Structural Studies of the Influenza Genome and its Interactions with Human Innate Immune Factors

Jack D. Whitehead

Division of Structural Biology (Strubi),
Sir William Dunn School of Pathology,
and
Brasenose College

University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity Term 2025

Table of Contents

Table of Figures

Table of Tables

Abstract

Structural Studies of the Influenza Genome and its Interactions with Human Innate Immune Factors

Jack David Whitehead, Brasenose College, University of Oxford

DPhil in Cellular Structural Biology

Trinity Term, 2025

Influenza viruses cause an estimated 3-5 million severe cases of contagious respiratory disease and regularly lead to up to 650,000 deaths annually according to the World Health Organisation, posing a significant burden to global public health systems. Despite decades of research, our structural understanding of key viral components remains incomplete. Central to the viral lifecycle is the transcription and replication of the segmented negative sense viral RNA genome, packaged into the viral ribonucleoprotein complex (vRNP). The vRNP comprises one copy of viral RNA corresponding to one genome segment, one copy of the heterotrimeric influenza polymerase complex, and multiple copies of the viral nucleoprotein. The nucleoprotein serves the dual purpose of protecting viral RNA from host immune recognition, while also facilitating genome transcription and replication, making it both essential for viral function and a prime target for the host innate immune response.

The mode of RNA and immune factor binding to the nucleoprotein remains structurally unresolved. Further, despite significant effort both in this laboratory and in others, the native structure of the vRNP remains unknown.

During this DPhil, I have aimed to answer these questions using biochemical, biophysical, and structural techniques. In this thesis, I present the structure of a previously-unsolved avian influenza A virus nucleoprotein using cryo-electron microscopy single particle analysis. I biochemically and biophysically characterise the interaction of two human innate immune factors with susceptible avian influenza A virus nucleoproteins and show that the interaction of these immune factors is selective for avian strains. I describe novel approaches to solving vRNP structure, a method of coating influenza virus vRNPs with binding factors of interest, and present a low-resolution 3D reconstruction of the vRNP.

The findings presented in this thesis advance our understanding of influenza virus architecture, host-pathogen interactions, and demonstrate the successful application of a number of structural techniques to the study of structural virology.

Acknowledgements

Firstly, I would like to thank the three main supervisors I have had the pleasure of working with – Jon Grimes, Ervin Fodor, and Loïc Carrique. Thank you for your support and supervision. Thank you for the freedom you've given me to try different things in the lab, and thank you especially for your support in these past few months. I'm sorry they were quite stressful!

Special thanks go to Cecilia Rocchi, without whom I would have struggled to get through these past two months. Thank you for being so wonderful, so kind, so supportive, and so generous with your time and expertise. I will never forget the selfless effort and support you have provided me when I've been stuck.

I would like to thank the Wellcome Trust and Brasenose College for their generous financial support through the DPhil, and to all the staff, students, and scientists both at the Sir William Dunn School of Pathology and in the Division of Structural Biology for their support of my work.

To my former supervisors and mentors, Max Renner and Jeremy Keown, thank you for your support and mentorship. Thanks also to Dan Hurdiss and Harry Williams for fruitful discussions and support of my work.

To Bob Siegel, thank you for your mentorship and ongoing support through my university career. Your enthusiasm and knowledge of virology truly inspired me to take this path, and has helped me to get to this stage.

To Jamie Auxillos, thank you for your support and mentorship through my career in research. You were the first to teach me the basics of lab work and inspire a joy in lab science, and you have been a wonderful mentor ever since.

To the unsung heroes – Jane Sharps, Margaret Jones, and Louise Samson. Thank you for your laboratory and administrative support. You have made the DPhil a much smoother experience than it otherwise would have been, and I've always felt supported by you. Thank you.

To a select few important teachers I had the pleasure of being taught by in school – Stuart Gledhill, Joanne Ormisher, Lyndsey Berry, and Tony McGuinness. Thank you for all the extra effort you made to inspire me to learn and to be curious about Biology and the World around me more broadly. Thank you for taking the time to answer all of my questions back then – you helped to cultivate in me the curiosity which still exists today.

To my friend and mentor, Elspeth Garman. Thank you for your support and guidance over the past five years. You have been one of my strongest advocates and have helped me through when things have been hard. I am eternally grateful for your unwavering support, and I vow never to forget to include a scale bar!

To my friends and colleagues at Stanford House – Autumn Tull, Sharon Scott, Kim Marsh, Tom Cooper, Alaba Angole, and Daniel Gerrard. Thank you all for your support, joy, and patience in my three years of working as a junior dean for Stanford. My involvement in Stanford throughout my eight years in Oxford has shaped my experience of this city, and more broadly, my life. I will never forget all of the great experiences we have shared.

To all the friends I've made along the way – Carys Williams, Misha Le Claire, Alison Rep, Cecilia Rocchi, Alaa Baazaoui, Amanda Zhu, Patrick McCubbin, Kuang-Yu Chen, Alasdair Hood, Pooja Gupta, and so many other people I haven't mentioned. Thank you for all of the love, support, and generosity you have shown me, and for all the experiences we've shared. I treasure all the memories.

To those special friends who have been checking up on me these past two months when I needed it the most – Joe Lever, Jaymin Shah, Lizzie Jones, Dillon Lim, Vrinda Vasavada, Priyadarshini Chatterjee, and Malhar Khushu. Thank you so much for your support. Knowing you were there for me through this made all the difference.

To the friends I've written with across various Oxford libraries – Carys Williams, Alison Rep, Jake Smith, and Helena Watson. I'm so very grateful to have spent the last few months surrounded by a wonderful support network of suffering PhD friends. Thank you for your support… the end is in sight.

To Miles Graham, one of my biggest supporters, thank you for your tireless patience with me, and for all of the help with computational work. Thanks for all the runs, fish and chip dinners, and for having my back these past eight years. It really means the World.

To Carys Williams, my best friend in Strubi, who has been there with me from the start. Thank you for your friendship and support these past four years. Thank you for all the coffee trips, Chester Arms dinners, and walks with Winnie. You have been a true and constant and support in the DPhil, and I don't think I could have got through it without someone as strong, wonderful, and inspiring as you.

To my parents, for their unwavering love and support of me through the DPhil and always. I will forever cherish your love and support, and try with every fibre of my being to make you proud. Thank you.

And finally, to Matt, my long-suffering, long-term, long-distance partner who has been by my side through thick and thin. Thank you for your patience, your love, and your support of me these past years (and especially in these last two months!) Thank you from the bottom of my heart.

Dedication

I dedicate this thesis to my Dad, Eric. Thank you for teaching me to be curious. Thank you for all that you have done and continue to do for me. I am forever grateful.

« Τά καλά κόποις κτώνται »

– Traditional Greek Proverb

Declaration

I declare that the contents of this thesis are a product of my own work except where explicitly stated. This work was undertaken between October 2021 and September 2025 under the supervision of Prof. Jonathan Grimes of the Division of Structural Biology, and Prof. Ervin Fodor of the Sir William Dunn School of Pathology. The length of this thesis does not exceed the word limit for a DPhil thesis submitted to the University of Oxford, Medical Sciences Division.

Jack David Whitehead

September 2025

Abbreviations

Å	Ångström (10^{-10} m)
Ab	Antibody
ANP32	Acidic Nuclear Phosphoprotein 32
ATP	Adenosine Triphosphate
BSA	Bovine Serum Albumin
BTN3A3	Butyrophilin Subfamily 3 Member A3
CC	Cross-Correlation
CRM1	Chromosomal Region Maintenance 1
cRNA	Complementary RNA
cRNP	Complementary Ribonucleoprotein
cryo-EM SPA	Cryogenic Electron Microscopy Single Particle Analysis
cryo-ET STA	Cryogenic Electron Tomography Sub-Tomogram Averaging
CTD	C-Terminal Domain
DLS	Dynamic Light Scattering
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl Sulfoxide $(CH_3)_2SO$
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol $(CH(OH)CH_2SH)_2$
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid $[CH_2N(CH_2CO_2H)_2]_2$
EM	Electron Microscopy

EMDB	Electron Microscopy Data Bank
ER	Endoplasmic Reticulum
ET	Electron Tomography
FCS	Foetal Calf Serum
FluPol	Influenza Polymerase
Gb	Gluebody
GST	Glutathione S-Transferase
GTP	Guanosine Triphosphate
HA	Haemagglutinin
HCl	Hydrochloric Acid
HEK	Human Embryonic Kidney Cell
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine (conventionally in relation to a poly-histidine tag)
HIV	Human Immunodeficiency Virus
HPAI	Highly Pathogenic Avian Influenza
Hpi	Hours Post-Infection
IAV	Influenza A Virus
IFITM3	Interferon-Induced Transmembrane Protein 3
IFN	Interferon
IgG	Immunoglobulin G
IMAC	Immobilised Metal Affinity Chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kb	Kilobase
kDa	Kilodalton

keV	Kiloelectron Volt
LB	Lysogeny or Luria Broth
LCAR	Low-Complexity Acidic Region
LLG	Log Likelihood Gain
LRT	Lower Respiratory Tract
M1	Matrix Protein 1
M2	Matric Protein 2
m7G	7-Methylguanosine
Mb	Megabody
MBCS	Multi-Basic Cleavage Site
MBP	Maltose Binding Protein
MDBK	Madin-Darby Bovine Kidney Cell
$MgCl_2$	Magnesium Chloride
MTOC	Microtubule Organising Centre
mM	Millimolar
MOI	Multiplicity of Infection
MR	Molecular Replacement
mRNA	Messenger RNA
MxA	Myxovirus Resistance Protein 1
MxB	Myxovirus Resistance Protein 2
N	Nucleoprotein
NA	Neuraminidase
NaCl	Sodium Chloride
Nb	Nanobody

Nm	Nanometre
NaOH	Sodium Hydroxide
NEP	Nuclear Export Protein
NLS	Nuclear Localisation Signal
nM	Nanomolar
NPC	Nuclear Pore Complex
Ntd	Nucleotide
NT60	Influenza A/Northern Territory/60/1968 (H3N2) Virus
NP	Nucleoprotein
Ni-NTA	Nickel – Nitrilotriacetic Acid
OTG	Octylthioglucoside
P3	Polymerase 3 (PA equivalent in influenza C virus)
PA	Polymerase Acidic Protein
PB1	Polymerase Basic Protein 1
PB2	Polymerase Basic Protein 2
PBS	Phosphate Buffered Saline
PBS-T	Phosphate Buffered Saline with added Tween® 20 Detergent
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PEG	Polyethylene Glycol
pH	$-\log_{10} c$, where c is $[H^+]$ in mol l^{-1}
poly(A)	Poly-adenosine tail
poly(U)	Poly-uracil tail
PRR	Pattern Recognition Receptor

Rab11	Ras-Related Nuclear Protein 11
RIG-I	Retinoic acid-inducible gene I
RMSD	Root Mean Squared Deviation
RNA	Ribonucleic Acid
RNAPII	RNA Polymerase II
RNP	Ribonucleoprotein
Rnase	Ribonuclease
rpm	Revolutions Per Minute
RSV	Respiratory Syncytial Virus
s	Second
SDS-PAGE	Sodium Dodecyl Sulfate – Polyacrylamide Gel Electrophoresis
SEC	Size-Exclusion Chromatography
SEC-MALS	Size-Exclusion Chromatography – Multi-Angle Light Scattering
SOC	Super Optimal Broth Medium
TAE	Tris-Acetate-EDTA Buffer
TB	Terrific Broth
TEM	Transmission Electron Microscopy
TFZ	Translation Function Z-Score
Tris	Tris(hydroxymethyl)aminomethane $(HOCH_2)_3CNH_2$
Tk	Influenza A/turkey/Turkey/1/2005 (H5N1) Virus
URT	Upper Respiratory Tract
UV	Ultra-Violet
vRNA	Viral RNA
vRNP	Viral Ribonucleoprotein

WCL	Whole Cell Lysate
WSN	Wilson Smith neurotropic Influenza A/WSN/1933 (H1N1) Virus
YB-1	Y Box Binding Protein 1

1. Introduction

1.1. Negative Sense RNA Virus Genome Architecture

Negative sense RNA viruses (NSVs), or the *Negarnaviricota*, pose significant disease burdens to the human population. The *Negarnaviricota* can be conveniently split into those members who replicate in the host cell cytoplasm and those who replicate in the host cell nucleus. Additionally, the *Negarnaviricota* can be further split into members which have one linear (non-segmented) genome (the *Haploviricotina*), and those which have a segmented viral genome (the *Polyploviricotina*). This is displayed as a table below (table 1.1).

Table 1.1: Diagram of *Negarnaviricota* Classification.

The Negarnaviricota can be conveniently split into members who replicate in the host cell cytoplasm, and those who replicate in the host cell nucleus. This classification can be further elaborated by splitting based on members who have one linear genome segment versus multiple segments. Example viruses are given in each category. The examples listed below are displayed in their appropriate section.

***Negarnaviricota* Classification**	**Cytoplasmic Replication**	**Nuclear Replication**
Non-Segmented Genome (*Haploviricotina*)	Rabies virus, Measles virus, Respiratory syncytial virus, Ebola virus	Borna Disease virus
Segmented Genome (*Polyploviricotina*)	Lassa virus, Rift Valley fever virus	Influenza virus

Both genome types include virus families with constituent viruses of concern to human health. The *Haploviricotina* include families such as the *Rhabdoviridae* (including rabies virus which causes a contagious disease of the central nervous system with an over 99 % case fatality rate), the *Paramyxoviridae* (including measles virus, causing a highly infectious respiratory disease and regularly over 100,000 deaths annually (Minta, 2024), the *Pneumoviridae* (including respiratory syncytial virus (RSV),

the causative agent of a highly transmissible respiratory disease which leads to over 100,000 deaths annually (Li et al., 2022), the *Bornaviridae* (including borna disease virus, the etiologic agent of a small number of fatal encephalitis cases in humans each year), and the *Filoviridae* (e.g. Ebola virus, which causes often fatal haemorrhagic fever in humans).

The *Polyploviricotina* include the *Arenaviridae* (which include Lassa virus, the pathogen responsible for a fatal haemorrhagic fever), the *Phenuiviridae* (including Rift Valley Fever Virus, RVFV, the causative agent of the potentially fatal Rift Valley Fever), and the *Orthomyxoviridae* (including influenza virus), which causes seasonal epidemics and occasional pandemics of contagious respiratory disease, causing up to 650,000 deaths annually (Ruigrok et al., 2011).

One major challenge faced by viruses in general is the recognition of their genome by the host cell immune response. As will be introduced later, the human innate immune response is primed to recognise foreign RNA which is present in the cell. Given this and the potential for RNA degradation by host nucleases, such as Rnase L (Karasik and Guydosh, 2024), the virus must develop a packaging mechanism to protect the viral genome. To address both innate immune recognition of viral RNA, and nuclease-mediated degradation of viral RNA, across the *Negarnaviricota*, a similar packaging system has evolved. The viral RNA genome is bound and protected by multiple copies of an RNA-binding nucleoprotein (often abbreviated to either N, or NP). The broad packaging mechanism is similar across the phylum, with a sequence-independent

RNA binding mechanism which binds to the negatively-charged RNA phosphodiester backbone.

Given this degenerate RNA binding, across the *Negarnaviricota*, the RNA binding groove is seen to be lined with positively charged residues. It is often experimentally favourable for any cellular RNA present in an expression system to bind to the nucleoprotein, which can lead to the spontaneous formation of regularly-sized oligomers. In one paper, this was seen to lead to the regular arrangement of either a 9-mer, 10-mer, 11-mer, 12-mer, or short helical particle of human metapneumovirus (HMPV) nucleoprotein expressed and purified from *E. coli* (Whitehead et al., 2023).

As a general rule of nucleoprotein binding to RNA across the NSVs, RNA is bound in a groove lined with positively-charged residues which is formed between the N and C portions of the nucleoprotein. However, the number of RNA nucleotides bound per nucleoprotein molecule varies widely across the NSVs (table 1.2) (Luo et al., 2020; Sabsay and te Velthuis, 2024).

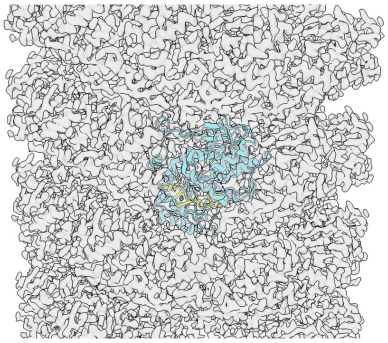
Table 1.2: Number of Nucleotides Bound to *Negarnaviricota* Nucleoproteins.

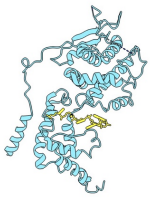
The number of nucleotides bound to each nucleoprotein monomer as observed through structural biology analysis of a number of viral examples. The number of nucleotides bound, example viruses, and references are provided.

Reported Number of Nucleotides Bound per Nucleoprotein	Examples	References
6	Measles Virus, Ebola Virus, Human Metapneumovirus	(Gutsche et al., 2015; Sugita et al., 2018; Whitehead et al., 2023)
7	Respiratory Syncytial Virus	(Tawar et al., 2009)
9	Rabies Virus, Vesicular Stomatitis Virus	(Albertini et al., 2006; Green et al., 2006; Jenni et al., 2022)
11	La Crosse Virus, Leanyer Virus	(Reguera et al., 2013; Niu et al., 2013)
24	Influenza A Virus	(Ortega et al., 2000; Tang et al., 2021)

Oligomeric Assembly **Monomeric Nucleoprotein**

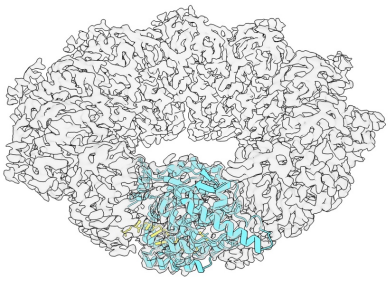
(A)

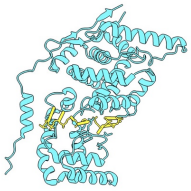




Measles virus

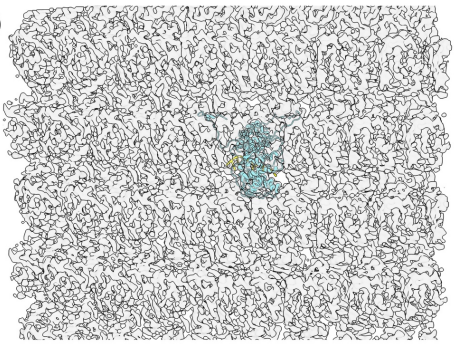
(B)





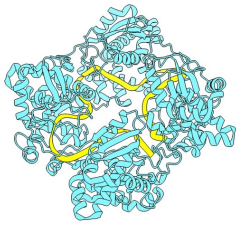
Human metapneumovirus

(C)



Vesicular stomatitis virus

(D)



Leanyer Orthobunyavirus

(E)

Influenza A virus

Figure 1.1: Structures of RNA-Bound Nucleocapsids and Nucleoproteins from the *Negarnaviricota*.

High-resolution structures of RNA-bound nucleocapsids (or higher order assemblies, as deposited on the Protein Data Bank (PDB)). (Left) The higher order oligomeric assembly deposited on the PDB is displayed. In the cases of panels A – C, this is accompanied by the cryo-EM density map as deposited on the Electron Microscopy Data Bank (EMDB) displayed in grey at a transparency of 30% with one copy of the nucleoprotein shown in blue with RNA bound displayed in yellow. In the case of panel D, this is accompanied by the tetrameric assembly of the nucleoprotein as solved by X-ray crystallography, which is thought to be physiologically relevant. In the case of panel E, no higher-order assembly is displayed, as the dimeric assembly deposited on the PDB is not thought to be physiologically relevant. (Right) One copy of the nucleoprotein coloured in blue is displayed with the RNA-binding groove facing to the right. The bound RNA is displayed in the RNA binding groove as atoms in the ladder form. (A) The helical structure of the measles virus nucleocapsid (Gutsche et al., 2015) (PDB: 4UFT; EMDB: 2867). (B) The decameric ring structure of the human metapneumovirus nucleocapsid (Whitehead et al., 2023) (PDB: 8PDL; EMDB: 17613). (C) The helical structure of the vesicular stomatitis virus nucleocapsid (Jenni et al., 2022) (PDB: 7UMK; EMDB: 26602). (D) The structure of the Leanyer Orthobunyavirus nucleoprotein in complex with ssRNA (Niu et al., 2013) (PDB: 4J1G). (E) The structure of the influenza A virus nucleoprotein in complex with three nucleotides of RNA (Tang et al., 2021) (PDB: 7DXP). This structure was deposited as a non-physiological dimer likely the result of crystal packing, and as such, is only displayed as a monomer.

In published literature, the packaged genome structure of the *Haploviricotina* is commonly referred to as a 'nucleocapsid', whereas the packaged genome structure of the *Polyploviricotina* is commonly referred to as a 'ribonucleoprotein' (RNP), or viral ribonucleoprotein (vRNP). This reflects the tightly packaged, monopartite packaging of the genome of non-segmented negative sense RNA viruses, versus the comparatively flexible and heterogeneous packaging of genome segments of segmented negative sense RNA viruses, as is observed when such genome segments are studied by electron microscopy (Modrego et al., 2023; Sabsay and te Velthuis, 2024) (figure 1.2).

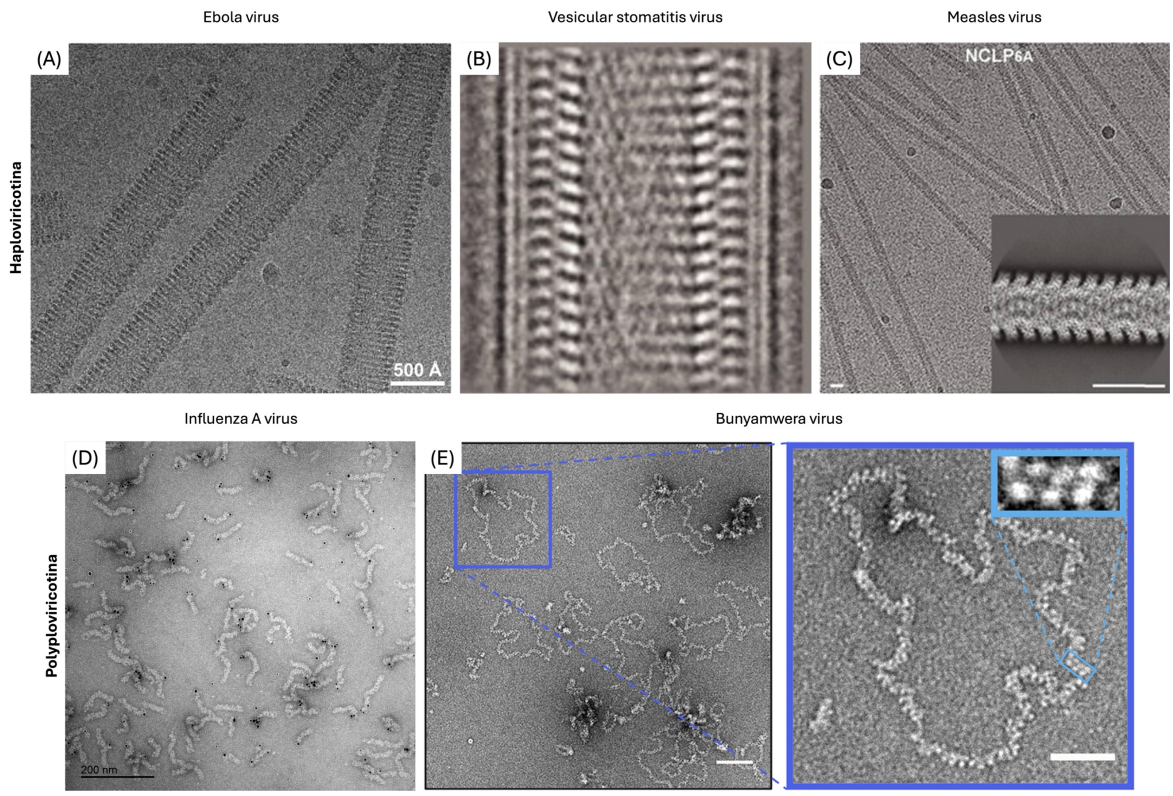


Figure 1.2: Electron microscopy images of negative sense RNA virus genome segments.

