are shown, with the colour scheme indicated by the legend inset. The curves are plotted using the average values calculated from the duplicate. The experiment was repeated by replacing Fab 17C7 with Fab RVC20 to investigate the competition between 1112-1 and RVC20 for RV-G spikes, and once more with only 1112-1 alone. The lower OD₄₀₅ signals for IgG 1112-1 binding to RV-G in the presence of Fab 17C7, as compared to those obtained with IgG 1112-1 alone, suggest that there may be spatial steric hindrance between the IgG and Fab. The results also indicate that 1112-1 and

RVC20 likely target epitopes that are not mutually accessible. Note that the experiment included in this figure was performed by Luke Thorley from Douglas group (Jenner Institute).

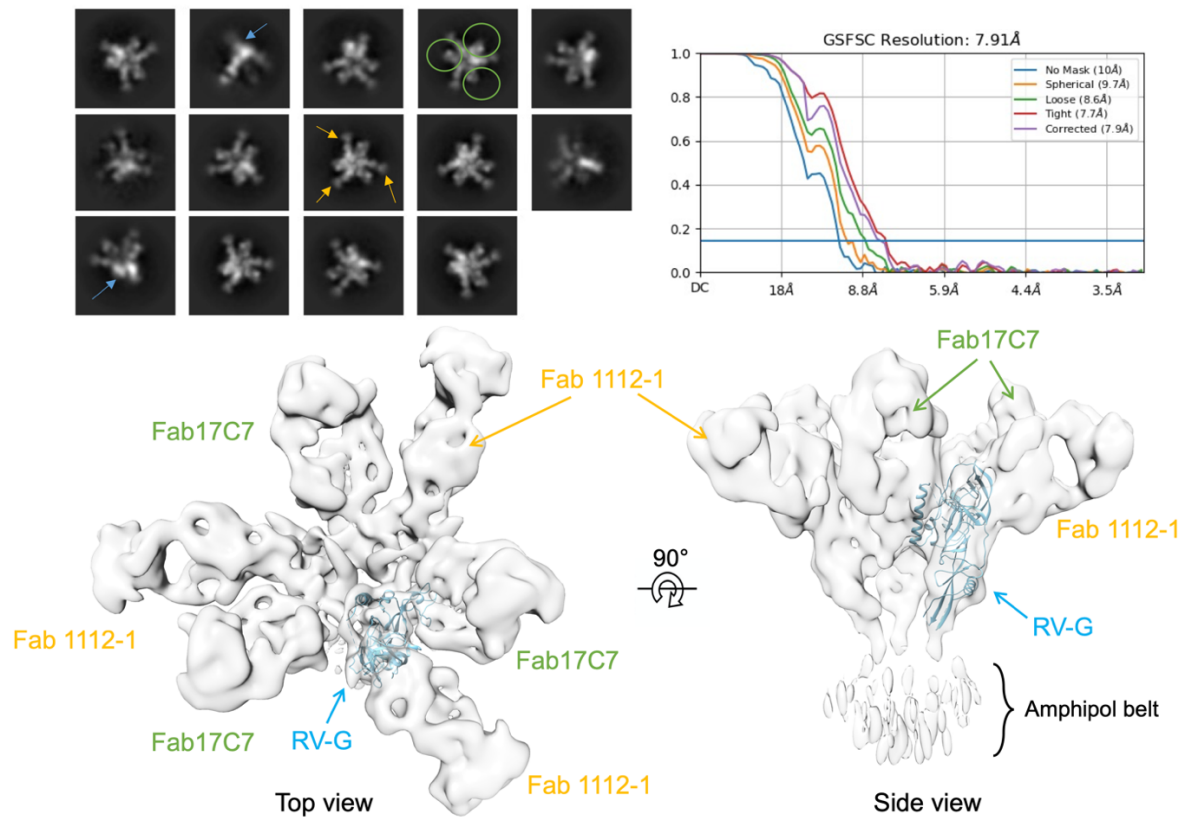


Fig. 5.6. Initial cryo-EM analysis of full-length RV-G in complex with Fabs 1112-1 and 17C7. (Top left) 2D class averages derived following collection of an initial dataset on a Glacios 200 kV electron microscope. The positions of the Fab 17C7, Fab 1112-1, and RV-G are indicated with green circles, yellow arrows, and blue arrows, respectively. (Top right) A gold-standard Fourier shell correlation (GSFSC) plot indicating an estimated resolution of approximately 7.9-Å for the initial dataset of approximately 44,719 particles using a gold-standard refinement procedure with a FSC = 0.143 criterion (blue horizontal line). Different FSC curves with no mask, spherical mask, loose mask, and tight mask applied to the half maps are also shown. Data was processed with cryoSPARC. (Bottom) The resulting 3D reconstruction (C3 symmetry applied), calculated with non-uniform refinement, with the crystal structure of the basic-pH RV-G ectodomain (PDB ID: 6LGX, truncated to Lys1–Ile373) fitted into the volume for illustrative purposes (fit using Chimera). Fabs 1112-1, 17C7, RV-G, and amphipol regions are labelled and indicated with yellow, green, blue, and black arrows, respectively. The visibility of all six Fabs, the three FD of RV-G spike, and the amphipol belt, along with the absence of “smearing” artifact as seen in Fig 5.2, 5.3, and 5.4, demonstrates promising sample and cryo-EM map quality.

(v) E559. NAb E559 recognises antigenic site II on RV-G and exhibits potent and broad RABV neutralisation spectrum, and therefore constitute an important candidate for inclusion in a mAb-based cocktail. E559 has been shown to demonstrate protective activity (67% survival rate) against *in vivo* lethal RABV challenge (hamsters) (van Dolleweerd et al., 2014). Furthermore, engineering of E559 for plant-based expression system enables rapid and high level of protein production, resolving the issue of limited availability often associated with the use of human and equine RIG (Tsekoa et al., 2016; van Dolleweerd et al., 2014).

Here, Fab E559 was complexed with full-length trimeric RV-G. Unlike nAb 1112-1 which binds to the PHD of RV-G in its prefusion state, the 3D reconstruction of RV-G–Fab E559 complex shows that the antibody recognises the RV-G in an extended conformation (**Fig. 5.7**). Fitting of the basic-pH crystal structure of RV-G ectodomain (PDB ID: 6LGX) into the calculated map reveals that the map is substantially taller than the crystal structure (**Fig. 5.7**). Fitting of the acidic-pH crystal structure of RV-G ectodomain (PDB ID: 6LGW), on the other hand, suggests that the map represents RV-G in an intermediate or a postfusion conformation (**Fig. 5.7**). Notably, preparation of the RV-G sample was consistently performed with buffers at pH 7.4 and therefore, any observed conformational changes in RV-G was likely due to Fab E559 binding. As the anisotropic $\sim 6.56\text{-}\text{\AA}$ map does not permit detailed interpretation of the RV-G–Fab E559 recognition mode, the molecular mechanism of E559-mediated neutralisation warrants further structural investigations.

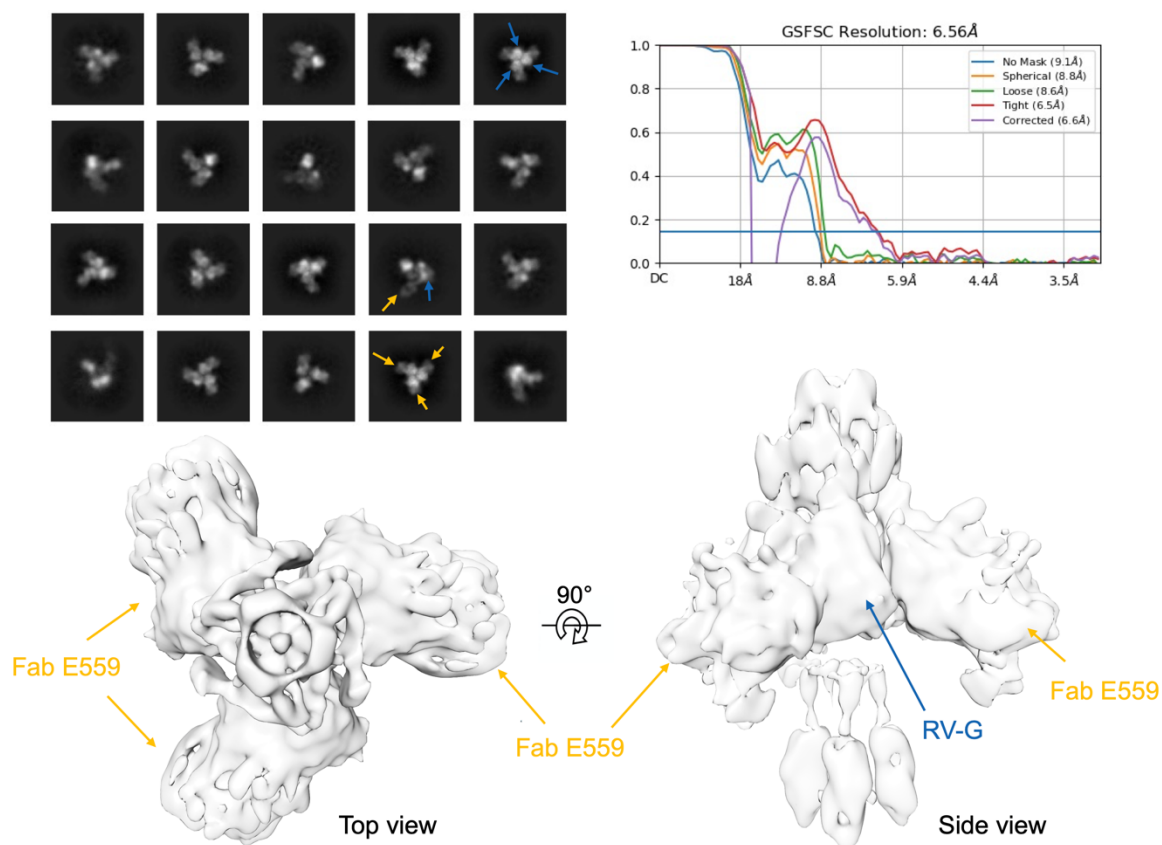


Fig. 5.7. Initial cryo-EM analysis of full-length RV-G in complex with the Fab of a neutralising mAb, termed E559. (*Top left*) 2D class averages derived following collection of an initial dataset on a Glacios 200 kV electron microscope. The positions of the Fab E559 and RV-G are indicated with yellow and blue arrows, respectively. (*Top right*) A gold-standard Fourier shell correlation (GSFSC) plot indicating an estimated resolution of approximately 6.6-Å for the initial dataset of 67,357 particles using a gold-standard refinement procedure with a FSC = 0.143 criterion (blue horizontal line). Different FSC curves with no mask, spherical mask, loose mask, and tight mask applied to the half maps are also shown. Data was processed with cryoSPARC. (*Bottom*) The resulting 3D reconstruction (C3 symmetry applied), calculated with non-uniform refinement, shown as a 3D volume. As the orientation of the structure cannot be discerned from the reconstructed map, no fitting of the RV-G crystal structures was performed. Fab E559 and RV-G are labelled and indicated with yellow and blue arrows, respectively.

5.2.3.2 High-Resolution Titan Krios Data Set Processing

As detailed in sections 5.2.2 and 5.2.3, cryo-EM grids containing freshly purified sample of RV-G–Fab 17C7–Fab 1112-1 were prepared using a Vitrobot. Sample and grid quality were assessed using Glacios microscope (see section 5.2.3.1). For high-resolution data collection, cryo-EM images were recorded for ~66 h using the in-house 300 kV Titan Krios microscope, equipped with Gatan K2 detector and GIF energy filter. Cryo-EM SPA data processing was performed using cryoSPARC v3 software (Punjani et al., 2017). Data collection and refinement statistics are detailed in **Table 5.1**.

Initially, three *ab initio* models were generated, two of which were selected as templates for 3D classification (heterogeneous refinement) and refinement (non-uniform refinement) with C3 symmetry applied. The resolutions of the resulting cryo-EM maps, as indicated by the FSC plots, were estimated to be 3.85-Å and 4.01-Å, respectively. Following local CTF refinement, the cryo-EM map improved from 3.85-Å to 3.68-Å. When visualised using the Chimera software, however, the map displayed “smearing effect” artifact, such as those observed in **Fig. 5.2, 5.3, 5.4, and 5.7**, suggesting that the samples suffered from orientation bias problem. As the Glacios initial data set did not detect such issue, it is likely that the 3.68-Å map contained particles that were not sufficiently homogeneous.

Subsequently, an initial 3D model reconstruction was repeated with five new *ab initio* models generated, in which all five were selected for heterogeneous and non-uniform refinements. The resulting cryo-EM maps confirmed the heterogeneous mixture of RV-G spike with varying Fab occupancy. Indeed, out of the five reconstructed maps, two contain 4-Fab bound RV-G (three 17C7 and one 1112-1), two contain 5-Fab bound RV-G (three 17C7 and two 1112-1), and one is fully bound with three copies of Fab 17C7 and three copies Fab 1112-1. The latter was selected for further refinement, where the resolution was improved from 4.14-Å (C3 symmetry) to 3.65-Å after local CTF refinement and motion correction.

The 3.65-Å 3D map was back-projected from all orientations to generate 2D images, which were then used as templates for another round of particle picking. The particles were subjected to two rounds of 2D classification to remove junk particles. Five new *ab initio* models were generated, which served as 3D templates for heterogeneous and non-uniform refinements. Out of the five reconstructed maps, one represented non-sensible blob and therefore, the corresponding junk particles were discarded. The remaining four maps comprise a 3-Fab bound, a 4-Fab bound, a 5-Fab bound, and a 6-Fab bound RV-G structure, respectively. As the resolutions and qualities of the 5-Fab (4.10-Å, C1 symmetry) and 6-Fab (4.48-Å, C1 symmetry) bound RV-G maps were the most promising, the particles that reconstructed these two maps were pooled and subjected to another round of heterogeneous refinement with only two 3D classes: 5-Fab and 6-Fab bound RV-G. The 3D maps were subsequently refined independently. No symmetry (C1) was applied for the 5-Fab bound RV-G map, and the FSC-estimated resolution is 3.88-Å after local CTF refinement and motion correction. The 6-Fab bound RV-G map was refined with C3 symmetry applied, and the FSC-estimated resolution is 3.80-Å after local CTF refinement and motion correction. The cryo-EM data processing workflow is summarised in **Fig. 5.8**.

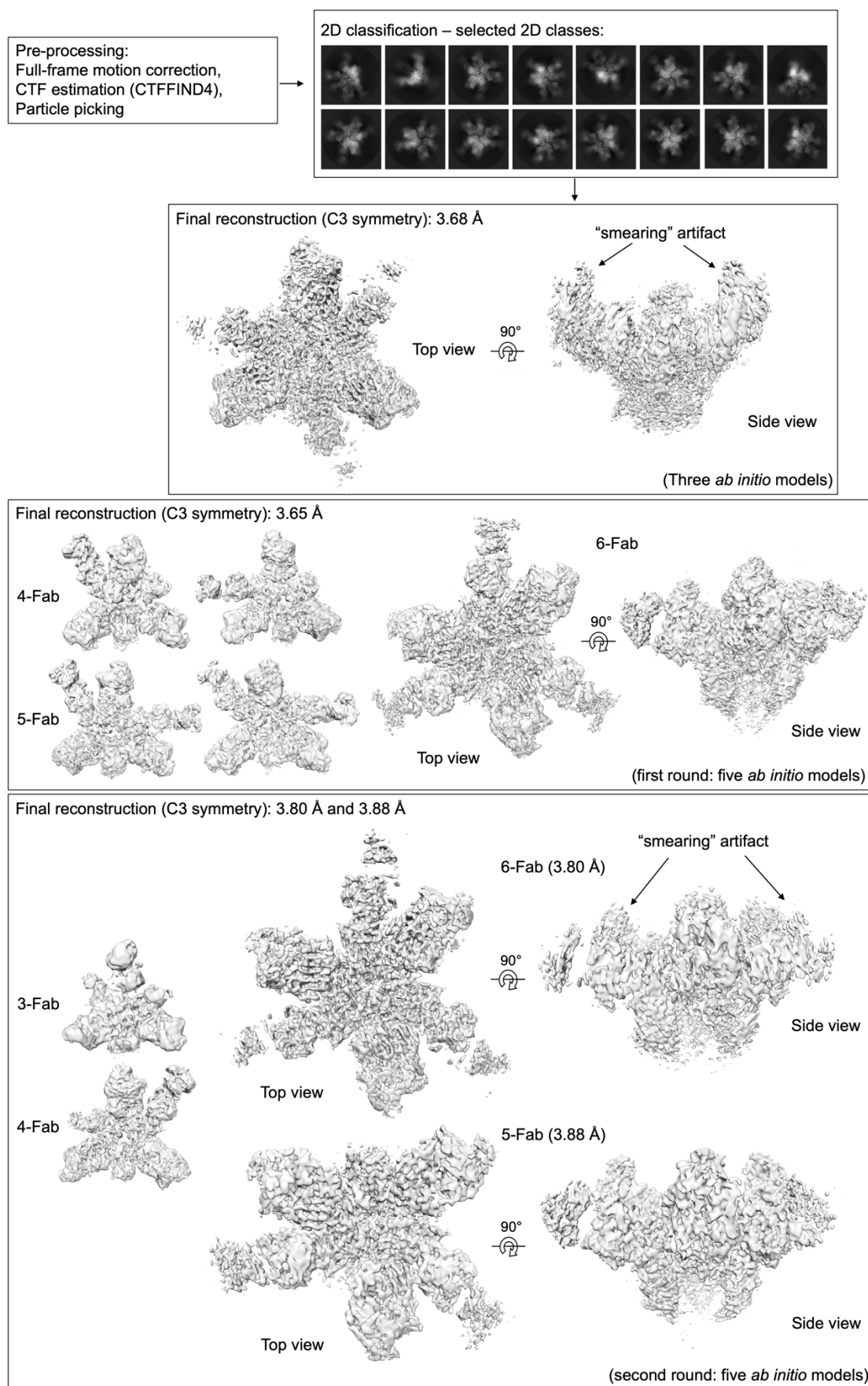


Fig. 5.8. Cryo-EM data processing workflow for the RV-G–Fab 17C7–Fab 1112-1 complex. Recorded movie frames were aligned for motion correction. CTF estimation was

performed using CTFFIND4. Following particle picking and extraction, 2D classification was performed to select 2D classes with clear protein features, minimal noise, and varied particle orientation. The selected particles were first used to generate three *ab initio* models. Following heterogeneous refinement, non-uniform refinement, and local CTF refinement, the cryo-EM map was reconstructed to 3.68-Å (613,730 particles) with C3 symmetry applied, as calculated using gold-standard FSC approach. The map exhibits “smearing” artifact, an indication of particle orientation bias, likely a result of heterogeneous particles used in reconstruction. Therefore, the refinement process was repeated with five *ab initio* models. Two 4-Fab, two 5-Fab, and one 6-Fab bound RV-G maps were generated. The 6-Fab bound structure was subsequently refined to 3.65-Å (180,591 particles), with C3 symmetry applied. The structure was back-projected as 2D templates for another round of particle picking and 2D classification. 3D reconstruction from the five new *ab initio* models yielded a junk class (not shown), a 3-Fab, a 4-Fab, a 5-Fab, and a 6-Fab bound RV-G cryo-EM maps. Particles from the 5-Fab and 6-Fab bound RV-G maps were pooled and classified to further ensure homogeneous particles were used for final 3D reconstructions. The resulting classified maps were refined to 3.88-Å (5-Fab; C1 symmetry, 258,945 particles) and 3.80-Å (6-Fab; C3 symmetry, 187,840 particles). As the 6-Fab bound RV-G map exhibits “smearing” artifact which affects its interpretability, the 5-Fab bound RV-G cryo-EM map was selected for further analysis.

The presence of RV-G–Fab complexes with partial Fab 1112-1 but full Fab 17C7 occupancy suggests that the affinity of nAb 1112-1 for the RV-G spike may be weaker than that of the nAb 17C7. While previously published literature reported a dissociation constant (K_D) of 13 nM for the nAb 17C7, the binding affinity of nAbs 1112-1 for the RV-G spike has not yet been reported. Therefore, this ongoing study will aim to quantify the binding affinity of these nAbs for the RV-G spike. Inspection of the two maps revealed that the map of the 3.88-Å 5-Fab bound complex has a better map quality than that of the 3.80-Å 6-Fab bound RV-G complex. Indeed, the cryo-EM map of the 6-Fab-bound RV-G complex displays the “smearing effect” artifact that is associated with particle preferential orientation, affecting its interpretability (**Fig. 5.8**). On the other hand, the cryo-EM map of the 5-Fab bound RV-G complex does not contain the “smearing effect” artifact (**Fig. 5.9**). Additionally, the constant region of Fab 1112-1 is better resolved in the 5-Fab map than the 6-Fab map. Since no information is missing in the 5-Fab map, including both the Fabs 17C7 and 1112-1 epitopes, the RV-G–(Fab 17C7)₃–(Fab 1112-1)₂ map was selected for further analysis.

Rigid fitting of the crystal structures of basic-pH RV-G ectodomain (PDB ID: 6LGX) (Yang et al., 2020) and prefusion VSV-G ectodomain (PDB ID: 5I2S) (Roche et al., 2007) indicates that the reconstructed cryo-EM map represents a RV-G–Fab complex in prefusion conformation (**Fig. 5.10**). Interestingly, the C-terminal loop of the basic-pH RV-G ectodomain crystal structure adopts an upward tilting conformation (Yang et al., 2020). Similarly, the crystal structure of the CHAV-G in an intermediate conformation (PDB ID: 5MDM) shows a C-terminal loop that pivots away from the FD (Baquero et al., 2017). Therefore, the basic-pH crystal structure of RV-G was proposed to represent an early intermediate state observed during the pre-to-post-fusion conformational transition. Indeed, inspection of this region in the trimeric, prefusion crystal structure of VSV-G shows that the C-terminal loop points and travels down towards the FD and membrane region (Roche et al., 2007). In addition, the reconstructed cryo-EM map of RV-G–Fab complex shows a continuous density for the C-terminal loop towards the FD and membrane region (**Fig. 5.10**). Therefore, our structural data represent a RV-G spike model that is more representative of the native prefusion spike (**Fig. 5.10**), complimenting the crystal structures of putative intermediate and post-fusion RV-G ectodomains. In order to obtain a better fitting of the crystal structure into the cryo-EM map, the C-terminal loop of the basic-pH RV-G structure was truncated (**Fig. 5.9**). Furthermore, the map contains region that was previously unresolved in the crystal structure, including linker residues Gln256–Leu276. However, the membrane-proximal β -sheet of the FD and transmembrane helices are less well resolved due to a noisier map at that amphipol region.