(A) Cryo-EM micrograph of recombinantly produced Ebola Virus nucleocapsid particles (Su et al., 2018). (B) Representative 2D class average of helical Vesicular Stomatitis Virus nucleocapsid within a viral particle (Ge et al., 2010). (C) Representative micrograph of Measles Virus nucleocapsid like structure (NCLP6A) with a representative 2D class average shown as an inset. Scale bar = 20 nm (Desfosses et al., 2019). (D) Representative negative stain electron microscopy micrograph of virion-derived influenza vRNPs with a nanogold particle (black) bound to influenza polymerase-targeting megabodies (Zhu, 2024). (E) Representative negative stain electron microscopy micrograph of Bunyamwera virus-derived RNPs with an area of apparent helical organisation shown as an inset. Scale bars equal to 100 nm in the main figures and 50 nm in the upper right inset (Hopkins et al., 2022). Note the clear helical organisation of non-segmented viral genome segments (A-C) compared to the more heterogeneous ultrastructure of segmented viral genome segments (D, E).

Much less is known of the genome segment structure of segmented negative sense RNA viruses due to this observed increased sample heterogeneity, with structures of nucleoproteins solved largely without RNA bound, and in low-order oligomeric and monomeric states.

Among the segmented negative sense RNA viruses, influenza A virus represents a promising model system to address gaps in our structural understanding of segmented negative sense RNA virus genome architecture and nucleoprotein – host factor interactions. Given its significant global health burden, its well-characterised molecular biology, wealth of resources available for the study of the virus, and the relative lack of high-resolution structural information on the native vRNP architecture and nucleoprotein – immune factor interactions, this thesis focuses specifically on the influenza A virus nucleoprotein.

An improved understanding of the structural basis of nucleoprotein function will provide valuable insights into influenza A virus genome organisation and the molecular mechanisms of viral immune evasion, while also establishing structural principles that may be broadly applicable to other segmented negative sense RNA viruses.

1.2. Influenza Virus

1.2.1. Nomenclature and Classification

Influenza viruses are a group of enveloped, pleomorphic viruses in the family *Orthomyxoviridae*, comprising of four specific genera: influenza A, B, C, and D. They are all segmented, negative sense single stranded RNA viruses (-ssRNA). Influenza A and B have eight genome segments, while influenza C and D have seven genome segments, which encode a variable number of gene products.

The difference in segment number is a result of influenza C and D viruses possessing only one membrane glycoprotein, hemagglutinin-esterase-fusion (HEF) (Li et al., 2021), whereas influenza A and B possess two membrane glycoproteins, hemagglutinin (HA), and neuraminidase (NA) (Long et al., 2019).

The different genera have varied natural hosts and host species ranges. Influenza A virus is the most diverse in this sense. Influenza A virus circulates in a waterfowl reservoir (infecting wild aquatic birds including but not limited to ducks, geese and other shorebirds) and can infect animals such as pigs, domestic and wild birds, horses, dogs, cats, seals, whales, and other mammals including humans.

Influenza A virus is the most diverse of the four influenza viruses and is the predominant causative agent of influenza in humans, being known to cause seasonal epidemics and occasional pandemics (Taubenberger and Kash, 2010).

Influenza B virus circulates mostly in humans, although infections of seals have additionally been reported. It has two main lineages, Victoria and Yamagata, and is known to cause seasonal epidemics in humans alongside influenza A virus.

Influenza C virus circulates in humans and pigs, and can infect cattle, pigs, sheep, and goats. It is generally known to cause milder respiratory illness compared to influenza A and B viruses.

Finally, influenza D virus circulates in cattle and pigs, and can infect cattle, pigs, sheep, and goats. There is however very limited serological evidence to suggest that influenza D virus infects humans (Bailey et al., 2018; Long et al., 2019; Javanian et al., 2021).

Given Influenza A virus's vast diversity and pandemic potential, it is the only genus of influenza to be extensively categorised based on its surface antigens (though influenza B virus is also classified based on its surface glycoproteins, but only into two lineages: Victoria and Yamagata). In influenza A virus, there are 18 different HA subtypes (H1 – H18) and 11 NA subtypes (N1 – N11), with the most commonly-circulating human strains of influenza currently being H1N1 and H3N2 (Long et al., 2019).

1.2.2. Epidemiology and Clinical Manifestations

All influenza viruses cause an infectious respiratory disease of varying severity depending on a number of factors including the patient's age and comorbidities, previous exposure to and immunological memory of specific influenza proteins, and adaptive mutations within the virus. Adaptive mutations include those which can increase viral replication (e.g. PB2 E627K, which facilitates more efficient influenza A replication in mammalian hosts) (Zhang et al., 2014; MacCosham et al., 2025), diminish the ability of the host to effectively restrict viral infection (e.g. MxA sensitivity mutations in influenza A nucleoprotein) (Mänz et al., 2013; Ashenberg et al., 2017), or facilitate the virus spreading systemically around the host organism by the addition of a multibasic cleavage site (Goot et al., 2003; Bailey et al., 2018; De Marco et al., 2023; Sacristán et al., 2024).

Influenza A and B viruses are the most relevant to human health, putting great financial and logistical burden on global health systems. The 1918 pandemic influenza alone causing a conservatively estimated 50 million deaths, which exceeded those as a direct result of the first world war (Johnson and Mueller, 2002; Spreeuwenberg et al., 2018).

Despite the advances of modern medicine, seasonal vaccination drives, and surveillance mechanisms, it is still estimated that there are around 1 billion cases of influenza per year, 3 – 5 million of which are severe, and 290,000 – 650,000 of these lead to death (Iuliano et al., 2018). All documented cases of pandemics caused by influenza have been a result of influenza A viruses (1918 (H1N1), 1957 (H2N2), 1968 (H3N2) and 2009 (H1N1)) (Monto and Fukuda, 2020).

The World Health Organisation (WHO) funds a Global Influenza Programme (GIP) and the Global Influenza Surveillance and Response System (GISRS) which acts to collect and analyse virological and epidemiological influenza surveillance data. Based on these data, they recommend specific compositions to vaccine manufacturers to ensure the best protection of vulnerable populations worldwide (https://www.who.int/teams/global-influenza-programme), with over 18 million influenza vaccines administered in the UK between September 2024 to February 2025 (https://www.england.nhs.uk/statistics/statistical-work-areas/flu-vaccinations/).

1.2.3. Evolution

Adaptation of the virus to the host is critical for successful influenza replication. In broad terms, influenza can undergo two key processes to evolve – antigenic shift, and antigenic drift (Kim et al., 2018).

Antigenic drift is the slow accumulation of adaptive mutations across both influenza surface proteins as a result of point mutations introduced by erroneous replication by the influenza polymerase. This is key for the virus to develop beneficial mutations which may aid in its adaptation to a new host. As a process across the entire viral genome, this can be referred to as genetic, or evolutionary drift.

Antigenic shift arises as a result of the segmented nature of the influenza genome. In a co-infection event – where multiple influenza A viruses may infect a given host cell – entire genome segments can be exchanged between viruses. Antigenic shift between two virus strains in a co-infection event can result in 4 (2^2) different possible combinations of HA and NA. Across the whole genome, such reassortment events – provided they result in a progeny virus with all eight segments of the influenza A viral genome – can lead to 256 (2^8) different potential genotypes which could be produced from a theoretical co-infection with two influenza A viral strains, yielding up to a theoretical 254 ($2^8 - 2$) recombinant genotypes.

The implications of genetic reassortment are drastic when compared to genetic drift alone. Reassortment has dire implications for the potential of influenza A viruses to

undergo host-switching and rapidly accelerated evolution. Given such a novel progeny virion, there may be little to no recognition of reassorted antigens, which gives the resulting virus population a significant advantage in terms of host immune evasion.

1.2.4. Structure and Components of Infectious Virions

The influenza A virion is enveloped and pleomorphic, displaying multiple morphologies including spherical (with diameter of around 100 nm) and filamentous (with some particles being more than 20 μm in length). Notably, filamentous influenza A virus (IAV) particles are consistently found in human and animal isolates (Neumann et al., 2009; Nakajima et al., 2010; Arai et al., 2019), however spherical particles are found to be more prevalent in lab-adapted strains, where serial passaging of clinical isolates either in embryonated chicken eggs or cell culture leads to the loss of filamentous morphology (Peterl et al., 2025). Each infectious influenza A virion has eight viral ribonucleoprotein (vRNP) segments which encode for one or more proteins, described below (table 1.3).

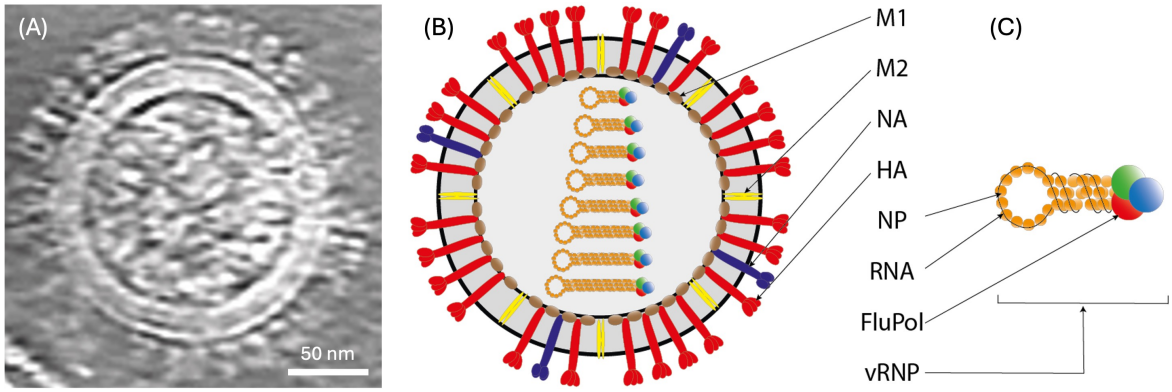


Figure 1.3: Structure of the Influenza A Virion.

(A) Cryo-ET tomogram slice of an influenza A virus particle (A/WSN/1933 (H1N1)). (B) Cartoon diagram of an influenza A virus particle with key proteins annotated. (C) Cartoon diagram of an influenza A virus viral ribonucleoprotein, with key constituent proteins and RNA labelled. M1; matrix protein, M2; matrix-2 protein (ion channel), NA; neuraminidase, HA; hemagglutinin, NP; nucleoprotein, RNA; ribonucleic acid, FluPol; Influenza Polymerase complex (made of the PB1; polymerase-basic 1, PB2; polymerase-basic 2, and PA; polymerase acidic subunits), vRNP; viral ribonucleoprotein.

Table 1.3: Genome Segments of the Influenza A Virus.

A table detailing the different genome segments, their length in nucleotides, the protein products from each segment, the protein product length in amino acids and approximate molecular weight in kDa, and the known functions of the protein products (Bouvier and Palese, 2008).

Segment	Length (nt)	Major Protein Product	Length (aa)	Approximate Molecular Weight (kDa)	Known Function(s)
1	2341	PB2	759	83.5	Viral polymerase subunit – cap snatching
2	2341	PB1	757	83.3	Viral polymerase subunit-catalytic core
		PB1-F2	87	9.6	Virulence factor and immune modulator
3	2233	PA	716	78.8	Viral polymerase subunit – endonuclease
		PA-X	41 / 61	4.5 / 6.7	Endonuclease in host genome shut-off factor
4	1175	HA	566	62.3	Surface glycoprotein receptor in binding and entry. Virulence and immunogenicity factor
5	1565	NP	498	54.8	Viral RNA-binding protein in the vRNP
6	1409	NA	454	49.9	Surface glycoprotein key in virion budding and spread. Cleaves sialic acid moieties from mucins. Immunogenicity factor
7	1027	M1	252	27.7	Structural protein that forms the viral capsid
		M2	97	10.7	Structural protein that forms an ion channel key for viral uncoating and the release of vRNPs in the host cell. Yet uncharacterised role in vRNP export
8	809	NS1	230	25.3	Multifunctional modulator of the host immune response
		NEP	121	13.3	Mediator of vRNP export and RNA synthesis

1.2.5. Viral Lifecycle

Infection and Internalisation

Influenza a virus is able to infect cells of the upper respiratory tract (URT) in humans. This is due to the affinity of HA for sialyated proteins: that is, proteins modified with a sialic acid (SA) moiety, in the respiratory tract. Avian influenza viruses have evolved a greater affinity for α-2,3 sialyated surface proteins (which are only present in the human lower respiratory tract (LRT)). A hallmark of a human-adapted influenza A virus are mutations to the HA which enable the virus to bind to α-2,6 sialyated surface proteins, which occur in the URT (Kumlin et al., 2008; Gamblin et al., 2021; Zhao and Pu, 2022).

Once the virus has bound to the host cell, it is internalised via receptor-mediated endocytosis into an early endosome. The endosome is gradually acidified by an endosomal membrane proton V-ATPase, in the canonical endocytic pathway (Huotari and Helenius, 2011; Scott et al., 2014; Hu et al., 2015). The endosome moves along microtubule tracks towards the perinuclear region. The drop in pH leads to a conformational change in HA, exposing a hydrophobic fusion peptide which is used to insert into the endosomal membrane, either as a transmembrane helix or a peripheral helical hairpin (Lorieau et al., 2010; Lousa and Soares, 2021) leading to the fusion of the endosomal and viral membranes. Concomitantly, the viral M2 ion channel facilitates the acidification of the virion. This leads to the disruption of the M1 matrix layer (Fontana and Steven, 2013) in preparation for release of the vRNPs into the cytoplasm.

Nuclear Import

Once in the cytosol, host importins – specifically importins α and β1 (Gabriel et al., 2011; Chou et al., 2013) – recognise the nuclear localisation signals (NLSs) on the viral nucleoprotein and transport the vRNPs into the host nucleus (Bouvier and Palese, 2008).

Transcription and Replication

Once the vRNPs have been imported into the host cell nucleus, they undergo both transcription, into mRNAs, and a two-stage replication process; first into positive-sense complementary RNA (cRNA) packaged by nucleoprotein and influenza polymerase into a complementary ribonucleoprotein (cRNP), whose sole function as a nuclear-resident ribonucleoprotein complex is as a replicative intermediate for the replication of more viral RNA (vRNA). Similarly, nascent vRNA is packaged by nucleoprotein and influenza polymerase into nascent vRNPs, which are then exported to the cytoplasm for progeny virus particle production.

Viral transcription occurs in early infection. The virus produces 5' 7-methyl-guanosine (m7G) capped and 3' polyadenylated (polyA) mRNA to replicate the mRNA production in eukaryotes so as not to trigger the host immune response. It does this by the well-understood 'cap snatching' mechanism to obtain an m7G cap (Plotch et al., 1981; Reich et al., 2014) and produces a poly(A) tail by a polymerase stuttering mechanism on a poly-U tail of vRNA close to the 5' terminus (Poon et al., 1999; Walker and Fodor, 2019). Nascent viral mRNA transcripts are then exported out of the nucleus and bound

by poly(A) binding protein 1, which recruits the mRNA transcript to the ribosome (de Rozières et al., 2022). Viral proteins required for v/cRNP production are then reimported into the nucleus via their NLSs.

Viral replication occurs in two distinct phases – 'primary replication' which refers to the synthesis of cRNA from vRNA, and 'secondary replication' which refers to the synthesis of vRNA from cRNA. cRNA is a copy of vRNA without an m7G cap or poly(A) tail. In primary replication, two influenza polymerases are bridged together by one molecule of host acidic nuclear phosphoprotein 32 (ANP32) (Carrique et al., 2020) (figure 1.4), with one polymerase complex replicating the viral genome into cRNA, and the other encapsidating the nascent cRNA into a cRNP. In secondary replication, a similar process happens in reverse, with a slightly different priming procedure and specific polymerase conformations (York et al., 2013, p.201; Fangzheng Wang et al., 2022; Zhu et al., 2023).

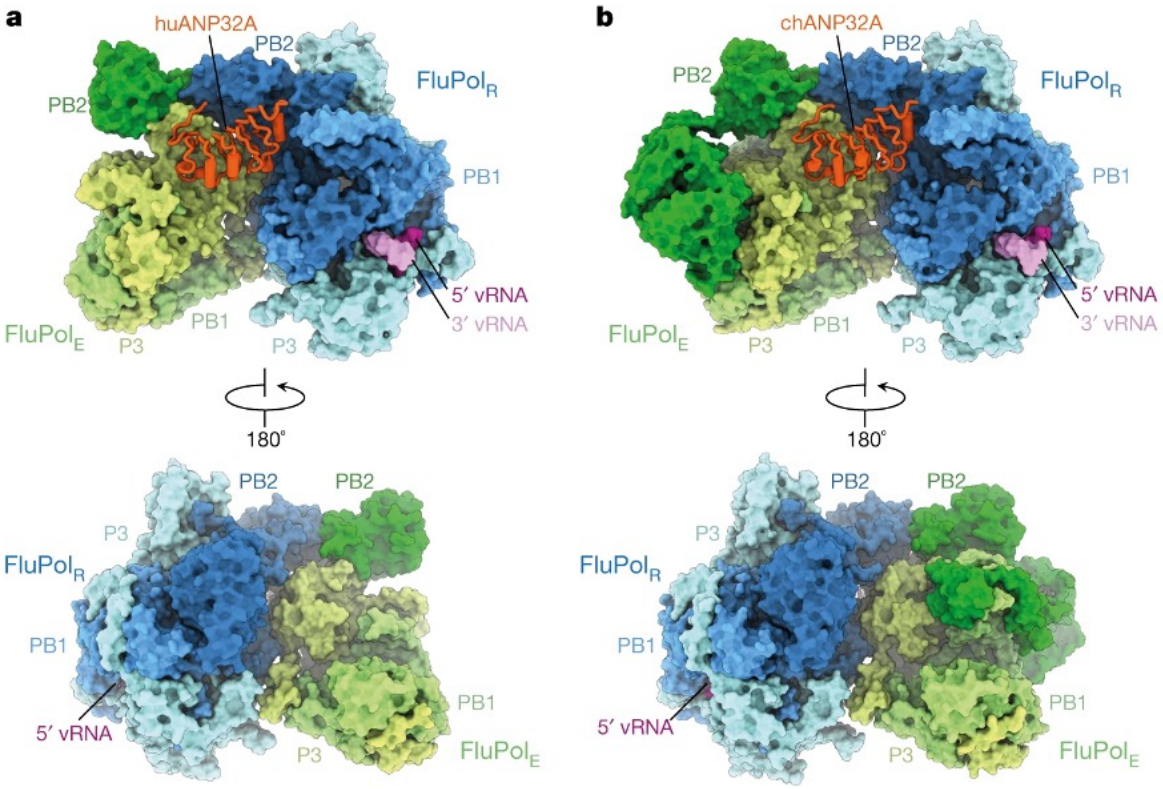


Figure 1.4: Host ANP32A Mediates the Assembly of the Influenza Virus Replicase.

Structures solved by cryo-electron microscopy (cryo-EM) single particle analysis (SPA) displaying influenza C polymerase heterotrimers with bound human ANP32A (huANP32A) (a) and chicken ANP32A (chANP32A) (b). ANP32A is displayed in cartoon form in orange in both panels. A dimer of influenza C polymerase (FluPol) is displayed as spherical atoms. The two distinct conformers of influenza polymerase displayed are the encapsidating (FluPol$_E$, shades of green) and the replicating (FluPol$_R$, shades of blue) conformer. Bound vRNA is additionally displayed in purple. P3 is the influenza C virus equivalent of influenza A virus PA. Figure taken from (Carrique et al., 2020).

Nuclear Export and vRNP Trafficking

After sufficient replication has occurred, progeny vRNPs, but not cRNPs, are exported from the host cell nucleus via the chromosomal region maintenance protein 1 (CRM1) pathway (Huang et al., 2013) likely via nuclear export protein (NEP) binding both M1 and the influenza polymerase (Akarsu et al., 2003; Shimizu et al., 2011; Brunotte et al., 2014). Interestingly, the same site on the influenza polymerase is occupied by both ANP32 and NEP (Alison Rep, personal communication), which could suggest a potential competition mechanism which is dependent on local protein concentration in the nucleus as a function of the time post infection, whereby, with more NEP present

in the host nucleus, NEP out-competes ANP32 for binding of the influenza polymerase, leading to selective vRNP export at later timepoints. NEP mRNA is derived from inefficient splicing of NS1 mRNA and has thus been described as a 'molecular clock', as NEP therefore accumulates later in infection, supporting this mechanism of a temporal switch to vRNP export.

Once progeny vRNPs have been exported into the cytoplasm, they are transported to the plasma membrane by hijacked host machinery. Y-Box binding protein 1 (YB-1) is involved in the regulation of cellular transcription and translation, and in influenza A infection was found to be required for the interaction of vRNPs with microtubules around the microtubule organising centre (MTOC), from which vRNPs enter the vesicular trafficking system (Kawaguchi et al., 2012). The human Ras-related protein Rab11a is a member of the small GTPase family of proteins which are required for the subsequent trafficking of vRNPs towards the plasma membrane through recycling endosomes (Amorim et al., 2011; Hutchinson and Fodor, 2013; Amorim, 2019). More recently, Rab11a has also been implicated in binding the PB2 subunit of the influenza polymerase complex (Veler et al., 2022). *In situ* cryo-electron tomography work has also been carried out to unpack the molecular basis of this process, in which it has been shown that vRNPs additionally interact with HA-coated membranes in a Rab11a-dependent manner (Dr. Alasdair Hood, personal communication) (Wachsmuth-Melm et al., 2025).

Assembly of the Influenza A Virus Genome

Each viral particle requires all eight viral genome segments in order to carry out a successful infection. This suggests that there is a highly orchestrated mechanism by which influenza virus assembles each of its constituent vRNPs for loading into a progeny virus particle. The best understood mechanism of this occurring is by inter-vRNP segment RNA-RNA interactions guiding the formation of a canonical "7+1" arrangement of vRNPs in the newly forming viral particle (Dadonaite et al., 2019). This 7+1 arrangement has been observed by multiple groups using electron microscopy techniques (Noda et al., 2006; Fournier et al., 2012; Lakdawala et al., 2014). The M1 layer polymerises around this vRNP core (Peukes et al., 2020), which is likely the key determinant of influenza virion morphology (Peterl et al., 2025), and aids in the budding process giving shape to the nascent viral particle alongside HA and NA (Jin et al., 1997; Chlanda et al., 2015).

During and after membrane scission, the virally-produced sialidase NA – which is distributed throughout the viral membrane but found to be concentrated at the tip of progeny virus particles – cleaves sialic acid linkages between host cell surface receptors and viral HA glycoproteins. This prevents re-attachment of newly budded virions to the host cell surface, facilitating the release of the nascent virions into the extracellular environment. NA continues to function after budding, by preventing viral aggregation (Palese et al., 1974; Morris et al., 1999; McAuley et al., 2019) and aiding viral movement through sialic acid-rich environments such as mucous found in the respiratory tract (Yang et al., 2014; McAuley et al., 2019), promoting viral spread to new target cells (Bouvier and Palese, 2008; Chlanda et al., 2015).