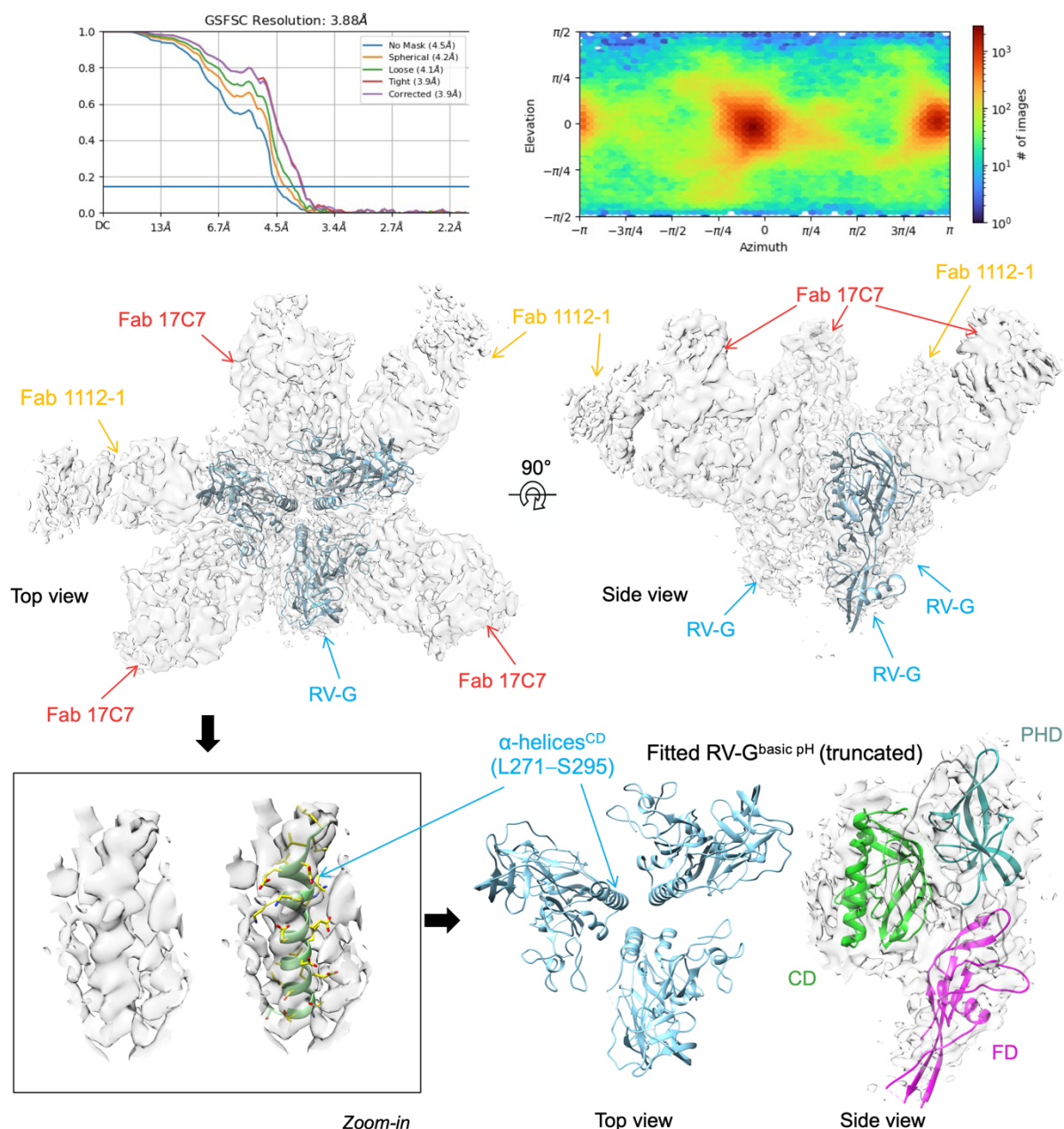


Fig. 5.9. Initial cryo-EM analysis of the higher-order assembly of the RV-G-(Fab 17C7)₃-(Fab 1112-1)₂ complex. (*Top left*) A gold-standard Fourier shell correlation (GSFSC) plot indicating an estimated resolution of approximately 3.9-Å for the dataset of 258,945 particles using a gold-standard refinement procedure with a FSC = 0.143 criterion (blue horizontal line). Different FSC curves with no mask, spherical mask, loose mask, and tight mask applied to the half maps are also shown. Data was processed with cryoSPARC. (*Top right*) Angular distribution calculated in cryoSPARC for particle projections. Heat map shows number of particles for each viewing angle (less = blue, more = red). (*Middle row*) The resulting 3D reconstruction (no symmetry applied: C1), calculated with non-uniform refinement, with the crystal structure of the basic-pH RV-G ectodomain (PDB ID: 6LGX, truncated to Lys1-Ile373) fitted into the volume (fit using Chimera). The cryo-EM map is shown as a translucent white surface. The crystal structure is shown as a cartoon and coloured blue. The map regions corresponding to Fab 17C7 and Fab 1112-1 are indicated and labelled

accordingly. Three copies of the basic-pH RV-G ectodomain crystal structure is fitted in the top view but only one copy is fitted in the side view for visual clarity. (*Bottom row*) A zoom-in panel of cryo-EM map (*left*) corresponding to the CD α -helix (Leu271–Ser295) that forms the trimerisation interface. The helix is displayed as a green cartoon and fitted into the map (*right*) with the side chains shown as yellow sticks and constituent nitrogen and oxygen atoms coloured blue and red, respectively. On the right side of the panel, only the fitted basic-pH RV-G ectodomain crystal structures are shown to demonstrate the trimerisation interface (top view) composed by the long α -helices of the CD. The fitting of a single protomer is also shown (side view) with the cryo-EM map around the molecule (8-Å radius) displayed. The PHD, CD, and FD are labelled and coloured in cyan, green, and magenta, respectively.

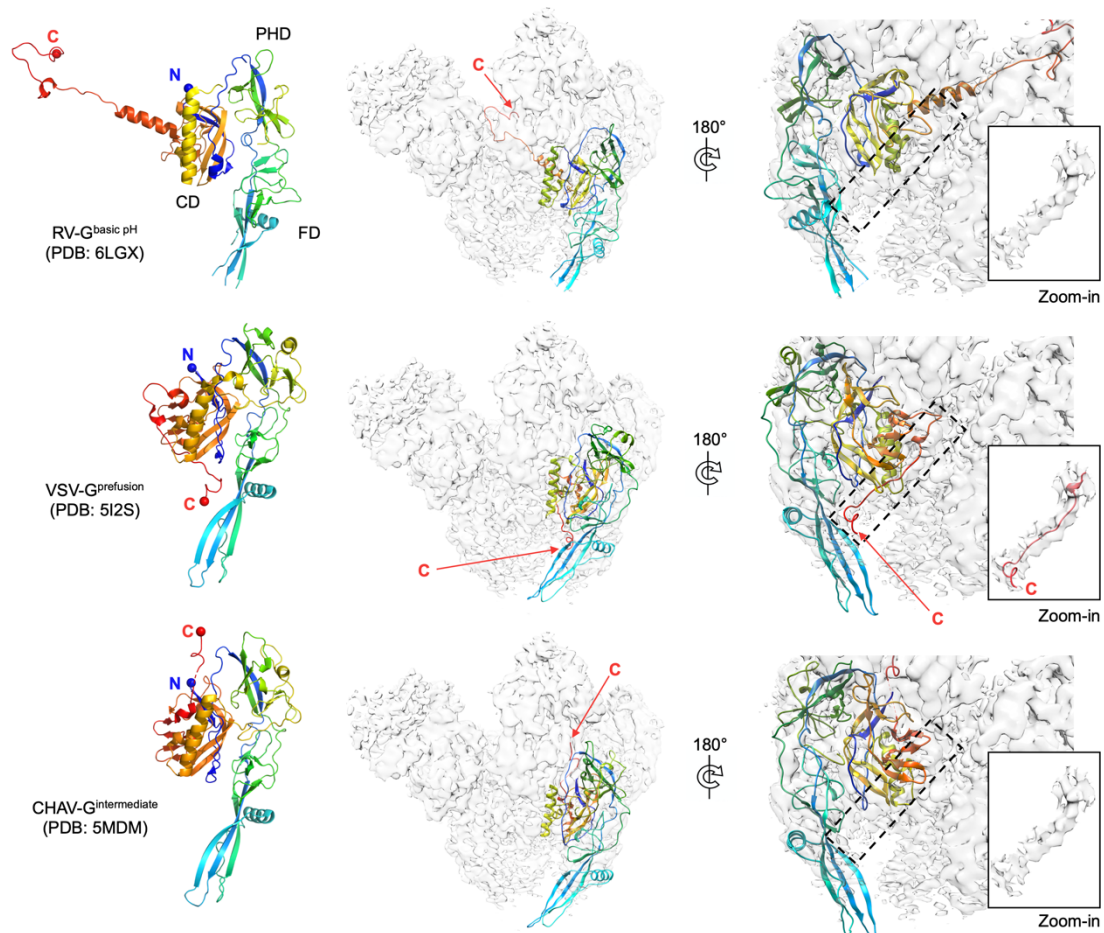


Fig. 5.10. Rigid fitting of RV-G, VSV-G, and CHAV-G ectodomain crystal structures into the cryo-EM-derived 5-Fab bound full-length trimeric RV-G map. (*Left column*) The crystal structures of basic-pH RV-G (PDB ID: 6LGX), prefusion VSV-G (PDB ID: 5I2S), and CHAV-G in an intermediate conformation (PDB ID: 5MDM) are shown in cartoon representation and coloured as a rainbow ramped from N- (blue) to C-terminus (red). The N- and C-termini are shown as blue and red spheres, respectively. (*Middle and right columns*) Rigid fitting of the crystal structures of RV-G^{basic pH}, VSV-G^{prefusion}, and CHAV-G^{intermediate} ectodomain into the cryo-EM map of full-length trimeric RV-G–(Fab 17C7)₃–(Fab 1112-1)₂. Only one copy of the G protein is fitted for visual clarity. The crystal structures are shown as previously described. The cryo-EM map is shown as a translucent white surface. The C-termini

of the structures are labelled. 180°-rotated zoom-in panels of the C-terminal loop for each structure fitting are shown. The map within the dashed-line box is isolated and displayed separately. The results demonstrate that the experimental cryo-EM map represents a trimeric RV-G spike in the prefusion conformation.

The observed higher-order trimer via rigid fitting of the RV-G crystal structure reveals that the CD of each RV-G protomer is positioned towards the trimerisation axis, with its long-spanning α -helix (Leu271–Ser295) forming part of the trimerisation interface (**Fig. 5.9**). Fab 17C7 recognises the RV-G spike by engaging the surface of CD that is solvent exposed in between the RV-G protomers, while Fab 1112-1 binds to the membrane-distal PHD. Each leg of the fitted RV-G is composed of the FD with the fusion loops pointing toward the membrane (amphipol) region. Together with VSV-G and HCMV gB, RV-G is the third full-length trimeric structure available for class III fusogen and the first full-length trimeric lyssaviral-G structure. Despite low sequence identities (~18%), comparison of the fitted RV-G ectodomains with the prefusion trimeric VSV-G revealed that the overall tripodal arrangement is highly similar, suggesting that lyssaviruses and vesiculoviruses have evolutionarily conserved membrane-fusion mechanisms. As the atomic model building for the RV-G trimer is still underway, structural comparison between the prefusion structures is not currently available. We anticipate that the structure of the RV-G–Fab will serve as a template for understanding the lyssaviral G architecture and the molecular basis of antibody-mediated immune response, and as a molecular blueprint for future structural investigation of trimeric RV-G fusogenic rearrangements and RV-G–receptor interactions.

5.3 SUMMARY AND DISCUSSION

In collaboration with Dr Alexander Douglas at Jenner Institute, we aimed to characterise the structure of the native full-length trimeric RV-G spike to delineate the molecular basis of

antibody-mediated neutralisation of RABV and to provide a structural template for vaccine design that is enabled by the production of stable RV-G spike. Cryo-EM SPA approach was used to obtain 3.80–3.88-Å cryo-EM maps of RV-G in complex with Fab 17C7 and Fab 1112-1 (**Fig. 5.9**). The maps show that the RV-G spike adopts a prefusion conformation with a tripodal arrangement, as demonstrated by rigid fitting of the crystal structures of basic-pH RV-G ectodomain (Yang et al., 2020) and prefusion VSV-G ectodomain (Roche et al., 2007) (**Fig. 5.10**). Identification of the trimerisation interface composed by the long α -helix in the CD provides a molecular template for designing mutations that can stabilise the trimeric RV-G ectodomain in the prefusion conformation, which is otherwise metastable and could transit among several different conformational states. Production of stabilised prefusion RV-G ectodomain without transmembrane region will eliminate the need of detergent solubilisation and reduce the costs and complications associated with membrane protein production and purification. Furthermore, a stabilised RV-G construct will serve as a more homogeneous sample for high-resolution structural studies and as a better vaccine design.

The cryo-EM maps further reveal that the RV-G spike can be recognised by three copies of Fab 17C7 and three copies of Fab 1112-1 simultaneously, indicative of the presence of more than one key site of vulnerability on the RV-G spike and is reflective of the importance of nAb-based PEP cocktails. Previous characterisation of anti-RABV nAbs highlighted the utilisation of different nAb-mediated neutralisation mechanisms, such as by preventing receptor binding (e.g. nAb 523-11) (Yang et al., 2020) and fusogenic rearrangement of RV-G (e.g. nAb RVC-20) (Hellert et al., 2020). Here, cryo-EM determination of RV-G spike in complex with Fabs 17C7 and 1112-1 will provide the first structural insights into the molecular determinants of potent and broad RABV neutralisation mechanisms utilised by these two nAbs. We observe that Fab 17C7 targets the antigenic site III located on the CD while Fab 1112-1 binds to the antigenic site II located on the PHD. Targeting of the major antigenic sites located in the

membrane-distal region of the RV-G by the nAbs corroborates with the structure of the prefusion RV-G spike, as the CD and PHD are the most protruded and exposed domains. Thus far, no antigenic site has been mapped to the membrane-proximal FD, suggesting that the domain is well shielded by the PHD and CD on the mature virion.

Structural data of RV-G in the presence of Fab E559 suggests that the nAb promotes structural rearrangement of the spike from the prefusion to an intermediate or a post-fusion conformation (**Fig. 5.7**). While the molecular basis of E559-mediated RABV neutralisation warrants further investigation, we postulate that E559 may lock the RV-G spike in an intermediate or a post-fusion conformation and therefore, render it incompatible for receptor binding. The ability of RV-G to reversibly rearrange from pre- to post-fusion conformation highlights the presence of an equilibrium state between the conformations on the RABV virion (Yang et al., 2020). Thus, the inclusion of nAbs, such as E559, that likely promotes RV-G into an intermediate or the post-fusion state constitutes an important element in nAb-based PEP therapeutic, as the E559-mediated antigenic change in RV-G spikes on the virion surface may likely disrupt host receptor recognition.

In *toto*, we have successfully derived a well-expressing RV-G construct and identified purification conditions that allowed production of trimeric, full-length RV-G to sufficient homogeneity. Moreover, cryo-EM SPA of RV-G in the presence of Fabs 17C7 and 1112-1 demonstrated that the purified protein is a tractable target for high-resolution structural analysis and that the nAbs target the trimeric membrane-distal region of the RV-G spike. As RABV constitutes an important human and animal pathogen, the information revealed in this study will be of broad academic and commercial interest. In particular, our data may lead us to structure-guided engineering of RV-G in a stabilised prefusion state, which may in turn lead to the development of more effective RABV vaccines. Our structural insights on nAb-targeting epitopes on RV-G may also allow the optimisation of pre-existing therapeutic nAb-based PEP,

resulting in lower dose regimes and more affordable and efficacious products for RABV-endemic areas.

6 CONCLUDING REMARKS AND FUTURE WORK

6.1 MAMMARENAVIRUS

The continual discovery of novel mammarenaviruses raises concerns for the public health and economic impacts of these viruses with unknown pathogenic potential. Characterisation of their origins, cross-species transmission potential, host cell entry mechanism, and associated therapeutic countermeasures is therefore of paramount importance. Here, I studied the GP1 glycoprotein of LORV, a novel mammarenavirus isolated from *Bandicota* rat in Thailand in 2016 (Blasdell et al., 2016). Sequence analysis indicates that LORV is closely related to other OW mammarenaviruses, including the pathogenic LASV (**Fig. 3.1**). Comparison of the structures of monomeric LORV GP1 and monomeric LASV and MORV GP1s shows that the protein fold of the receptor-binding glycoprotein is highly similar (**Fig. 3.4**). Furthermore, the high structural similarity between the basic and acidic pH structures of LORV GP1 suggests that the pH-derived alteration of electrostatic surface potential does not lead to any appreciable local conformational changes.

Structural superimposition of monomeric LORV GP1 with LASV GP1 associated with the cognate GP2 subunit revealed dramatic structural rearrangement of the GP1 molecule (**Fig. 3.5**). Indeed, while the architecture of the central β -sheet remains considerably conserved, reorientations of the GP1 termini and α -helices render the monomeric molecule α -DG-binding incompetent. This observation is supported by the mapping of the putative α -DG-binding residues onto the monomeric OW GP1 subunit where spatial declustering of the residues with more than 15 Å apart is observed (**Fig. 3.6**). In addition, mapping of LASV LAMP1 binding site onto LORV GP1 revealed substantial compositional and conformational differences in the LAMP1-binding histidine triad, strand β 5, and loop L7 (**Fig. 3.7**). In consistent with a previous structural study of MORV GP1 (Israeli et al., 2017), the mapping analysis suggests that LORV

GP1 does not utilise LAMP1. Further experiment testing LAMP1 recognition by LORV GP1 is required to confirm this hypothesis. In addition, SBPA of available OW mammarenaviral GP1 structures demonstrates the bifurcation of the structures into two distinct classes, termed ‘GP2-free’ and ‘GP2-associated’ (**Fig. 3.9**). Antigenic deconfiguration associated with the structural changes in GP1 from the GP2-associated to GP2-free conformation may render the molecule as an immunological decoy. Indeed, ineffective humoral immune response during early LASV infection have been reported along with the detection of shed LASV GP1 in patient sera (Branco et al., 2010).

To date, a number of OW and NW GP1 structures have been characterised (**Table 1.1**). Structural comparison of GP2-free OW GP1, GP2-associated OW GP1, and GP2-free clade B and D NW GP1 showed that the latter two structural classes are more similar to each other than with GP2-free OW GP1, and that the conformations displayed by the structures of GP2-associated OW GP1 and GP2-free NW GP1 likely resemble those presented on the mature virion surface (**Fig. 1.11** and **3.5**). However, the structures of NW GP1 from clade A and C have yet to be reported. Like OW GP1s, NW clade C GP1s recognise α -DG. Therefore, the question remains if the structures of NW clade C OLV and LATV GP1s will resemble those of OW or NW GP1s. In addition, the receptor for NW clade A such as PICV remains unknown. In order to understand the structural lability of the mammarenaviral GP1 and to predict their receptor tropism, structure determination of NW clade A and C GP or GP1 constitutes an important part in the future mammarenaviral research.

Previous mass-spectrometry glycan analysis of OW LASV GP reported by Watanabe *et al.* in 2018 (**Fig. 1.13**) and the glycan mapping analysis of LORV GP performed in this thesis (**Fig. 3.8**) demonstrates that the proteinaceous surface of the OW GP spike is covered by a dense glycan shield, rendering this therapeutically important molecule an immunologically resistant target. The abundance of N-linked glycosylation on the LORV GP, which surpasses

that of LASV, suggests that LORV may also be an “evasion strong” virus (Burton, 2017; Watanabe et al., 2018) with heightened resistance to antibody-mediated neutralisation. Notably, mapping of the position of these putative N-linked glycan sites reveals the close spatial association of Asn65^{LORV} (not present in LASV: Glu67^{LASV}) to an epitope recognised by a LASV neutralising antibody, 37.7H (Hastie et al., 2017). If N-linked glycans are present at this site, it is likely that the glycans may help to shield this region of the GP surface. Thus far, no similar analysis has been performed for NW mammarenaviral GP. In the future, I suggest that NW GP to be subjected to a similar glycan compositional and mapping analysis. The structural data will be integral to the overview of glycan-mediated shielding of mammarenaviral GP, identification of sites that may be most vulnerable to the antibody immune response, and the creation of a humoral immunity-based vaccine capable of mimicking the GP spikes inclusive of the native-like glycosylation presented on the mature virion surface.

Characterisation of the nAb epitopes on the mammarenaviral GP spikes is crucial in identifying the minimal subcomponent that is capable of eliciting protective humoral immune response. Previous studies have revealed that most human nAbs require both LASV GP1 and GP2 subunits (Hastie et al., 2017; Robinson et al., 2016; Zeltina and Bowden, 2017). Indeed, the combined structural data of GP conformation and dense glycan shield presented in this thesis provide a rationale to this observation, indicating that the previously reported OW GP1–GP2 structures present a native-like conformation, and that the nAb epitopes often include the GP2 subunit, which is more sequentially and structurally conserved and contains a lower glycan density than the GP1 subunit. To investigate the antigenic surface of NW mammarenaviral GP1, in this thesis, I have structurally characterised the nAb epitope of Fab JUN1 and Fab MAC1 that target JUNV GP1 and MACV GP1, respectively.

Structure analysis of JUNV GP1–Fab JUN1 and MACV GP1–Fab MAC1 demonstrates that the nAbs target the TfR1 RBS and neutralise through disruption of hTfR1

recognition (**Fig. 4.3, 4.4, 4.8**). Comparison of the two structures with previously reported GP1–nAb complexes reveals that despite originating from a unique germline, nAb JUN1 utilises a binding mode common to all characterised infection and vaccine-elicited JUNV-specific nAbs (e.g. OD01, GD01, CR1-28), which involves mimicking TfR1 binding through the insertion of a tyrosine residue into the RBS (**Fig. 4.3–4.5**). In contrast, nAb MAC1 undergoes a distinct mode of antigen recognition that contrasts that used by the reported anti-JUNV nAbs and the only other characterised anti-MACV nAb (CR1-07), providing evidence for varied modes of GP1-specific recognition amongst New World arenaviruses more broadly (**Fig. 4.8, 4.9, 4.13**). Combined, these data provide a structure-guided rationale for the requirement of nAb cross-reactivity against the NW GP1s, which involves antigenic differences (e.g. MACV GP1 loop 10) and the ability to engage a sufficiently large and conserved surface. To augment the effort in creating bnAb-based cocktails, future work involving identification and characterisation of NW GP2 and/or GP1–GP2 nAb epitopes are essential. Furthermore, immunogen design will greatly benefit from the structures of NW GP1–GP2 and GP, which are currently unavailable.

The development of therapeutic mAbs targeting SARS-CoV-2 (Deb et al., 2021; Ning et al., 2021), respiratory syncytial virus (Gilman et al., 2016; Goodwin et al., 2018), Ebola virus (Bornholdt et al., 2016; Corti et al., 2016; Rijal et al., 2019), and influenza virus (Chai et al., 2017; Nakamura et al., 2013) demonstrates the important role nAbs can play in the treatment of viral infectious diseases. The work involving the characterisation of neutralising epitopes on NW mammarenavirus surface presented in this thesis provides a platform for the development of mAb therapy against these biomedically important pathogens. Future endeavours towards this goal will likely involve in the isolation of a larger library of nAbs for comparative assessment of efficacy. Several approaches such as mouse/transgenic mouse vaccination, phage display, and single mouse/human B cell sorting can be employed (Lu et al., 2020).

Screening for binding and neutralisation *in vitro* followed by *in vivo* characterisation of the mAbs (e.g. efficacy, pharmacokinetics, safety, and cellular toxicity) using appropriate animal models are essential before the treatment can be clinically trialled in humans. These essential assays may be complemented by structural analysis to help identify synergetic combinations of nAbs that target distinct and non-competing sites on the NW mammarenaviral GP surface. Furthermore, the production cost, product stability, storage condition, dosage, and treatment regime for the application are important considerations for the development for these biologics. A humanised mAb (hu199) has shown to be highly efficacious in guinea pig and NHP models of AHF, suggesting that mAbs may be viable for AHF therapy (Zeitlin et al., 2021). In addition to making the manufacture of mAbs economically feasible, the licensure expense associated with their clinical use can be a hurdle, especially for rare, endemic viral diseases with high morbidity and mortality that occupy relatively small commercial markets (Kelley, 2009; Shaughnessy, 2012). Therefore, it will be important to coordinate this developmental work with local policy makers and regulators in regions at risk to ensure the best possible outcome for those at risk (Zeitlin et al., 2021).

While high-resolution structure determination of recombinant mammarenaviral GP1 or GP1–GP2 in the presence of Fab fragments of nAbs provide molecular-level insights into the nAb-mediated neutralisation, such data do not give information about the mode of whole virus engagement and the occupancy of the nAbs on the mature virions. For instance, the crystal structure of LASV GP1–GP2 in complex with Fab 37.7H (PDB ID: 5VK2) (Hastie et al., 2017) shows that the Fab fragment mainly engages the GP2 subunit at an angle that is parallel to the viral membrane (**Fig. 6.1A**). A cryo-ET and STA experiment may help to illuminate on how the spaces between the GP spikes on the virion surface may accommodate the nAbs and the average number of nAbs required to neutralise the virus. Indeed, I have begun the produce LASV VLP using a previously established stable cell line developed by Dr. Thomas Strecker

at Philipps University of Marburg (see **section 2.4**). Cryo-EM grids containing purified VLP and Fab 37.7H were prepared and screened using the Glacios microscope (**Fig. 6.1B**). While the ongoing COVID-19 pandemic has disrupted this experiment, the successful production of LASV VLP indicates that this study is a highly tractable project.

In *toto*, the structural data on mammarenaviral GP1 presented in this thesis highlights the differential conformational plasticity and immunogenicity of OW and NW mammarenaviral GP spikes. Creation of OW GP-based vaccines will require the incorporation of both GP1 and GP2 subunits and reproduction of native-like glycosylation in order to present epitopes that are natively accessible while GP1-targeting nAb-based therapeutics have proven to be effective against NW mammarenaviruses. However, sequence and structural variation of the NW GP1 subunits may limit the efficacious of GP1-specific nAbs against different NW viruses. Therefore, I envisage that nAb-based cocktail consisting GP1-, GP2-, and GP1–GP2-targeting bnAbs will likely constitute one of the most promising treatments against NW mammarenavirus infection.

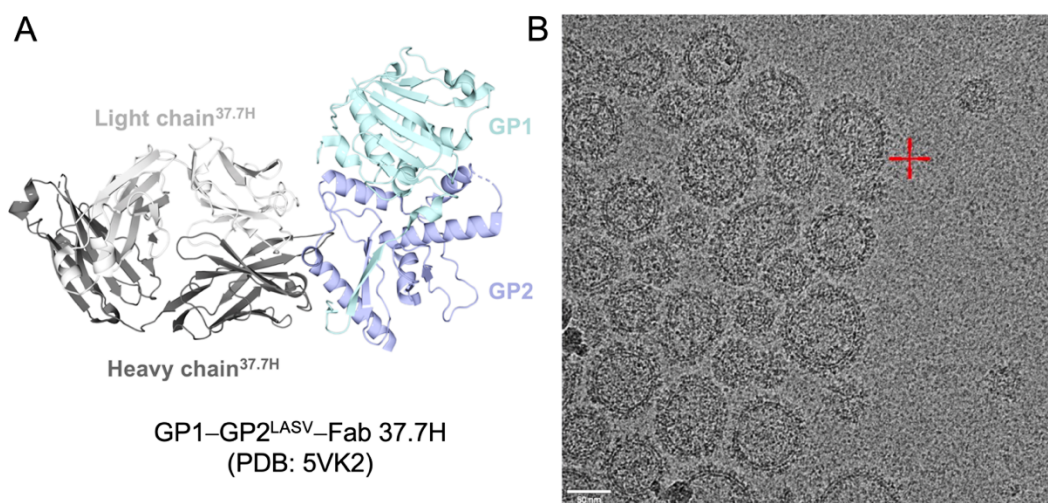


Fig. 6.1. Preparation of LASV GP-displayed VLPs lays a groundwork for cryo-ET investigation of mAb recognition of LASV GP (A) The crystal structure of LASV GP1–GP2 in complex with Fab 37.7H (PDB ID: 5VK2). The structure is shown in cartoon representation with GP1 coloured light blue, GP2 in purple blue, Fab 37.7H heavy chain in dark grey, and Fab 37.7H light chain in light grey. Unresolved regions are shown as dashed lines. (B) A cryo-EM image of LASV VLPs in the presence of Fab 37.7H recorded using Glacios microscope. Purified LASV VLPs and Fab 37.7H were mixed and incubated for 30 min prior to cryo-EM

grid preparation. 3 μ L of VLP–Fab mixture was then pipetted onto Quantifoil® R 2/1 or C-Flat™ R 2/2 cryo-EM grids. Cryo-plunge freezing was performed using Vitrobot. The 2D cryo-EM images did not provide clear evidence of Fab 37.7H binding to the LASV GP displayed on the VLP surface. Therefore, cryo-ET and STA approaches will be required to ascertain the antigen recognition mode utilised by Fab 37.7H on the VLP surface. *Scale bar* = 50 nm.

6.2 RABIES VIRUS

The high cost and dose administration associated with the current RABV vaccine and PEP treatment represents a huge obstacle to WHO’s ‘Zero by 30’ programme, a global plan to eliminate dog-mediated rabies by 2030 (Consortium, 2019). Currently, approximately 60,000 deaths occur annually due to rabies (Fooks et al., 2017). The current RABV vaccines use purified cell cultured-based vaccine and have proven to be safe and efficacious in treating rabies and preventing death when administered together with RIG (Gerber et al., 2020; Moulenat et al., 2020). 3–4 vaccine doses are often required to offer full protection. Here, to explore the design of a glycoprotein-based vaccine, I structurally characterised the full length, trimeric RV-G using cryo-EM SPA method.

The glycoprotein has demonstrated particle orientation bias on cryo-EM grids (**Fig. 5.2–5.4**), a phenomenon that is often caused by the adsorption of protein particles at the air-water interface. I resolved this issue by trying different RV-G-targeting Fab fragments and determined that Fab 17C7 and Fab 1112-1 were the best combination capable of alleviating the preferential orientation issue (**Fig. 5.6**). Subsequently, high-resolution cryo-EM images of RV-G in complex with Fabs 17C7 and 1112-1 were collected using a Titan Krios microscope equipped with a K2 camera. The cryo-EM map at approximately 3.9-Å estimated by gold-standard FSC approach revealed the higher-order trimeric assembly of RV-G and the epitopes targeted by Fabs 17C7 and 1112-1 (**Fig. 5.9**).

While the model building is currently ongoing, the cryo-EM map informs several important characteristics of the structure. Fitting of the basic-pH RV-G ectodomain crystal

structure (PDB ID: 6LGX) (Yang et al., 2020) into the cryo-EM map shows that the trimerisation interface is composed by the long α -helices of the CD, with the PHD located at the most protruded and exposed region and the FD at the most membrane-proximal region of the RV-G spike (**Fig. 5.9**). Identification of the trimerisation interface will provide a structural framework for designing trimeric RV-G immunogen stabilised in prefusion conformation by inserting stabilising mutations at this interface. Furthermore, the stabilised construct may allow the removal of transmembrane region and eliminate the need of detergent-based protein purification protocol, reducing costs and complications associated with membrane protein production and purification, leading to the production of highly homogeneous sample for high-resolution structural studies and vaccine candidate. The cryo-EM map also provides the first structural insights into the recognition mode of the nAbs in the context of higher order trimeric RV-G assembly. Fab 17C7 is observed to recognise the antigenic site III located on the CD while Fab 1112-1 binds to the antigenic site II located on the PHD, indicative of the presence of more than one key site of vulnerability on RV-G surface. High-resolution visualisation of these epitopes will inform discussion about the risk of failure of the PEP due to antigenic variation as well as the development and optimisation of other nAb-based PEP cocktails that target two or more antigenic sites.

In addition of designing a stabilised prefusion RV-G construct, the RV-G construct described above paves the way for biophysical quantification and structure determination of RV-G–receptor interactions, which will reveal the RBS on RV-G spikes. Notably, ectodomain regions for p75NTR (Tuffereau et al., 1998), NCAM1 (Thoulouze et al., 1998), and mGluR2 (Wang et al., 2018) have been previously shown to be amenable to recombinant production and purification. High resolution structural analysis of virus-receptor pairings will provide molecular-level information about the specificity underlying host-cell recognition, and further aid the identification of epitopes vulnerable to RBS-targeting nAbs. This work will clarify the

role of RV-G during host cell attachment, providing a structural model for how RABV is capable of recognising human cells during host cell entry.

Low pH-induced structural rearrangement of RV-G ectodomain and other class III fusogens have been previously shown (Backovic et al., 2009; Baquero et al., 2015; Belot et al., 2020; Chandramouli et al., 2015; Heldwein et al., 2006; Roche et al., 2006; Yang et al., 2020). A recent structural study of a RV-G ectodomain protomer demonstrates the conformational differences between acidic- and basic-pH structures (Yang et al., 2020), consistent with class-III fusogenic activity. However, the structural basis for how the full-length trimeric RV-G spike undergoes these rearrangements remains unknown and is key for understanding the coverage of conformational landscape by the molecule. Future work including cryo-EM analysis of acidified full-length RV-G will validate proposed models of class III viral fusion. Interestingly, initial cryo-EM analysis of RV-G in complex with Fab E559 (**Fig. 5.7**) suggests that the nAb promotes structural rearrangement of the spike from a prefusion to an intermediate or a post-fusion conformation. High-resolution cryo-EM analysis of RV-G–Fab E559 complex will be highly beneficial not only for the identification of a structural state adopted by the RV-G, but also for the characterisation of the varied modes by which nAbs can neutralise RABV.

In conclusion, the structural data of RV-G and anti-RV-G Fabs presented in this thesis will illuminate the structure and functionality of RV-G, a key therapeutic target. In particular, the data may lead to structure-guided RV-G immunogen design, stabilised in prefusion conformation, as a more effective RABV vaccine. Furthermore, the structure provides a molecular blueprint necessary for the investigations of RV-G–receptor interactions and RV-G fusogenic activity. In addition to the potential academic and commercial values of the therapeutic insights derived from this work, this project will also expand our understanding and appreciation of the structures and functionality of class III viral fusogens.

7 REFERENCES

Abraham, J., Corbett, K.D., Farzan, M., Choe, H., Harrison, S.C., 2010. Structural basis for receptor recognition by New World hemorrhagic fever arenaviruses. *Nat Struct Mol Biol* 17, 438-444.

Abraham, J., Kwong, J.A., Albarino, C.G., Lu, J.G., Radoshitzky, S.R., Salazar-Bravo, J., Farzan, M., Spiropoulou, C.F., Choe, H., 2009. Host-species transferrin receptor 1 orthologs are cellular receptors for nonpathogenic new world clade B arenaviruses. *PLoS Pathog* 5, e1000358.

Abrescia, N.G., Bamford, D.H., Grimes, J.M., Stuart, D.I., 2012. Structure unifies the viral universe. *Annu Rev Biochem* 81, 795-822.

Abudurexiti, A., Adkins, S., Alioto, D., Alkhovsky, S.V., Avšič-Županc, T., Ballinger, M.J., Bente, D.A., Beer, M., Bergeron, É., Blair, C.D., Briese, T., Buchmeier, M.J., Burt, F.J., Calisher, C.H., Cháng, C., Charrel, R.N., Choi, I.R., Clegg, J.C.S., de la Torre, J.C., de Lamballerie, X., Dèng, F., Di Serio, F., Digiaro, M., Drebot, M.A., Duàn, X., Ebihara, H., Elbeaino, T., Ergünay, K., Fulhorst, C.F., Garrison, A.R., Gão, G.F., Gonzalez, J.J., Groschup, M.H., Günther, S., Haenni, A.L., Hall, R.A., Hepojoki, J., Hewson, R., Hú, Z., Hughes, H.R., Jonson, M.G., Junglen, S., Klempa, B., Klingström, J., Kòu, C., Laenen, L., Lambert, A.J., Langevin, S.A., Liu, D., Lukashevich, I.S., Luò, T., Lǔ, C., Maes, P., de Souza, W.M., Marklewitz, M., Martelli, G.P., Matsuno, K., Mielke-Ehret, N., Minutolo, M., Mirazimi, A., Moming, A., Mühlbach, H.P., Naidu, R., Navarro, B., Nunes, M.R.T., Palacios, G., Papa, A., Pauvolid-Corrêa, A., Pawęska, J.T., Qiáo, J., Radoshitzky, S.R., Resende, R.O., Romanowski, V., Sall, A.A., Salvato, M.S., Sasaya, T., Shěn, S., Shí, X., Shirako, Y., Simmonds, P., Sironi, M., Song, J.W., Spengler, J.R., Stenglein, M.D., Sū, Z., Sūn, S., Táng, S., Turina, M., Wáng, B., Wáng, C., Wáng, H., Wáng, J., Wèi, T., Whitfield, A.E., Zerbini, F.M., Zhāng, J., Zhāng, L., Zhāng, Y., Zhang, Y.Z., Zhāng, Y., Zhou, X., Zhū, L., Kuhn, J.H., 2019. Taxonomy of the order Bunyavirales: update 2019. *Arch Virol* 164, 1949-1965.

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., Terwilliger, T.C., 2002. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* 58, 1948-1954.

Agnihothram, S.S., York, J., Nunberg, J.H., 2006. Role of the stable signal peptide and cytoplasmic domain of G2 in regulating intracellular transport of the Junín virus envelope glycoprotein complex. *J Virol* 80, 5189-5198.

Albariño, C.G., Bird, B.H., Chakrabarti, A.K., Dodd, K.A., White, D.M., Bergeron, E., Shrivastava-Ranjan, P., Nichol, S.T., 2011. Reverse genetics generation of chimeric

infectious Junin/Lassa virus is dependent on interaction of homologous glycoprotein stable signal peptide and G2 cytoplasmic domains. *J Virol* 85, 112-122.

AlQuraishi, M., 2019. AlphaFold at CASP13. *Bioinformatics* 35, 4862-4865.

Amanat, F., Duehr, J., Oestereich, L., Hastie, K.M., Ollmann Saphire, E., Krammer, F., 2018. Antibodies to the Glycoprotein GP2 Subunit Cross-React between Old and New World Arenaviruses. *mSphere* 3.

Argenta, F.F., Hepojoki, J., Smura, T., Szirovicza, L., Hammerschmitt, M.E., Driemeier, D., Kipar, A., Hetzel, U., 2020. Identification of Reptarenaviruses, Hartmaniviruses, and a Novel Chuvirus in Captive Native Brazilian Boa Constrictors with Boid Inclusion Body Disease. *J Virol* 94.

Aricescu, A.R., Lu, W., Jones, E.Y., 2006. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr D Biol Crystallogr* 62, 1243-1250.

Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., Notredame, C., 2006. Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res* 34, W604-608.

Auperin, D., Dimock, K., Cash, P., Rawls, W.E., Leung, W.C., Bishop, D.H., 1982a. Analyses of the genomes of prototype pichinde arenavirus and a virulent derivative of Pichinde Munchique: evidence for sequence conservation at the 3' termini of their viral RNA species. *Virology* 116, 363-367.

Auperin, D.D., Compans, R.W., Bishop, D.H., 1982b. Nucleotide sequence conservation at the 3' termini of the virion RNA species of New World and Old World arenaviruses. *Virology* 121, 200-203.

Auperin, D.D., Galinski, M., Bishop, D.H., 1984a. The sequences of the N protein gene and intergenic region of the S RNA of pichinde arenavirus. *Virology* 134, 208-219.

Auperin, D.D., Romanowski, V., Galinski, M., Bishop, D.H., 1984b. Sequencing studies of pichinde arenavirus S RNA indicate a novel coding strategy, an ambisense viral S RNA. *J Virol* 52, 897-904.

Backovic, M., Jardetzky, T.S., 2011. Class III viral membrane fusion proteins. *Adv Exp Med Biol* 714, 91-101.

Backovic, M., Longnecker, R., Jardetzky, T.S., 2009. Structure of a trimeric variant of the Epstein-Barr virus glycoprotein B. *Proc Natl Acad Sci U S A* 106, 2880-2885.

Bagcchi, S., 2020. Lassa fever outbreak continues across Nigeria. *Lancet Infect Dis* 20, 543.

Bamford, D.H., 2003. Do viruses form lineages across different domains of life? *Res Microbiol* 154, 231-236.

Bamford, D.H., Grimes, J.M., Stuart, D.I., 2005. What does structure tell us about virus evolution? *Curr Opin Struct Biol* 15, 655-663.

Baquero, E., Albertini, A.A., Raux, H., Abou-Hamdan, A., Boeri-Erba, E., Ouldali, M., Buonocore, L., Rose, J.K., Lepault, J., Bressanelli, S., Gaudin, Y., 2017. Structural intermediates in the fusion-associated transition of vesiculovirus glycoprotein. *EMBO J* 36, 679-692.

Baquero, E., Albertini, A.A., Raux, H., Buonocore, L., Rose, J.K., Bressanelli, S., Gaudin, Y., 2015. Structure of the low pH conformation of Chandipura virus G reveals important features in the evolution of the vesiculovirus glycoprotein. *PLoS Pathog* 11, e1004756.

Bausch, D.G., Demby, A.H., Coulibaly, M., Kanu, J., Goba, A., Bah, A., Condé, N., Wurtzel, H.L., Cavallaro, K.F., Lloyd, E., Baldet, F.B., Cissé, S.D., Fofona, D., Savané, I.K., Tolno, R.T., Mahy, B., Wagoner, K.D., Ksiazek, T.G., Peters, C.J., Rollin, P.E., 2001. Lassa fever in Guinea: I. Epidemiology of human disease and clinical observations. *Vector Borne Zoonotic Dis* 1, 269-281.

Bederka, L.H., Bonhomme, C.J., Ling, E.L., Buchmeier, M.J., 2014. Arenavirus stable signal peptide is the keystone subunit for glycoprotein complex organization. *mBio* 5, e02063.

Belot, L., Ouldali, M., Roche, S., Legrand, P., Gaudin, Y., Albertini, A.A., 2020. Crystal structure of Mokola virus glycoprotein in its post-fusion conformation. *PLoS Pathog* 16, e1008383.

Benmansour, A., Leblois, H., Coulon, P., Tuffereau, C., Gaudin, Y., Flamand, A., Lafay, F., 1991. Antigenicity of rabies virus glycoprotein. *J Virol* 65, 4198-4203.

Beyer, W.R., Pöppelau, D., Garten, W., von Laer, D., Lenz, O., 2003. Endoproteolytic processing of the lymphocytic choriomeningitis virus glycoprotein by the subtilase SKI-1/S1P. *J Virol* 77, 2866-2872.

Blasdell, K.R., Duong, V., Eloit, M., Chretien, F., Ly, S., Hul, V., Deubel, V., Morand, S., Buchy, P., 2016. Evidence of human infection by a new mammarenavirus endemic to Southeastern Asia. *Elife* 5.

Bodewes, R., Kik, M.J.L., Raj, V.S., Schapendonk, C.M.E., Haagmans, B.L., Smits, S.L., Osterhaus, A., 2013. Detection of novel divergent arenaviruses in boid snakes with inclusion body disease in The Netherlands. *J Gen Virol* 94, 1206-1210.

Bolken, T.C., Laquerre, S., Zhang, Y., Bailey, T.R., Pevear, D.C., Kickner, S.S., Sperzel, L.E., Jones, K.F., Warren, T.K., Amanda Lund, S., Kirkwood-Watts, D.L., King, D.S., Shurtleff, A.C., Guttieri, M.C., Deng, Y., Bleam, M., Hruby, D.E., 2006. Identification and characterization of potent small molecule inhibitor of hemorrhagic fever New World arenaviruses. *Antiviral Res* 69, 86-97.

Bonhomme, C.J., Capul, A.A., Lauron, E.J., Bederka, L.H., Knopp, K.A., Buchmeier, M.J., 2011. Glycosylation modulates arenavirus glycoprotein expression and function. *Virology* 409, 223-233.

Bonhomme, C.J., Knopp, K.A., Bederka, L.H., Angelini, M.M., Buchmeier, M.J., 2013. LCMV glycosylation modulates viral fitness and cell tropism. *PLoS One* 8, e53273.

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

Borenstein-Katz, A., Shulman, A., Hamawi, H., Leitner, O., Diskin, R., 2019. Differential Antibody-Based Immune Response against Isolated GP1 Receptor-Binding Domains from Lassa and Junin Viruses. *J Virol* 93.

Borio, L., Inglesby, T., Peters, C.J., Schmaljohn, A.L., Hughes, J.M., Jahrling, P.B., Ksiazek, T., Johnson, K.M., Meyerhoff, A., O'Toole, T., Ascher, M.S., Bartlett, J., Breman, J.G., Eitzen, E.M., Jr., Hamburg, M., Hauer, J., Henderson, D.A., Johnson, R.T., Kwik, G., Layton, M., Lillibridge, S., Nabel, G.J., Osterholm, M.T., Perl, T.M., Russell, P., Tonat, K., 2002. Hemorrhagic fever viruses as biological weapons: medical and public health management. *Jama* 287, 2391-2405.