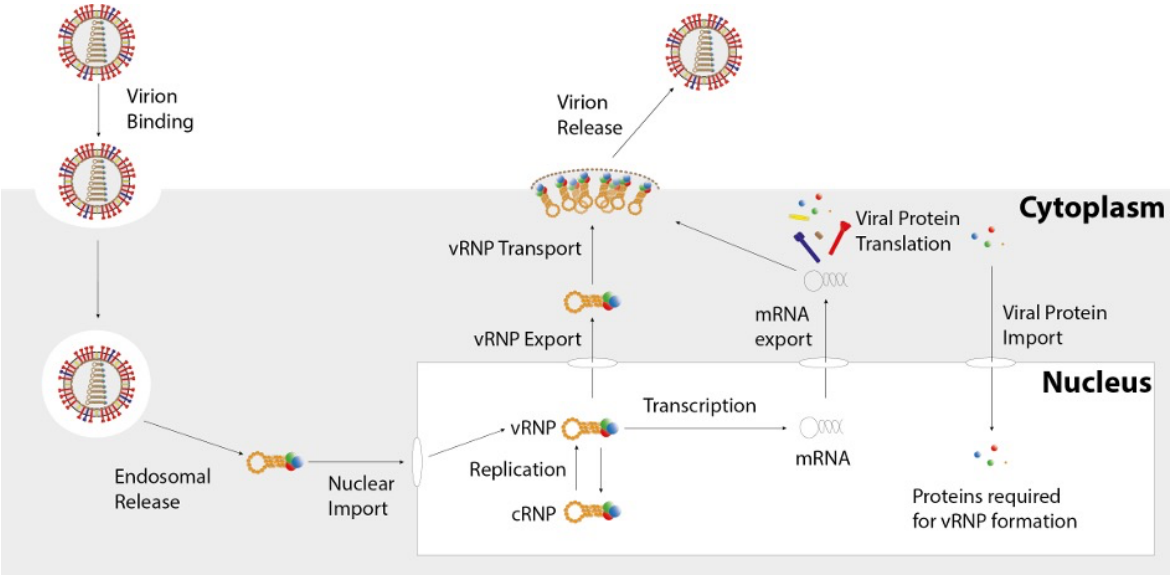


Figure 1.5: Overview of the Influenza A Virus Lifecycle.

Virion Binding: HA surface proteins bind to sialic acid (SA) moieties on host cell surface proteins. The virus then fuses with the host cell membrane and is taken up into an endosome. Endosomal Release: vRNPs are released into the cytosol upon M2-mediated acidification of the viral core. Nuclear Import: vRNPs are imported into the nucleus through nuclear pore complexes (NPCs). Transcription: vRNPs associate with host RNA Polymerase II (RNAPII). mRNA Export: mRNA is exported into the cytoplasm for viral protein translation. Viral Protein Translation: mRNA transcripts of viral proteins direct the translation of all viral proteins necessary for replication. Proteins Viral Protein Import: Proteins involved in vRNP production (PB1, PB2, PA, NP) are re-imported into the nucleus alongside other viral proteins required for nuclear export. Replication: Positive sense complementary ribonucleoproteins (cRNPs) are produced as a replicative intermediate for de novo vRNP production. vRNP Export: Progeny vRNPs are exported to the cytoplasm. vRNP Transport: vRNPs are transported to the plasma membrane where they assemble new virion particles with other viral proteins which were translated earlier. Virion Release: Progeny influenza virions assemble at the plasma membrane, curve by the M1 protein forming a rounded core upon oligomerisation, and bud, with NA cleaving SA moieties to prevent virion reattachment.

1.3. Components of Influenza Virus Replication

1.3.1. Influenza Virus Polymerase

The influenza virus RNA-dependent RNA polymerase complex has three constituent protein subunits. In order of increasing size, these are PA, including the viral endonuclease domain, approximately 78.8 kDa; PB1, the catalytic core of the polymerase approximately 83.3 kDa; and PB2, which includes the cap snatching domain of the influenza polymerase, approximately 83.5 kDa. Together they form an

approximately 250 kDa heterotrimeric complex, which has approximate dimensions of 115 x 100 x 75 Å (Wandzik et al., 2021) (figure 1.6). The overall structure of the complex is a globular core which contains the catalytic domains of the polymerase, and two protruding domains, the PA endonuclease domain and the PB2 cap-binding domain. The complex adopts multiple conformations as it must carry out diverse functions. Over the past eleven years, structures of influenza polymerases from all four influenza genera: A (Reich et al., 2014) (PDB: 4WSB); B (Reich et al., 2014) (PDB: 4WRT); C (Hengrung et al., 2015) (PDB: 5D98); and D (Peng et al., 2019) (PDB: 6KUJ) have been solved, initially solely by X-ray crystallography, but since 2019, also routinely by cryo-electron microscopy to comparable resolution (sub-4 Å). There are multiple states of influenza polymerase which have been solved, including apo, pre-initiation, transcriptase, encapsidase, elongation, recycling, and promoter bound.

These structures have elucidated a mechanism for vRNA binding to the polymerase. Both terminal ends of the vRNA bind to structurally conserved regions of the polymerase. The 5' end is tightly bound to the polymerase, and forms a characteristic hook formed by the PB1 and PA (P3 in influenza C virus) polymerase domains, whereas the 3' end occupies two different binding sites depending on the conformational state of the polymerase – either active in transcription or replication, or in the resting state. (Pflug et al., 2014; Fan et al., 2019; Carrique et al., 2020).

Table 1.4: Annotation of Select Previously-Solved Influenza Polymerase Structures.

This table is used for the annotation of the below figure, with PDB codes of the corresponding structures, the strain of influenza they originate from, the reported resolution they are solved to, the method used to solve the structure (X-ray crystallography (X-ray) or cryo-electron microscopy (cryo-EM)), the conformation they are resolved in, and the reference paper the structure was published in.

PDB Code	Panel in Figure 1.6	Strain	Resolution (Å)	Method	Conformation	Reference
SD98	B	C/Johannesburg/1/1966	3.90	X-ray	Apo	(Hengrung et al., 2015)
6QCS	C	B/Memphis/13/2003	3.10	Cryo-EM	Pre-Initiation	(Kouba et al., 2019)
6F5P	D	C/Johannesburg/1/1966	4.14	X-ray	Transcriptase	(Martin et al., 2018)
8RN1	E	B/Memphis/13/2003	3.64	Cryo-EM	Encapsidase	(Arragain et al., 2024)
6T0V	F	A/Little Yellow-Shouldered Bat/060/2010 (H17N10)	3.02	Cryo-EM	Elongation Complex	(Wandzik et al., 2021)

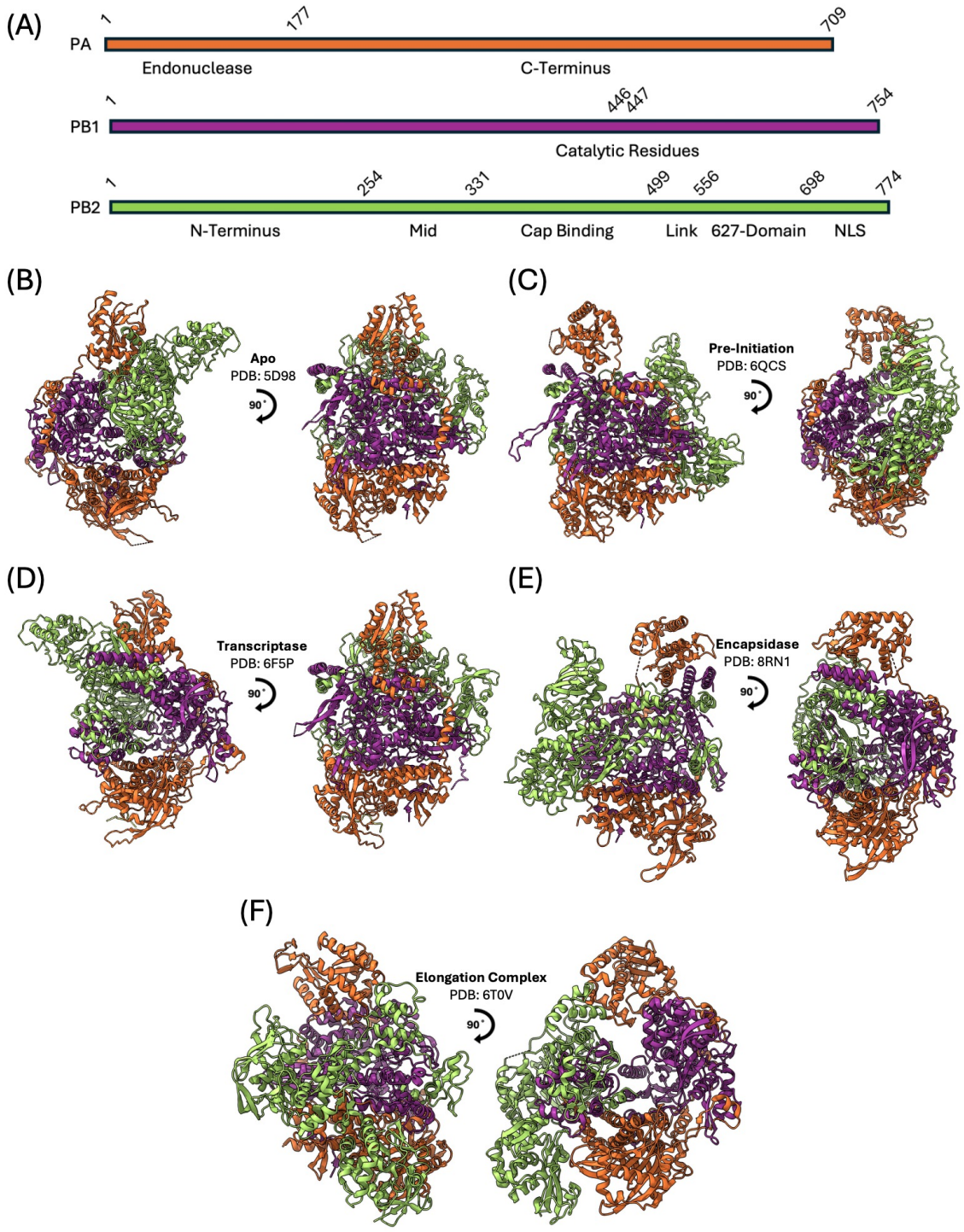


Figure 1.6: Structural Overview of the Influenza Polymerase.

The heterotrimeric influenza polymerase complex displayed as a linear peptide for each constituent protein, with domain boundaries taken from a recent review (Zhu et al., 2023) (A). Throughout the figure, PA is displayed in orange, PB1 in purple, and PB2 in green. (B) Influenza C apo conformation (Hengrung et al., 2015) (PDB: 5D98), (C) Influenza B pre-initiation conformation (Kouba et al., 2019) (PDB: 6QCS), (D) Influenza C transcriptase conformation (Martin et al., 2018) (PDB: 6F5P), (E) Influenza B encapsidase conformation (Arragain et al., 2024) (PDB: 8RN1), (F) Influenza A elongation complex conformation (Wandzik et al., 2021) (PDB: 6T0V).

1.3.2. Influenza Virus Nucleoprotein

Influenza nucleoprotein packages the influenza genome into the vRNP, facilitating the transport of the genome into and out of the cell. Each individual nucleoprotein monomer is estimated to bind to 24 nucleotides of RNA (Ortega et al., 2000; Ye et al., 2006). With this estimate, the longest vRNP, segment 1, encoding for PB2, is expected to contain 97 nucleoprotein molecules, and the shortest, segment 8, encoding for NS1, is expected to contain 34 nucleoprotein molecules. RNA binding occurs in a degenerate manner and is facilitated by a broad positive patch in the RNA binding groove, formed by the so-called G1 and G2 loops (Ng et al., 2008; Fangzheng Wang et al., 2022).

The nucleoprotein readily undergoes low-order oligomerisation (from monomers to dimers, trimers, and tetramers) - even without the addition of RNA - by insertion of a flexible tail loop into neighbouring protomers (Ye et al., 2006). Residues R416 of the tail loop and E339 within the body of the nucleoprotein have been demonstrated to be key determinants of oligomerisation, where when R416 is mutated to alanine, the propensity of oligomers forming is much reduced. This phenomenon is further observed when residue E339 is also mutated to alanine (Ye et al., 2006). Upon the addition of RNA, the wild-type nucleoprotein undergoes higher order oligomerisation (Labaronne et al., 2016), and even pseudo-vRNP formation given specific N-terminal truncations (Chenavier et al., 2025).

There are multiple binding partners of the nucleoprotein. First, the polymerase (Peng et al., 2025), whose binding is likely mediated both through the association of vRNA with both the nucleoprotein and the polymerase, but also via the PB2 subunit (Poole et al., 2004). High-resolution structures of such interactions remain to be elucidated. The nucleoprotein is also known to interact with the C-terminal low complexity acidic region (LCAR) of ANP32 (Wang et al., 2022) via the G1 and G2 domains of the protein (figure 1.7, panels A, B), which is suggested to be a mode of nucleoprotein recruitment to the active polymerase, whereupon binding of RNA outcompetes the LCAR binding to the G1 and G2 domains. Various studies have described a functional association of the influenza nucleoprotein with the influenza polymerase complex in an RNA-dependent manner through co-purification, co-immuno precipitation, and SPR studies. Specifically, the PB2 '627 domain' and the arginine-rich cleft formed by the G1 and G2 loops of the nucleoprotein (Ng, Chan, et al., 2012; Turrell et al., 2013; Szeto et al., 2020).

The nucleoprotein, being the most abundant protein in the vRNP, if not the virus as a whole, is also the target of immune factors, including MxA and BTN3A3, introduced below, but also of small molecule antivirals (Kao et al., 2010).

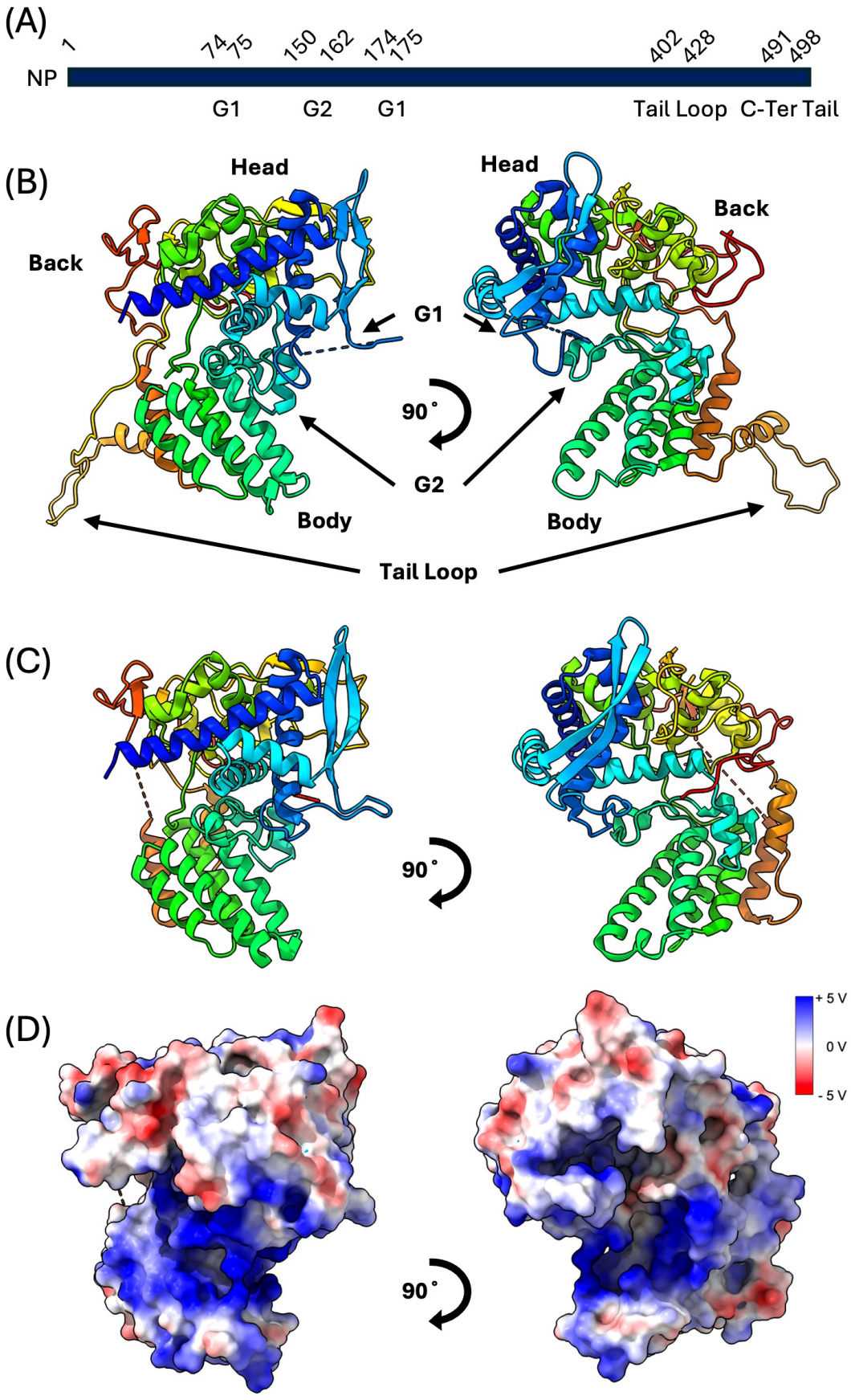

(A)
1
74
75
150
162
174
175
402
428
491
498
NP
G1
G2
G1
Tail Loop
C-Ter Tail
(B)
Head
Back
G1
90˚
G2
Head
Back
Body
Body
Tail Loop
(C)
90˚
(D)
90˚
+5 V
0 V
-5 V

Figure 1.7: Structural Overview of the Influenza Nucleoprotein.

(A) Key regions and residues are annotated onto a linear representation of the influenza nucleoprotein. Domain boundaries are taken from recent work (Wang et al., 2022). (B) A crystal structure of an influenza A nucleoprotein (wild type) (Ng et al., 2008) (PDB: 2Q06) coloured in chainbow, with the G1, G2, and tail loops annotated. Additionally annotated are the 'body' (N-terminal), 'head' (C-terminal), and 'back' regions of the nucleoprotein commonly referred to in literature. (C) A crystal structure of an influenza A nucleoprotein (R416A, monomeric mutant) (Knight et al., 2021) (PDB: 7NT8) coloured in chainbow. (D) The above nucleoprotein structure coloured by surface electrostatics with a spectrum of red-to-white-to-blue. The central RNA binding groove of the nucleoprotein is highly positively charged.

1.3.3. The Viral Ribonucleoprotein Complex (vRNP)

The nucleoprotein forms a homo-oligomeric double helix in the context of the vRNP which has been characterised at the ultrastructural level by both negative stain electron microscopy and cryo-electron microscopy. Unlike other RNA viruses (e.g. human metapneumovirus (Whitehead et al., 2023), the oligomeric form of the nucleoprotein is highly flexible, and increased heterogeneity arises from the differing lengths of the vRNPs according to segment and vRNA length, which makes high-resolution structural determination of vRNPs challenging using conventional structural biology techniques.

In terms of structure, the vRNP contains varying numbers of nucleoprotein molecules according to the length of the specific vRNA molecule – as stated above, this is from approximately 34 copies of the nucleoprotein in the shortest segment, segment 8, which encodes for NS1, to approximately 97 copies of the nucleoprotein in the longest segment, segment 1, encoding for PB2. In the resting state of the vRNP, that which is present in virions and in cytoplasmic trafficking, the 5' and 3' ends of the vRNA are thought to be held in place by the polymerase. Low-resolution structural information confirms the existence of the polymerase predominantly at one end of the vRNP in its

resting state (Arranz et al., 2012; Moeller et al., 2012), especially by localising the polymerase using recombinant labelling with a nanogold tag (Moeller et al., 2012) or immunolabelling with a nanogold-tagged megabody (Zhu, 2024). Newer, higher resolution information from reconstituted influenza D virus RNPs has also confirmed the existence of the polymerase in the middle of the RNP, suggesting a mechanism by which the polymerase translocates down the RNP during transcription (Peng et al., 2025) (figure 1.8, panels B and C).

Various studies have been carried out which have reported on the oligomeric state of recombinantly-expressed nucleoproteins, where the purified nucleoprotein was incubated with RNAs of varying length for the production of low-order oligomers or vRNP-like complexes of influenza nucleoprotein (Ruigrok and Baudin, 1995; Ye et al., 2006; Ng, Lam, et al., 2012) or the closely related *Orthomyxoviridae* member, thogotovirus (THOV) (Dick et al., 2024). In the last year, great progress has been made in determining the high-resolution structure of one such pseudo-vRNP, where a 14-mer N-terminal deletion of the nucleoprotein, and incubation with an 18-mer RNA were found to form an RNP-like assembly with the canonical right-handed, antiparallel double helix of nucleoprotein, including major and minor grooves (Chenavier et al., 2023).

High speed atomic force microscopy (hsAFM) has also been used to study the ultrastructure of the vRNP (Nakano et al., 2021; Chenavier et al., 2023), with this technique being especially useful for low-resolution structural data on the native state of vRNPs and conformational dynamics during transcription *in vitro*. The handedness

of RNPs can also be determined using AFM, as the sample is observed directly. This information is ambiguous in cryo-EM until an atomic resolution structure has been solved, as depth information is lost when a 2D projection image is taken from a 3D object. By both modern AFM and tomographic methods, the vRNP has been confirmed to be a right-handed helix (Nakano et al., 2021; Misha Le Claire, personal communication).

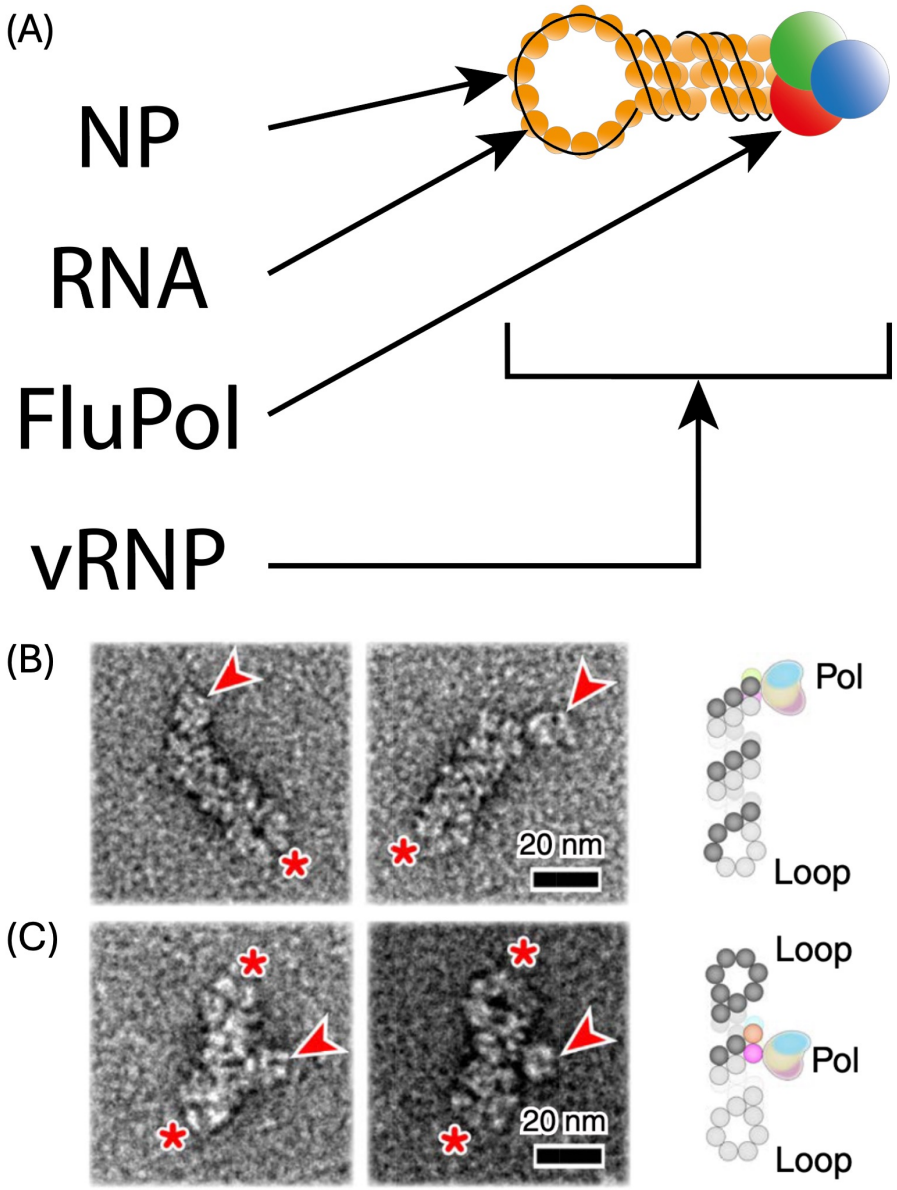


Figure 1.8: vRNP Ultrastructure.

(A) Cartoon diagram of the vRNP, with the nucleoprotein (NP), viral RNA (RNA), and influenza polymerase (FluPol) labelled. (B) and (C) Negative stain EM images (left, centre) of example vRNPs, and cartoon models (right) showing the relative positions of nucleoprotein and polymerase, with the position of the polymerase annotated by an arrow and the opposite loop end(s) annotated with a star. (B) The influenza polymerase is often found located at one end of the vRNP, with the opposite end containing a loop of nucleoprotein molecules. (C) In this study, vRNPs were also found with the polymerase located mid-way across the vRNP. Panels (B) and (C) are adapted form Peng et al., 2025.

1.4. The Human Innate Immune Response to Viral Infection

1.4.1. General Overview of Human Innate Antiviral Immune Responses

Successful influenza infection requires the virus to overcome multiple layers of host immune defenses, beginning with the innate immune system's rapid detection and response mechanisms. The innate immune system employs a rigorous surveillance network centred around Pattern Recognition Receptors (PRRs), which serve as the first line of targeted molecular defense against viral infection.

The host deploys three major classes of PRRs to detect viral infection; each poised to monitor different cellular compartments and recognise distinct viral signatures. Toll-Like Receptors (TLRs) are employed both on the surface of the host cell and in the endosomal compartments of the host cell, detecting viral RNA through characteristic features such as extensive RNA secondary structures, unmethylated CpG motifs, or non-self-proteins present in viruses (Koyama et al., 2008; Takeuchi and Akira, 2009; Thompson et al., 2011).

In the cytosol, RIG-I Like Receptors (RLRs) function as RNA helicases that specifically recognise foreign RNA signatures, particularly the 5'-triphosphate structures found on viral RNAs – a molecular signature absent from host RNAs, which instead bear 5' 7-methylguanosine (m7G) caps (Koyama et al., 2008; Takeuchi and Akira, 2008; Thompson et al., 2011).

Complementing these RNA-sensing mechanisms, nucleotide-binding oligomerisation domain (NOD)-Like Receptors (NLRs) monitor for both direct viral components and indirect indicators of cellular damage, such as ionic concentration perturbations, ultimately triggering inflammasome formation and subsequent inflammatory responses (Aoshi et al., 2011).

Despite the diverse target species repertoires and cellular locations occupied, all PRR classes converge on a common signalling paradigm: oligomerisation-dependent activation of downstream signalling cascades. Oligomerisation of the initial PRR triggers the downstream activation of key transcription factors, including Nuclear Factor Kappa-light-chain-enhancer of activated B cells (NF-κB), Interferon Regulatory Factor 3 (IRF3), and Activator Protein-1 (AP-1). The convergence of these three pathways ensures robust signal amplification in response to molecular signatures of viral infection, and the activation of the adaptive immune response.

The result of PRR signalling is the production of type I interferons (IFN-α and IFN-β), which establish an antiviral state in the host. The interferon response leads to the

upregulation of interferon-stimulated genes (ISGs), creating a comprehensive antiviral state that targets multiple stages of the viral lifecycle including viral entry, nuclear import, protein translation, genome replication, and virion assembly (Schoggins, 2019).

Pattern Recognition Receptors (PRRs) make up one of the first molecular barriers to infection. Their specific targets are diverse, but as a general rule, their downstream signalling cascades result in a similar set of transcription factor protein complexes being activated, such as NF-κB, IRF3, and AP-1. This signalling cascade and transcription factor activation leads to the upregulation of other innate antiviral immune proteins, and the triggering the adaptive immune system.

1.4.2. Specific Innate Immune Responses against Influenza Virus

The innate immune response poses a significant barrier to influenza virus infection and represents a major evolutionary constraint limiting cross-species transmission. Along with species-specific differences in cellular receptors, tissue tropism, and adaptive immune recognition, innate immunity creates a selective pressure which drives influenza virus evolution. Only viral variants capable of evading, supressing, or overwhelming these host defense mechanisms can establish productive infections in new host species.

Retinoic Acid-Inducible Gene I (RIG-I) is one such specific innate immune factor active against influenza virus. RIG-I recognises non-self RNA structures in the host cytoplasm, recognising short dsRNA, blunt-end dsRNA, and ssRNA with 5'-triphosphate structures. Additionally, the panhandle structure of the viral promoter is known to be sufficient to activate RIG-I signalling. Binding to a non-self RNA triggers conformational changes which facilitate interactions with downstream signalling molecules, and subsequent type I interferon and antiviral gene production (Rehwinkel et al., 2010; Liu et al., 2015; Rehwinkel and Gack, 2020).

The E3 ubiquitin ligase, TRIM25, is known to recognise and bind to viral RNA in the cytoplasm and subsequently ubiquitinate RIG-I to facilitate type I interferon production via the RIG-I pathway (Álvarez et al., 2024). It is also understood that influenza NS1 (and also various SARS-CoV-2 proteins) binds to and suppresses RIG-I ubiquitination, to downregulate the RIG-I pathway (Gack et al., 2009; Koliopoulos et al., 2018). TRIM25 additionally targets vRNPs in the cell nucleus to inhibit RNA chain elongation independent of its ubiquitin ligase activity (Meyerson et al., 2017).

Interferon-Induced Transmembrane Protein 3 (IFITM3) is another potent innate immune factor active against influenza virus infection in cells. IFITM3 associates with cell membranes and alters membrane curvature, lipid composition, and membrane fluidity. This has been shown to increase the energy barrier for fusion pore formation, which leads to viral degradation in lysosomes rather than release of the viral genome into the cytoplasm. It has been suggested that deficiencies in IFITM3, which are

prevalent in some populations, provide viruses including influenza an opportunity to enter and more efficiently adapt to humans (Klein et al., 2023; Denz et al., 2024).

1.5. Human Myxovirus Resistance Protein 1 (MxA)

1.5.1. Human Myxovirus Resistance Protein Family

One family of interferon-stimulated genes active in the antiviral response are the Myxovirus Resistance Protein family (Mx). The Mx family of proteins are found in diverse lifeforms, across animals, fungi, plants, and other eukaryotic lineages (Langley et al., 2025). Their expression is not constitutive but tightly regulated by type I and type III interferon signalling (Holzinger et al., 2007; Haller and Kochs, 2011; Matzinger et al., 2013). This may be because cells become sensitised to apoptotic stimuli upon MxA expression (Mibayashi et al., 2002).

Most mammals have two Mx protein family genes, MX1, encoding the antiviral protein MxA, and MX2, encoding the antiviral protein MxB. Birds, however, only have one gene encoding an Mx family protein, whereas fish have up to seven paralogs (Boudinot et al., 2016; Langley et al., 2025).

Across all species, Mx proteins demonstrate potent antiviral activity against a broad spectrum of viruses. In humans, MxA was first found to restrict influenza A virus (Zürcher et al., 1992; Matzinger et al., 2013; Haller et al., 2015), as inbred mice were observed to survive an otherwise lethal dose of influenza A virus (Lindenmann, 1962; Lindenmann, 2005). Further work has confirmed that it is indeed MxA which is

responsible for this phenotype, as human MxA was found to protect mice which lacked functional interferon systems (Hefti et al., 1999). Since this first observation, MxA has been found to restrict infection by viruses including vesicular stomatitis virus (Zürcher et al., 1992), measles virus (Schneider-Schaulies et al., 1994), La Crosse Virus (Hefti et al., 1999), and hepatitis B virus (Li et al., 2012) among others. The human MxA paralog MxB, however, restricts retroviruses, namely HIV (Goujon et al., 2013; Kane et al., 2013).

Despite conserved antiviral activity across species, MxA demonstrates species-specific localisation. In humans and swine, MxA is localised in the cytoplasm, whereas the murine Mx1 homolog is localised to the nucleus (Kolb et al., 1992; Zürcher et al., 1992). The differential localisation of MxA and its homologues is important in antiviral action, as it was shown that murine Mx1 was able to inhibit different steps in the influenza virus replication cycle (due to its nuclear localisation) than human MxA (Pavlovic et al., 1992).

1.5.2. Structure and Function of MxA

The Mx family proteins, MxA and MxB, are members of the dynamin superfamily of proteins (Ramachandran and Schmid, 2018; Jimah and Hinshaw, 2019) with 56 % amino acid identity and 70 % amino acid similarity (Matreyek et al., 2014). Structurally, the two proteins share a broadly conserved architecture (figure 1.9, below). The two proteins contain an N-terminal bundle signalling element (BSE) and GTPase domain (G domain). The role of the GTPase domain is to bind and hydrolyse GTP, leading to

structural changes conferred by the BSE across the whole protein (Rennie et al., 2014). The C-terminal half of the protein is predominantly composed of the stalk domain, which is the key determinant of homo-oligomerisation (Gao et al., 2010). Within the stalk domain is a loop (L4), which is the key determinant of antiviral activity for MxA (von der Malsburg et al., 2011; Mitchell et al., 2012; Patzina et al., 2014), but not MxB (Matreyek et al., 2014). The importance of the L4 loop in MxA antiviral action was demonstrated by removing the L4 loop from murine Mx1 proteins, which was shown to abolish antiviral activity against influenza A virus (Mitchell et al., 2012; Verhelst et al., 2015).

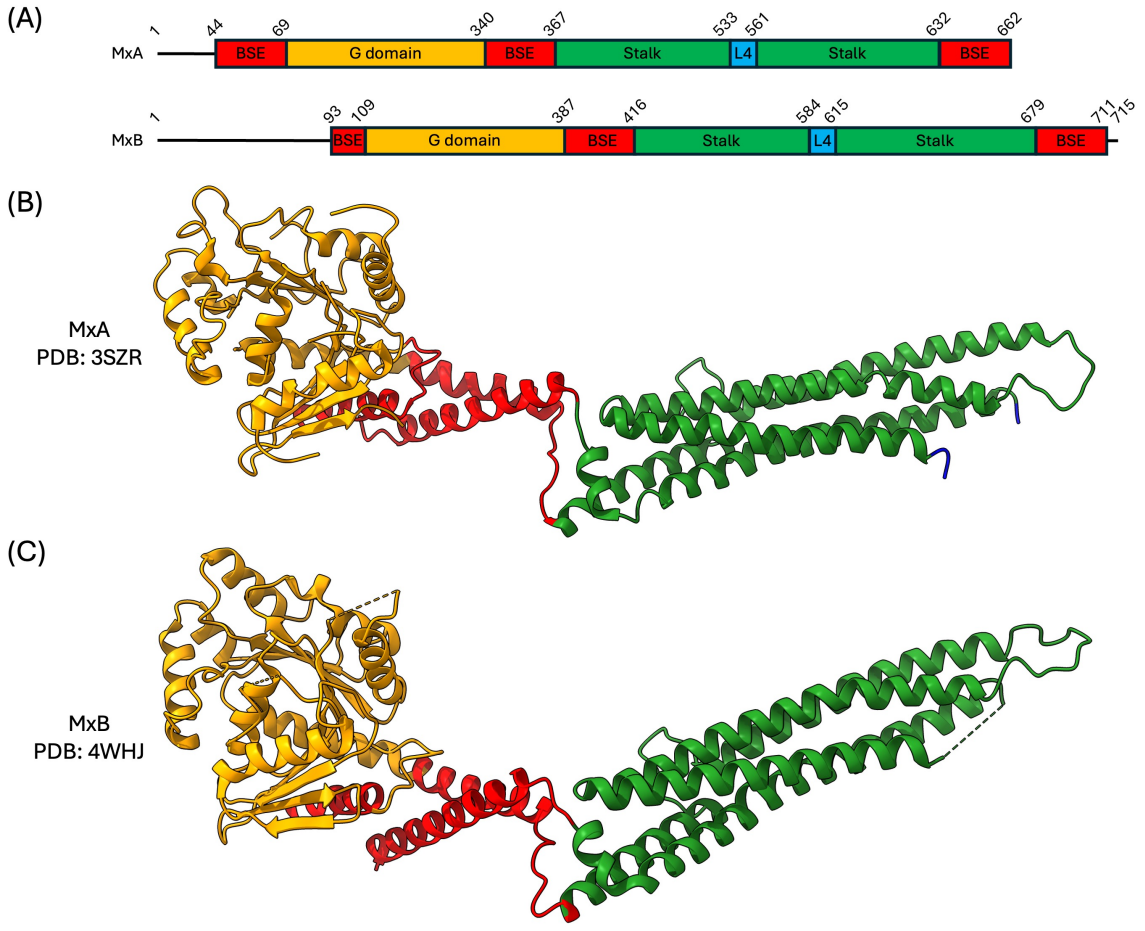


Figure 1.9: Structure and Domain Architecture of Mx Family Proteins, MxA and MxB.

(A) Linear domain architecture of MxA and MxB, with all major domains annotated approximately – Bundle Signalling Element (BSE) (red), GTPase (G) domain (orange), Stalk domain (green), L4 loop (L4) (blue, only partially resolved in MxA). (B) Crystal structure of MxA (Gao, von der Malsburg, et al., 2011) (PDB: 3SZR) coloured by domain. (C) Crystal structure of MxB (Fribourgh et al., 2014) (PDB: 4WHJ). Domain numbering was taken from published work for MxA (Gao, von der Malsburg, et al., 2011) and for MxB (Fribourgh et al., 2014).

MxA homo-oligomerisation involves inter- and intra-molecular domain interactions, those being predominantly stalk-stalk and BSE-stalk. These interactions have been shown to be required for detectable antiviral function, and mutations in these regions limiting either higher order oligomerisation or structural flexibility about the hinge regions between the G domain and BSE, and the BSE and the stalk have been shown to be detrimental to antiviral action (Gao et al., 2010), but GTP binding, and not hydrolysis, is required for antiviral action (Schwemmle et al., 1995; Janzen et al., 2000).

There is structural evidence of MxA homo-oligomerisation (figure 1.10, panel A, below) (Kochs et al., 2002), whereby MxA is observed to homo-oligomerise to varying orders as a function of salt concentration. There is also structural evidence of MxA homo-oligomerisation in a manner analogous to dynamin on negatively charged liposomes (figure 1.10, panel B, below) (von der Malsburg et al., 2011). In this work, the authors also claim that this interaction is likely due to the L4 loop, and within the L4 loop, a stretch of four positively-charged lysine residues (K554 – K557), as the same tubulation of liposomes was not observed when positively charged liposomes were tested. The L4 loop was suggested as a point of interaction with the liposome due to cleavage experiments of the protein when bound to its liposome target and immuno-staining with antibodies targeted to specific parts of the MxA protein which would or

would not be present in the cleaved version of MxA, and because of the canonical 'T-bar shape' in front views obtained by cryo-EM.

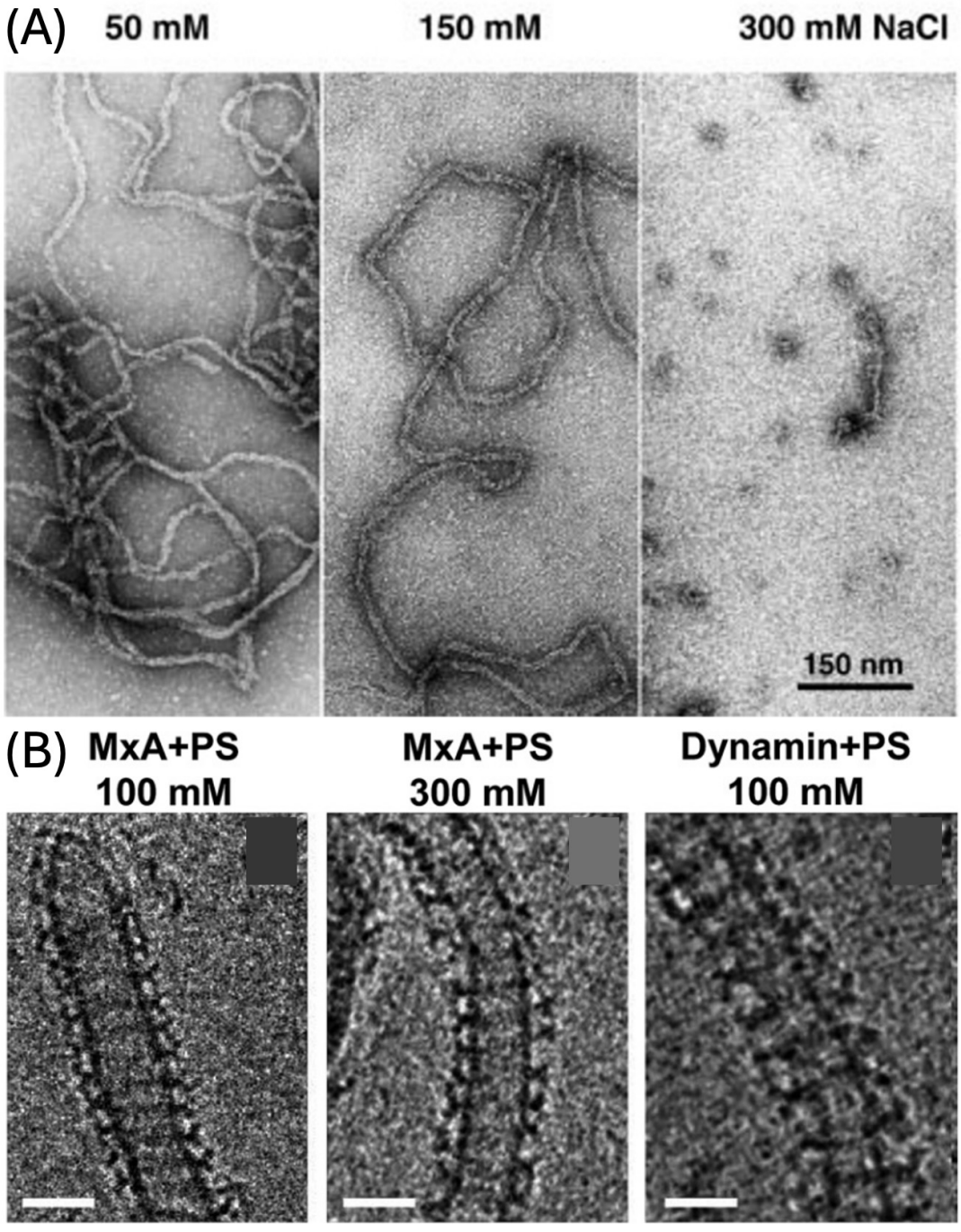


Figure 1.10: MxA Homo-Oligomerisation.

(A) Self-assembly of MxA into filaments is highly controlled by salt concentration (Kochs et al., 2002). (B) Assembly of MxA (left, centre) and dynamin (right) onto negatively charged phosphatidylserine-formed liposomes. Scale bar 50 nm. (von der Malsburg et al., 2011).

There have been no structural data presented illustrating the mode of MxA binding to the influenza nucleoprotein. There has, however, been a wealth of biochemical data which suggests that the mode of binding requires the L4 loop (Nigg and Pavlovic, 2015; Sun, 2017). Provided enough MxA is present in the cytoplasm at infection (likely due to the interferon response from neighbouring infected cells), MxA is able to inhibit the transport of vRNPs into the nucleus at early timepoints of infection (Xiao et al., 2013), likely by assembling into higher-order oligomers around the vRNP. MxA is also able to inhibit nuclear import of newly translated nucleoprotein molecules, sequestering them in the cytoplasm to inhibit replication, likely in lower-order oligomers, up to approximately tetramers (Nigg and Pavlovic, 2015). Through recent work, another mode of inhibition has been demonstrated, that being the cytoplasmic sequestration of newly synthesised influenza vRNPs being trafficked from the nucleus to the plasma membrane (McKellar et al., 2024).

The expected mode of action of MxA against influenza nucleoprotein is therefore to form low-order oligomers to bind nascent low-order oligomeric nucleoproteins (Nigg and Pavlovic, 2015). Since MxA is expected to homo-oligomerise into rings (Gao et al., 2010; Gao et al., 2011) with an approximately 14 – 24 nm central hole in a 13 – 16-meric ring (Gao et al., 2010; Haller et al., 2010) the current accepted model is that MxA is forms rings around vRNPs, whereby the avidity of interaction given multiple MxA monomer L4 loops each contacting a nucleoprotein molecule should be greater than compared to lower order oligomers of MxA binding monomeric or low-order oligomeric nucleoprotein, forming a more-stable inhibitory complex. There is currently no structural information of MxA bound to influenza nucleoprotein, either in monomeric,

low-order oligomeric, or high-order oligomeric form, therefore these hypotheses are yet to be confirmed.

Likewise, in MxB literature, there have been no structures to date published which illustrate the mode of MxB binding to its HIV capsid target. There has, however, been one paper which proposed a conceptual binding model of MxB to the HIV capsid protein (CA) informed only by the structure of the MxB dimer solved by X-ray crystallography (Fribourgh et al., 2014).

1.5.3. Influenza A Virus Nucleoprotein Sensitivity to MxA

MxA restriction poses a significant species barrier to the influenza A virus lifecycle in humans. For example, human MxA was found to be a potent interspecies barrier against the novel bat-derived influenza A virus H18N11 (Ciminski et al., 2019). In general, influenza A viruses of avian origin are considered to be sensitive to MxA restriction, whereas influenza A viruses adapted to humans are not (Haller and Kochs, 2011; Zimmermann et al., 2011; Deeg et al., 2017), as avian Mx proteins lack anti-influenza A virus activity (Benfield et al., 2008; Schusser et al., 2011). This so-called MxA escape is mediated by adaptive mutations in the viral nucleoprotein, initially to the detriment of viral fitness, before stabilising mutations accommodate for the required escape mutations (figure 1.11, below) (Mänz et al., 2013; Deeg et al., 2017; Dornfeld et al., 2019; Haller and Kochs, 2020). These mutations have been shown to come at the cost of efficient vRNP nuclear import (Götz et al., 2016), which demonstrates the importance of overcoming this restriction for the efficient

replication and assembly pathway of influenza virus. The hallmark escape mutations of MxA resistance are generally considered to be R100V/I, L283P/S, and F313Y (Haller et al., 2015; Götz et al., 2016).

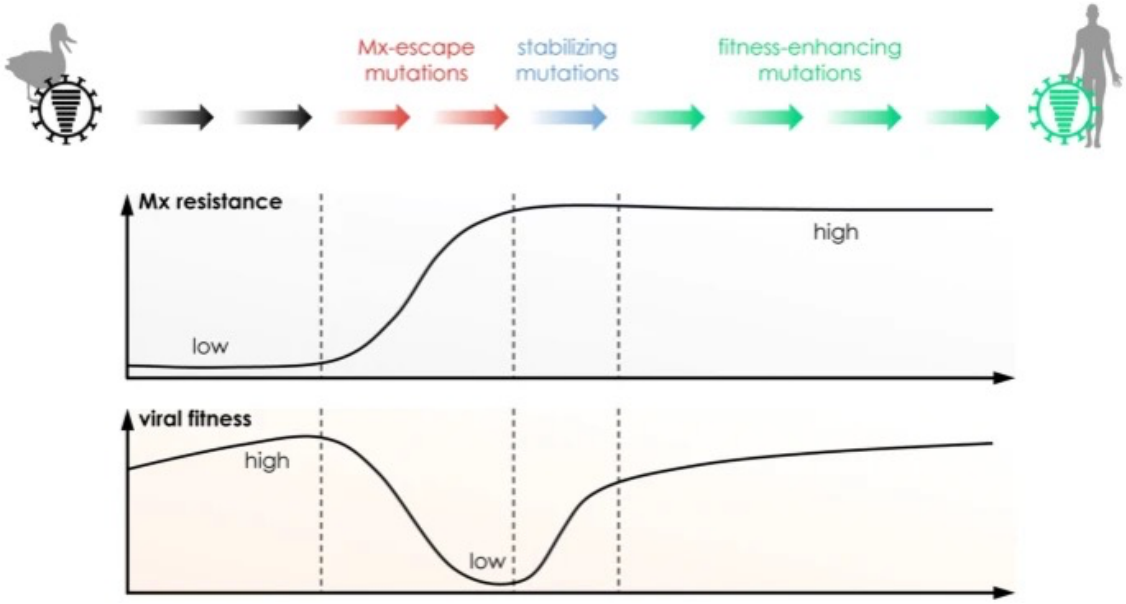


Figure 1.11: Mx Resistance and Viral Fitness.