Bornholdt, Z.A., Turner, H.L., Murin, C.D., Li, W., Sok, D., Souders, C.A., Piper, A.E., Goff, A., Shamblin, J.D., Wollen, S.E., Sprague, T.R., Fusco, M.L., Pommert, K.B., Cavacini, L.A., Smith, H.L., Klempner, M., Reimann, K.A., Krauland, E., Gerngross, T.U., Wittrup, K.D., Saphire, E.O., Burton, D.R., Glass, P.J., Ward, A.B., Walker, L.M., 2016. Isolation of potent neutralizing antibodies from a survivor of the 2014 Ebola virus outbreak. *Science* 351, 1078-1083.

Both, G.W., Banerjee, A.K., Shatkin, A.J., 1975. Methylation-dependent translation of viral messenger RNAs in vitro. *Proc Natl Acad Sci U S A* 72, 1189-1193.

Bowden, T.A., Crispin, M., Graham, S.C., Harvey, D.J., Grimes, J.M., Jones, E.Y., Stuart, D.I., 2009. Unusual molecular architecture of the machupo virus attachment glycoprotein. *J Virol* 83, 8259-8265.

Bowden, T.A., Jones, E.Y., Stuart, D.I., 2011. Cells under siege: viral glycoprotein interactions at the cell surface. *J Struct Biol* 175, 120-126.

Bowen, M.D., Peters, C.J., Nichol, S.T., 1997. Phylogenetic analysis of the Arenaviridae: patterns of virus evolution and evidence for cospeciation between arenaviruses and their rodent hosts. *Mol Phylogenet Evol* 8, 301-316.

Braakman, I., van Anken, E., 2000. Folding of viral envelope glycoproteins in the endoplasmic reticulum. *Traffic* 1, 533-539.

Braberg, H., Webb, B.M., Tjioe, E., Pieper, U., Sali, A., Madhusudhan, M.S., 2012. SALIGN: a web server for alignment of multiple protein sequences and structures. *Bioinformatics* 28, 2072-2073.

Branco, L.M., Grove, J.N., Moses, L.M., Goba, A., Fullah, M., Momoh, M., Schoepp, R.J., Bausch, D.G., Garry, R.F., 2010. Shedding of soluble glycoprotein 1 detected during acute Lassa virus infection in human subjects. *Virol J* 7, 306.

Briknarová, K., Thomas, C.J., York, J., Nunberg, J.H., 2011. Structure of a zinc-binding domain in the Junin virus envelope glycoprotein. *J Biol Chem* 286, 1528-1536.

Brouillette, R.B., Phillips, E.K., Ayithan, N., Maury, W., 2017. Differences in Glycoprotein Complex Receptor Binding Site Accessibility Prompt Poor Cross-Reactivity of Neutralizing Antibodies between Closely Related Arenaviruses. *J Virol* 91.

Brunker, K., Mollentze, N., 2018. Rabies Virus. *Trends Microbiol* 26, 886-887.

Brunotte, L., Kerber, R., Shang, W., Hauer, F., Hass, M., Gabriel, M., Lelke, M., Busch, C., Stark, H., Svergun, D.I., Betzel, C., Perbandt, M., Günther, S., 2011. Structure of the Lassa virus nucleoprotein revealed by X-ray crystallography, small-angle X-ray scattering, and electron microscopy. *J Biol Chem* 286, 38748-38756.

Buchmeier, M., Adam, E., Rawls, W.E., 1974. Serological evidence of infection by Pichinde virus among laboratory workers. *Infect Immun* 9, 821-823.

Buchmeier, M., de la Torre, J.C., Peters, C.J., 2013. Arenaviridae, in: D.M. Knipe, P.M.H., J.I. Cohen, D.E. Griffin, R.A. Lamb, M.A. Martin, V.R. Racaniello, and B. Roizman (Ed.), *In*

Fields Virology. Wolters Kluwer Heath/Lippincott Williams & Wilkins, Philadelphia, PA, pp. 1283-1303.

Buchmeier, M.J., 2002. Arenaviruses: protein structure and function. *Curr Top Microbiol Immunol* 262, 159-173.

Burrell, C.J., Howard, C.R., Murphy, F.A., 2017. Chapter 30 - Arenaviruses, in: Burrell, C.J., Howard, C.R., Murphy, F.A. (Eds.), *Fenner and White's Medical Virology* (Fifth Edition). Academic Press, London, pp. 425-436.

Burri, D.J., da Palma, J.R., Kunz, S., Pasquato, A., 2012. Envelope glycoprotein of arenaviruses. *Viruses* 4, 2162-2181.

Burri, D.J., da Palma, J.R., Seidah, N.G., Zanotti, G., Cendron, L., Pasquato, A., Kunz, S., 2013a. Differential recognition of Old World and New World arenavirus envelope glycoproteins by subtilisin kexin isozyme 1 (SKI-1)/site 1 protease (S1P). *J Virol* 87, 6406-6414.

Burri, D.J., Pasquato, A., da Palma, J.R., Igonet, S., Oldstone, M.B., Kunz, S., 2013b. The role of proteolytic processing and the stable signal peptide in expression of the Old World arenavirus envelope glycoprotein ectodomain. *Virology* 436, 127-133.

Burton, D.R., 2017. What Are the Most Powerful Immunogen Design Vaccine Strategies? Reverse Vaccinology 2.0 Shows Great Promise. *Cold Spring Harb Perspect Biol* 9.

Buthlezi, S.G., Dirr, H.W., Chakauya, E., Chikwamba, R., Martens, L., Tsekoa, T.L., Stoychev, S.H., Vandermarliere, E., 2016. The Lyssavirus glycoprotein: A key to cross-immunity. *Virology* 498, 250-256.

Candurra, N.A., Maskin, L., Damonte, E.B., 1996. Inhibition of arenavirus multiplication in vitro by phenothiazines. *Antiviral Res* 31, 149-158.

Cantaert, T., Borand, L., Kergoat, L., Leng, C., Ung, S., In, S., Peng, Y., Phoeun, C., Hing, C., Taing, C.N., Saman, M., Ong, S., Mey, C., Choeung, R., Ly, S., Dussart, P., Bourhy, H., Tarantola, A., 2019. A 1-week intradermal dose-sparing regimen for rabies post-exposure prophylaxis (RESIST-2): an observational cohort study. *Lancet Infect Dis* 19, 1355-1362.

Cao, W., Henry, M.D., Borrow, P., Yamada, H., Elder, J.H., Ravkov, E.V., Nichol, S.T., Compans, R.W., Campbell, K.P., Oldstone, M.B., 1998. Identification of alpha-dystroglycan as a receptor for lymphocytic choriomeningitis virus and Lassa fever virus. *Science* 282, 2079-2081.

Capul, A.A., de la Torre, J.C., Buchmeier, M.J., 2011. Conserved residues in Lassa fever virus Z protein modulate viral infectivity at the level of the ribonucleoprotein. *J Virol* 85, 3172-3178.

Capul, A.A., Perez, M., Burke, E., Kunz, S., Buchmeier, M.J., de la Torre, J.C., 2007. Arenavirus Z-glycoprotein association requires Z myristoylation but not functional RING or late domains. *J Virol* 81, 9451-9460.

Carrillo-Bustamante, P., Nguyen, T.H.T., Oestereich, L., Günther, S., Guedj, J., Graw, F., 2017. Determining Ribavirin's mechanism of action against Lassa virus infection. *Scientific Reports* 7, 11693.

Casabona, J.C., Levingston Macleod, J.M., Loureiro, M.E., Gomez, G.A., Lopez, N., 2009. The RING domain and the L79 residue of Z protein are involved in both the rescue of nucleocapsids and the incorporation of glycoproteins into infectious chimeric arenavirus-like particles. *J Virol* 83, 7029-7039.

CDC, 2014. Lymphocytic Choriomeningitis Signs and Symptoms.

Chai, N., Swem, L.R., Park, S., Nakamura, G., Chiang, N., Estevez, A., Fong, R., Kamen, L., Kho, E., Reichelt, M., Lin, Z., Chiu, H., Skippington, E., Modrusan, Z., Stinson, J., Xu, M., Lupardus, P., Ciferri, C., Tan, M.W., 2017. A broadly protective therapeutic antibody against influenza B virus with two mechanisms of action. *Nat Commun* 8, 14234.

Chandramouli, S., Ciferri, C., Nikitin, P.A., Caló, S., Gerrein, R., Balabanis, K., Monroe, J., Hebner, C., Lilja, A.E., Settembre, E.C., Carfi, A., 2015. Structure of HCMV glycoprotein B in the postfusion conformation bound to a neutralizing human antibody. *Nature Communications* 6, 8176.

Charrel, R.N., Coutard, B., Baronti, C., Canard, B., Nougairede, A., Frangeul, A., Morin, B., Jamal, S., Schmidt, C.L., Hilgenfeld, R., Klempa, B., de Lamballerie, X., 2011. Arenaviruses and hantaviruses: from epidemiology and genomics to antivirals. *Antiviral Res* 90, 102-114.

Charrel, R.N., de Lamballerie, X., 2003. Arenaviruses other than Lassa virus. *Antiviral Res* 57, 89-100.

Charrel, R.N., de Lamballerie, X., Emonet, S., 2008. Phylogeny of the genus Arenavirus. *Curr Opin Microbiol* 11, 362-368.

Chen, J., Noble, A.J., Kang, J.Y., Darst, S.A., 2019. Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: Bacterial RNA polymerase and CHAPSO. *J Struct Biol X* 1.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66, 12-21.

Chenik, M., Schnell, M., Conzelmann, K.K., Blondel, D., 1998. Mapping the interacting domains between the rabies virus polymerase and phosphoprotein. *J Virol* 72, 1925-1930.

Choe, H., Jemielity, S., Abraham, J., Radoshitzky, S.R., Farzan, M., 2011. Transferrin receptor 1 in the zoonosis and pathogenesis of New World hemorrhagic fever arenaviruses. *Curr Opin Microbiol* 14, 476-482.

Choi, K.H., 2012. Viral polymerases. *Adv Exp Med Biol* 726, 267-304.

Clark, L.E., Mahmutovic, S., Raymond, D.D., Dilanyan, T., Koma, T., Manning, J.T., Shankar, S., Levis, S.C., Briggiler, A.M., Enria, D.A., Wucherpfennig, K.W., Paessler, S., Abraham, J., 2018. Vaccine-elicited receptor-binding site antibodies neutralize two New World hemorrhagic fever arenaviruses. *Nat Commun* 9, 1884.

Clegg, J.C., 2002. Molecular phylogeny of the arenaviruses. *Curr Top Microbiol Immunol* 262, 1-24.

Cohen, F.S., Melikyan, G.B., 2004. The energetics of membrane fusion from binding, through hemifusion, pore formation, and pore enlargement. *J Membr Biol* 199, 1-14.

Cohen-Dvashi, H., Amon, R., Agans, K.N., Cross, R.W., Borenstein-Katz, A., Mateo, M., Baize, S., Padler-Karavani, V., Geisbert, T.W., Diskin, R., 2020. Rational design of universal immunotherapy for TfR1-tropic arenaviruses. *Nat Commun* 11, 67.

Cohen-Dvashi, H., Cohen, N., Israeli, H., Diskin, R., 2015. Molecular Mechanism for LAMP1 Recognition by Lassa Virus. *J Virol* 89, 7584-7592.

Cohen-Dvashi, H., Israeli, H., Shani, O., Katz, A., Diskin, R., 2016. Role of LAMP1 Binding and pH Sensing by the Spike Complex of Lassa Virus. *J Virol* 90, 10329-10338.

Cohen-Dvashi, H., Kilimnik, I., Diskin, R., 2018. Structural basis for receptor recognition by Lujo virus. *Nat Microbiol* 3, 1153-1160.

Consortium, W.R.M., 2019. Zero human deaths from dog-mediated rabies by 2030: perspectives from quantitative and mathematical modelling. *Gates Open Res* 3, 1564.

Cornu, T.I., Feldmann, H., de la Torre, J.C., 2004. Cells Expressing the RING Finger Z Protein Are Resistant to Arenavirus Infection. *Journal of Virology* 78, 2979.

Corpet, F., 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* 16, 10881-10890.

Corti, D., Misasi, J., Mulangu, S., Stanley, D.A., Kanekiyo, M., Wollen, S., Ploquin, A., Doria-Rose, N.A., Staupe, R.P., Bailey, M., Shi, W., Choe, M., Marcus, H., Thompson, E.A., Cagigi, A., Silacci, C., Fernandez-Rodriguez, B., Perez, L., Sallusto, F., Vanzetta, F., Agatic, G., Cameroni, E., Kisalu, N., Gordon, I., Ledgerwood, J.E., Mascola, J.R., Graham, B.S., Muyembe-Tamfun, J.J., Trefry, J.C., Lanzavecchia, A., Sullivan, N.J., 2016. Protective monotherapy against lethal Ebola virus infection by a potentially neutralizing antibody. *Science* 351, 1339-1342.

Cosset, F.L., Marianneau, P., Verney, G., Gallais, F., Tordo, N., Pêcheur, E.I., ter Meulen, J., Deubel, V., Bartosch, B., 2009. Characterization of Lassa virus cell entry and neutralization with Lassa virus pseudoparticles. *J Virol* 83, 3228-3237.

Coto, C.E., Rey, E., Parodi, A.S., 1967. Tacaribe virus infection of guinea-pig. Virus distribution, appearance of antibodies and immunity against Junin virus infection. *Arch Gesamte Virusforsch* 20, 81-86.

Crispin, M., Ward, A.B., Wilson, I.A., 2018. Structure and Immune Recognition of the HIV Glycan Shield. *Annu Rev Biophys* 47, 499-523.

Crispin, M., Zeltina, A., Zitzmann, N., Bowden, T.A., 2016. Native functionality and therapeutic targeting of arenaviral glycoproteins. *Curr Opin Virol* 18, 70-75.

da Palma, J.R., Burri, D.J., Oppliger, J., Salamina, M., Cendron, L., de Laureto, P.P., Seidah, N.G., Kunz, S., Pasquato, A., 2014. Zymogen activation and subcellular activity of subtilisin kexin isozyme 1/site 1 protease. *J Biol Chem* 289, 35743-35756.

Dan-Nwafor, C.C., Ipadeola, O., Smout, E., Ilori, E., Adeyemo, A., Umeokonkwo, C., Nwidi, D., Nwachukwu, W., Ukponu, W., Omabe, E., Anaebonam, U., Igwenyi, N., Igbodo, G., Eteng, W., Uzoma, I., Saleh, M., Agboeze, J., Mutbam, S., de Gooyer, T., Short, R., Aniaku, E., Onoh, R., Ogah, E., Nguku, P., Oladejo, J., Peter, C., Ojo, O., Ihekweazu, C., 2019. A cluster of nosocomial Lassa fever cases in a tertiary health facility in Nigeria: Description and lessons learned, 2018. *Int J Infect Dis* 83, 88-94.

De Benedictis, P., Minola, A., Rota Nodari, E., Aiello, R., Zecchin, B., Salomoni, A., Foglierini, M., Agatic, G., Vanzetta, F., Lavenir, R., Lepelletier, A., Bentley, E., Weiss, R., Cattoli, G., Capua, I., Sallusto, F., Wright, E., Lanzavecchia, A., Bourhy, H., Corti, D., 2016.

Development of broad-spectrum human monoclonal antibodies for rabies post-exposure prophylaxis. *EMBO Mol Med* 8, 407-421.

Deb, P., Molla, M.M.A., Saif-Ur-Rahman, K.M., 2021. An update to monoclonal antibody as therapeutic option against COVID-19. *Biosafety and Health* 3, 87-91.

Demogines, A., Abraham, J., Choe, H., Farzan, M., Sawyer, S.L., 2013. Dual host-virus arms races shape an essential housekeeping protein. *PLoS Biol* 11, e1001571.

Dietzschold, B., Kao, M., Zheng, Y.M., Chen, Z.Y., Maul, G., Fu, Z.F., Rupprecht, C.E., Koprowski, H., 1992. Delineation of putative mechanisms involved in antibody-mediated clearance of rabies virus from the central nervous system. *Proc Natl Acad Sci U S A* 89, 7252-7256.

Doores, K.J., 2015. The HIV glycan shield as a target for broadly neutralizing antibodies. *Febs j* 282, 4679-4691.

Downs, W.G., Anderson, C.R., Spence, L., Aitken, T.H., Greenhall, A.H., 1963. Tacaribe virus, a new agent isolated from *Artibeus* bats and mosquitoes in Trinidad, West Indies. *Am J Trop Med Hyg* 12, 640-646.

Eichler, R., Lenz, O., Strecker, T., Eickmann, M., Klenk, H.D., Garten, W., 2003. Identification of Lassa virus glycoprotein signal peptide as a trans-acting maturation factor. *EMBO Rep* 4, 1084-1088.

Eichler, R., Lenz, O., Strecker, T., Eickmann, M., Klenk, H.D., Garten, W., 2004. Lassa virus glycoprotein signal peptide displays a novel topology with an extended endoplasmic reticulum luminal region. *J Biol Chem* 279, 12293-12299.

Elagoz, A., Benjannet, S., Mammarbassi, A., Wickham, L., Seidah, N.G., 2002. Biosynthesis and cellular trafficking of the convertase SKI-1/S1P: ectodomain shedding requires SKI-1 activity. *J Biol Chem* 277, 11265-11275.

Elbein, A.D., Tropea, J.E., Mitchell, M., Kaushal, G.P., 1990. Kifunensine, a potent inhibitor of the glycoprotein processing mannosidase I. *J Biol Chem* 265, 15599-15605.

Elegheert, J., Behiels, E., Bishop, B., Scott, S., Woolley, R.E., Griffiths, S.C., Byrne, E.F.X., Chang, V.T., Stuart, D.I., Jones, E.Y., Siebold, C., Aricescu, A.R., 2018. Lentiviral transduction of mammalian cells for fast, scalable and high-level production of soluble and membrane proteins. *Nat Protoc* 13, 2991-3017.

Emonet, S.F., de la Torre, J.C., Domingo, E., Sevilla, N., 2009. Arenavirus genetic diversity and its biological implications. *Infect Genet Evol* 9, 417-429.

Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126-2132.

Enria, D.A., Briggiler, A.M., Fernandez, N.J., Levis, S.C., Maiztegui, J.I., 1984. Importance of dose of neutralising antibodies in treatment of Argentine haemorrhagic fever with immune plasma. *Lancet* 2, 255-256.

Enria, D.A., Briggiler, A.M., Sánchez, Z., 2008. Treatment of Argentine hemorrhagic fever. *Antiviral Res* 78, 132-139.

Enserink, M., 2000. Emerging diseases. New arenavirus blamed for recent deaths in California. *Science* 289, 842-843.

Escalera-Antezana, J.P., Rodriguez-Villena, O.J., Arancibia-Alba, A.W., Alvarado-Arnez, L.E., Bonilla-Aldana, D.K., Rodríguez-Morales, A.J., 2020. Clinical features of fatal cases of Chapare virus hemorrhagic fever originating from rural La Paz, Bolivia, 2019: A cluster analysis. *Travel Med Infect Dis* 36, 101589.

Eschli, B., Quirin, K., Wepf, A., Weber, J., Zinkernagel, R., Hengartner, H., 2006. Identification of an N-terminal trimeric coiled-coil core within arenavirus glycoprotein 2 permits assignment to class I viral fusion proteins. *J Virol* 80, 5897-5907.

Falkowska, E., Le, K.M., Ramos, A., Doores, K.J., Lee, J.H., Blattner, C., Ramirez, A., Derking, R., van Gils, M.J., Liang, C.H., McBride, R., von Bredow, B., Shivatare, S.S., Wu, C.Y., Chan-Hui, P.Y., Liu, Y., Feizi, T., Zwick, M.B., Koff, W.C., Seaman, M.S., Swiderek, K., Moore, J.P., Evans, D., Paulson, J.C., Wong, C.H., Ward, A.B., Wilson, I.A., Sanders, R.W., Poignard, P., Burton, D.R., 2014. Broadly neutralizing HIV antibodies define a glycan-dependent epitope on the prefusion conformation of gp41 on cleaved envelope trimers. *Immunity* 40, 657-668.

Fedeli, C., Torriani, G., Galan-Navarro, C., Moraz, M.L., Moreno, H., Gerold, G., Kunz, S., 2018. Axl Can Serve as Entry Factor for Lassa Virus Depending on the Functional Glycosylation of Dystroglycan. *J Virol* 92.

Fehling, S.K., Lennartz, F., Strecker, T., 2012. Multifunctional nature of the arenavirus RING finger protein Z. *Viruses* 4, 2973-3011.

Felsenstein, J., 1989. PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics* 5, 164-166.

- Fenner, F., 1976. Classification and nomenclature of viruses. Second report of the International Committee on Taxonomy of Viruses. *Intervirology* 7, 1-115.
- Ferron, F., Weber, F., de la Torre, J.C., Reguera, J., 2017. Transcription and replication mechanisms of Bunyaviridae and Arenaviridae L proteins. *Virus Res* 234, 118-134.
- Fichet-Calvet, E., Rogers, D.J., 2009. Risk maps of Lassa fever in West Africa. *PLoS Negl Trop Dis* 3, e388.
- Finke, S., Mueller-Waldeck, R., Conzelmann, K.K., 2003. Rabies virus matrix protein regulates the balance of virus transcription and replication. *J Gen Virol* 84, 1613-1621.
- Fischer, S.A., Graham, M.B., Kuehnert, M.J., Kotton, C.N., Srinivasan, A., Marty, F.M., Comer, J.A., Guarner, J., Paddock, C.D., DeMeo, D.L., Shieh, W.J., Erickson, B.R., Bandy, U., DeMaria, A., Jr., Davis, J.P., Delmonico, F.L., Pavlin, B., Likos, A., Vincent, M.J., Sealy, T.K., Goldsmith, C.S., Jernigan, D.B., Rollin, P.E., Packard, M.M., Patel, M., Rowland, C., Helfand, R.F., Nichol, S.T., Fishman, J.A., Ksiazek, T., Zaki, S.R., 2006. Transmission of lymphocytic choriomeningitis virus by organ transplantation. *N Engl J Med* 354, 2235-2249.
- Fisher, C.R., Streicker, D.G., Schnell, M.J., 2018. The spread and evolution of rabies virus: conquering new frontiers. *Nature Reviews Microbiology* 16, 241-255.
- Fisher-Hoch, S.P., McCormick, J.B., 2001. Towards a human Lassa fever vaccine. *Rev Med Virol* 11, 331-341.
- Flanagan, M.L., Oldenburg, J., Reignier, T., Holt, N., Hamilton, G.A., Martin, V.K., Cannon, P.M., 2008. New world clade B arenaviruses can use transferrin receptor 1 (TfR1)-dependent and -independent entry pathways, and glycoproteins from human pathogenic strains are associated with the use of TfR1. *J Virol* 82, 938-948.
- Fooks, A.R., Cliquet, F., Finke, S., Freuling, C., Hemachudha, T., Mani, R.S., Müller, T., Nadin-Davis, S., Picard-Meyer, E., Wilde, H., Banyard, A.C., 2017. Rabies. *Nat Rev Dis Primers* 3, 17091.
- Frame, J.D., Baldwin, J.M., Jr., Gocke, D.J., Troup, J.M., 1970. Lassa fever, a new virus disease of man from West Africa. I. Clinical description and pathological findings. *Am J Trop Med Hyg* 19, 670-676.
- Franze-Fernández, M.T., Zetina, C., Iapalucci, S., Lucero, M.A., Bouissou, C., López, R., Rey, O., Daheli, M., Cohen, G.N., Zakin, M.M., 1987. Molecular structure and early events in the replication of Tacaribe arenavirus S RNA. *Virus Res* 7, 309-324.

Freed, E.O., 2002. Viral late domains. *J Virol* 76, 4679-4687.

Fu, Z.F., Zheng, Y., Wunner, W.H., Koprowski, H., Dietzschold, B., 1994. Both the N- and the C-terminal domains of the nominal phosphoprotein of rabies virus are involved in binding to the nucleoprotein. *Virology* 200, 590-597.

Fulhorst, C.F., Charrel, R.N., Weaver, S.C., Ksiazek, T.G., Bradley, R.D., Milazzo, M.L., Tesh, R.B., Bowen, M.D., 2001. Geographic distribution and genetic diversity of Whitewater Arroyo virus in the southwestern United States. *Emerg Infect Dis* 7, 403-407.

Fuller-Pace, F.V., Southern, P.J., 1988. Temporal analysis of transcription and replication during acute infection with lymphocytic choriomeningitis virus. *Virology* 162, 260-263.

Furuichi, Y., LaFiandra, A., Shatkin, A.J., 1977. 5'-Terminal structure and mRNA stability. *Nature* 266, 235-239.

Furuta, Y., Takahashi, K., Fukuda, Y., Kuno, M., Kamiyama, T., Kozaki, K., Nomura, N., Egawa, H., Minami, S., Watanabe, Y., Narita, H., Shiraki, K., 2002. In vitro and in vivo activities of anti-influenza virus compound T-705. *Antimicrob Agents Chemother* 46, 977-981.

Furuta, Y., Takahashi, K., Kuno-Maekawa, M., Sangawa, H., Uehara, S., Kozaki, K., Nomura, N., Egawa, H., Shiraki, K., 2005. Mechanism of action of T-705 against influenza virus. *Antimicrob Agents Chemother* 49, 981-986.

Garcin, D., Kolakofsky, D., 1990. A novel mechanism for the initiation of Tacaribe arenavirus genome replication. *J Virol* 64, 6196-6203.

Garcin, D., Kolakofsky, D., 1992. Tacaribe arenavirus RNA synthesis in vitro is primer dependent and suggests an unusual model for the initiation of genome replication. *J Virol* 66, 1370-1376.

Gaudin, Y., Tuffereau, C., Segretain, D., Knossow, M., Flamand, A., 1991. Reversible conformational changes and fusion activity of rabies virus glycoprotein. *Journal of Virology* 65, 4853.

Gerber, F., Tetchi, M., Kallo, V., Léchenne, M., Hattendorf, J., Bonfoh, B., Zinsstag, J., 2020. RABIES IMMUNOGLOBULIN: Brief history and recent experiences in Côte d'Ivoire. *Acta Trop* 211, 105629.

Giesen, A., Gniel, D., Malerczyk, C., 2015. 30 Years of rabies vaccination with Rabipur: a summary of clinical data and global experience. *Expert Rev Vaccines* 14, 351-367.

Gilman, M.S., Castellanos, C.A., Chen, M., Ngwuta, J.O., Goodwin, E., Moin, S.M., Mas, V., Melero, J.A., Wright, P.F., Graham, B.S., McLellan, J.S., Walker, L.M., 2016. Rapid profiling of RSV antibody repertoires from the memory B cells of naturally infected adult donors. *Sci Immunol* 1.

Gluska, S., Zahavi, E.E., Chein, M., Gradus, T., Bauer, A., Finke, S., Perlson, E., 2014. Rabies Virus Hijacks and accelerates the p75NTR retrograde axonal transport machinery. *PLoS Pathog* 10, e1004348.

Goeijenbier, M., Wagenaar, J., Goris, M., Martina, B., Henttonen, H., Vaheri, A., Reusken, C., Hartskeerl, R., Osterhaus, A., Van Gorp, E., 2013. Rodent-borne hemorrhagic fevers: under-recognized, widely spread and preventable - epidemiology, diagnostics and treatment. *Crit Rev Microbiol* 39, 26-42.

Goncalves, A.R., Moraz, M.L., Pasquato, A., Helenius, A., Lozach, P.Y., Kunz, S., 2013. Role of DC-SIGN in Lassa virus entry into human dendritic cells. *J Virol* 87, 11504-11515.

González, P.H., Cossio, P.M., Arana, R., Maiztegui, J.I., Laguens, R.P., 1980. Lymphatic tissue in Argentine hemorrhagic fever. Pathologic features. *Arch Pathol Lab Med* 104, 250-254.

Goodwin, E., Gilman, M.S.A., Wrapp, D., Chen, M., Ngwuta, J.O., Moin, S.M., Bai, P., Sivasubramanian, A., Connor, R.I., Wright, P.F., Graham, B.S., McLellan, J.S., Walker, L.M., 2018. Infants Infected with Respiratory Syncytial Virus Generate Potent Neutralizing Antibodies that Lack Somatic Hypermutation. *Immunity* 48, 339-349.e335.

Gowen, B.B., Wong, M.H., Jung, K.H., Sanders, A.B., Mendenhall, M., Bailey, K.W., Furuta, Y., Sidwell, R.W., 2007. In vitro and in vivo activities of T-705 against arenavirus and bunyavirus infections. *Antimicrob Agents Chemother* 51, 3168-3176.

Graham, S.C., Assenberg, R., Delmas, O., Verma, A., Gholami, A., Talbi, C., Owens, R.J., Stuart, D.I., Grimes, J.M., Bourhy, H., 2008. Rhabdovirus matrix protein structures reveal a novel mode of self-association. *PLoS Pathog* 4, e1000251.

Grande-Pérez, A., Martin, V., Moreno, H., de la Torre, J.C., 2016. Arenavirus Quasispecies and Their Biological Implications, in: Domingo, E., Schuster, P. (Eds.), *Quasispecies: From Theory to Experimental Systems*. Springer International Publishing, Cham, pp. 231-275.

Grant, A., Seregin, A., Huang, C., Kolokoltsova, O., Brasier, A., Peters, C., Paessler, S., 2012. Junín virus pathogenesis and virus replication. *Viruses* 4, 2317-2339.

Groseth, A., Wolff, S., Strecker, T., Hoenen, T., Becker, S., 2010. Efficient budding of the tacaribe virus matrix protein z requires the nucleoprotein. *J Virol* 84, 3603-3611.

Guichard, P., Krell, T., Chevalier, M., Vaysse, C., Adam, O., Ronzon, F., Marco, S., 2011. Three dimensional morphology of rabies virus studied by cryo-electron tomography. *Journal of Structural Biology* 176, 32-40.

Guo, L., Liu, S., Song, J., Han, L., Zhang, H., Wu, C., Wang, C., Zhou, H., Wang, J., 2020. Seroprevalence of Wenzhou virus in China. *Biosaf Health* 2, 152-156.

Guo, Y., Duan, M., Wang, X., Gao, J., Guan, Z., Zhang, M., 2019. Early events in rabies virus infection-Attachment, entry, and intracellular trafficking. *Virus Res* 263, 217-225.

Haist, K., Ziegler, C., Botten, J., 2015. Strand-Specific Quantitative Reverse Transcription-Polymerase Chain Reaction Assay for Measurement of Arenavirus Genomic and Antigenomic RNAs. *PLoS One* 10, e0120043.

Hallam, H.J., Hallam, S., Rodriguez, S.E., Barrett, A.D.T., Beasley, D.W.C., Chua, A., Ksiazek, T.G., Milligan, G.N., Sathiyamoorthy, V., Reece, L.M., 2018a. Baseline mapping of Lassa fever virology, epidemiology and vaccine research and development. *NPJ Vaccines* 3, 11.

Hallam, S.J., Koma, T., Maruyama, J., Paessler, S., 2018b. Review of Mammarenavirus Biology and Replication. *Front Microbiol* 9, 1751.

Halldorsson, S., Li, S., Li, M., Harlos, K., Bowden, T.A., Huiskonen, J.T., 2018. Shielding and activation of a viral membrane fusion protein. *Nature Communications* 9, 349.

Harrison, S.C., 2008. Viral membrane fusion. *Nature Structural & Molecular Biology* 15, 690-698.

Harty, R.N., Paragas, J., Sudol, M., Palese, P., 1999. A proline-rich motif within the matrix protein of vesicular stomatitis virus and rabies virus interacts with WW domains of cellular proteins: implications for viral budding. *J Virol* 73, 2921-2929.

Hass, M., Westerkofsky, M., Müller, S., Becker-Ziaja, B., Busch, C., Günther, S., 2006. Mutational analysis of the lassa virus promoter. *J Virol* 80, 12414-12419.

Hastie, K.M., Igonet, S., Sullivan, B.M., Legrand, P., Zandonatti, M.A., Robinson, J.E., Garry, R.F., Rey, F.A., Oldstone, M.B., Saphire, E.O., 2016a. Crystal structure of the prefusion surface glycoprotein of the prototypic arenavirus LCMV. *Nat Struct Mol Biol* 23, 513-521.

Hastie, K.M., Liu, T., Li, S., King, L.B., Ngo, N., Zandonatti, M.A., Woods, V.L., Jr., de la Torre, J.C., Saphire, E.O., 2011. 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.

Hastie, K.M., Saphire, E.O., 2018. Lassa virus glycoprotein: stopping a moving target. *Curr Opin Virol* 31, 52-58.

Hastie, K.M., Zandonatti, M., Liu, T., Li, S., Woods, V.L., Jr., Saphire, E.O., 2016b. Crystal Structure of the Oligomeric Form of Lassa Virus Matrix Protein Z. *J Virol* 90, 4556-4562.

Hastie, K.M., Zandonatti, M.A., Kleinfelter, L.M., Heinrich, M.L., Rowland, M.M., Chandran, K., Branco, L.M., Robinson, J.E., Garry, R.F., Saphire, E.O., 2017. Structural basis for antibody-mediated neutralization of Lassa virus. *Science* 356, 923-928.

Heckman, K.L., Pease, L.R., 2007. Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat Protoc* 2, 924-932.

Heldwein, E.E., Lou, H., Bender, F.C., Cohen, G.H., Eisenberg, R.J., Harrison, S.C., 2006. Crystal structure of glycoprotein B from herpes simplex virus 1. *Science* 313, 217-220.

Hellert, J., Buchrieser, J., Larrous, F., Minola, A., de Melo, G.D., Soriaga, L., England, P., Haouz, A., Telenti, A., Schwartz, O., Corti, D., Bourhy, H., Rey, F.A., 2020. Structure of the prefusion-locking broadly neutralizing antibody RVC20 bound to the rabies virus glycoprotein. *Nat Commun* 11, 596.

Hepojoki, J., Hepojoki, S., Smura, T., Szirovicza, L., Dervas, E., Prähauser, B., Nufer, L., Schraner, E.M., Vapalahti, O., Kipar, A., Hetzel, U., 2018. Characterization of Haartman Institute snake virus-1 (HISV-1) and HISV-like viruses-The representatives of genus Hartmanivirus, family Arenaviridae. *PLoS Pathog* 14, e1007415.

Hepojoki, J., Salmenperä, P., Sironen, T., Hetzel, U., Korzyukov, Y., Kipar, A., Vapalahti, O., 2015. Arenavirus Coinfections Are Common in Snakes with Boid Inclusion Body Disease. *J Virol* 89, 8657-8660.

Hetzel, U., Sironen, T., Laurinmäki, P., Liljeroos, L., Patjas, A., Henttonen, H., Vaheri, A., Artelt, A., Kipar, A., Butcher, S.J., Vapalahti, O., Hepojoki, J., 2013. Isolation, identification, and characterization of novel arenaviruses, the etiological agents of boid inclusion body disease. *J Virol* 87, 10918-10935.

Hirabayashi, Y., Oka, S., Goto, H., Shimada, K., Kurata, T., Fisher-Hoch, S.P., McCormick, J.B., 1988. An imported case of Lassa fever with late appearance of polyserositis. *J Infect Dis* 158, 872-875.

Hoenen, T., Kolesnikova, L., Becker, S., 2007. Recent advances in filovirus- and arenavirus-like particles. *Future Virology* 2, 193-203.

Holm, L., Sander, C., 1993. Protein structure comparison by alignment of distance matrices. *J Mol Biol* 233, 123-138.

Hulseberg, C.E., Fénéant, L., Szymańska, K.M., White, J.M., 2018. Lamp1 Increases the Efficiency of Lassa Virus Infection by Promoting Fusion in Less Acidic Endosomal Compartments. *mBio* 9.

Huson, D.H., Richter, D.C., Rausch, C., Dezulian, T., Franz, M., Rupp, R., 2007. Dendroscope: An interactive viewer for large phylogenetic trees. *BMC Bioinformatics* 8, 460.

Huttner, A., Dayer, J.A., Yerly, S., Combescure, C., Auderset, F., Desmeules, J., Eickmann, M., Finckh, A., Goncalves, A.R., Hooper, J.W., Kaya, G., Krähling, V., Kwilas, S., Lemaître, B., Matthey, A., Silvera, P., Becker, S., Fast, P.E., Moorthy, V., Kieny, M.P., Kaiser, L., Siegrist, C.A., 2015. The effect of dose on the safety and immunogenicity of the VSV Ebola candidate vaccine: a randomised double-blind, placebo-controlled phase 1/2 trial. *Lancet Infect Dis* 15, 1156-1166.

Igonet, S., Vaney, M.C., Vonrhein, C., Bricogne, G., Stura, E.A., Hengartner, H., Eschli, B., Rey, F.A., 2011. X-ray structure of the arenavirus glycoprotein GP2 in its postfusion hairpin conformation. *Proc Natl Acad Sci U S A* 108, 19967-19972.

Irie, T., Shimazu, Y., Yoshida, T., Sakaguchi, T., 2007. The YLDL sequence within Sendai virus M protein is critical for budding of virus-like particles and interacts with Alix/AIP1 independently of C protein. *J Virol* 81, 2263-2273.

Israeli, H., Cohen-Dvashi, H., Shulman, A., Shimon, A., Diskin, R., 2017. Mapping of the Lassa virus LAMP1 binding site reveals unique determinants not shared by other old world arenaviruses. *PLoS Pathog* 13, e1006337.

Iwasaki, M., Cubitt, B., Sullivan, B.M., de la Torre, J.C., 2016. The High Degree of Sequence Plasticity of the Arenavirus Noncoding Intergenic Region (IGR) Enables the Use of a Nonviral Universal Synthetic IGR To Attenuate Arenaviruses. *J Virol* 90, 3187-3197.

Jackson, A.C., 2018. Rabies: a medical perspective. *Rev Sci Tech* 37, 569-580.

Jacome, R., Becerra, A., Ponce de Leon, S., Lazcano, A., 2015. Structural Analysis of Monomeric RNA-Dependent Polymerases: Evolutionary and Therapeutic Implications. *PLoS One* 10, e0139001.

Jae, L.T., Brummelkamp, T.R., 2015. Emerging intracellular receptors for hemorrhagic fever viruses. *Trends in Microbiology* 23, 392-400.

Jae, L.T., Raaben, M., Herbert, A.S., Kuehne, A.I., Wirchnianski, A.S., Soh, T.K., Stubbs, S.H., Janssen, H., Damme, M., Saftig, P., Whelan, S.P., Dye, J.M., Brummelkamp, T.R., 2014. Virus entry. Lassa virus entry requires a trigger-induced receptor switch. *Science* 344, 1506-1510.

Jay, M.T., Glaser, C., Fulhorst, C.F., 2005. The arenaviruses. *J Am Vet Med Assoc* 227, 904-915.

Jemielity, S., Wang, J.J., Chan, Y.K., Ahmed, A.A., Li, W., Monahan, S., Bu, X., Farzan, M., Freeman, G.J., Umetsu, D.T., Dekruyff, R.H., Choe, H., 2013. TIM-family proteins promote infection of multiple enveloped viruses through virion-associated phosphatidylserine. *PLoS Pathog* 9, e1003232.

Johnson, K.M., McCormick, J.B., Webb, P.A., Smith, E.S., Elliott, L.H., King, I.J., 1987. Clinical virology of Lassa fever in hospitalized patients. *J Infect Dis* 155, 456-464.

Jones, B.A., Grace, D., Kock, R., Alonso, S., Rushton, J., Said, M.Y., McKeever, D., Mutua, F., Young, J., McDermott, J., Pfeiffer, D.U., 2013. Zoonosis emergence linked to agricultural intensification and environmental change. *Proc Natl Acad Sci U S A* 110, 8399-8404.

Kauffman, F.H., Goldmann, B.J., 1986. Rabies. *Am J Emerg Med* 4, 525-531.

Kelley, B., 2009. Industrialization of mAb production technology: the bioprocessing industry at a crossroads. *MAbs* 1, 443-452.

Kentsis, A., Gordon, R.E., Borden, K.L., 2002. Control of biochemical reactions through supramolecular RING domain self-assembly. *Proc Natl Acad Sci U S A* 99, 15404-15409.

Kenyon, R.H., Green, D.E., Eddy, G.A., Peters, C.J., 1986. Treatment of junin virus-infected guinea pigs with immune serum: development of late neurological disease. *J Med Virol* 20, 207-218.

Kerr, S.A., Jackson, E.L., Lungu, O.I., Meyer, A.G., Demogines, A., Ellington, A.D., Georgiou, G., Wilke, C.O., Sawyer, S.L., 2015. Computational and Functional Analysis of the Virus-Receptor Interface Reveals Host Range Trade-Offs in New World Arenaviruses. *J Virol* 89, 11643-11653.

Kielian, M., Rey, F.A., 2006. Virus membrane-fusion proteins: more than one way to make a hairpin. *Nature Reviews Microbiology* 4, 67-76.

Kiley, M.P., Lange, J.V., Johnson, K.M., 1979. Protection of rhesus monkeys from Lassa virus by immunisation with closely related Arenavirus. *Lancet* 2, 738.

Kilgore, P.E., Ksiazek, T.G., Rollin, P.E., Mills, J.N., Villagra, M.R., Montenegro, M.J., Costales, M.A., Paredes, L.C., Peters, C.J., 1997. Treatment of Bolivian hemorrhagic fever with intravenous ribavirin. *Clin Infect Dis* 24, 718-722.

Knobel, D.L., Cleaveland, S., Coleman, P.G., Fèvre, E.M., Meltzer, M.I., Miranda, M.E., Shaw, A., Zinsstag, J., Meslin, F.X., 2005. Re-evaluating the burden of rabies in Africa and Asia. *Bull World Health Organ* 83, 360-368.

Koma, T., Patterson, M., Huang, C., Seregin, A.V., Maharaj, P.D., Miller, M., Smith, J.N., Walker, A.G., Hallam, S., Paessler, S., 2016. Machupo Virus Expressing GPC of the Candid#1 Vaccine Strain of Junin Virus Is Highly Attenuated and Immunogenic. *J Virol* 90, 1290-1297.

Konagurthu, A.S., Whisstock, J.C., Stuckey, P.J., Lesk, A.M., 2006. MUSTANG: a multiple structural alignment algorithm. *Proteins* 64, 559-574.

Kranzusch, P.J., Schenk, A.D., Rahmeh, A.A., Radoshitzky, S.R., Bavari, S., Walz, T., Whelan, S.P., 2010. Assembly of a functional Machupo virus polymerase complex. *Proc Natl Acad Sci U S A* 107, 20069-20074.

Krissinel, E., Henrick, K., 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* 60, 2256-2268.

Krissinel, E., Henrick, K., 2007. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372, 774-797.

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* 35, 1547-1549.

Kunz, S., 2009. Receptor binding and cell entry of Old World arenaviruses reveal novel aspects of virus–host interaction. *Virology* 387, 245-249.

Kunz, S., Rojek, J.M., Kanagawa, M., Spiropoulou, C.F., Barresi, R., Campbell, K.P., Oldstone, M.B., 2005. Posttranslational modification of alpha-dystroglycan, the cellular receptor for arenaviruses, by the glycosyltransferase LARGE is critical for virus binding. *J Virol* 79, 14282-14296.

Lafon, M., 2005. Rabies virus receptors. *J Neurovirol* 11, 82-87.

- Lafon, M., Wiktor, T.J., Macfarlan, R.I., 1983. Antigenic sites on the CVS rabies virus glycoprotein: analysis with monoclonal antibodies. *J Gen Virol* 64 (Pt 4), 843-851.
- Larson, R.A., Dai, D., Hosack, V.T., Tan, Y., Bolken, T.C., Hraby, D.E., Amberg, S.M., 2008. Identification of a Broad-Spectrum Arenavirus Entry Inhibitor. *Journal of Virology* 82, 10768.
- Laskowski, R.A., Swindells, M.B., 2011. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* 51, 2778-2786.
- Lavanya, M., Cuevas, C.D., Thomas, M., Cherry, S., Ross, S.R., 2013. siRNA screen for genes that affect Junín virus entry uncovers voltage-gated calcium channels as a therapeutic target. *Sci Transl Med* 5, 204ra131.
- Lavergne, A., de Thoisy, B., Tirera, S., Donato, D., Bouchier, C., Catzefflis, F., Lacoste, V., 2016. Identification of lymphocytic choriomeningitis mammarenavirus in house mouse (*Mus musculus*, Rodentia) in French Guiana. *Infection, Genetics and Evolution* 37, 225-230.
- Lawrence, C.M., Ray, S., Babyonyshev, M., Galluser, R., Borhani, D.W., Harrison, S.C., 1999. Crystal structure of the ectodomain of human transferrin receptor. *Science* 286, 779-782.
- Le, S.Q., Gascuel, O., 2008. An Improved General Amino Acid Replacement Matrix. *Molecular Biology and Evolution* 25, 1307-1320.
- Lee, A.M., Rojek, J.M., Spiropoulou, C.F., Gundersen, A.T., Jin, W., Shaginian, A., York, J., Nunberg, J.H., Boger, D.L., Oldstone, M.B., Kunz, S., 2008. Unique small molecule entry inhibitors of hemorrhagic fever arenaviruses. *J Biol Chem* 283, 18734-18742.
- Lentz, T.L., Burrage, T.G., Smith, A.L., Crick, J., Tignor, G.H., 1982. Is the acetylcholine receptor a rabies virus receptor? *Science* 215, 182-184.
- Lentz, T.L., Burrage, T.G., Smith, A.L., Tignor, G.H., 1983. The acetylcholine receptor as a cellular receptor for rabies virus. *Yale J Biol Med* 56, 315-322.
- Lenz, O., ter Meulen, J., Klenk, H.D., Seidah, N.G., Garten, W., 2001. The Lassa virus glycoprotein precursor GP-C is proteolytically processed by subtilase SKI-1/S1P. *Proc Natl Acad Sci U S A* 98, 12701-12705.
- Levingston Macleod, J.M., D'Antuono, A., Loureiro, M.E., Casabona, J.C., Gomez, G.A., Lopez, N., 2011. Identification of two functional domains within the arenavirus nucleoprotein. *J Virol* 85, 2012-2023.