The evolution of avian influenza virus nucleoprotein when crossing the species barrier into humans requires adaptive mutations which facilitate Mx escape. This comes at a cost to viral fitness, which is later improved by stabilising mutations which help facilitate higher viral fitness in the new-found host. Taken from Götz et al., 2016.

The earliest recorded instance of human MxA resistance is the 1918 H1N1 pandemic (Reid et al., 2004; Mänz et al., 2013), where the cluster of mutations associated with MxA resistance was found to be conserved in all descendent strains of seasonal and pandemic influenza viruses. This further demonstrates the importance of MxA escape for the successful spillover of avian influenza viruses in human hosts. In the 1918 pandemic, there were three waves of pandemic influenza, with the second being the most severe pandemic wave by deaths known to mankind (Taubenberger and Morens, 2006). This was likely brought about by a number of factors: trench warfare,

hospitalisation in open wards, and crowded troop transport across oceans among other factors all contributed to a high density of sick young patients facilitating the incubation and adaptation of a number of pathogens to an otherwise vulnerable young population (Oxford et al., 2002). Within the context of the 1918 pandemic, MxA escape may have indeed been the key determinant in the severity of the second wave. Our understanding of the molecular nature of nucleoprotein evolution was significantly hampered by the predominance of second-wave isolates – notably the Brevig Mission sample, isolated by Taubenberger and colleagues from an Inuit victim of the pandemic buried in Alaskan permafrost (Taubenberger and Morens, 2008) – greatly informing the first whole genome sequence of an influenza virus published in 2005 (Tumpey et al., 2005), but with no information on how the viral nucleoprotein evolved between the first and second pandemic waves. However, in 2022, sequencing results of lung autopsy samples collected during the first wave of the 1918 pandemic in Germany were published (Patrono et al., 2022). This study identified two residues which are highly conserved in avian influenza A viruses but are absent in second-wave 1918 pandemic sequences. These are G16 (compared to D16 in MxA-resistant strains) and L283 (compared to P283 in MxA-resistant strains), further informing our understanding of the molecular determinants of MxA resistance in the context of the 1918 pandemic.

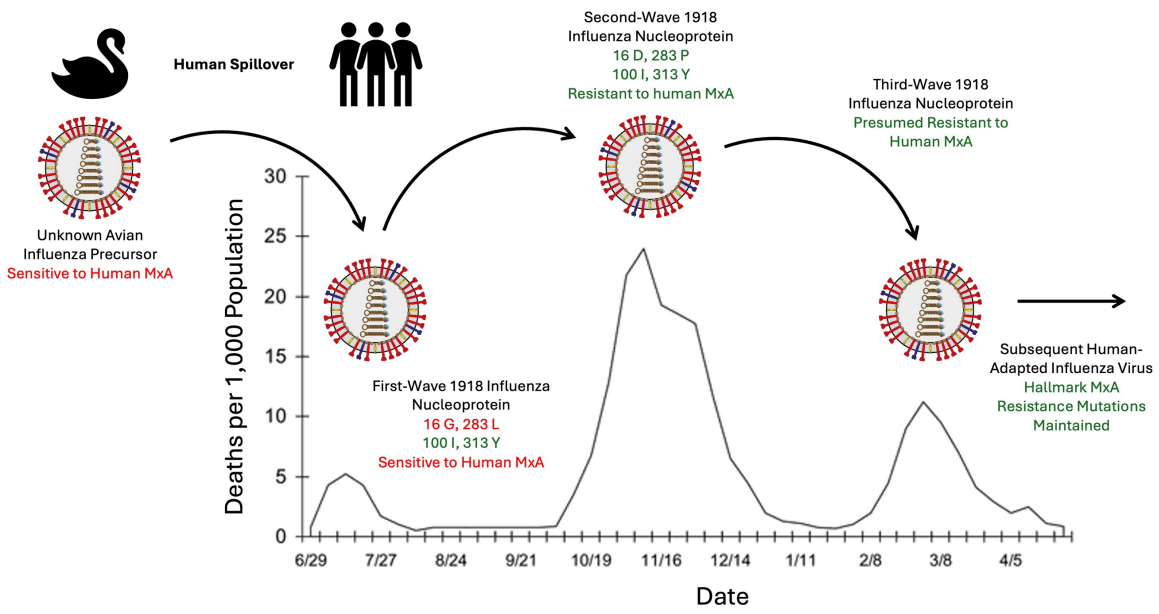


Figure 1.12: Graphical Representation of Inter-Wave Evolution of 1918 Pandemic Influenza A Virus.

Graph of mortality against time due to the 1918 influenza pandemic in the United Kingdom. Overlayed is a predicted influenza A virus precursor to the 1918 pandemic virus, the first, second, and third wave viruses, and a comment on subsequent human-adapted influenza A viruses, with comments on their respective MxA sensitivity residues and an overall comment on sensitivity or resistance to human MxA, where sensitive comments and residues are annotated in red and resistance comments and residues are annotated in green. Graph adapted from Taubenberger and Morens, 2006.

1.6. Butyrophilin Subfamily 3 Member A3 (BTN3A3)

1.6.1. Butyrophilin Subfamily 3 Protein Family

Butyrophilin (BTN) family proteins belong to the immunoglobulin (Ig) superfamily of transmembrane proteins. Butyrophilin subfamily 1 member A1 (BTN1A1) was the first of the family to be identified and was discovered during investigation of milk fat globule formation in the lactating mammary tissues of cows. It is expressed on the apical surface of secretory cells in the mammary gland epithelium during the expulsion of newly synthesised milk (Franke et al., 1981; Ogg et al., 2004; Robenek et al., 2006).

Upon further investigation, the human butyrophilin family was found to have 13 members which all have a highly conserved domain organisation: an N-terminal, extracellular IgV domain followed by an IgC domain, then a transmembrane domain, and a juxtamembrane domain. Finally, some BTN family members also contain an intracellular B30.2/SPRY domain (Afrache et al., 2012; Rhodes et al., 2016).

BTN family proteins belong to a family of co-stimulatory molecules and have strong homology to the B7 family of costimulatory receptor molecules, with a shared extracellular domain architecture, and the addition of the intracellular B30.2 domain (Arnett and Viney, 2014; Lebrero Fernández, 2016).

Given this noted homology with B7 family proteins, it was later shown that the BTN family have a number of immunomodulatory functions (Afrache et al., 2012; Harly et al., 2012; Perfetto et al., 2013; Arnett and Viney, 2014; Gu et al., 2018), with the BTN3A sub-family of proteins being known to modulate specific T-cell subset activity (Arnett and Viney, 2014; Blazquez et al., 2018). The extracellular domains of the BTN3A sub-family demonstrate 95 % sequence identity (Blazquez et al., 2018), although the intracellular domains differ; BTN3A1 and BTN3A3 contain a B30.2 domain, whereas BTN3A2 does not. It has been additionally shown that major differences between the BTN3A subfamily members lie in their juxtamembrane domain, beyond the transmembrane domain but before the B30.2/SPRY domain, where differences in the juxtamembrane region are seen to affect target binding (Karunakaran et al., 2023).

The intracellular B30.2/SPRY domain is found in many human proteins and is additionally involved in various innate immune pathways (D'Cruz et al., 2013). One such example is the B30.2/SPRY domain common in members of the tripartite motif (TRIM) family of proteins (Ozato et al., 2008; Tomar and Singh, 2015).

Members of the TRIM family have a variety of functions in regulating the human immune response, including as viral restriction factors (e.g. TRIM5α), and in immune signalling (e.g. TRIM21). In these cases, among others, the B30.2/SPRY domain is known to act as a pattern-recognition receptor, involved in oligomerisation and ligand binding (Rhodes et al., 2016). For example, the B30.2/SPRY domain of TRIM5α is known to bind to the capsid of HIV (Pertel et al., 2011; D'Cruz et al., 2013), and the B30.2/SPRY domain of TRIM21 is known to bind the Fc domain of human IgG (Rhodes and Trowsdale, 2007). In all subsequent sections of this thesis, the BTN3A3 B30.2/SPRY domain will be referred to as the B30.2 domain for simplicity.

At the inception of this work, no structure existed of full-length BTN3A3. However, recent studies using cryo-EM have demonstrated the formation of BTN3A1:BTN3A2 heterodimers, and their association to BTN2A1 homodimers in lipid monolayers as a mechanism for T-cell receptor activation (Zhang et al., 2025) (PDB: 9JQQ). No literature has been published on the full-length structure of BTN3A3, but one structure has been deposited on the Protein Data Bank of a similar (but unannotated) complex of BTN2A1, BTN3A1 and BTN3A3 which shows a very similar domain organisation, with two BTN dimers interacting via both the IgV and B30.2 domains (PDB: 8ZYR). For this reason,

the annotated structure (Zhang et al., 2025) (PDB: 9JQQ) is displayed below (figure 1.13).

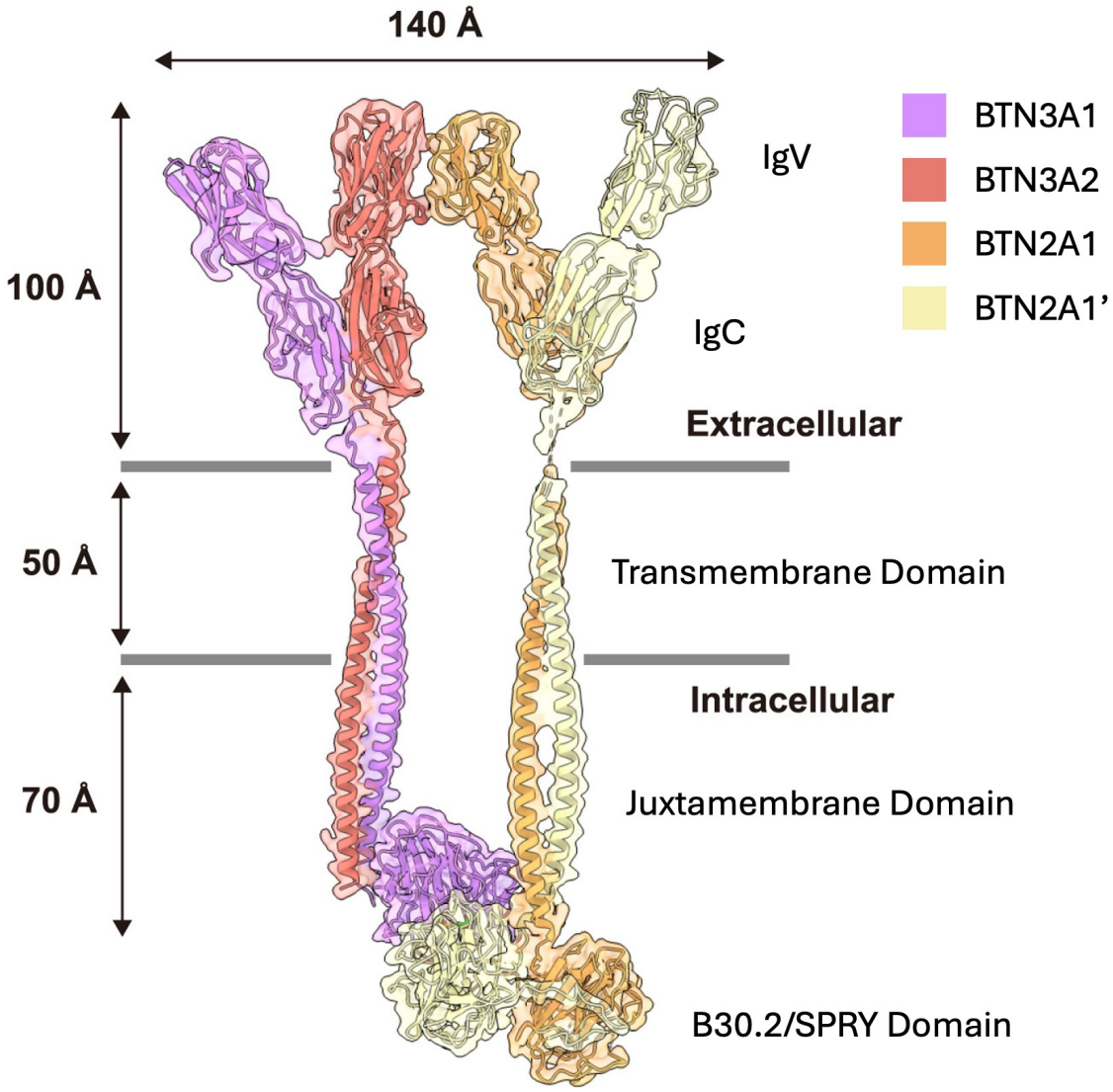


Figure 1.13: Full-Length Structure of Human BTN3A1, BTN3A2, and BTN2A1 in Complex.

A heterodimer of BTN3A1 and BTN3A2 bound to a homodimer of BTN2A1. Key domains are annotated, and each copy of the full-length BTN family protein is depicted in different colours as annotated. Figure adapted from Zhang et al., 2025.

While the immunomodulatory functions of BTN family proteins were initially characterised in the context of T-cell regulation, recent advances have revealed an additional role in intrinsic antiviral immunity, namely by BTN3A3's selective restriction

of avian influenza viruses. Similarly to MxA, BTN3A3 is also an interferon-stimulated gene (Petric et al., 2023; Pinto et al., 2023), positioning it as another component of the interferon-induced antiviral response.

In 2023, Pinto and colleagues screened a library of interferon-induced proteins for antiviral activity in cultured human cell lines. They demonstrated that BTN3A3 restricts avian-, but not human-adapted, influenza A viruses, posing another innate immune species barrier to effective spillover into humans (Pinto et al., 2023). In the same work, they found that BTN3A3 was mainly present in the nucleus, but the molecular mode of action to inhibit avian influenza A viruses remains elusive.

1.6.2. Influenza A Virus Nucleoprotein Sensitivity to BTN3A3

Pinto and colleagues identified key nucleoprotein residues that determine BTN3A3 sensitivity: BTN3A3-sensitive viral genotypes possess nucleoproteins with residues 52Y and 313F, while BTN3A3-resistant viral genotypes possess nucleoproteins with residues 52N/H/Q and 313Y/V. Notably, successful transmission of avian influenza A virus to humans was associated with the acquisition of these resistance mutations.

The residues are located in close proximity on the same surface-exposed patch of the nucleoprotein (figure 1.14). While residue 313 is also a determinant of MxA sensitivity, Pinto and colleagues reasoned that BTN3A3 restriction occurs in a MxA-independent manner, having different dependencies on the two residues highlighted, and affecting different stages of the viral lifecycle.

BTN3A3 specifically targets avian influenza virus at the level of genome replication, rather than in early infection steps. Time-course analysis revealed that viral proteins of the vRNP (PB2, PB1, PA, and the nucleoprotein) accumulated equally in both cytoplasmic and nuclear mock-treated and BTN3A3-expressing cells at 45 minutes post-infection, indicating normal virus entry and vRNP nuclear import. However, pronounced differences emerged by 90 minutes post-infection, and were further amplified by 6 hours post-infection, but only for avian influenza viruses.

By using cycloheximide to block protein synthesis (limiting the virus to primary transcription), BTN3A3 had no effect on viral mRNA levels. In contrast, without cycloheximide treatment, BTN3A3 caused a 100-fold reduction in avian influenza mRNA levels by 6 hours post-infection, while not affecting levels of mRNA earlier in infection. Importantly, avian influenza virus carrying the resistance mutation F313Y was unaffected by BTN3A3 in these assays, confirming the specificity of the restriction mechanism. Taken together, this suggests that BTN3A3 restriction of sensitive influenza viruses occurs at the level of viral genome replication.

Despite BTN3A3 being a transmembrane protein (Afrache et al., 2012), Pinto and colleagues found by using immunostaining and confocal microscopy that constitutively expressed BTN3A3 in hBEC3-KT cells is mainly localised to the cell nucleus, thus explaining the ability of BTN3A3 to inhibit nuclear viral replication processes. It is known that differing localisation of the butyrophilin family of proteins occurs across different cell lines and as a result of different protein variants (Arnett

and Viney, 2014; Karunakaran et al., 2023). However, exact knowledge of the localisation of BTN3A3 in physiologically relevant cells, and the mechanism of influenza virus restriction is yet to be fully determined. Improved molecular knowledge of BTN3A3's proposed antiviral mechanism is key for better understanding species barriers to influenza transmission.

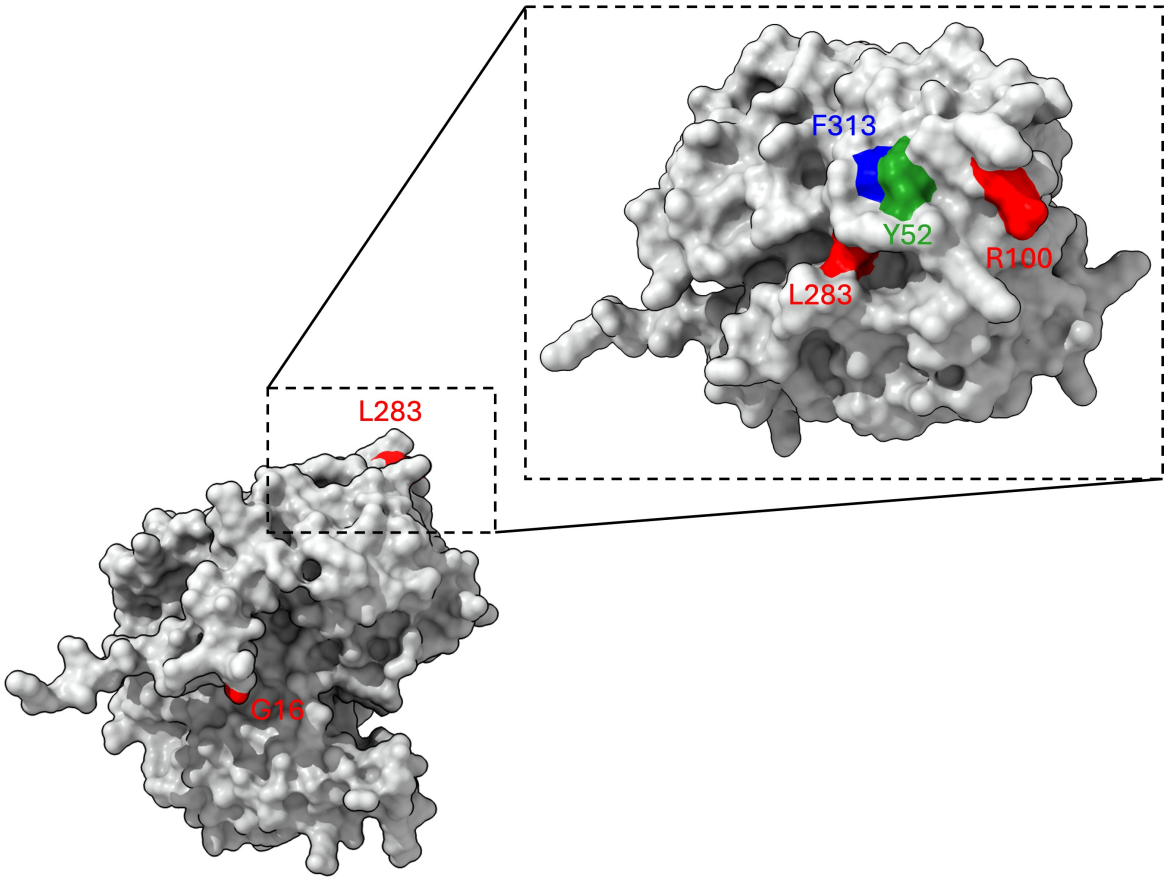


Figure 1.14: Surface View of Influenza A Nucleoprotein Immune Factor Sensitivity Residues.

AlphaFold3 prediction of S09 NP R416A in a grey coloured surface view with sensitivity residues for MxA (in red), BTN3A3 (in green) and both (in blue) highlighted and annotated.

1.7. Thesis Aims

Our current understanding of innate immune restriction relies heavily on biochemical and cell biology approaches, which are useful to identify binding partners and to map regions of interaction, but cannot reveal the atomic details that govern specificity, affinity, and resistance mechanisms. Without high-resolution structural data of these interactions to support the proposed models, mechanistic questions of influenza virology remain unanswered, which limits both our basic understanding of host-pathogen evolution, and our ability to develop rational therapeutic interventions.

MxA is proposed to inhibit nucleoprotein trafficking by forming oligomeric rings around vRNPs while also binding to individual nucleoproteins in lower-order oligomeric states. However, no structural data of a MxA – nucleoprotein complex in any form exists to validate this model or explain how nucleoprotein resistance mutations disrupt binding. Similarly, the restriction mechanism proposed for BTN3A3 remains poorly understood at the molecular level. While immunofluorescence studies suggest BTN3A3 can localise to the nucleus during infection, the mechanism for this translocation of a transmembrane protein is yet unclear, and the mode of interaction between the influenza nucleoprotein and BTN3A3 has not been demonstrated. The absence of structural data of these host-pathogen interfaces limits our mechanistic understanding of how restriction occurs in non-human-adapted strains of influenza A virus, and also limits our capacity to predict escape mutations, to develop rationally-designed therapeutics, and to understand why restriction does not occur in human-adapted influenza A virus strains.

The structural study of nucleoprotein-immune factor interactions presents a number of technical challenges. X-ray crystallography approaches are limited by the difficulty of obtaining stable, homogeneous complexes of these often transient, low-affinity interaction complexes. While monomeric nucleoprotein structures have been readily solved by X-ray crystallography, the small size of the nucleoprotein (≈ 55 kDa) and its flexibility make it challenging to study by single particle cryo-EM without assembly into larger complexes.

The determination of native vRNP structure is a great challenge for structural biology techniques given the complex's overarching heterogeneity, yet this may be the required state of the nucleoprotein to ensure immune factor binding. Where there is immune factor binding, heterogeneity in sample due to flexible, or low-affinity protein interactions would deem X-ray crystallography unsuitable for structure determination, and where there is not complete coating of the vRNP, signal would likely be averaged out by cryo-EM.

Both innate immune factor and viral RNA interactions with the nucleoprotein pose fundamental, unanswered questions in the field, and they are of great clinical importance and potential therapeutic promise. By understanding the mechanism of immune factor binding to the influenza nucleoprotein, a greater basic biology understanding and potential mechanisms for clinical interventions would be unlocked. Added to this, an improved understanding of the mechanisms of antiviral action would facilitate better understanding of the importance of specific mutations in the virus through global surveillance efforts.

This thesis aims to overcome these limitations by implementing novel structural approaches to investigate the influenza A nucleoprotein, its innate immune factor interactions, and its mechanism of binding to viral RNA. The specific objectives are as follows.

In chapter 3, methods to study influenza nucleoprotein structure, I aim to determine the high-resolution structure of a model avian influenza nucleoprotein and to establish tractable systems for studying nucleoprotein and binding partner complexes by single particle cryo-EM. Establishing a binding platform will enable routine structural characterisation of nucleoprotein interactions partners, allowing dissection of their interaction interfaces. This will aid in determination of the molecular basis of binding partner interactions and may support future efforts to study interaction partners.