Li, K., Lin, X.D., Wang, W., Shi, M., Guo, W.P., Zhang, X.H., Xing, J.G., He, J.R., Wang, K., Li, M.H., Cao, J.H., Jiang, M.L., Holmes, E.C., Zhang, Y.Z., 2015. Isolation and characterization of a novel arenavirus harbored by Rodents and Shrews in Zhejiang province, China. *Virology* 476, 37-42.

Li, S., Sun, Z., Pryce, R., Parsy, M.L., Fehling, S.K., Schlie, K., Siebert, C.A., Garten, W., Bowden, T.A., Strecker, T., Huiskonen, J.T., 2016. Acidic pH-Induced Conformations and LAMP1 Binding of the Lassa Virus Glycoprotein Spike. *PLoS Pathog* 12, e1005418.

Liu, P., Yang, J., Wu, X., Fu, Z.F., 2004. Interactions amongst rabies virus nucleoprotein, phosphoprotein and genomic RNA in virus-infected and transfected cells. *J Gen Virol* 85, 3725-3734.

Liu, Y., Heim, K.P., Che, Y., Chi, X., Qiu, X., Han, S., Dormitzer, P.R., Yang, X., 2021. Prefusion structure of human cytomegalovirus glycoprotein B and structural basis for membrane fusion. *Science Advances* 7, eabf3178.

Loureiro, M.E., Wilda, M., Levingston Macleod, J.M., D'Antuono, A., Foscaldi, S., Marino Buslje, C., Lopez, N., 2011. Molecular determinants of arenavirus Z protein homo-oligomerization and L polymerase binding. *J Virol* 85, 12304-12314.

Lu, R.-M., Hwang, Y.-C., Liu, I.J., Lee, C.-C., Tsai, H.-Z., Li, H.-J., Wu, H.-C., 2020. Development of therapeutic antibodies for the treatment of diseases. *Journal of Biomedical Science* 27, 1.

Lukashevich, I.S., Carrion, R., Jr., Salvato, M.S., Mansfield, K., Brasky, K., Zapata, J., Cairo, C., Goicochea, M., Hoosien, G.E., Ticer, A., Bryant, J., Davis, H., Hammamieh, R., Mayda, M., Jett, M., Patterson, J., 2008. Safety, immunogenicity, and efficacy of the ML29 reassortant vaccine for Lassa fever in small non-human primates. *Vaccine* 26, 5246-5254.

Lukashevich, I.S., Pushko, P., 2016. Vaccine platforms to control Lassa fever. *Expert Rev Vaccines* 15, 1135-1150.

Maes, P., Alkhovsky, S.V., Bào, Y., Beer, M., Birkhead, M., Brieze, T., Buchmeier, M.J., Calisher, C.H., Charrel, R.N., Choi, I.R., Clegg, C.S., de la Torre, J.C., Delwart, E., DeRisi, J.L., Di Bello, P.L., Di Serio, F., Digiaro, M., Dolja, V.V., Drosten, C., Druciarek, T.Z., Du, J., Ebihara, H., Elbeaino, T., Gergerich, R.C., Gillis, A.N., Gonzalez, J.J., Haenni, A.L., Hepojoki, J., Hetzel, U., Hồ, T., Hóng, N., Jain, R.K., Jansen van Vuren, P., Jin, Q., Jonson, M.G., Junglen, S., Keller, K.E., Kemp, A., Kipar, A., Kondov, N.O., Koonin, E.V., Kormelink, R., Korzyukov, Y., Krupovic, M., Lambert, A.J., Laney, A.G., LeBreton, M., Lukashevich, I.S., Marklewitz, M., Markotter, W., Martelli, G.P., Martin, R.R., Mielke-Ehret, N., Mühlbach, H.P., Navarro, B., Ng, T.F.F., Nunes, M.R.T., Palacios, G., Pawęska, J.T., Peters, C.J., Plyusnin, A., Radoshitzky, S.R., Romanowski, V., Salmenperä, P., Salvato, M.S., Sanfaçon, H., Sasaya, T., Schmaljohn, C., Schneider, B.S., Shirako, Y., Siddell, S.,

Sironen, T.A., Stenglein, M.D., Storm, N., Sudini, H., Tesh, R.B., Tzanetakis, I.E., Uppala, M., Vapalahti, O., Vasilakis, N., Walker, P.J., Wáng, G., Wáng, L., Wáng, Y., Wèi, T., Wiley, M.R., Wolf, Y.I., Wolfe, N.D., Wú, Z., Xú, W., Yang, L., Yāng, Z., Yeh, S.D., Zhāng, Y.Z., Zhèng, Y., Zhou, X., Zhū, C., Zirkel, F., Kuhn, J.H., 2018. Taxonomy of the family Arenaviridae and the order Bunyavirales: update 2018. *Arch Virol* 163, 2295-2310.

Mahmutovic, S., Clark, L., Levis, S.C., Briggiler, A.M., Enria, D.A., Harrison, S.C., Abraham, J., 2015. Molecular Basis for Antibody-Mediated Neutralization of New World Hemorrhagic Fever Mammarenaviruses. *Cell Host Microbe* 18, 705-713.

Maiztegui, J.I., Fernandez, N.J., de Damilano, A.J., 1979. Efficacy of immune plasma in treatment of Argentine haemorrhagic fever and association between treatment and a late neurological syndrome. *Lancet* 2, 1216-1217.

Maiztegui, J.I., McKee, K.T., Jr., Barrera Oro, J.G., Harrison, L.H., Gibbs, P.H., Feuillade, M.R., Enria, D.A., Briggiler, A.M., Levis, S.C., Ambrosio, A.M., Halsey, N.A., Peters, C.J., 1998. Protective efficacy of a live attenuated vaccine against Argentine hemorrhagic fever. AHF Study Group. *J Infect Dis* 177, 277-283.

Marq, J.B., Kolakofsky, D., Garcin, D., 2010. Unpaired 5' ppp-nucleotides, as found in arenavirus double-stranded RNA panhandles, are not recognized by RIG-I. *J Biol Chem* 285, 18208-18216.

Marston, D.A., Banyard, A.C., McElhinney, L.M., Freuling, C.M., Finke, S., de Lamballerie, X., Müller, T., Fooks, A.R., 2018. The lyssavirus host-specificity conundrum-rabies virus-the exception not the rule. *Curr Opin Virol* 28, 68-73.

Martin, A.C., Thornton, J.M., 1996. Structural families in loops of homologous proteins: automatic classification, modelling and application to antibodies. *J Mol Biol* 263, 800-815.

Martin, P., Jensen, D.M., 2008. Ribavirin in the treatment of chronic hepatitis C. *J Gastroenterol Hepatol* 23, 844-855.

Martin, V.K., Droniou-Bonzom, M.E., Reignier, T., Oldenburg, J.E., Cox, A.U., Cannon, P.M., 2010. Investigation of clade B New World arenavirus tropism by using chimeric GP1 proteins. *J Virol* 84, 1176-1182.

Martinez, M.G., Bialecki, M.A., Belouzard, S., Cordo, S.M., Candurra, N.A., Whittaker, G.R., 2013. Utilization of human DC-SIGN and L-SIGN for entry and infection of host cells by the New World arenavirus, Junin virus. *Biochem Biophys Res Commun* 441, 612-617.

Martinez, M.G., Cordo, S.M., Candurra, N.A., 2007. Characterization of Junin arenavirus cell entry. *J Gen Virol* 88, 1776-1784.

Martinez, M.G., Forlenza, M.B., Candurra, N.A., 2009. Involvement of cellular proteins in Junin arenavirus entry. *Biotechnol J* 4, 866-870.

Martínez-Sobrido, L., de la Torre, J.C., 2016. Reporter-Expressing, Replicating-Competent Recombinant Arenaviruses. *Viruses* 8.

Mateo, M., Reynard, S., Carnec, X., Journeaux, A., Baillet, N., Schaeffer, J., Picard, C., Legras-Lachuer, C., Allan, R., Perthame, E., Hillion, K.H., Pietrosemoli, N., Dillies, M.A., Barrot, L., Vallve, A., Barron, S., Fellmann, L., Gaillard, J.C., Armengaud, J., Carbonnelle, C., Raoul, H., Tangy, F., Baize, S., 2019. Vaccines inducing immunity to Lassa virus glycoprotein and nucleoprotein protect macaques after a single shot. *Sci Transl Med* 11.

Mavrakis, M., Iseni, F., Mazza, C., Schoehn, G., Ebel, C., Gentzel, M., Franz, T., Ruigrok, R.W., 2003. Isolation and characterisation of the rabies virus N degrees-P complex produced in insect cells. *Virology* 305, 406-414.

May, E.R., Armen, R.S., Mannan, A.M., Brooks, C.L., 3rd, 2010. The flexible C-terminal arm of the Lassa arenavirus Z-protein mediates interactions with multiple binding partners. *Proteins* 78, 2251-2264.

McCormick, J.B., Fisher-Hoch, S.P., 2002. Lassa fever. *Curr Top Microbiol Immunol* 262, 75-109.

McCormick, J.B., King, I.J., Webb, P.A., Scribner, C.L., Craven, R.B., Johnson, K.M., Elliott, L.H., Belmont-Williams, R., 1986. Lassa fever. Effective therapy with ribavirin. *N Engl J Med* 314, 20-26.

McCormick, J.B., Webb, P.A., Krebs, J.W., Johnson, K.M., Smith, E.S., 1987. A prospective study of the epidemiology and ecology of Lassa fever. *J Infect Dis* 155, 437-444.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J., 2007. Phaser crystallographic software. *J Appl Crystallogr* 40, 658-674.

Mebatsion, T., Weiland, F., Conzelmann, K.K., 1999. Matrix protein of rabies virus is responsible for the assembly and budding of bullet-shaped particles and interacts with the transmembrane spike glycoprotein G. *J Virol* 73, 242-250.

Mehand, M.S., Al-Shorbaji, F., Millett, P., Murgue, B., 2018. The WHO R&D Blueprint: 2018 review of emerging infectious diseases requiring urgent research and development efforts. *Antiviral Res* 159, 63-67.

Mendenhall, M., Russell, A., Juelich, T., Messina, E.L., Smee, D.F., Freiberg, A.N., Holbrook, M.R., Furuta, Y., de la Torre, J.C., Nunberg, J.H., Gowen, B.B., 2011a. T-705 (favipiravir) inhibition of arenavirus replication in cell culture. *Antimicrob Agents Chemother* 55, 782-787.

Mendenhall, M., Russell, A., Smee, D.F., Hall, J.O., Skirpstunas, R., Furuta, Y., Gowen, B.B., 2011b. Effective oral favipiravir (T-705) therapy initiated after the onset of clinical disease in a model of arenavirus hemorrhagic Fever. *PLoS Negl Trop Dis* 5, e1342.

Messina, E.L., York, J., Nunberg, J.H., 2012. Dissection of the role of the stable signal peptide of the arenavirus envelope glycoprotein in membrane fusion. *J Virol* 86, 6138-6145.

Meulen, J., Badusche, M., Satoguina, J., Strecker, T., Lenz, O., Loeliger, C., Sakho, M., Koulemou, K., Koivogui, L., Hoerauf, A., 2004. Old and New World arenaviruses share a highly conserved epitope in the fusion domain of the glycoprotein 2, which is recognized by Lassa virus-specific human CD4+ T-cell clones. *Virology* 321, 134-143.

Meyer, B.J., de la Torre, J.C., Southern, P.J., 2002. Arenaviruses: genomic RNAs, transcription, and replication. *Curr Top Microbiol Immunol* 262, 139-157.

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

Mills, J.N., Ellis, B.A., Childs, J.E., McKee, K.T., Jr., Maiztegui, J.I., Peters, C.J., Ksiazek, T.G., Jahrling, P.B., 1994. Prevalence of infection with Junin virus in rodent populations in the epidemic area of Argentine hemorrhagic fever. *Am J Trop Med Hyg* 51, 554-562.

Möller, S., Croning, M.D., Apweiler, R., 2001. Evaluation of methods for the prediction of membrane spanning regions. *Bioinformatics* 17, 646-653.

Monath, T.P., Maher, M., Casals, J., Kissling, R.E., Cacciapuoti, A., 1974. Lassa fever in the Eastern Province of Sierra Leone, 1970-1972. II. Clinical observations and virological studies on selected hospital cases. *Am J Trop Med Hyg* 23, 1140-1149.

Moreno, H., Grande-Pérez, A., Domingo, E., Martín, V., 2012. Arenaviruses and lethal mutagenesis. Prospects for new ribavirin-based interventions. *Viruses* 4, 2786-2805.

Morens, D.M., Folkers, G.K., Fauci, A.S., 2004. The challenge of emerging and re-emerging infectious diseases. *Nature* 430, 242-249.

Morin, B., Coutard, B., Lelke, M., Ferron, F., Kerber, R., Jamal, S., Frangeul, A., Baronti, C., Charrel, R., de Lamballerie, X., Vonnheim, C., Lescar, J., Bricogne, G., Günther, S., Canard, B., 2010. The N-terminal domain of the arenavirus L protein is an RNA endonuclease essential in mRNA transcription. *PLoS Pathog* 6, e1001038.

Moulenat, T., Petit, C., Bosch Castells, V., Houillon, G., 2020. Purified Vero Cell Rabies Vaccine (PVRV, Verorab®): A Systematic Review of Intradermal Use Between 1985 and 2019. *Trop Med Infect Dis* 5.

Mrzljak, A., Novak, R., Pandak, N., Tabain, I., Franusic, L., Barbic, L., Bogdanic, M., Savic, V., Mikulic, D., Pavicic-Saric, J., Stevanovic, V., Vilibic-Cavlek, T., 2020. Emerging and neglected zoonoses in transplant population. *World J Transplant* 10, 47-63.

Müller, T., Dietzschold, B., Ertl, H., Fooks, A.R., Freuling, C., Fehlner-Gardiner, C., Kliemt, J., Meslin, F.X., Franka, R., Rupprecht, C.E., Tordo, N., Wanderler, A.I., Kieny, M.P., 2009. Development of a mouse monoclonal antibody cocktail for post-exposure rabies prophylaxis in humans. *PLoS Negl Trop Dis* 3, e542.

Nagarajan, T., Marissen, W., Rupprecht, C., 2014. Monoclonal antibodies for the prevention of rabies: theory and clinical practice. *Antibody Technology Journal* 4, 1-12.

Nagarajan, T., Rupprecht, C.E., Dessain, S.K., Rangarajan, P.N., Thiagarajan, D., Srinivasan, V.A., 2008. Human monoclonal antibody and vaccine approaches to prevent human rabies. *Curr Top Microbiol Immunol* 317, 67-101.

Nakamura, G., Chai, N., Park, S., Chiang, N., Lin, Z., Chiu, H., Fong, R., Yan, D., Kim, J., Zhang, J., Lee, W.P., Estevez, A., Coons, M., Xu, M., Lupardus, P., Balazs, M., Swem, L.R., 2013. An in vivo human-plasmablast enrichment technique allows rapid identification of therapeutic influenza A antibodies. *Cell Host Microbe* 14, 93-103.

Naydenova, K., Russo, C.J., 2017. Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy. *Nat Commun* 8, 629.

Ng, W.M., Stelfox, A.J., Bowden, T.A., 2020. Unraveling virus relationships by structure-based phylogenetic classification. *Virus Evol* 6, veaa003.

Nguyen, M.N., Tan, K.P., Madhusudhan, M.S., 2011. CLICK--topology-independent comparison of biomolecular 3D structures. *Nucleic Acids Res* 39, W24-28.

Ning, L., Abagna, H.B., Jiang, Q., Liu, S., Huang, J., 2021. Development and application of therapeutic antibodies against COVID-19. *Int J Biol Sci* 17, 1486-1496.

Nunberg, J.H., York, J., 2012. The curious case of arenavirus entry, and its inhibition. *Viruses* 4, 83-101.

Oksanen, H.M., Pietilä, M.K., Sencilo, A., Atanasova, N.S., Roine, E., Bamford, D.H., 2012. Virus Universe: Can It Be Constructed from a Limited Number of Viral Architectures, in: Witzany, G. (Ed.), *Viruses: Essential Agents of Life*. Springer Netherlands, Dordrecht, pp. 83-105.

Oldenburg, J., Reignier, T., Flanagan, M.L., Hamilton, G.A., Cannon, P.M., 2007. Differences in tropism and pH dependence for glycoproteins from the Clade B1 arenaviruses: implications for receptor usage and pathogenicity. *Virology* 364, 132-139.

Olschewski, S., Cusack, S., Rosenthal, M., 2020. The Cap-Snatching Mechanism of Bunyaviruses. *Trends Microbiol* 28, 293-303.

Olschläger, S., Flatz, L., 2013. Vaccination strategies against highly pathogenic arenaviruses: the next steps toward clinical trials. *PLoS Pathog* 9, e1003212.

Oppliger, J., Torriani, G., Herrador, A., Kunz, S., 2016. Lassa Virus Cell Entry via Dystroglycan Involves an Unusual Pathway of Macropinocytosis. *J Virol* 90, 6412-6429.

Ortiz, A.R., Strauss, C.E., Olmea, O., 2002. MAMMOTH (matching molecular models obtained from theory): an automated method for model comparison. *Protein Sci* 11, 2606-2621.

Ortiz-Riaño, E., Cheng, B.Y., de la Torre, J.C., Martínez-Sobrido, L., 2011. 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. *J Virol* 85, 13038-13048.

Paessler, S., Walker, D.H., 2013. Pathogenesis of the viral hemorrhagic fevers. *Annu Rev Pathol* 8, 411-440.

Pancera, M., Shahzad-Ul-Hussan, S., Doria-Rose, N.A., McLellan, J.S., Bailer, R.T., Dai, K., Loesgen, S., Louder, M.K., Staupe, R.P., Yang, Y., Zhang, B., Parks, R., Eudailey, J., Lloyd, K.E., Blinn, J., Alam, S.M., Haynes, B.F., Amin, M.N., Wang, L.X., Burton, D.R., Koff, W.C., Nabel, G.J., Mascola, J.R., Bewley, C.A., Kwong, P.D., 2013. Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16. *Nat Struct Mol Biol* 20, 804-813.

Papageorgiou, N., Spiliopoulou, M., Nguyen, T.V., Vaitopoulou, A., Laban, E.Y., Alvarez, K., Margiolaki, I., Canard, B., Ferron, F., 2020. Brothers in Arms: Structure, Assembly and Function of Arenaviridae Nucleoprotein. *Viruses* 12.

- Parsy, M.L., Harlos, K., Huiskonen, J.T., Bowden, T.A., 2013. Crystal structure of Venezuelan hemorrhagic fever virus fusion glycoprotein reveals a class 1 postfusion architecture with extensive glycosylation. *J Virol* 87, 13070-13075.
- Pasqual, G., Rojek, J.M., Masin, M., Chatton, J.Y., Kunz, S., 2011. Old world arenaviruses enter the host cell via the multivesicular body and depend on the endosomal sorting complex required for transport. *PLoS Pathog* 7, e1002232.
- Pasquato, A., Cendron, L., Kunz, S., 2018. Cleavage of the Glycoprotein of Arenaviruses, in: Böttcher-Friebertshäuser, E., Garten, W., Klenk, H.D. (Eds.), *Activation of Viruses by Host Proteases*. Springer International Publishing, Cham, pp. 47-70.
- Passmore, L.A., Russo, C.J., 2016. Specimen Preparation for High-Resolution Cryo-EM. *Methods Enzymol* 579, 51-86.
- Patterson, M., Grant, A., Paessler, S., 2014. Epidemiology and pathogenesis of Bolivian hemorrhagic fever. *Curr Opin Virol* 5, 82-90.
- Peng, R., Xu, X., Jing, J., Wang, M., Peng, Q., Liu, S., Wu, Y., Bao, X., Wang, P., Qi, J., Gao, G.F., Shi, Y., 2020. Structural insight into arenavirus replication machinery. *Nature* 579, 615-619.
- Perez, M., Craven, R.C., de la Torre, J.C., 2003. The small RING finger protein Z drives arenavirus budding: implications for antiviral strategies. *Proc Natl Acad Sci U S A* 100, 12978-12983.
- Perez, M., de la Torre, J.C., 2003. Characterization of the genomic promoter of the prototypic arenavirus lymphocytic choriomeningitis virus. *J Virol* 77, 1184-1194.
- Perez, M., Greenwald, D.L., de la Torre, J.C., 2004. Myristoylation of the RING finger Z protein is essential for arenavirus budding. *J Virol* 78, 11443-11448.
- Peters, C.J., Jahrling, P.B., Liu, C.T., Kenyon, R.H., McKee, K.T., Jr., Barrera Oro, J.G., 1987. Experimental studies of arenaviral hemorrhagic fevers. *Curr Top Microbiol Immunol* 134, 5-68.
- Pfau, C.J., 1996. Arenaviruses. *Medical Microbiology*.
- Pinschewer, D.D., Perez, M., de la Torre, J.C., 2005. Dual role of the lymphocytic choriomeningitis virus intergenic region in transcription termination and virus propagation. *J Virol* 79, 4519-4526.

Poirot, O., Suhre, K., Abergel, C., O'Toole, E., Notredame, C., 2004. 3DCoffee@igs: a web server for combining sequences and structures into a multiple sequence alignment. *Nucleic Acids Res* 32, W37-40.

Polyak, S.J., Zheng, S., Harnish, D.G., 1995. 5' termini of Pichinde arenavirus S RNAs and mRNAs contain nontemplated nucleotides. *J Virol* 69, 3211-3215.

Pryce, R., Ng, W.M., Zeltina, A., Watanabe, Y., El Omari, K., Wagner, A., Bowden, T.A., 2018. Structure-based classification defines the discrete conformational classes adopted by the arenaviral GP1. *J Virol*.

Punjani, A., Rubinstein, J.L., Fleet, D.J., Brubaker, M.A., 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* 14, 290-296.