In chapter 4, nucleoprotein interaction partners of the human innate immune system, I aim to biophysically and structurally characterise the molecular basis of MxA and BTN3A3 restriction through determining the selective binding of these immune factors to avian influenza A nucleoproteins, and to determine their structures. This will aid in our understanding of nucleoprotein interaction partners of the human innate immune system, by comparing interaction between susceptible avian influenza strains and (expected) resistant human-adapted influenza strains. Since no structures of either immune factor have been solved in complex with the influenza nucleoprotein, structural elucidation will greatly improve our understanding of the mode of immune factor binding to susceptible influenza nucleoproteins.

In chapter 5, methods to study influenza vRNP structure and its interaction with nucleoprotein binding partners, I aim to resolve the structure of the nucleoprotein and influenza polymerase in the native vRNP and determine how MxA assembles around intact vRNPs to achieve restriction of influenza A virus replication. At the time of this work's inception, determination of native influenza vRNP structure had not been achieved, and questions of the nature of RNA binding to the nucleoprotein in the context of the vRNP, and the mode of nucleoprotein-nucleoprotein interactions between helical strands had not been determined. The mode of MxA binding to the vRNP has also not been determined, and remains a significant gap in our knowledge of the mode of innate immune factor function.

2. Materials and Methods

2.1. Molecular Biology and Microbiology Methods

2.1.1. Preparation of Solutions and Media

Buffers and solutions used in experiments outlined in this thesis were sterilised either by autoclaving at 121˚C for 90 minutes or by filtration through a 0.22 µm filter (Merck-Millipore). Bacterial growth media, glassware and plasticware were autoclaved at 121˚C for 90 minutes prior to use. Luria-Broth (LB) and Terrific Broth (TB) were prepared according to manufacturers' instructions (Sigma-Aldrich).

2.1.2. Transformation of Chemically Competent *E. coli*

Pre-aliquoted chemically competent *E. coli* (DH5α, BL21 Rosetta (DE3) or KRX strains, as required) were thawed from long-term storage at -80˚C and aliquoted into 25 µL aliquots in Eppendorf tubes on ice before adding 25 ng of plasmid DNA. The bacteria were incubated with the DNA for 20 minutes prior to a 45 second heat shock in a 42˚C water bath. The bacteria were then incubated back on ice for a further 2 minutes prior to the addition of 200 µL of SOC medium (Super Optimal broth with Catabolite repression, Thermo Fisher). The resulting culture was incubated in a 37˚C incubator shaking at 800 rpm for 1 hour and then plated onto LB-agar plates with appropriate antibiotics to select for transformant bacteria.

2.1.3. Cultivation and Storage of *E. coli*

Overnight cultures of *E. coli* from stab cultures, glycerol stocks or individual colonies on LB-agar plates were grown in small volumes (≤ 200 mL) of LB medium supplemented with appropriate antibiotics (ampicillin or kanamycin where required) in a 37˚C incubator shaking at 800 rpm for 12 hours. For long-term storage, glycerol stocks of transformed *E. coli* strains were made by flash freezing cultures with 25 % (v/v) glycerol and stored at -80˚C.

2.1.4. Preparation of Plasmid DNA from *E. coli*

Plasmid DNA from transformed *E. coli* was prepared from overnight cultures of bacteria. Either a single colony from an LB-agar plate, a scraping of a glycerol stock, or a scraping from a stab culture was taken with a small pipette tip and used to inoculate a minimum of 5 mL of LB medium with appropriate antibiotics. The culture was then grown overnight in a 37˚C incubator shaking at 800 rpm. The subsequent day, plasmid DNA was extracted using the QIAprep Miniprep Kit (Qiagen). The sequences of all resulting plasmid DNA, regions of interest or inserts were verified using Sanger sequencing or with Oxford Nanopore sequencing (Source Biosciences).

2.1.5. DNA and RNA Oligonucleotides

In this thesis, all DNA oligonucleotides were purchased from Sigma-Aldrich and subsequently dissolved in nuclease-free water (Ambion) and diluted to a concentration of 100 mM. All RNA oligonucleotides were purchased from IDT and

subsequently dissolved in nuclease-free water (Ambion) and diluted to a concentration of 100 mM.

2.1.6. Plasmids

Plasmids used in this thesis, their origin, and their purpose are outlined below (table 2.1).

Table 2.1: Table of plasmids used in this thesis, their origin, and their purpose.

Plasmid Name	Plasmid Origin	Plasmid Purpose
Influenza Nucleoproteins (all pGEX-6P-1 unless stated)		
pCAG_S09_NP_Wt	Donated by Stephanie Williams (Taubenberger / Fodor labs)	Cloning insert into pGEX bacterial expression vectors
Borna_N	Donated by Jeremy Keown	Recipient vector for pGEX cloning
S09_NP_Wt	Cloned in-house from pCAG_S09_NP_Wt insert and pGEX-6P-1_Borna_N backbone	Expression of avian S09 (A/Green-Winged Teal/Ohio/175/1966) strain influenza A nucleoprotein
S09_NP_R416A	Cloned in-house from pGEX-6P-1_S09_NP_Wt	Expression of monomeric avian S09 (A/Green-Winged Teal/Ohio/175/1966) strain influenza A nucleoprotein
Tk_NP_R416A	Donated by Cecilia Rocchi	Expression of monomeric H5N1(A/turkey/Turkey/1/2005) influenza A nucleoprotein
NT60_NP_R416A	Donated by Haitian Fan	Expression of monomeric NT60 (A/Northern Territory/60/1968) influenza A nucleoprotein
Human Innate Immune Factors		
pGEX-6P-1_MxA	Donated by Haitian Fan	Expression of human MxA (GST tagged)
pET151_D_TOPO_Viet_PB2C	Donated by Hana Veler	Recipient bacterial expression vector
pET151_D_TOPO_MxA	Cloned in-house from pGEX-6P-1_MxA insert and pET151_D_TOPO_Viet_PB2C backbone	Expression of human MxA (N-terminal His-tagged)
pET29b(+)_BTN3A3_B30.2	Purchased from Twist Biosciences	Expression of human BTN3A3 B30.2 domain (N-terminal His-tagged)
Engineered Nanobodies		
Nanobodies		
pHEN6_Nb	Donated by Florian Schmidt (Schmidt et al., 2016)	Panel of influenza nucleoprotein-binding nanobodies for bacterial expression

pMESy4_Nb8210	Donated by Amanda (Zihan) Zhu (Keown et al., 2022)	Influenza polymerase (PA-linker)-binding nanobody for bacterial expression
Gluebodies		
pNIC_MBP_Gb	Donated by Gangshun Yi	Recipient backbone vector for cloning of gluebody fragments (MBP-tagged)
pNIC_NHStIIT_Gb	Donated by Gangshun Yi	Recipient backbone vector for cloning of gluebody fragments (Strep-tagged)
VHH_Gb	Purchased from Twist Biosciences	Blunt-end gene fragments of nanobodies engineered from above Schmidt origin nanobodies for cloning into pNIC bacterial expression vectors
Megabodies		
pMESP23NO_Mb	Donated by Amanda (Zihan) Zhu	Bacterial expression of NP-binding Megabodies
Virus Reverse Genetics		
pHW2000-PB2_WSN	Fodor Laboratory (Hoffmann et al., 2000)	Bidirectional mammalian expression plasmid (producing both mRNA for protein translation and negative-sense vRNA) for of influenza A/WSN/1933 PB2 protein, and PB2 vRNA
pHW2000-PB1_WSN	Fodor Laboratory (Hoffmann et al., 2000)	Bidirectional mammalian expression plasmid of influenza A/WSN/1933 PB1 protein, and PB1 vRNA
pHW2000-PA_WSN	Fodor Laboratory (Hoffmann et al., 2000)	Bidirectional mammalian expression plasmid of influenza A/WSN/1933 PA protein, and PA vRNA
pHW2000-HA_WSN	Fodor Laboratory (Hoffmann et al., 2000)	Bidirectional mammalian expression plasmid of influenza A/WSN/1933 HA protein, and HA vRNA
pHW2000-NP_WSN	Fodor Laboratory (Hoffmann et al., 2000)	Bidirectional mammalian expression plasmid of influenza A/WSN/1933 NP protein, and NP vRNA
pHW2000-NA_WSN	Fodor Laboratory (Hoffmann et al., 2000)	Bidirectional mammalian expression plasmid of influenza A/WSN/1933 NA protein, and NA vRNA
pHW2000-M_WSN	Fodor Laboratory (Hoffmann et al., 2000)	Bidirectional mammalian expression plasmid of influenza A/WSN/1933 M1 and M2 protein, and M1 and M2 vRNA
pHW2000-NS_WSN	Fodor Laboratory (Hoffmann et al., 2000)	Bidirectional mammalian expression plasmid of influenza A/WSN/1933 NS1 and NEP proteins, and NS vRNA
pCAGGS-NP_S09	Stephanie Williams	Mammalian expression plasmid of influenza S09 (A/Green-Winged Teal/Ohio/175/1966) NP protein
pHH21-NP_S09	Stephanie Williams	Mammalian expression plasmid of influenza S09 (A/Green-Winged Teal/Ohio/175/1966) NP vRNA

2.1.7. Synthetic Genes

Synthetic genes used in this thesis were ordered from Twist Bioscience either as gene fragments (Gluebody genes) or as codon optimised genes in bacterial expression plasmids.

2.1.8. Polymerase Chain Reaction (PCR)

PCR reactions for various purposes were performed using Q5 High-Fidelity DNA Polymerase (New England Biolabs, NEB). Reactions were performed in 50 µL volumes using 0.2 mL PCR tubes with 100 ng of plasmid DNA, 1.25 µL of each 10 µM forward and reverse DNA primer, 0.5 µL of 10 mM dNTPs, 1.5 µL of 100 % (v/v) DMSO, 0.25 µL of Q5 High-Fidelity DNA Polymerase, and 10 µL of 5 x Q5 Reaction Buffer (NEB). The remaining volume was made up to 50 µL with nuclease-free water (Ambion). Typically, the thermal cycling protocol was used as shown below (table 2.2). For particularly difficult PCR reactions, a touchdown PCR protocol was used in the annealing step (as shown), or else an annealing temperature of 47˚C was used without a gradient.

Table 2.2: Typical Thermal Cycling Protocol used for PCR Reactions in this Thesis.

Step	Temperature	Duration	Cycles
Initial Denaturation	98˚C	30 s	x1
Denaturation	98˚C	10 s	x30
Annealing	47 – 67˚C (ΔT = - 5˚ / step)	30 s	
Extension	72˚C	30 s / kb of DNA	
Final Extension	72˚C	10 m	x1
Infinite Hold	12˚C	Infinite	x1

2.1.9. Gibson Assembly

PCR was carried out (as above) to linearise the recipient vector. PCR was also carried out (as above) to amplify the fragment(s) to prepare them for insertion into the linearised vector. Resulting PCR products were visualised on an agarose gel and successful PCR reaction products extracted from the gel using the QIAquick Gel Extraction Kit (Qiagen) as per the manufacturers' instructions. The NEB Hi-Fi Gibson Assembly Kit (NEB) was used to assemble Gibson assembly products as per the manufacturers' instructions. Briefly, 50 ng insert fragment DNA was added to 50 ng linearised backbone DNA and 2.5 µL of Gibson Assembly Mix (Hi-Fi) was added. The remaining volume was made up to 5 µL with nuclease-free water (Ambion) and the resulting mixture was heated at 50˚C for 15 minutes. 2 µL of assembled product was used to transform competent *E. coli* cultures, and subsequent cultures were verified using Oxford Nanopore sequencing (Source Biosciences). In instances where this did not work on the first attempt, two subsequent attempts were made simultaneously where either a 2:1 or a 1:2 insert : vector ratio was used.

2.1.10. Agarose Gel Electrophoresis

Agarose electrophoresis gels were prepared by adding 1 % (w/v) of agarose powder to TAE buffer, pH 8.0 (Tris-acetate-EDTA; 1 mM EDTA, 40 mM Tris-NaOH, 20 mM acetic acid) and heating the mixture until it dissolved in a microwave. SYBR Safe DNA gel stain (Thermo Fisher) was added before gel pouring to be able to visualise DNA bands. Samples containing DNA were loaded onto the gel and separated sufficiently (e.g. 30 minutes at 100 V) and then visualised on an iBright Imaging System (Thermo Fisher).

2.2. Protein Production

2.2.1. Recombinant Protein Expression in *E. coli*

Unless otherwise stated, all proteins expressed in *E. coli* strains were grown overnight from glycerol stocks in 200 mL of LB medium with appropriate antibiotics in a 37˚C shaking incubator. On the day of large-scale bacterial growth, equal volumes of overnight starter culture were added to each flask. 2 L plastic flasks were used for large-scale growth with 500 mL of TB medium and appropriate antibiotics. OD_{600} was measured initially after 4 hours of growth and checked regularly thereafter until an OD_{600} of 0.8 was reached. At this point, protein expression was induced by the addition of IPTG to a final concentration of 1 mM. The temperature was decreased to 18˚C and cells were grown for 12 – 16 h. Bacterial pellets were harvested by centrifugation (3,500 x g, 15 minutes, 4˚C).

2.2.2. Purification of Influenza Nucleoprotein

Pellets of bacteria expressing GST-tagged influenza nucleoproteins were resuspended to homogeneity in 25 mL NP buffer (50 mM HEPES-NaOH (pH 7.5), 1 M NaCl, 10 % (v/v) glycerol) supplemented with 0.1 % Triton X-100, 5 mM DTT, 2 µL 50,000 U Supernuclease (SinoBiological), and one cOmplete EDTA-free Protease Inhibitor Cocktail tablet (Sigma). The resuspended bacteria were sonicated and centrifuged (35,000 g, 45 min, 4˚C). Clarified lysate was then either added to 1 mL / L (of culture) equilibrated GST 4B resin (Cytiva) for 2 h or bound to GST resin in a pre-packed GSTrap

HB column (Cytiva) for 2 column volumes Where free resin was used, it was collected post-binding by centrifugation (1,000 g, 5 min, 4˚C) and washed thoroughly with High Salt Wash Buffer (HSWB) (50 mM HEPES-NaOH (pH 7.5), 1.5 M NaCl, 10 % (v/v) glycerol) followed by overnight elution by the addition of 3C protease (produced in-house) or elution with 20 mM L-reduced glutathione depending on if the GST tag was further required. Lysates added to GSTrap columns were washed thoroughly with NP buffer and HSWB followed by elution with 20 mM L-reduced glutathione. Eluted protein was concentrated to ≤ 500 µL by means of spin concentration using equilibrated 30 kDa MWCO spin concentrators (Amicon, Sigma). Eluted protein was further purified by size exclusion chromatography (SEC). ≤ 500 µL of concentrated protein was injected into a 500 µL loop and SEC carried out using a Superdex S200 10/300 increase GL column (Cytiva) using an ÄKTA Pure machine (Cytiva). Fractions were collected and assessed for protein purity by means of SDS-PAGE. Suitable fractions were pooled, concentrated where necessary, flash frozen in liquid nitrogen, and stored at -80˚C.

2.2.3. Purification of Human MxA

Pellets of bacteria expressing His_6-tagged human MxA were resuspended to homogeneity in 25 mL MxA buffer A (50 mM HEPES-NaOH (pH 7.5), 1 M NaCl, 5 mM $MgCl_2$, 10 % (v/v) glycerol) supplemented with 0.1 % Triton X-100, 5 mM DTT, 2 µL 50,000 U Supernuclease (SinoBiological), and one cOmplete EDTA-free Protease Inhibitor Cocktail tablet (Sigma). The resuspended bacteria were sonicated and centrifuged (35,000 g, 45 min, 4˚C). Clarified lysate was then applied to Ni-NTA resin in a pre-packed HisTrap HB column (Cytiva) for 2 column volumes. Bound resin was

first washed thoroughly with MxA buffer B (20 mM HEPES-NaOH (pH 7.5), 800 mM NaCl, 5 mM $MgCl_2$, 10 % (v/v) glycerol, 20 mM Imidazole) and then MxA buffer C (20 mM HEPES-NaOH (pH 7.5), 400 mM NaCl, 5 mM $MgCl_2$, 10 % (v/v) glycerol, 20 mM Imidazole) before elution by a linear gradient of MxA buffer C to buffer E (20 mM HEPES-NaOH (pH 7.5), 400 mM NaCl, 5 mM $MgCl_2$, 10 % (v/v) glycerol, 300 mM Imidazole). Eluted protein was spin concentrated to ≤ 500 µL and applied to an S200 10/300 increase GL column for further purification by means of SEC. Fractions were taken and analysed by means of an SDS-PAGE gel. Fractions containing MxA were pooled and flash-frozen in liquid nitrogen, and stored at -80˚C. Where required, the presence of MxA was further validated by means of western blot.

2.2.4. Purification of Human BTN3A3 B30.2 Domain Protein

Pellets of bacteria expressing His_6-tagged human BTN3A3 B30.2 domain protein were resuspended to homogeneity in 25 mL B30.2 wash buffer (20 mM Tris-HCl (pH 8.0), 400 mM NaCl, 20 mM Imidazole) supplemented with 0.1 % Triton X-100, 5 mM DTT, 2 µL 50,000 U Supernuclease (SinoBiological), and one cOmplete EDTA-free Protease Inhibitor Cocktail tablet (Sigma). The resuspended bacteria were sonicated and centrifuged (35,000 g, 45 min, 4˚C). Clarified lysate was then applied to Ni-NTA resin in a pre-packed HisTrap HB column (Cytiva) for 2 column volumes. Bound resin was first washed thoroughly with B30.2 wash buffer before elution by a linear gradient of B30.2 wash buffer with additional imidazole (250 mM imidazole). Eluted protein was spin concentrated to ≤ 500 µL and applied to an S75 10/300 increase GL column equilibrated in SEC buffer (25 mM HEPES-NaOH (pH 7.5), 150 mM NaCl) for further

purification by means of SEC. Fractions were taken and analysed by means of an SDS-PAGE gel. Fractions containing B30.2 domain were pooled and flash-frozen in liquid nitrogen, and stored at -80˚C.

2.2.5. Purification of Nanobodies

Pellets of bacteria expressing His_6-tagged influenza nucleoprotein and polymerase nanobodies were resuspended to homogeneity in 50 mL PE buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 20 % sucrose, 0.4 mg/mL lysozyme) supplemented with one cOmplete EDTA-free Protease Inhibitor Cocktail tablet (Sigma) for lysis by means of osmolysis. The resuspended, lysed bacteria were sonicated and centrifuged (35,000 g, 45 min, 4˚C). The clarified lysate was supplemented to a final concentration of 5 mM $MgCl_2$ and 500 mM NaCl. 1.5 mL of Ni-NTA resin (Qiagen) equilibrated in wash buffer I (50 mM HEPES-NaOH (pH 7.5), 500 mM NaCl) was added per litre of bacterial cell culture grown. Protein was allowed to bind the Ni-NTA resin for 1 hour at 4˚C rotating. The bound resin was washed with 15 column volumes of wash buffer I, and then 15 column volumes of wash buffer II (50 mM HEPES-NaOH (pH 7.5), 500 mM NaCl, 20 mM imidazole). Purified protein was then sequentially eluted in 1 mL fractions using elution buffer (20 mM HEPES-NaOH (pH 7.5), 500 mM NaCl, 250 mM imidazole). Fractions containing pure nanobody were pooled and concentrated to ≤ 500 µL and applied to an S75 10/300 increase GL column equilibrated in SEC buffer (25 mM HEPES-NaOH (pH 7.5), 150 mM NaCl) for further purification by means of SEC. Fractions were taken and analysed by means of an SDS-PAGE gel. Fractions containing pure nanobodies were pooled and flash-frozen in liquid nitrogen, and stored at -80˚C.

2.2.6. Purification and Formation of di-Gluebodies

di-Gluebody constructs were either produced in an MBP-His_6 tag backbone (pNIC_MBP) or a streptavidin tag backbone (pNIC_NHS).

Pellets of bacteria expressing MBP-His_6-tagged influenza gluebodies were resuspended to homogeneity in 25 mL MBP wash buffer (10 mM HEPES-NaOH (pH 7.5), 500 mM NaCl, 5 % (v/v) glycerol) and supplemented with 0.1 % Triton X-100, 5 mM DTT, 2 µL 50,000 U Supernuclease (SinoBiological), and one cOmplete EDTA-free Protease Inhibitor Cocktail tablet (Sigma). The resuspended bacteria were sonicated and centrifuged (35,000 g, 45 min, 4˚C). 1.5 mL of Ni-NTA resin equilibrated in MBP wash buffer was added per litre of bacterial cell culture grown. Protein was allowed to bind the Ni-NTA resin for 2 hours at 4˚C rotating. The bound resin was washed thoroughly with MBP wash buffer supplemented with 30 mM imidazole. Pure protein was eluted from Ni-NTA resin in two 3 mL MBP wash buffer washes supplemented with 300 mM imidazole. The imidazole was removed by use of a PD-10 desalting column (Cytiva). MBP and His_6 tags were removed by the addition of 0.5 mg of TEV protease. The protein – protease mix was incubated overnight rotating in a 4˚C cold room, after which reverse-IMAC was carried out to remove the contaminating TEV protease and cleaved tags.

Pellets of bacteria expressing streptavidin-tagged influenza gluebodies were resuspended to homogeneity in 25 mL Strep wash buffer (100 mM Tris-HCl (pH 8.0),

400 mM NaCl, 5 % (v/v) glycerol, 100 mM KCl, 10 g/L glycine) and supplemented with 0.1 % Triton X-100, 5 mM DTT, 2 µL 50,000 U Supernuclease (SinoBiological), and one cOmplete EDTA-free Protease Inhibitor Cocktail tablet (Sigma). The resuspended bacteria were sonicated and centrifuged (35,000 g, 45 min, 4˚C). 1.5 mL of Streptactin-XT resin (IBA) equilibrated in Strep wash buffer was added per litre of bacterial cell culture grown. Protein was allowed to bind the streptavidin resin for 2 hours at 4˚C rotating. The bound resin was washed thoroughly with Strep wash buffer. Relatively pure protein was eluted from Streptavidin resin in three 3 mL washes supplemented with 10 mM, 25 mM and 50 mM biotin.

Purified gluebodies - both MPB-His_6-tagged and strep-tagged – then underwent dimerisation into di-Gluebodies. Gluebodies were first diluted to ≤ 0.5 mg/mL to avoid precipitation, and 350 µL of 20 mM HEPES-NaOH (pH 9.5) were added to every 3.5 mL of Gb solution and the resulting solution supplemented with a final concentration of 50 µM $CuSO_4$. The resulting solution was incubated in a 37˚C incubator for 30’ and left overnight at room temperature. The formation of di-Gluebodies was confirmed both by means of an SDS-PAGE gel, and (in the case of the streptavidin-tagged di-Gluebody) a western blot. Size-exclusion chromatography using an S75 10/300 increase GL column was carried out where mixed populations of gluebodies and di-gluebodies were observed. Fractions containing di-Gluebodies were pooled and flash-frozen in liquid nitrogen, and stored at -80˚C.

2.3. Protein Analysis

2.3.1. SDS-PAGE

Protein purity was confirmed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Briefly, proteins were incubated with Laemmli sample buffer (with added β-mercaptoethanol (reducing agent) unless explicitly stated) for 5 minutes at 95˚C. The denatured proteins were loaded onto 4 – 12.5 % precast gradient gels (Invitrogen) and separated by means of electrophoresis in a Mini Gel Tank (Invitrogen) using MES buffer (Invitrogen) until the dye running front reached the end of the gel. Unless otherwise stated, Precision Plus Dual Protein Ladder (Bio-Rad) was used to estimate molecular weight of analyte proteins. Proteins were visualised by staining with Coomassie Instant Blue protein stain (Abcam).

2.3.1.1. Silver Staining

In cases where protein concentration was below the sensitivity limit of Coomassie staining (e.g. when working with viral lysate or purified vRNPs), silver staining was instead used. In these cases, proteins were visualised by silver staining using SilverXpress (Invitrogen) according to the manufacturer's instructions.