Raaben, M., Jae, L.T., Herbert, A.S., Kuehne, A.I., Stubbs, S.H., Chou, Y.Y., Blomen, V.A., Kirchhausen, T., Dye, J.M., Brummelkamp, T.R., Whelan, S.P., 2017. NRP2 and CD63 Are Host Factors for Lujo Virus Cell Entry. *Cell Host Microbe* 22, 688-696.e685.

Radoshitzky, S.R., Abraham, J., Spiropoulou, C.F., Kuhn, J.H., Nguyen, D., Li, W., Nagel, J., Schmidt, P.J., Nunberg, J.H., Andrews, N.C., Farzan, M., Choe, H., 2007. Transferrin receptor 1 is a cellular receptor for New World haemorrhagic fever arenaviruses. *Nature* 446, 92-96.

Radoshitzky, S.R., Bào, Y., Buchmeier, M.J., Charrel, R.N., Clawson, A.N., Clegg, C.S., DeRisi, J.L., Emonet, S., Gonzalez, J.P., Kuhn, J.H., Lukashevich, I.S., Peters, C.J., Romanowski, V., Salvato, M.S., Stenglein, M.D., de la Torre, J.C., 2015. Past, present, and future of arenavirus taxonomy. *Arch Virol* 160, 1851-1874.

Radoshitzky, S.R., Kuhn, J.H., Spiropoulou, C.F., Albariño, C.G., Nguyen, D.P., Salazar-Bravo, J., Dorfman, T., Lee, A.S., Wang, E., Ross, S.R., Choe, H., Farzan, M., 2008. Receptor determinants of zoonotic transmission of New World hemorrhagic fever arenaviruses. *Proc Natl Acad Sci U S A* 105, 2664-2669.

Raju, R., Raju, L., Hacker, D., Garcin, D., Compans, R., Kolakofsky, D., 1990. Nontemplated bases at the 5' ends of Tacaribe virus mRNAs. *Virology* 174, 53-59.

Rao, S.T., Rossmann, M.G., 1973. Comparison of super-secondary structures in proteins. *J Mol Biol* 76, 241-256.

Ravanti, J., Bamford, D., Stuart, D.I., 2013. Automatic comparison and classification of protein structures. *J Struct Biol* 183, 47-56.

Ren, J., Wen, L., Gao, X., Jin, C., Xue, Y., Yao, X., 2009. DOG 1.0: illustrator of protein domain structures. *Cell Res* 19, 271-273.

Richmond, J.K., Baglolle, D.J., 2003. Lassa fever: epidemiology, clinical features, and social consequences. *Bmj* 327, 1271-1275.

Rijal, P., Elias, S.C., Machado, S.R., Xiao, J., Schimanski, L., O'Dowd, V., Baker, T., Barry, E., Mendelsohn, S.C., Cherry, C.J., Jin, J., Labbé, G.M., Donnellan, F.R., Rampling, T., Dowall, S., Rayner, E., Findlay-Wilson, S., Carroll, M., Guo, J., Xu, X.N., Huang, K.A., Takada, A., Burgess, G., McMillan, D., Popplewell, A., Lightwood, D.J., Draper, S.J., Townsend, A.R., 2019. Therapeutic Monoclonal Antibodies for Ebola Virus Infection Derived from Vaccinated Humans. *Cell Rep* 27, 172-186.e177.

Ritchie, G., Harvey, D.J., Stroehrer, U., Feldmann, F., Feldmann, H., Wahl-Jensen, V., Royle, L., Dwek, R.A., Rudd, P.M., 2010. Identification of N-glycans from Ebola virus glycoproteins by matrix-assisted laser desorption/ionisation time-of-flight and negative ion electrospray tandem mass spectrometry. *Rapid Commun Mass Spectrom* 24, 571-585.

Robert, X., Gouet, P., 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42, W320-324.

Robinson, J.E., Hastie, K.M., Cross, R.W., Yenni, R.E., Elliott, D.H., Rouelle, J.A., Kannadka, C.B., Smira, A.A., Garry, C.E., Bradley, B.T., Yu, H., Shaffer, J.G., Boisen, M.L., Hartnett, J.N., Zandonatti, M.A., Rowland, M.M., Heinrich, M.L., Martinez-Sobrido, L., Cheng, B., de la Torre, J.C., Andersen, K.G., Goba, A., Momoh, M., Fullah, M., Gbakie, M., Kanneh, L., Koroma, V.J., Fonnies, R., Jalloh, S.C., Kargbo, B., Vandi, M.A., Gbetuwa, M., Ikponmwosa, O., Asogun, D.A., Okokhere, P.O., Follarin, O.A., Schieffelin, J.S., Pitts, K.R., Geisbert, J.B., Kulakowski, P.C., Wilson, R.B., Happi, C.T., Sabeti, P.C., Geva, S.M., Khan, S.H., Grant, D.S., Geisbert, T.W., Saphire, E.O., Branco, L.M., Garry, R.F., 2016. Most neutralizing human monoclonal antibodies target novel epitopes requiring both Lassa virus glycoprotein subunits. *Nat Commun* 7, 11544.

Roche, S., Bressanelli, S., Rey, F.A., Gaudin, Y., 2006. Crystal structure of the low-pH form of the vesicular stomatitis virus glycoprotein G. *Science* 313, 187-191.

Roche, S., Rey, F.A., Gaudin, Y., Bressanelli, S., 2007. Structure of the prefusion form of the vesicular stomatitis virus glycoprotein G. *Science* 315, 843-848.

Rojek, J.M., Lee, A.M., Nguyen, N., Spiropoulou, C.F., Kunz, S., 2008a. Site 1 protease is required for proteolytic processing of the glycoproteins of the South American hemorrhagic fever viruses Junin, Machupo, and Guanarito. *J Virol* 82, 6045-6051.

- Rojek, J.M., Sanchez, A.B., Nguyen, N.T., de la Torre, J.C., Kunz, S., 2008b. Different mechanisms of cell entry by human-pathogenic Old World and New World arenaviruses. *J Virol* 82, 7677-7687.
- Roldán, J.S., Martínez, M.G., Forlenza, M.B., Whittaker, G.R., Candurra, N.A., 2016. Human transferrin receptor triggers an alternative Tacaribe virus internalization pathway. *Arch Virol* 161, 353-363.
- Rossmann, M.G., Argos, P., 1976. Exploring structural homology of proteins. *J Mol Biol* 105, 75-95.
- Rouvinski, A., Guardado-Calvo, P., Barba-Spaeth, G., Duquerroy, S., Vaney, M.C., Kikuti, C.M., Navarro Sanchez, M.E., Dejnirattisai, W., Wongwiwat, W., Haouz, A., Girard-Blanc, C., Petres, S., Shepard, W.E., Desprès, P., Arenzana-Seisdedos, F., Dussart, P., Mongkolsapaya, J., Screaton, G.R., Rey, F.A., 2015. Recognition determinants of broadly neutralizing human antibodies against dengue viruses. *Nature* 520, 109-113.
- Ruo, S.L., Mitchell, S.W., Kiley, M.P., Roumillat, L.F., Fisher-Hoch, S.P., McCormick, J.B., 1991. Antigenic relatedness between arenaviruses defined at the epitope level by monoclonal antibodies. *J Gen Virol* 72 (Pt 3), 549-555.
- Safronetz, D., Sogoba, N., Lopez, J.E., Maiga, O., Dahlstrom, E., Zivcec, M., Feldmann, F., Haddock, E., Fischer, R.J., Anderson, J.M., Munster, V.J., Branco, L., Garry, R., Porcella, S.F., Schwan, T.G., Feldmann, H., 2013. Geographic distribution and genetic characterization of Lassa virus in sub-Saharan Mali. *PLoS Negl Trop Dis* 7, e2582.
- Salazar, M., Yun, N.E., Poussard, A.L., Smith, J.N., Smith, J.K., Kolokoltsova, O.A., Patterson, M.J., Linde, J., Paessler, S., 2012. Effect of ribavirin on junin virus infection in guinea pigs. *Zoonoses Public Health* 59, 278-285.
- Sali, A., Blundell, T.L., 1990. Definition of general topological equivalence in protein structures. A procedure involving comparison of properties and relationships through simulated annealing and dynamic programming. *J Mol Biol* 212, 403-428.
- Sánchez, A.B., de la Torre, J.C., 2005. Genetic and biochemical evidence for an oligomeric structure of the functional L polymerase of the prototypic arenavirus lymphocytic choriomeningitis virus. *J Virol* 79, 7262-7268.
- Sarute, N., Ross, S.R., 2017. New World Arenavirus Biology. *Annu Rev Virol* 4, 141-158.
- Saunders, A.A., Ting, J.P., Meisner, J., Neuman, B.W., Perez, M., de la Torre, J.C., Buchmeier, M.J., 2007. Mapping the landscape of the lymphocytic choriomeningitis virus stable signal peptide reveals novel functional domains. *J Virol* 81, 5649-5657.

Scanlan, C.N., Pantophlet, R., Wormald, M.R., Ollmann Saphire, E., Stanfield, R., Wilson, I.A., Katinger, H., Dwek, R.A., Rudd, P.M., Burton, D.R., 2002. The broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2G12 recognizes a cluster of alpha1-->2 mannose residues on the outer face of gp120. *J Virol* 76, 7306-7321.

Scheres, S.H., Chen, S., 2012. Prevention of overfitting in cryo-EM structure determination. *Nat Methods* 9, 853-854.

Schlie, K., Maisa, A., Lennartz, F., Ströher, U., Garten, W., Strecker, T., 2010. Characterization of Lassa virus glycoprotein oligomerization and influence of cholesterol on virus replication. *J Virol* 84, 983-992.

Senior, A.W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., Židek, A., Nelson, A.W.R., Bridgland, A., Penadones, H., Petersen, S., Simonyan, K., Crossan, S., Kohli, P., Jones, D.T., Silver, D., Kavukcuoglu, K., Hassabis, D., 2020. Improved protein structure prediction using potentials from deep learning. *Nature* 577, 706-710.

Shaughnessy, A.F., 2012. Monoclonal antibodies: magic bullets with a hefty price tag. *Bmj* 345, e8346.

Shi, M., Lin, X.D., Chen, X., Tian, J.H., Chen, L.J., Li, K., Wang, W., Eden, J.S., Shen, J.J., Liu, L., Holmes, E.C., Zhang, Y.Z., 2018. The evolutionary history of vertebrate RNA viruses. *Nature* 556, 197-202.

Shimojima, M., Kawaoka, Y., 2012. Cell surface molecules involved in infection mediated by lymphocytic choriomeningitis virus glycoprotein. *J Vet Med Sci* 74, 1363-1366.

Shimojima, M., Stroher, U., Ebihara, H., Feldmann, H., Kawaoka, Y., 2012. Identification of cell surface molecules involved in dystroglycan-independent Lassa virus cell entry. *J Virol* 86, 2067-2078.

Shimon, A., Shani, O., Diskin, R., 2017. Structural Basis for Receptor Selectivity by the Whitewater Arroyo Mammarenavirus. *J Mol Biol* 429, 2825-2839.

Shindyalov, I.N., Bourne, P.E., 1998. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng* 11, 739-747.

Shtanko, O., Imai, M., Goto, H., Lukashevich, I.S., Neumann, G., Watanabe, T., Kawaoka, Y., 2010. A role for the C terminus of Mopeia virus nucleoprotein in its incorporation into Z protein-induced virus-like particles. *J Virol* 84, 5415-5422.

Sieczkarski, S.B., Whittaker, G.R., 2003. Differential requirements of Rab5 and Rab7 for endocytosis of influenza and other enveloped viruses. *Traffic* 4, 333-343.

Sievers, F., Higgins, D.G., 2014. Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol Biol* 1079, 105-116.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7, 539.

Simmonds, P., Adams, M.J., Benko, M., Breitbart, M., Brister, J.R., Carstens, E.B., Davison, A.J., Delwart, E., Gorbalenya, A.E., Harrach, B., Hull, R., King, A.M., Koonin, E.V., Krupovic, M., Kuhn, J.H., Lefkowitz, E.J., Nibert, M.L., Orton, R., Roossinck, M.J., Sabanadzovic, S., Sullivan, M.B., Suttle, C.A., Tesh, R.B., van der Vlugt, R.A., Varsani, A., Zerbini, F.M., 2017. Consensus statement: Virus taxonomy in the age of metagenomics. *Nat Rev Microbiol* 15, 161-168.

Simulundu, E., Mweene, A.S., Changula, K., Monze, M., Chizema, E., Mwaba, P., Takada, A., Ippolito, G., Kasolo, F., Zumla, A., Bates, M., 2016. Lujo viral hemorrhagic fever: considering diagnostic capacity and preparedness in the wake of recent Ebola and Zika virus outbreaks. *Rev Med Virol* 26, 446-454.

Singh, M.K., Fuller-Pace, F.V., Buchmeier, M.J., Southern, P.J., 1987. Analysis of the genomic L RNA segment from lymphocytic choriomeningitis virus. *Virology* 161, 448-456.

Sloan, S.E., Hanlon, C., Weldon, W., Niezgoda, M., Blanton, J., Self, J., Rowley, K.J., Mandell, R.B., Babcock, G.J., Thomas, W.D., Jr., Rupprecht, C.E., Ambrosino, D.M., 2007. Identification and characterization of a human monoclonal antibody that potently neutralizes a broad panel of rabies virus isolates. *Vaccine* 25, 2800-2810.

Sommerstein, R., Flatz, L., Remy, M.M., Malinge, P., Magistrelli, G., Fischer, N., Sahin, M., Bergthaler, A., Igonet, S., Ter Meulen, J., Rigo, D., Meda, P., Rabah, N., Coutard, B., Bowden, T.A., Lambert, P.H., Siegrist, C.A., Pinschewer, D.D., 2015. Arenavirus Glycan Shield Promotes Neutralizing Antibody Evasion and Protracted Infection. *PLoS Pathog* 11, e1005276.

Southern, P.J., Singh, M.K., Riviere, Y., Jacoby, D.R., Buchmeier, M.J., Oldstone, M.B., 1987. Molecular characterization of the genomic S RNA segment from lymphocytic choriomeningitis virus. *Virology* 157, 145-155.

Spiropoulou, C.F., Kunz, S., Rollin, P.E., Campbell, K.P., Oldstone, M.B., 2002. New World arenavirus clade C, but not clade A and B viruses, utilizes alpha-dystroglycan as its major receptor. *J Virol* 76, 5140-5146.

Stass, R., Ng, W.M., Kim, Y.C., Huiskonen, J.T., 2019. Structures of enveloped virions determined by cryogenic electron microscopy and tomography. *Adv Virus Res* 105, 35-71.

Stenglein, M.D., Jacobson, E.R., Chang, L.W., Sanders, C., Hawkins, M.G., Guzman, D.S., Drazenovich, T., Dunker, F., Kamaka, E.K., Fisher, D., Reavill, D.R., Meola, L.F., Levens, G., DeRisi, J.L., 2015. Widespread recombination, reassortment, and transmission of unbalanced compound viral genotypes in natural arenavirus infections. *PLoS Pathog* 11, e1004900.

Stenglein, M.D., Sanders, C., Kistler, A.L., Ruby, J.G., Franco, J.Y., Reavill, D.R., Dunker, F., Derisi, J.L., 2012. Identification, characterization, and in vitro culture of highly divergent arenaviruses from boa constrictors and annulated tree boas: candidate etiological agents for snake inclusion body disease. *mBio* 3, e00180-00112.

Strecker, T., Eichler, R., Meulen, J., Weissenhorn, W., Dieter Klenk, H., Garten, W., Lenz, O., 2003. Lassa virus Z protein is a matrix protein and sufficient for the release of virus-like particles [corrected]. *J Virol* 77, 10700-10705.

Stuart, D.I., Levine, M., Muirhead, H., Stammers, D.K., 1979. Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 Å. *J Mol Biol* 134, 109-142.

Tan, Y.Z., Baldwin, P.R., Davis, J.H., Williamson, J.R., Potter, C.S., Carragher, B., Lyumkis, D., 2017. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nature Methods* 14, 793-796.

Tani, H., Fukushi, S., Yoshikawa, T., Saijo, M., Morikawa, S., 2012. [Arenavirus infections]. *Uirusu* 62, 229-238.

ter Meulen, J., Badusche, M., Kuhnt, K., Doetze, A., Satoguina, J., Marti, T., Loeliger, C., Koulemou, K., Koivogui, L., Schmitz, H., Fleischer, B., Hoerauf, A., 2000. Characterization of human CD4(+) T-cell clones recognizing conserved and variable epitopes of the Lassa virus nucleoprotein. *J Virol* 74, 2186-2192.

Thoulouze, M.I., Lafage, M., Schachner, M., Hartmann, U., Cremer, H., Lafon, M., 1998. The neural cell adhesion molecule is a receptor for rabies virus. *J Virol* 72, 7181-7190.

Tsekoa, T.L., Lotter-Stark, T., Buthelezi, S., Chakauya, E., Stoychev, S.H., Sabeta, C., Shumba, W., Phahladira, B., Hume, S., Morton, J., Rupprecht, C.E., Steinkellner, H., Pauly, M., Zeitlin, L., Whaley, K., Chikwamba, R., 2016. Efficient In Vitro and In Vivo Activity of

Glyco-Engineered Plant-Produced Rabies Monoclonal Antibodies E559 and 62-71-3. *PLoS One* 11, e0159313.

Tuffereau, C., Bénéjean, J., Blondel, D., Kieffer, B., Flamand, A., 1998. Low-affinity nerve-growth factor receptor (P75NTR) can serve as a receptor for rabies virus. *EMBO J* 17, 7250-7259.

Urata, S., de la Torre, J.C., 2011. Arenavirus budding. *Adv Virol* 2011, 180326.

Urata, S., Noda, T., Kawaoka, Y., Morikawa, S., Yokosawa, H., Yasuda, J., 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.

Urata, S., Noda, T., Kawaoka, Y., Yokosawa, H., Yasuda, J., 2006. Cellular factors required for Lassa virus budding. *J Virol* 80, 4191-4195.

Urata, S., Yasuda, J., 2012. Molecular mechanism of arenavirus assembly and budding. *Viruses* 4, 2049-2079.

van Dolleweerd, C.J., Teh, A.Y., Banyard, A.C., Both, L., Lotter-Stark, H.C., Tsekoa, T., Phahladira, B., Shumba, W., Chakauya, E., Sabeta, C.T., Gruber, C., Fooks, A.R., Chikwamba, R.K., Ma, J.K., 2014. Engineering, expression in transgenic plants and characterisation of E559, a rabies virus-neutralising monoclonal antibody. *J Infect Dis* 210, 200-208.

Vela, E., 2012. Animal models, prophylaxis, and therapeutics for arenavirus infections. *Viruses* 4, 1802-1829.

Vela, E.M., Bowick, G.C., Herzog, N.K., Aronson, J.F., 2008. Genistein treatment of cells inhibits arenavirus infection. *Antiviral Res* 77, 153-156.

Veliziotis, I., Roman, A., Martiny, D., Schuldt, G., Claus, M., Dauby, N., Van den Wijngaert, S., Martin, C., Nasreddine, R., Perandones, C., Mahieu, R., Swaan, C., Van Praet, S., Konopnicki, D., Morales, M.A., Malvy, D., Stevens, E., Dechamps, P., Vlieghe, E., Vandenberg, O., Günther, S., Gérard, M., 2020. Clinical Management of Argentine Hemorrhagic Fever using Ribavirin and Favipiravir, Belgium, 2020. *Emerg Infect Dis* 26, 1562-1566.

Volpon, L., Osborne, M.J., Borden, K.L., 2008. NMR assignment of the arenaviral protein Z from Lassa fever virus. *Biomol NMR Assign* 2, 81-84.

Volpon, L., Osborne, M.J., Capul, A.A., de la Torre, J.C., Borden, K.L., 2010. Structural characterization of the Z RING-eIF4E complex reveals a distinct mode of control for eIF4E. *Proc Natl Acad Sci U S A* 107, 5441-5446.

Walker, D.H., McCormick, J.B., Johnson, K.M., Webb, P.A., Komba-Kono, G., Elliott, L.H., Gardner, J.J., 1982. Pathologic and virologic study of fatal Lassa fever in man. *Am J Pathol* 107, 349-356.

Walker, P.J., Blasdel, K.R., Calisher, C.H., Dietzgen, R.G., Kondo, H., Kurath, G., Longdon, B., Stone, D.M., Tesh, R.B., Tordo, N., Vasilakis, N., Whitfield, A.E., Ictv Report, C., 2018. ICTV Virus Taxonomy Profile: Rhabdoviridae. *J Gen Virol* 99, 447-448.

Walker, P.J., Kongsuwan, K., 1999. Deduced structural model for animal rhabdovirus glycoproteins. *J Gen Virol* 80 (Pt 5), 1211-1220.

Walter, T.S., Diprose, J.M., Mayo, C.J., Siebold, C., Pickford, M.G., Carter, L., Sutton, G.C., Berrow, N.S., Brown, J., Berry, I.M., Stewart-Jones, G.B., Grimes, J.M., Stammers, D.K., Esnouf, R.M., Jones, E.Y., Owens, R.J., Stuart, D.I., Harlos, K., 2005. A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization. *Acta Crystallogr D Biol Crystallogr* 61, 651-657.

Wang, J., Wang, Z., Liu, R., Shuai, L., Wang, X., Luo, J., Wang, C., Chen, W., Wang, X., Ge, J., He, X., Wen, Z., Bu, Z., 2018. Metabotropic glutamate receptor subtype 2 is a cellular receptor for rabies virus. *PLoS Pathog* 14, e1007189.

Wang, W., Zhou, Z., Zhang, L., Wang, S., Xiao, G., 2016. Structure-function relationship of the mammarenavirus envelope glycoprotein. *Virol Sin* 31, 380-394.

Wanzeck, K., Boyd, K.L., McCullers, J.A., 2011. Glycan shielding of the influenza virus hemagglutinin contributes to immunopathology in mice. *Am J Respir Crit Care Med* 183, 767-773.

Watanabe, Y., Berndsen, Z.T., Raghwani, J., Seabright, G.E., Allen, J.D., Pybus, O.G., McLellan, J.S., Wilson, I.A., Bowden, T.A., Ward, A.B., Crispin, M., 2020. Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. *Nat Commun* 11, 2688.

Watanabe, Y., Bowden, T.A., Wilson, I.A., Crispin, M., 2019. Exploitation of glycosylation in enveloped virus pathobiology. *Biochimica et biophysica acta. General subjects*.

Watanabe, Y., Raghwani, J., Allen, J.D., Seabright, G.E., Li, S., Moser, F., Huiskonen, J.T., Strecker, T., Bowden, T.A., Crispin, M., 2018. Structure of the Lassa virus glycan shield provides a model for immunological resistance. *Proc Natl Acad Sci U S A*.

Weissenbacher, M.C., Calello, M.A., Merani, M.S., Rodriguez, M., McCormick, J.B., 1986. Therapeutic Effect of the Antiviral Agent Ribavirin in Junin Virus Infection of Primates. *Journal of Medical Virology* 20, 261-267.

White, J.M., Delos, S.E., Brecher, M., Schornberg, K., 2008. Structures and mechanisms of viral membrane fusion proteins: multiple variations on a common theme. *Crit Rev Biochem Mol Biol* 43, 189-219.

Whitt, M.A., Buonocor, L., Prehaud, C., Rose, J.K., 1991. Membrane fusion activity, oligomerization, and assembly of the rabies virus glycoprotein. *Virology* 185, 681-688.

Wilda, M., Lopez, N., Casabona, J.C., Franze-Fernandez, M.T., 2008. Mapping of the tacaribe arenavirus Z-protein binding sites on the L protein identified both amino acids within the putative polymerase domain and a region at the N terminus of L that are critically involved in binding. *J Virol* 82, 11454-11460.

Winter, G., Lobley, C.M., Prince, S.M., 2013. Decision making in xia2. *Acta Crystallogr D Biol Crystallogr* 69, 1260-1273.

Wolf, Y.I., Kazlauskas, D., Iranzo, J., Lucia-Sanz, A., Kuhn, J.H., Krupovic, M., Dolja, V.V., Koonin, E.V., 2018. Origins and Evolution of the Global RNA Virome. *MBio* 9.

Wright, K.E., Salvato, M.S., Buchmeier, M.J., 1989. Neutralizing epitopes of lymphocytic choriomeningitis virus are conformational and require both glycosylation and disulfide bonds for expression. *Virology* 171, 417-426.

Wu, Z., Lu, L., Du, J., Yang, L., Ren, X., Liu, B., Jiang, J., Yang, J., Dong, J., Sun, L., Zhu, Y., Li, Y., Zheng, D., Zhang, C., Su, H., Zheng, Y., Zhou, H., Zhu, G., Li, H., Chmura, A., Yang, F., Daszak, P., Wang, J., Liu, Q., Jin, Q., 2018. Comparative analysis of rodent and small mammal viromes to better understand the wildlife origin of emerging infectious diseases. *Microbiome* 6, 178.

Wunner, W.H., 2007. 2 - Rabies Virus, in: Jackson, A.C., Wunner, W.H. (Eds.), *Rabies* (Second Edition). Academic Press, Oxford, pp. 23-68.

Wunner, W.H., Conzelmann, K.-K., 2013. Chapter 2 - Rabies Virus, in: Jackson, A.C. (Ed.), *Rabies* (Third Edition). Academic Press, Boston, pp. 17-60.

Wycoff, K., Maclean, J., Belle, A., Yu, L., Tran, Y., Roy, C., Hayden, F., 2015. Anti-infective immunoadhesins from plants. *Plant Biotechnol J* 13, 1078-1093.

- Xing, J., Ly, H., Liang, Y., 2015. The Z proteins of pathogenic but not nonpathogenic arenaviruses inhibit RIG-I-like receptor-dependent interferon production. *J Virol* 89, 2944-2955.
- Yang, F., Lin, S., Ye, F., Yang, J., Qi, J., Chen, Z., Lin, X., Wang, J., Yue, D., Cheng, Y., Chen, Z., Chen, H., You, Y., Zhang, Z., Yang, Y., Yang, M., Sun, H., Li, Y., Cao, Y., Yang, S., Wei, Y., Gao, G.F., Lu, G., 2020. Structural Analysis of Rabies Virus Glycoprotein Reveals pH-Dependent Conformational Changes and Interactions with a Neutralizing Antibody. *Cell Host Microbe* 27, 441-453.e447.
- Yaro, C.A., Kogi, E., Opara, K.N., Batiha, G.E.-S., Baty, R.S., Albrakati, A., Altalbawy, F.M.A., Etuh, I.U., Oni, J.P., 2021. Infection pattern, case fatality rate and spread of Lassa virus in Nigeria. *BMC Infectious Diseases* 21, 149.
- Ye, J., Ma, N., Madden, T.L., Ostell, J.M., 2013. IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res* 41, W34-40.
- Ye, Y., Godzik, A., 2004. FATCAT: a web server for flexible structure comparison and structure similarity searching. *Nucleic Acids Res* 32, W582-585.
- York, J., Dai, D., Amberg, S.M., Nunberg, J.H., 2008. pH-induced activation of arenavirus membrane fusion is antagonized by small-molecule inhibitors. *J Virol* 82, 10932-10939.
- York, J., Nunberg, J.H., 2006. Role of the stable signal peptide of Junín arenavirus envelope glycoprotein in pH-dependent membrane fusion. *J Virol* 80, 7775-7780.
- York, J., Nunberg, J.H., 2007a. Distinct requirements for signal peptidase processing and function in the stable signal peptide subunit of the Junín virus envelope glycoprotein. *Virology* 359, 72-81.
- York, J., Nunberg, J.H., 2007b. A novel zinc-binding domain is essential for formation of the functional Junín virus envelope glycoprotein complex. *J Virol* 81, 13385-13391.
- York, J., Romanowski, V., Lu, M., Nunberg, J.H., 2004. The signal peptide of the Junín arenavirus envelope glycoprotein is myristoylated and forms an essential subunit of the mature G1-G2 complex. *J Virol* 78, 10783-10792.
- Zeitlin, L., Cross, R.W., Geisbert, J.B., Borisevich, V., Agans, K.N., Prasad, A.N., Enterlein, S., Aman, M.J., Bornholdt, Z.A., Brennan, M.B., Campbell, L., Kim, D., Mlakar, N., Moyer, C.L., Pauly, M.H., Shestowsky, W., Whaley, K.J., Fenton, K.A., Geisbert, T.W., 2021. Therapy for Argentine hemorrhagic fever in nonhuman primates with a humanized monoclonal antibody. *Proceedings of the National Academy of Sciences* 118, e2023332118.

Zeitlin, L., Geisbert, J.B., Deer, D.J., Fenton, K.A., Bohorov, O., Bohorova, N., Goodman, C., Kim, D., Hiatt, A., Pauly, M.H., Velasco, J., Whaley, K.J., Altmann, F., Gruber, C., Steinkellner, H., Honko, A.N., Kuehne, A.I., Aman, M.J., Sahandi, S., Enterlein, S., Zhan, X., Enria, D., Geisbert, T.W., 2016. Monoclonal antibody therapy for Junin virus infection. *Proc Natl Acad Sci U S A* 113, 4458-4463.

Zeltina, A., Bowden, T.A., 2017. Human antibody pieces together the puzzle of the trimeric Lassa virus surface antigen. *Nat Struct Mol Biol* 24, 559-560.

Zeltina, A., Krumm, S.A., Sahin, M., Struwe, W.B., Harlos, K., Nunberg, J.H., Crispin, M., Pinschewer, D.D., Doores, K.J., Bowden, T.A., 2017. Convergent immunological solutions to Argentine hemorrhagic fever virus neutralization. *Proc Natl Acad Sci U S A* 114, 7031-7036.

Zhang, G., Fu, Z.F., 2012. Complete Genome Sequence of a Street Rabies Virus from Mexico. *Journal of Virology* 86, 10892.

Zhao, S., Musa, S.S., Fu, H., He, D., Qin, J., 2020. Large-scale Lassa fever outbreaks in Nigeria: quantifying the association between disease reproduction number and local rainfall. *Epidemiol Infect* 148, e4.

Ziegler, C.M., Eisenhauer, P., Bruce, E.A., Weir, M.E., King, B.R., Klaus, J.P., Kremmentsov, D.N., Shirley, D.J., Ballif, B.A., Botten, J., 2016. The Lymphocytic Choriomeningitis Virus Matrix Protein PPXY Late Domain Drives the Production of Defective Interfering Particles. *PLoS Pathog* 12, e1005501.

Zong, M., Fofana, I., Choe, H., 2014. Human and host species transferrin receptor 1 use by North American arenaviruses. *J Virol* 88, 9418-942

8 APPENDIX

8.1 cDNA AND AMINO ACID SEQUENCES OF FAB FRAGMENTS

Fab JUN1 (Heavy chain)

cagatccagttggtgcagtctgtagagaccgggggaggcttggtgaggcctggaaattct
Q I Q L V Q S V E T G G G L V R P G N S
ctgaaactctcctgtgttacctcgggattcactttcagtaactaccaaatgcactggctt
L K L S C V T S G F T F S N Y Q M H W L
cgccagcctccaggggaagaggctggaatggattgctgtgaattacagtcaaattctgataat
R Q P P G K R L E W I A V I T V K S D N
tatggagcaaattatgtagagtctgtgaaaggcagattcgccatttcaagagatgattca
Y G A N Y V E S V K G R F A I S R D D S
aaaagcagtgtctacctagagatgaacagattaagagaggaagacactgccacttatttt
K S S V Y L E M N R L R E E D T A T Y F
tgtagtagatcggggatctatgatggttactacgcctatgctatggactactgggggcaa
C S R S G I Y D G Y Y A Y A M D Y W G Q
ggaacctcagtcaccgtctcgagcgccaccaccaaggggcccatctgtctatccactggcc
G T S V T V S S A T T K G P S V Y P L A
cctggatctgctgccccaaactaactccatggtgaccctgggatgcctggtcaagggtat
P G S A A Q T N S M V T L G C L V K G Y
ttccctgagccagtgcagtgacctggaactctggatccctgtccagcggtgtgcacacc
F P E P V T V T W N S G S L S S G V H T
ttcccagctgtcctgcagtctgacctctacactctgagcagctcagtgactgtcccctcc
F P A V L Q S D L Y T L S S S V T V P S
agcacctggcccagccagaccgtcacctgcaacgttgcccacccggccagcagcaccaag
S T W P S Q T V T C N V A H P A S S T K
gtggacaagaaaattgtgcccagggttgt
V D K K I V P R D C

5' forward primer: gcgACCGGTcagatccagttggtgcagtctgta

3' reverse complementary primer: gcgGGTACCacaatccctgggcacaattt

“gcg” is a GC-overhang inserted to facilitate restriction endonuclease digestion.

ACCGGT and GGTACC are recognised by AgeI and KpnI restriction enzymes, respectively.

The heavy chain construct was cloned into pHLsec expression vector and expressed in HEK293T cells together with the light chain at a 1:1 molar ratio.

Fab JUN1 (Light chain)

agtgttgatgatgacccagtcctcaaaaattcatgtccacttcagtaggagacagggtcagc
S V V M T Q S Q K F M S T S V G D R V S
atcacctgcaaggccagtcacaaattgtgggtacttctgtagcctgggtatcaacagaaagca
I T C K A S Q I V G T S V A W Y Q Q K A
ggacaatctcctaaactactgattttattgggcctctacccggcacactggagtcctgat
G Q S P K L L I Y W A S T R H T G V P D
cgcttcacagccgggtggatctgggacagatttcactctcaccattaccaatgtgcagtct
R F T A G G S G T D F T L T I T N V Q S
gaagacttggcagattatttctgtcagcaatatgccacctatcctctcacgttcggctcg
E D L A D Y F C Q Q Y A T Y P L T F G S
gggaccaagctggagctgaaacgtacggatgctgcaccaactgtatccatcttcccacca
G T K L E L K R T D A A P T V S I F P P
tccagtgcagcagttaacatctggaggtgcctcagtcgtgtgcttcttgaacaacttctac
S S E Q L T S G G A S V V C F L N N F Y
cccaaagacatcaatgtcaagtgggaagattgatggcagtgaaacgacaaaatggcgtcctg
P K D I N V K W K I D G S E R Q N G V L
aacagttggactgatcaggacagcaaagacagcacctacagcatgagcagcaccctcacg
N S W T D Q D S K D S T Y S M S S T L T
ttgaccaaggacgagtatgaacgacataacagctatacctgtgaggccactcacaagaca
L T K D E Y E R H N S Y T C E A T H K T
tcaacttcacccattgtcaagagcttcaacaggaatgagtggtga
S T S P I V K S F N R N E C -

5' forward primer: gcgACCGGTagtgttgatgatgacccagtcctca

3' reverse complementary primer: gcgCTCGAGtcaacactcattcctgttgaagctc

“gcg” is a GC-overhang inserted to facilitate restriction endonuclease digestion.

ACCGGT and CTCGAG are recognised by AgeI and XhoI restriction enzymes, respectively.

The light chain construct was cloned into pHLsec expression vector and expressed in HEK293T cells together with the heavy chain at a 1:1 molar ratio.

Fab MAC1 (Heavy chain)

cagggtccaactgcagcagtcctggggctgagctggtgaagcctggggcctcagtgaagatt
Q V Q L Q Q S G A E L V K P G A S V K I
tcctgcaaagcttctggctacgcattcggtagccactggatgaactgggtgaagcagagg
S C K A S G Y A F G S H W M N W V K Q R
cctggaaagggctcttgagtggattggacagatttatcctggagatgggtgataactaac
P G K G L E W I G Q I Y P G D G D T N Y
aacggaaaggttcaagggtaaggccacactgactgcagacaaatcctccagcacagcctac
N G K F K G K A T L T A D K S S S T A Y
atgcagttcagcagcctgacctctgaggactctgcggtctatcttctgtgcaagagatgac
M Q F S S L T S E D S A V Y F C A R D D
tacggtacaagataactactttgactactggggccaaggcaccactctcacagtctcgagc
Y G T R Y Y F D Y W G Q G T T L T V S S
gccaccaccaagggcccatctgtctatccactggcccctggatctgctgcccactaac
A T T K G P S V Y P L A P G S A A Q T N
tccatggtgaccctgggatgcctgggtcaagggctatcttccctgagccagtgcagtgacc
S M V T L G C L V K G Y F P E P V T V T
tggaactctggatccctgtccagcgggtgtgcacaccttcccagctgtcctgcagtctgac
W N S G S L S S G V H T F P A V L Q S D
ctctacactctgagcagctcagtgcactgtcccctccagcacctggcccagccagaccgtc
L Y T L S S S V T V P S S T W P S Q T V
acctgcaacggttggccaccggccagcagcaccagggtggacaagaaaattgtgcccagg
T C N V A H P A S S T K V D K K I V P R
gattgt
D C

5' forward primer: gcgACCGGTcagggtccaactgcagcagt

3' reverse complementary primer: gcgGGTACCacaatccctgggcacaattt

“gcg” is a GC-overhang inserted to facilitate restriction endonuclease digestion.

ACCGGT and GGTACC are recognised by AgeI and KpnI restriction enzymes, respectively.

The heavy chain construct was cloned into pHLsec expression vector and expressed in HEK293T cells together with the light chain at a 1:1 molar ratio.

Fab MAC1 (Light chain)

gacatccagatgacacagactacatcctcgctgtctgcctctctgggagacagagtctcc
D I Q M T Q T T S S L S A S L G D R V S
atcagttgcagggctagtcaggacattaacaattatttaaactgggtaccagcaaaaacca
I S C R A S Q D I N N Y L N W Y Q Q K P
gatggaactgtttaaactcctgatccactacacatcaagattacgttcaggagtcccatca
D G T V K L L I H Y T S R L R S G V P S
aggttcagtggtcagtggttttgaacagattattctctcaccattaccaacctggaacaa
R F S G S G F G T D Y S L T I T N L E Q
gaagatatgtccacttacttttgccaacagggtaaaacgcttcctctcacgttcggtgct
E D I A T Y F C Q Q G K T L P L T F G A
ggcacaaagctggaaatcaaacgtacggatgctgcaccaactgtatccatcttcccacca
G T K L E I K R T D A A P T V S I F P P
tccagtgagcagttaacatctggaggtgcctcagtcgtgtgcttcttgaacaacttctac
S S E Q L T S G G A S V V C F L N N F Y
cccaaagacatcaatgtcaagtggagattgatggcagtgaaacgacaaaatggcgctcctg
P K D I N V K W K I D G S E R Q N G V L
aacagttggactgatcaggacagcaaagacagcacctacagcatgagcagcaccctcacg
N S W T D Q D S K D S T Y S M S S T L T
ttgaccaaggacgagtatgaacgacataacagctatacctgtgaggccactcacaagaca
L T K D E Y E R H N S Y T C E A T H K T
tcaacttcacccattgtcaagagcttcaacaggaatgagtggtga
S T S P I V K S F N R N E C -

5' forward primer: gcgACCGGTgacatccagatgacacagactacatc

3' reverse complementary primer: gcgCTCGAGtcaacactcattcctgttgaagctc

“gcg” is a GC-overhang inserted to facilitate restriction endonuclease digestion.

ACCGGT and CTCGAG are recognised by AgeI and XhoI restriction enzymes, respectively.

The light chain construct was cloned into pHLsec expression vector and expressed in HEK293T cells together with the heavy chain at a 1:1 molar ratio.

Fab 17C7 (Heavy chain)

gaggtgcagctggtggagtctgggggaggcgtggtccagcctgggaggtccctgagactc
E V Q L V E S G G G V V Q P G R S L R L
tcctgtgcagcctctggattcaccttcagtacctatgctatgcactgggtccgccaggct
S C A A S G F T F S T Y A M H W V R Q A
ccaggcaaggggctggagtgggtggcagttgtatcatatgatggacgcactaaagactac
P G K G L E W V A V V S Y D G R T K D Y
gcagactccgtgaagggccgattcacctctccagagacaattccaagaacacgctgtat
A D S V K G R F T I S R D N S K N T L Y
ctgcaaatgaacagcctgagaactgaggacacggctgtgtatcttctgtgagagagagg
L Q M N S L R T E D T A V Y F C A R E R
ttctctggggcctactttgactactggggccagggaaccctgggtcacctctcctcagcc
F S G A Y F D Y W G Q G T L V T V S S A
aaaacgacacccccatctgtctatccactggccccctggatctgctgcccactaactcc
K T T P P S V Y P L A P G S A A Q T N S
atgggtgaccctgggatgcctgggtcaagggtctatttccctgagccagtgacagtgacctg
M V T L G C L V K G Y F P E P V T V T W
aactctggatccctgtccagcgggtgtgcacaccttcccagctgtcctgcagtctgacctc
N S G S L S S G V H T F P A V L Q S D L
tacactctgagcagctcagtgactgtccccctccagcacctggcccagcgagaccgtcacc
Y T L S S S V T V P S S T W P S E T V T
tgcaacgttgcccacccggccagcagcaccaagggtggacaagaaaattgtgcccagggat
C N V A H P A S S T K V D K K I V P R D
tgtggt
C G

The heavy chain construct was cloned into pTT expression vector containing twin-strep tag using In-Fusion cloning and expressed in Expi293 cells together with the light chain at a 1:1 molar ratio.

Gene cloning was performed by Luke Thorley from Douglas group at the Jenner Institute.

Fab 17C7 (Light chain)

gaaattgtgttgacacagtcctccagccaccctgtctttgtctccaggggaaagagccacc
E I V L T Q S P A T L S L S P G E R A T
ctctcctgcagggccagtcagagtggttagcagctacttagcctggtaccaacagaaacct
L S C R A S Q S V S S Y L A W Y Q Q K P
ggccaggctcccaggctcctcatctatgatgcatccaacagggccactggcatcccagcc
G Q A P R L L I Y D A S N R A T G I P A
aggttcagtggtcagtggttctgggacagacttcactctcaccatcagcagcctagagcct
R F S G S G S G T D F T L T I S S L E P
gaagattttgcagttttattcctgtcagcagcgtaacaactggcctcccacttttcggcgga
E D F A V Y S C Q Q R N N W P P T F G G
gggaccaaggtggagatcaaacgggctgatgctgcaccaactgtatccatcttcccacca
G T K V E I K R A D A A P T V S I F P P
tccagtgcagcagttaacatctggaggtgcctcagtcgtgtgcttcttgaacaacttctac
S S E Q L T S G G A S V V C F L N N F Y
cccaaagacatcaatgtcaagtgggaagattgatggcagtgaaacgacaaaatggcgctcctg
P K D I N V K W K I D G S E R Q N G V L
aacagttggactgatcaggacagcaaagacagcacctacagcatgagcagcaccctcacg
N S W T D Q D S K D S T Y S M S S T L T
ttgaccaaggacgagtatgaacgacataacagctatacctgtgaggccactcacaagaca
L T K D E Y E R H N S Y T C E A T H K T
tcaacttcacccatcgtcaagagcttcaacaggaatgagtggttaa
S T S P I V K S F N R N E C -

The light chain construct was cloned into pOPINVL expression vector using In-Fusion cloning and expressed in Expi293 cells together with the heavy chain at a 1:1 molar ratio.

Gene cloning was performed by Luke Thorley from Douglas group at the Jenner Institute.

8.2 LIST OF PUBLICATIONS

Weng M. Ng*, Mehmet Sahin, Stefanie A. Krumm, Jeffrey Seow, Antra Zeltina, Karl Harlos, Guido Paesen, Daniel D. Pinschewer, Katie J. Doores, Thomas A. Bowden. 2021. Contrasting modes of New World arenavirus neutralization by immunization-elicited monoclonal antibodies. *Manuscript submitted*.

Carl Graham, Jeffrey Seow, Isabella Huettner, Hataf Khan, Neophytos Kouphou, Sam Acors, Helena Winstone, Suzanne Pickering, Rui Pedro Galao, Maria Jose Lista, Jose M. Jimenez-Guardeno, Adam G Laing, Yin Wu, Magdalene Joseph, Luke Muir, **Weng M. Ng**, Helen M. E. Duyvesteyn, Yuguang Zhao, Thomas A. Bowden, Manu Shankar-Hari, Annachiara Rosa, Peter Cherepanov, Laura E. McCoy, Adrian C. Hayday, Stuart J. D. Neil, Michael H. Malim, Katie J. Doores. 2021. Neutralization potency of monoclonal antibodies recognizing dominant and subdominant epitopes on SARS-CoV-2 Spike is impacted by the B.1.1.7 variant. *Immunity*. 54(6):1276-1289.e6.

Weng M. Ng*, Alice J. Stelfox, Thomas A. Bowden. 2020. Unraveling virus relationships by structure-based phylogenetic classification. *Virus Evol.* 6(1):veaa003.

Robert Stass*, **Weng M. Ng***, Young Chan Kim*, Juha T. Huiskonen. 2019. Structures of enveloped virions determined by cryogenic electron microscopy and tomography. *Adv Virus Res.* 105:35–71.

Rhys Pryce*, **Weng M. Ng***, Antra Zeltina, Yasunori Watanabe, Kamel El Omari, Armin Wagner, Thomas A. Bowden. 2018. Structure-Based Classification Defines the Discrete Conformational Classes Adopted by the Arenaviral GP1. *J Virol.* 93(1):e01048-18.

* first / co-first authors