2.3.2. Western Blot

Where detection of proteins was required by antibodies to confirm protein identity, proteins were first separated by means of SDS-PAGE and then transferred onto a

nitrocellulose membrane using a Mini Blot Module (Invitrogen) following the manufacturer's instructions. After a 1 h transfer run at 10 V, the membrane was incubated in blocking solution for 1 h (5 % milk in phosphate-buffered saline (PBS) + 0.2 % TWEEN-20, PBS-T). The membrane was then washed three times with PBS-T before primary antibody incubation, with the required dilution of primary antibody depending on instructions from antibody manufacturer. The primary antibody was incubated with the membrane with shaking for 1 h. The bound membrane was then washed twice with PBS-T and then again twice with PBS. Where a secondary antibody was required, the secondary antibody was added and incubated for 1 h, following which the membrane was washed twice with PBS-T and twice with PBS. Clarity Enhanced Chemiluminescence (ECL) substrate (Bio-Rad) was prepared and incubated on the membrane for 5 minutes. The chemiluminescent film was then immediately imaged using an iBright system (Invitrogen).

2.3.3. Size-Exclusion Chromatography (SEC)

SEC was carried out using specified SEC buffers on equilibrated Superdex 200 Increase 10/300 GL, Superdex 75 Increase 10/300 GL, or in the case of analytical SEC, Superdex 200 Increase 5/150 GL columns using an ÄKTA pure chromatography system (Cytiva) at 4˚C with a flow rate at 0.5 mL/min (0.25 mL/min for the analytical Superdex 200 Increase 5/150 GL). In every case, column backpressure was monitored and maintained below 1.5 mPa by the ÄKTA pure system's variable flow rate. Between 250 and 1000 µL of sample was applied to the column for SEC, and 100 – 250 µL for analytical SEC. 280 nm UV absorbance was plotted against elution volume (mL) in

resulting SEC traces processed using Unicorn 7 (Cytiva) and GraphPad Prism 10 (GraphPad). 500 µL fractions were collected in an isocratic column elution of 1.5 column volumes in 500 µL volume in the case of SEC, and 70 µL volume in the case of analytical SEC. Fractions of interest were analysed by means of SDS-PAGE.

2.3.3.1. Size-Exclusion Chromatography – Multi-Angle Light Scattering (SEC-MALS)

SEC-MALS was performed by Dr. David Staunton at the Department of Biochemistry, University of Oxford. A Superdex 200 HR 10/300 column (GE) was equilibrated in specific SEC buffers and sample flown at a flow rate of 0.5 mL/min using a Shimadzu liquid chromatography system. Elution was monitored by on-line light scattering (DAWN – HELEOS 8+, Wyatt Technology), differential refractive index (Optilab T – rEX, Wyat Technology), and UV (SPD – 20A, Shimadzu) detectors. Proteins were prepared at stated concentrations and data were analysed using the ASTRA software package v6 (Wyatt Technology).

2.3.4. Protein Interaction Analyses

2.3.4.1. Resin-Based Pull-Down Studies

Where resin-based pull-down studies were carried out, the choice of resin used for ligand immobilisation was based on ligand tag – Ni-NTA resin (Quiagen) for His_6-tagged proteins, and Glutathione Sepharose 4B resin (Cytiva) for GST-tagged proteins. Resin was equilibrated with binding buffer, rotating at 4˚C for 30 minutes and collected by

centrifugation (800 g, 3 minutes, 4˚C) and supernatant tipped off into waste. Ligand was added and the total volume in the Eppendorf tube made up to 1 mL with selected binding buffer. Binding was allowed to occur for 1 hour, rotating at 4˚C and bound resin collected by centrifugation. Flow-through was collected and SDS-PAGE samples made. Analyte was added in between 1-5 fold molar excess, the volume made back up to 1 mL with binding buffer, and analyte binding was allowed to occur for a further hour. Bound ligand sample was collected by centrifugation as above and flowthrough fraction taken for SDS-PAGE analysis. The bound resin was washed three times with binding buffer in 1 mL volumes. Bound resin samples were taken for SDS-PAGE analysis.

2.3.4.2. SEC-Based Interaction Studies

Where SEC was used to evaluate interaction partners, each individual protein was run on a Superdex 200 Increase 5/150 GL column at 0.25 mL/min to obtain a 'blank' peak in order to identify unbound protein sample (without fraction collection). Putative interaction partners were then incubated for 30 minutes at 4˚C in SEC buffer (25 mM HEPES-NaOH (pH 7.0), 150 mM NaCl). 100 – 250 µL of incubated sample were then applied to the column and a 1.5 column volumes isocratic elution followed. 280 nm UV absorbance was plotted against elution volume (mL) in resulting SEC traces processed using Unicorn 7 (Cytiva) and GraphPad Prism 10 (GraphPad). Fractions of 70 µL were collected and fractions of interest identified, and SDS-PAGE gels run to analyse the proteins present in these samples.

2.3.4.3. Bio-Layer Interferometry (BLI)

Quantitative and qualitative analyses of the interaction kinetics between recombinant his-tagged immune factors (MxA, BTN3A3 B30.2 domain) and monomeric GST-tagged influenza virus nucleoproteins (S09, Tk, NT60) were performed using an Octet Red 96e system (ForteBio). Proteins were diluted in a standard BLI assay buffer (20 mM HEPES-NaOH (pH 7.5), 0.2 % BSA, 0.2 % Tween-20, 2 mM DTT and 150 mM NaCl). A standard concentration (20 µg/mL) of ligand was immobilised on Anti-GST BLI tips (Sartorius) and a series of five 1:2 serial dilutions of analyte (from 60 µg/mL of analyte to 3.75 µg/mL of analyte) were flown over the tips to determine binding kinetics. Where non-specific binding of analyte to tip was detected, an optimisation screen was run with increased BSA and salt concentrations to determine the best conditions for interaction kinetics determination with minimal background binding. Where non-specific binding of analyte was present, a blank experiment was run in tandem with no ligand but varying analyte concentrations according to the concentration of analyte in the tandem experiment to determine non-specific binding response. This was then subtracted from the tandem experiment to remove background binding signal. The final condition used in each experiment – reflected in the assay buffer – was determined by considering the effect of salt and on the interaction. BLI curves were fitted using a 1:1 binding model and Kds were calculated using ForteBio Data Analysis software. Salt concentrations were varied as stated in experiments.

2.4. Viral Cell Culture

2.4.1. Influenza Virus Reverse Genetics

Influenza A viruses were generated by means of the 8-plasmid system (Hoffmann et al., 2000) using the wild type WSN (Influenza A/WSN/1933 (H1N1)) reverse genetics plasmids (pHW2000-WSN-PB1/PB2/PA/NP/HA/NA/M/NS). Where chimeric WSN backbone S09 (A/Green-Winged Teal/Ohio/175/1966) NP virus was generated, 7 WSN plasmids were used (the above, minus pHW2000-WSN-NP) and instead, pCAGGS-NP_S09 and pHH21-NP_S09 were used to produce the chimeric avian NP-containing virus. Approximately 0.6 x 10^6 HEK293T cells were transfected with 0.5 µg of each plasmid using Lipofectamine 2000 Transfection Reagent (Thermo Scientific) following the manufacturer's instructions and plated in a 6-well plate. The cells were incubated at 37˚C, 5 % CO_2 for 6 h before replacing DMEM media supplemented with 10 % (v/v) FCS with fresh DMEM media supplemented with 0.5 % (v/v) FCS. The cells were incubated for an additional 48 h before supernatant was collected, and cell debris removed by brief centrifugation at 2,800 x g (Hettich ROTINA 420, rotor 5624). The virus stock was stored at -80˚C until further use.

2.4.2. Influenza Virus Propagation

Strains of influenza were propagated by infecting 80 % confluent MDBK cells at a multiplicity of infection (MOI) of 0.01. Supernatant was collected 48 h post-infection (hpi), and cell debris was removed by brief centrifugation at 2,800 x g (Hettich ROTINA

420, rotor 5624) for 10 minutes at 4˚C. Virus stock was aliquoted and stored at -80˚C until further use. Viral titres were measured by a standard plaque assay.

2.4.3. Virion-Derived Influenza Virus vRNP Isolation

Virion-derived vRNPs were prepared as per previous published literature (Vreede and Brownlee, 2007). Infected MDBK cell culture supernatant 48 hpi was harvested and cell debris removed by brief centrifugation at 2,800 x g (Hettich ROTINA 420, rotor 5624) for 10 minutes at 4˚C. A 5 mL cushion of 30 % sucrose in resuspension buffer (10 mM Tris-HCl, pH 7.4, 100 mM NaCl, 1 mM EDTA) was added to thick-walled ultracentrifuge tubes and clarified infected cell supernatant was carefully added on top of the sucrose cushion and the resultant mixture was separated by ultracentrifugation in an SW32 rotor at 100,000 x g for 1 h 30 at 4˚C. The supernatant and cushion were aspirated and 80 µL of chilled resuspension buffer added (10 mM Tris-HCl (pH 7.4), 100 mM NaCl, 1 mM EDTA) and the tubes were sealed and incubated at 4˚C overnight for virion resuspension. The virion suspension was added into a microfuge tube and a 1:1 ratio of 2 x disruption buffer was added to lyse virions (200 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10 mM $MgCl_2$, 2 % Triton X-100, 10 % (v/v) glycerol, 1 % IGEPAL). The resulting solution was incubated at 31˚C for 30 minutes with vigorous shaking in a thermomixer to ensure complete virion lysis. 1 mL of each 70 %, 50 %, 40 %, and 33 % glycerol in NM buffer (50 mM Tris-HCl (pH 7.4), 150 mM NaCl) were added to a thin-walled ultracentrifuge tube to form a glycerol gradient. 1 mL of lysed virus was added on top of the glycerol gradient and balanced. The gradient was spun in a Beckman Ultracentrifuge using an SW 55 rotor (200,000 x g, 3 h 45 min, 4˚C). A hole was pierced

with a thin needle in the base of the ultracentrifuge tube and approximately 250 µL fractions collected into chilled Eppendorf tubes. These were vortexed and SDS-PAGE samples taken for silver staining. The fractions were stored at -20˚C until further use. In some cases, vRNPs were further concentrated using an Optimax-XP bench top ultracentrifuge and a TLA 100.3 fixed angle rotor (with 3.5 mL volume thick-walled polycarbonate tubes. vRNPs were concentrated by pelleting vRNP-containing fractions for 4 h at 500,000 x g at 4˚C. vRNPs were resuspended overnight at 4˚C in 25 µL 20 mM Tris-HCl (pH 8.0), 150 mM NaCl.

2.4.3.1. Influenza Virus vRNP Coating with Binding Partners

In some cases, after the virion lysis step outlined above, recombinantly produced vRNP binding partners (Nb170, MxA) were added in excess. Their binding to vRNPs was validated by SDS-PAGE, silver staining and western blot as appropriate.

2.5. Structural Biology Techniques

2.5.1. Macromolecular Protein Crystallography

Crystallisation trials were carried out using 96-well vapour diffusion sitting drop format crystal trays (either Greiner Bio-One, or SWISSCI) in the first instance, using the full range of crystallisation screens available at the Division of Structural Biology (both Hampton Research screens and personalised screens established by Dr. Thomas Walter). Screening was performed by mixing 100 nL of reservoir solution with 100 nL of protein solution (between 5 – 10 mg/mL) in the central shelf of CrystalQuick 96-well

Greiner plates (Greiner) or similarly with SWISSCI 3 lens crystallisation plates (SWISSCI) using a Cartesian Technologies robotics system (Genomic Solutions) (Walter et al., 2005). Crystal screen reservoirs contained 95 µL of mother liquor. Crystallisation conditions were improved by carrying out optimisation screens by varying concentrations of key precipitation solutions in the mother liquor. To cryo-protect crystals, solutions of mother liquor were supplemented with 15 % (v/v) ethylene glycol and added to the crystal drop before mounting on crystal loops. Crystals were mounted onto appropriately sized crystal loops (MiTeGen) and rapidly flash-cooled by plunging into liquid nitrogen. All crystal diffraction data were collected at Diamond Light Source, Didcot, UK. Diffraction data were auto processed with xia2_dials (Gildea et al., 2022), autoPROC (Vonrhein et al., 2011), autoPROC+STARANISO (Vonrhein et al., 2018), and xia2_multiplex (Gildea et al., 2022). In general, programs built into the PHENIX suite (Phenix 2.21.2) (Liebschner et al., 2019) were used for structure solution, refinement, and analysis. AlphaFold3 models (Abramson et al., 2024) were used for initial molecular replacement, and manual model building was performed using Coot (0.9.6) (Emsley and Cowtan, 2004). More detailed information, including beamline and software information, as well as crystallographic tables, is found in the relevant sections in chapter 3.

2.5.2. Cryo-Electron Microscopy (cryo-EM)

Cryo-EM experiments were carried out at the Oxford Particle Imaging Centre (OPIC) unless stated. Unless stated otherwise, Dr. Loïc Carrique was present to carry out data collection.

As standard, EM grids were glow discharged and then mounted on a Vitrobot Mark IV (FEI). 3.5 µL of sample between 0.3 and 0.6 mg/mL as standard was applied to each grid and the grids were blotted for 5.5 seconds in a humidity-controlled environment (95 – 100 %) at 6˚C. Following blotting, grids were plunge frozen into a cryo-cooled liquid ethane cup and subsequently kept in liquid nitrogen. Screening of cryo-EM grids was carried out using a Thermo Fisher Glacios 200 keV Transmission Electron Microscope at OPIC, equipped with a Falcon 3 direct electron detector. Large dataset collection at either OPIC or the electron Bio-Imaging Centre (eBIC) was carried out on a Thermo Fisher Titan Krios 300 keV Transmission Electron Microscope equipped with a Falcon 4i direct electron detector and Selectris-X energy filter (in the case of OPIC), or a Gatan K3 direct electron detector (in the case of eBIC). Data were recorded using EPU software (Thermo Fisher) at defocus and magnification values were selected in a sample-dependent manner. Further data processing was mostly carried out in CryoSPARC (Punjani et al., 2017) but also to a limited extent in Relion 5.0 (Zivanov et al., 2018). More detailed information, including on specific data collection parameters and software used is found in the relevant sections of chapters 3 and 5.

2.5.2.1. Cryo-EM Structure Model Building

Initial structure models were generated using AlphaFold3 (Abramson et al., 2024) or ModelAngelo (Jamali et al., 2024), as stated. Programs built into the PHENIX suite (Phenix 2.21.2) (Liebschner et al., 2019) were used for structure refinement and model validation. Iterative rounds of manual model building and refinement were performed

using Coot (0.9.6) (Emsley and Cowtan, 2004) and PHENIX to yield the final reported structure. More detailed information is found in the relevant sections in chapter 3.

2.5.3. Functionalised Electron Microscopy Grids

2.5.3.1. Functionalisation of Electron Microscopy Grids

Functionalised EM grid preparation was carried out by Dr. Gangshun Yi.

DOPC (18:1 (Δ9-Cis) PC) (DOPC, Avanti) and DOGS (Avanti) was mixed to the final concentration of 2 mg/mL. Polymerase (Nb8210) or NP (Nb170) nanobodies were prepared in a buffer containing 50 mM HEPES-NaOH (pH 7.2), 150 mM NaCl. 28.4 µL of nanobody-containing solution of either 0.02 mg/mL or 0.06 mg/mL, as reported, was added into a Teflon well (on ice) and 0.9 µL of lipid mixture was added onto the surface of the nanobody solution. The Teflon block was incubated in a humidity chamber and incubated for 2 hours. Holey carbon grids (Quantifoil 1.2/1/3 or 2/2, without glow discharging) were laid onto the surface of the Teflon well (C-face down) for 2 minutes to establish a lipid monolayer on the grid. The grid was washed with three 120 µL drops of sample buffer to remove free nanobody. 5 – 10 µL of sample was added and incubated for 15 minutes prior to washing with one 120 µL drop of sample buffer to remove unbound particles. Grids were subsequently plunge frozen using a Thermo Fisher Vitrobot IV (time: 2.5 s, force: -20, temperature: 20˚C, humidity: 100 %). Screening was caried out on a Thermo Scientific Glacios Transmission Electron Microscope, operating at 200 keV. The images shown in the relevant sections of chapter 5 were captured at a pixel size of 2.0 Å/pixel and a total dose of 50 $e^-/Å^2$.

2.5.3.2. Immobilisation of Influenza Virus Genome Segments on Electron Microscopy Grids

Influenza A vRNP samples were pooled and added to holey carbon grids which had been charged with a lipid monolayer and a nanobody of choice. The grid was washed as above to remove unbound particles before blotting, plunge freezing and screening on a Thermo Fisher Glacios TEM. Upon positive assessment of grid and sample quality, grids were either collected on as a single particle cryo-EM dataset or tilt series collected for cryo-ET data processing.

2.5.4. Cryo-Electron Tomography (cryo-ET)

Cryo-ET data collection was carried out by Dr. Loïc Carrique and data processing was greatly aided by Misha Le Claire.

Cryo-ET data were collected using SerialEM – PACE-tomo (Eisenstein et al., 2023). A dose-symmetrical tilt scheme covering +/- 54˚ range with 2˚ steps was used to collect data. Tilt series were recorded at a nominal magnification of 105,000 x resulting in a pixel size of 1.187 Å, with a dose of 3 $e^-/Å^2$ per tilt, resulting in a total dose of 138.6 $e^-/Å^2$. Data were pre-processed using Warp (Tegunov and Cramer, 2019). EER files were motion corrected using MotionCorr2 within Warp (Zheng et al., 2022) using half map creation from frames for downstream denoising. CTF estimation was carried out using Warp and initial tilt series alignment and subsequent tomogram reconstruction was carried out using AreTomo (Zheng et al., 2022). Tomograms were then assessed for

quality in IMOD (Mastronarde and Held, 2017). Tomograms which had good sample but were not well aligned were realigned manually using patch-based tracking in Etomo (Mastronarde and Held, 2017). Tomograms were denoised using cryoCARE (Buchholz et al., 2019) and downstream sub-tomogram averaging was carried out using Relion 5.0 (Zivanov et al., 2018).

2.5.5. Structural Figures

Figures of structures were made in ChimeraX (Pettersen et al., 2021) unless explicitly stated, with protein models not solved in this work sourced from the Protein Data Bank (PDB) (www.rcsb.org) and electron density maps not solved in this work sourced from the Electron Microscopy Data Bank (www.ebi.ac.uk/emdb).

3. Methods to Study Influenza Nucleoprotein Structure

3.1. Introduction

3.1.1. Structural Studies of the Influenza Nucleoprotein

The biological mechanisms of proteins are intrinsically linked to their three-dimensional structure. This underlies the necessity of high-resolution structure determination in order to elucidate information on the molecular mechanisms of protein action. At 56 kDa, the influenza nucleoprotein is an ideal candidate for high-resolution structure determination by X-ray crystallography (XCR). This technique relies on the formation of well-ordered protein crystals to achieve atomic resolution structural information.

When initial crystallisation attempts prove unsuccessful, several established strategies can improve the success rate of further crystallisation trials, including flexible loop truncation (Malawski et al., 2006), buffer and additive optimisation for improved protein stability (Boivin et al., 2013), and systematic crystal screening approaches (Budziszewski et al., 2023). Provided that the protein forms an ordered crystal lattice capable of high-resolution X-ray diffraction, routine structural determination is achievable (Su et al., 2015; Maveyraud and Mourey, 2020). However, inherent protein structural heterogeneity can hinder the formation of protein crystals,

making successful crystallisation uncertain even for proteins with high sequence identity to proteins whose structures have been solved by XCR before.

Another now common approach to determine protein structure is single particle analysis (SPA) cryogenic electron microscopy (cryo-EM). SPA enables the structural determination of proteins in a near-native state. In SPA, particles are applied to a cryo-EM grid, plunge frozen into liquid ethane, and suspended in a thin layer of vitreous ice between 50 – 200 nm thick. The grid is loaded into a cryo transmission electron microscope under vacuum and data are collected. Data collection employs low-dose conditions (around 40 – 60 $e^-/Å^2$) distributed across multiple frames to enable motion correction while minimising radiation damage. Due to the low electron dose, there is a low signal to noise ratio (SNR) in the images obtained. Given an ideal sample, there will be random orientation of protein samples in the vitreous ice layer. This is useful to tackle the low SNR, as particles – assumed to be identical – are computationally averaged across their many different orientations, improving the SNR of the data. However, high-resolution 3D structure determination relies on high particle number and good sampling of all possible orientations. Two problems that commonly occur in SPA are preferential orientation and adsorption onto the liquid-air interface. These arise as proteins, especially those with exposed hydrophobic patches, readily adsorb at the liquid-air interface, suspending the protein in a preferred orientation, and so limiting the number of different views of the protein. This impacts the 3D reconstruction of the sample, leading to resolution anisotropy in the final map. These two challenges are overcome by a number of methods including obtaining thicker ice to allow the protein sample to occupy more orientations by blotting for less time or

with less force, by using detergents to coat the air-water interface and reduce protein adsorption, or by using support film-coated grids to adsorb proteins onto the support films (e.g. carbon, gold).

One key stipulation is that, while the final reconstructed map is often called an “electron density map”, it is in fact an “electron potential map” (or more correctly, a Coulomb potential map), which is distinct from the electron density maps produced by X-ray crystallography. In XCR, incident X-rays are scattered by the electrons of atoms in the crystalline sample, producing maps of electron density. In cryo-EM, incident electron beams are scattered by their interaction with the electrostatic potential field created by nuclei and electrons in the sample. Transmitted electrons are detected, and contrast is produced between areas with higher and lower electrostatic (Coulombic) potential, yielding a Coulomb potential map, commonly called a density map (Marques et al., 2019).

Cryo-EM SPA has the benefit of potentially enabling the structure determination of the different states of a protein in the acquired dataset. Heterogeneity in the acquired data, often due to differing conformational states of proteins, can be separated in 2D and 3D classification. This allows for selection of specific classes of particles which adopt a certain conformation without having polluting conformers of the protein or complex in the particle stack. The cost to this is that the greater the number of potential conformational states of the protein, or the greater the flexibility of the protein, the lower the numbers of particles will be in a given conformational class, and as such,

the harder it is to determine protein structure to a high-resolution given the lack of homogeneous data to average.

In 2022, fewer than 3 % of human protein entries in the Electron Microscopy Data Bank (EMDB) below 100 kDa achieved resolutions better than 4.0 Å (Wentinck et al., 2022) despite nearly 75 % of the proteins in the known human proteome being less than 50 kDa (Brocchieri and Karlin, 2005). A size of 38 kDa represents a likely theoretical minimum molecular weight for structure elucidation (Henderson, 1995). Current practical limitations restrict high-resolution determination of protein structure by cryo-EM SPA to proteins above approximately 50 kDa (Herzik et al., 2019). These constraints primarily arise from low signal-to-noise ratios, which can limit robust reconstructions. Consequently, high-resolution structure determination by cryo-EM SPA becomes challenging for proteins of a molecular weight below 100 kDa.

3.1.2. Current State of Knowledge of Influenza Nucleoproteins

The focus of this thesis is on the RNA-binding nucleoprotein of Influenza virus, introduced fully earlier, which can be found in multiple oligomeric states in a salt-dependent, or phosphorylation dependent manner (Chenavas et al., 2013). Crystal structures of recombinantly-expressed influenza A nucleoproteins have been solved both as monomers (Knight et al., 2021) (PDB: 7NT8), and as trimers (even with the putative R416A monomeric mutation included) (Ye et al., 2006) (PDB: 2IQH) (Ng et al., 2008) (PDB: 2Q06) (Chenavas et al., 2013) (PDB: 3ZDP); (Hanke et al., 2016) (PDB:

5TJW). Nucleoprotein oligomerisation can also be observed to a greater extent in both the absence and presence of RNA (Ye et al., 2006) using negative stain electron microscopy, with a bias for higher order oligomers with the addition of RNA (Ng et al., 2008). One tetrameric influenza B nucleoprotein structure has been reported using X-ray crystallography (Ng, Lam, et al., 2012) (PDB: 3TJ0), and one tetrameric influenza D nucleoprotein has also been similarly determined (Donchet et al., 2019) (PDB: 5N2U). Only two crystal structures exist of an influenza nucleoprotein (of any genus) with RNA bound, published in 2021 with three nucleotides bound to the putative RNA-binding groove (Tang et al., 2021) (PDB: 7DKG, 7DXP).

At the beginning of this work, structural information on influenza nucleoproteins was limited to low-resolution models of purified vRNP complexes (Coloma et al., 2020) (PDB: 6I54, 6I85). However, in 2025, two important studies were published which reported the high-resolution structure of the non-physiological pseudo-vRNP complex (Chenavier et al., 2025) (PDB: 9GAS), and the high-resolution structure of reconstituted influenza D RNP (Peng et al., 2025) (PDB: 9BWV). However, to date, no monomeric influenza nucleoprotein structure has been reported to high-resolution when studied using cryo-EM SPA. The challenges of determining the high resolution structure of such a small, flexible protein have been outlined above. However, this poses an interesting challenge and offers potential in studying the structure of monomeric nucleoproteins and their binding partners, provided a stable scaffold is able to be produced.

3.1.2.1. RNA Binding to the Influenza Nucleoprotein

The nucleoprotein has been recognised for decades to be important for influenza virus transcription and replication. While the influenza polymerase complex alone is able to efficiently transcribe short RNA templates, the nucleoprotein is required for the synthesis of longer vRNAs (Turrell et al., 2013). This requirement establishes the vRNP complex, rather than the naked vRNA, as the functional template for influenza virus transcription and replication (Bishop et al., 1971; Honda et al., 1988). The observation that nucleoprotein-associated RNA remains susceptible to Rnase treatment indicates that RNA binds to the nucleoprotein's solvent-exposed surface (Baudin et al., 1994).

Many investigations have helped to map nucleoprotein – RNA contacts residues. UV crosslinking and chemical fragmentation studies by Elton and colleagues identified key interaction-disrupting residues: W120, W139, R267, S314, W330, A332, W386, F412, and R416 (Elton et al., 1999). The first nucleoprotein crystal structure, solved by Ye and colleagues in 2006 (figure 3.1, panel A), revealed a prominent surface groove enriched in basic residues including R65, R150, R152, R156, R174, R175, R195, R199, R213, R214, R221, R236, R355, K357, R361 and R391. This positively charged groove was proposed to mediate RNA binding through electrostatic interactions with the negatively charged RNA phosphodiester backbone (Ye et al., 2006) (PDB: 2QIH). Surface Plasmon Resonance (SPR) studies by Ng and colleagues provided quantitative validation of these structural predictions (Ng et al., 2008). They found the residues R174, R175, and R221 in the NP-G1 groove to be essential for RNA binding by mutagenesis. Where these residues were mutated to alanine, there was almost a

complete abolition in RNA binding. Additionally, mutations of residues R150, R152, R156, and R162 in the NP-G2 groove reduced RNA affinity five-fold.

Comprehensive mutational analysis of conserved residues across the nucleoprotein found the following individual mutations to lead to nonviable mutant virus, which supports their proposed critical roles in RNA binding: D72A, G93A, K113A, Y148A, R150A, R152A, R156A, R174A, R195A, R199A, R208A, R213A, E254A, A260R, K273R, K325A, A337R, E339R, R355A, R361A, R387A, Q405A, F412A, R416A, F488A, and F489A. The tail loop region was shown to be indispensable for both replication and transcription (Li et al., 2009).

It is known that RNA binding occurs in a sequence non-specific manner (Williams et al., 2018) with the wild-type nucleoprotein exhibiting high-affinity binding in the low nanomolar range (Labaronne et al., 2016). However, monomeric nucleoprotein is known to bind RNA with much reduced affinity (Elton et al., 1999; Turrell et al., 2013; Chenavas et al., 2013), indicating that oligomerisation is prerequisite for RNA association.

Tang and colleagues solved the first structure of an influenza nucleoprotein with RNA bound (figure 3.1, panel I), revealing three nucleotides bound within the predicted binding groove given a tail-loop deletion (Δ402-428) and an additional K87E mutation. Using microscale thermophoresis, Tang and colleagues also demonstrated that a synthetic peptide of the C-terminal 8 residues of the nucleoprotein (DNAEEYDN)

competitively inhibits nucleoprotein-RNA interactions, reducing binding affinity eight-fold in tail-loop truncation mutants (Tang et al., 2021).

Structural analysis of monomeric nucleoprotein reveals that the C-terminus adopts a conformation that occludes the RNA-binding groove (Knight et al., 2021). This has also been observed in the monomeric, but not the oligomeric WSN R416A structure. It has been suggested that this mechanism reduces the positive charge in this groove, and thus, prohibit RNA binding (Chenavas et al., 2013). Hence, this model suggests the tail-loop as a critical regulatory element controlling nucleoprotein – RNA interactions through oligomerisation-dependent conformational switching.

3.1.2.2. Oligomerisation of the Influenza Nucleoprotein

Deletion of the conserved C-terminal loop of the influenza nucleoprotein (Δ402-428), or simply individual point mutations of either E339A, or R416A, have been shown to abolish influenza A nucleoprotein oligomerisation (Ye et al., 2006). This dependence on tail-loop mediated interactions establishes the tail-loop binding pocket as a compelling target for antiviral drug development (Taft et al., 2025).

Structural investigation by Ng and colleagues revealed that the nucleoprotein homo-oligomerisation proceeds through insertion of the tail-loop into a complementary groove within the body domain of the adjacent nucleoprotein (Ng et al., 2008).

Chan and colleagues elucidated the molecular determinants of this critical interaction, demonstrating that oligomerisation depends on a salt bridge formed by residues E339 in the body domain of one nucleoprotein, and R416A of the tail loop of the neighbouring nucleoprotein (Chan et al., 2010). Through systematic point mutagenesis and RNP reconstitution assays, they showed that the E339-R416 ion pair form an essential inter-subunit contact. Additionally, hydrophobic interactions mediated by residues V408, P410, L418, and P419 provide crucial stabilising contacts for nucleoprotein homo-oligomerisation. Disruption of these interactions through mutations E339A, R416A, and V408S, L418S, P419S resulted in complete loss of oligomerisation capacity while preserving the structural integrity of the tail-loop.

These findings demonstrate that tail-loop insertion into the neighbouring nucleoprotein is absolutely required for oligomerisation, and that individual point mutations in highly conserved residues can completely abolish this interaction despite maintaining overall structural architecture (Chan et al., 2010). The strict conservation and functional importance of these molecular contacts underscore their potential as targets for structure-based antiviral drug design.

The figure below (figure 3.1) and accompanying table (table 3.1) illustrate all structures of influenza nucleoproteins deposited in the PDB solved by either X-ray crystallography (the vast majority of such examples) or cryo-EM to a resolution better than 8.5 Å. Only one of the solved crystal structures was complexed with RNA, and in this structure (PDB: 7DXP), only one of the two copies of the nucleoprotein in the asymmetric unit has the three nucleotides of RNA bound. All of the structures solved by cryo-EM have

RNA bound, be they authentic influenza RNAs from a reconstituted RNP or purified vRNP in the case of the work published by Peng and colleagues (Peng et al., 2025), or synthetic 18-mer RNA fragments in the case of the work published by Chenavier and colleagues (Chenavier et al., 2025). In all cases, the solved nucleoprotein structures are highly similar in terms of overall protein structure, whereas the significant differences in protein structure come from flexible loops.

At the outset of this work, only structures in figure 3.1 panels A-G had been deposited in the PDB, and no structure of RNA bound to influenza nucleoprotein had been solved by either structural method presented, X-ray crystallography or cryo-EM SPA.

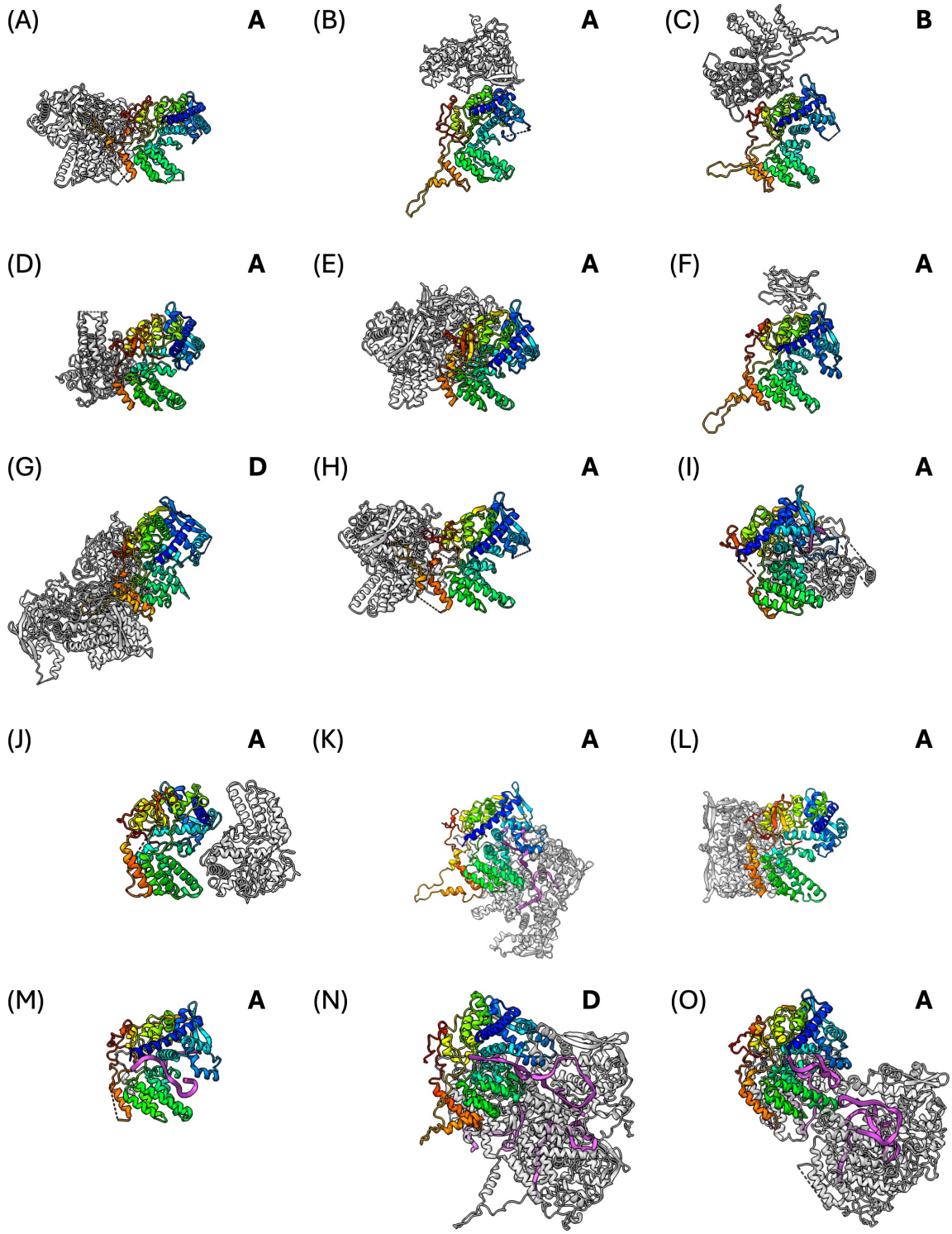


Figure 3.1: Published Structures Determined of Influenza Nucleoproteins.

This figure illustrates example structures of influenza nucleoproteins previously determined (and deposited in the PDB, solved to a resolution of better than 8.5 Å) to demonstrate the globally conserved structure of nucleoproteins across influenza virus genera. All nucleoproteins are shown in the oligomeric context of their deposition on the PDB – be that as a physiological oligomer or a result of crystal packing – with one copy of the nucleoprotein coloured using the rainbow colour command (N- blue, C- red), and any nucleotides shown in magenta. Other copies (if present) are coloured in grey at 50 % transparency with the head and body domains of the rainbow coloured nucleoprotein monomer

pointing to the right, and the coloured nucleoprotein aligned between figure panels. The specific genus of influenza (A, B, or D) is annotated in bold on the right hand side of each panel. Annotation of these structures is shown in table 3.1.

Table 3.1: Annotation of Published Influenza Nucleoprotein Structures.

This table is used for the annotation of the above figure, with PDB codes of the corresponding structures above, the source influenza strain, the resolution to which the structure was determined, the method by which the structural determination was carried out, the biologically-relevant oligomeric form of the structure, whether RNA was present in the structure, and the reference paper in which the structure was published.

PDB Code	Panel in figure 3.1	Strain (Mutations)	Resolution (Å)	Method	Oligomeric Form	RNA Present?	Reference
2IQH	A	A/WSN (H1N1) (wt)	3.20	X-ray	Trimer	No	(Ye et al., 2006)
2Q06	B	A/Hk (H5N1) (wt)	3.30	X-ray	Trimer	No	(Ng et al., 2008)
3TJ0	C	B/Hk (wt)	3.20	X-ray	Tetramer	No	(Ng, Lam, et al., 2012)
4IRY	D	A/WSN (H1N1) (Δ402-428)	2.80	X-ray	Monomer	No	(Ye et al., 2012)
3ZDP	E	A/WSN (H1N1) (R416A)	2.69	X-ray	Monomer	No	(Chenavas et al., 2013)
5TJW	F	A/WSN (H1N1) (wt)	3.23	X-ray	Trimer	No	(Hanke et al., 2016)
5N2U	G	D/bovine/France (wt)	2.40	X-ray	Tetramer	No	(Donchet et al., 2019)
6J1U	H	A/WSN (H1N1) (wt)	2.80	X-ray	Trimer	No	(Yang et al., 2021)
7DXP	I	A/Hk (H5N1) (Δ402-428, K87E)	2.30	X-ray	Monomer	Yes, three ntd's	(Tang et al., 2021)
7NT8	J	A/NT60 (H3N2) (R416A)	2.22	X-ray	Monomer	No	(Knight et al., 2021)
8PZQ	K	A/WSN (H1N1) (wt)	5.30	Cryo-EM	Reconstructed Trimer	Yes	(Chenavier et al., 2023)
8TWP	L	A/Aichi (H3N2) (Δ2-7, P283S, R416A)	2.90	X-ray	Monomer	No	(Yoon et al., 2024)
9GAS	M	A/WSN (H1N1) (Δ1-14)	3.08	Cryo-EM	Pseudo-vRNP Filament Monomer Reconstruction	Yes	(Chenavier et al., 2025)
9BWV	N	D/swine/Oklahoma (wt)	5.10	Cryo-EM	Reconstructed Tetramer	Yes	(Peng et al., 2025)
9BX1	O	A/Hk (H5N1) (wt)	8.00	Cryo-EM	Reconstructed Tetramer	Yes	(Peng et al., 2025)

3.1.3. Rationale of Influenza Nucleoprotein Strains Used

Nucleoproteins from multiple influenza A virus strains were used in this study. The rationale of their use is outlined in the table below (table 3.2) where two human-adapted strains of influenza, NT60 (A/Northern Territory/60/1968 (**H3N2**)), and WSN (A/Wilson-Smith/1933 (**H1N1**)) are noted for use in recombinant expression work (in the case of NT60), and live influenza A work (in the case of WSN). These two strains of influenza A nucleoprotein are known to have resistance mutations to both immune factors used and reported later in this thesis (MxA and BTN3A3), which is highlighted in the sequence alignment below (figure 3.2, panel B). Two avian influenza A strains were used in this study, S09 (A/Green-Winged Teal/Ohio/175/1966 (**H2N1**)) which was utilised for both recombinant protein expression and live virus work, and Tk (A/turkey/Turkey/1/2005 (**H5N1**)) which was used for recombinant protein expression. These two strains are of particular interest, as, being avian influenza strains, they carry the sensitivity mutations to both immune factors used. These mutations are highlighted in the sequence alignment below (figure 3.2, panel B). The sequence identity of these nucleoproteins is shown below (table 3.3), where it can be seen that the two most similar proteins are S09 nucleoprotein and Tk nucleoprotein, at 97.6 % sequence identity. Comparably, NT60 and WSN are 93.4 % sequence identical, with NT60 and S09 being 92.4 % sequence identical and NT60 and Tk being 92.0 % sequence identical. WSN is slightly more similar to the two avian nucleoproteins, with 94.6 % and 94.2 % sequence identity to S09 and Tk respectively.

Table 3.2: Nucleoproteins of Influenza A Virus Strains Used in this Thesis, and the Rationale behind their Use.

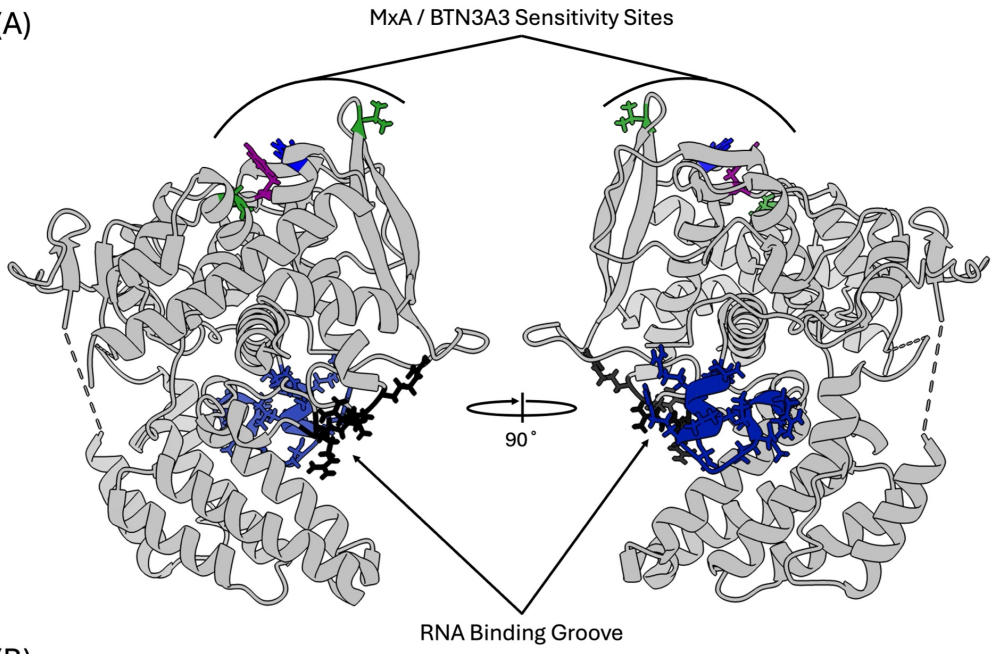
Influenza A Strain	Use Rationale
NT60 (A/Northern Territory/60/1968 (**H3N2**))	Standard human-adapted influenza A virus strain used for recombinant protein expression (resistant to MxA and BTN3A3 restriction).
WSN (A/Wilson-Smith/1933 (**H1N1**))	Standard live influenza A virus strain used in lab work (resistant to MxA and BTN3A3 restriction).
S09 (A/Green-Winged Teal/Ohio/175/1966 (**H2N1**))	Standard avian-adapted influenza A virus strain used for both recombinant protein expression and live virus work (sensitive to MxA and BTN3A3 restriction). No structure of a H2N1 avian nucleoprotein has yet been solved.
Tk (A/turkey/Turkey/1/2005 (**H5N1**))	Novel avian-adapted influenza A virus strain used for recombinant protein expression (sensitive to MxA and BTN3A3 restriction).

Table 3.3: Sequence Identity of Nucleoprotein Strains Used.

Sequence identity compared between NT60, WSN, S09, and Tk influenza A virus strain nucleoproteins used in this thesis. Protein sequence identity calculated by SIM (https://web.expasy.org/sim/) by direct comparison of amino acid code.

NT60	100 %			
WSN	93.4 %	100 %		
S09	92.4 %	94.6 %	100 %	
Tk	92.0 %	94.2 %	97.6 %	100 %
	NT60	**WSN**	**S09**	**Tk**

(A)


MxA / BTN3A3 Sensitivity Sites
90°
RNA Binding Groove

(B)

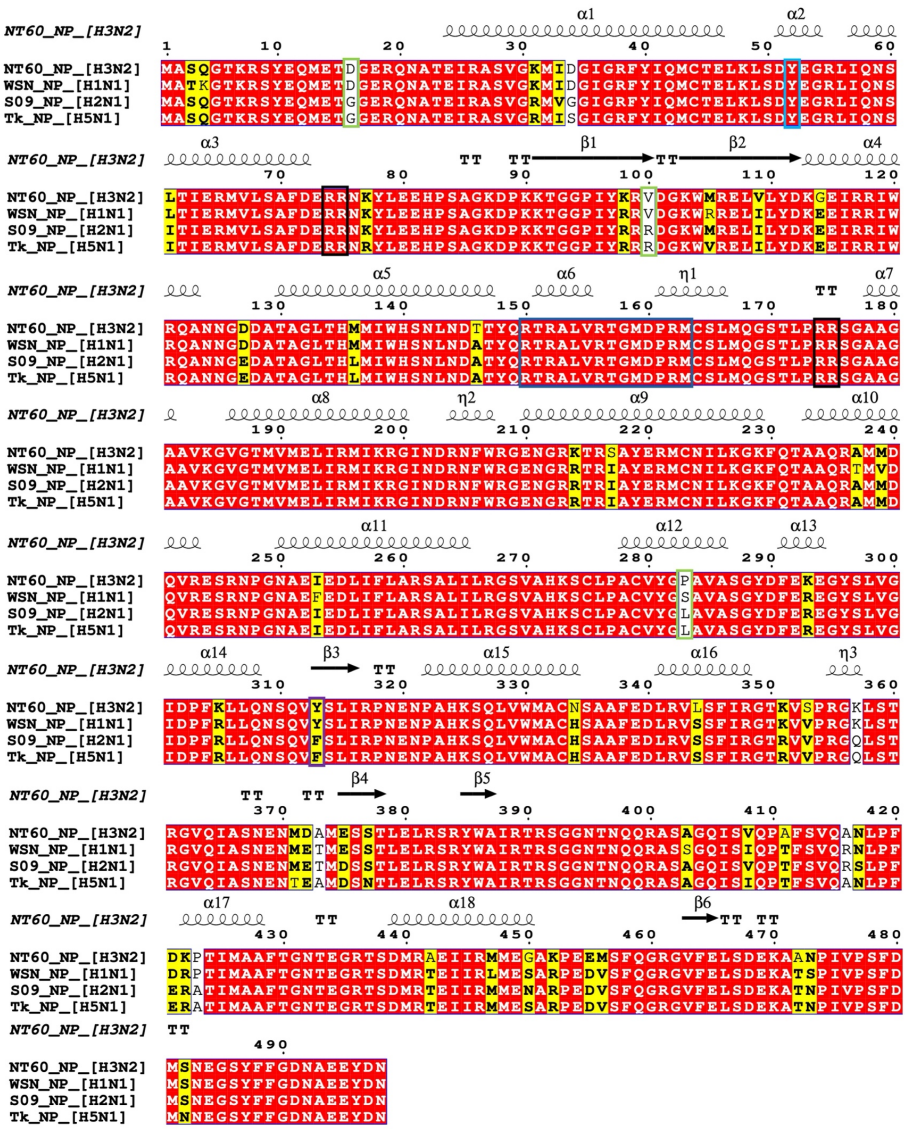

NT60_NP_[H3N2]
WSN_NP_[H1N1]
S09_NP_[H2N1]
Tk_NP_[H5N1]

Figure 3.2: Nucleoprotein Structure and Sequence Alignment.

(A) Structural diagram using a previously solved crystal structure of Influenza A NT60 nucleoprotein (Knight et al., 2021) (PBD: 7NT8) highlighting amino acids involved in the formation of the two RNA binding grooves, the reported BTN3A3, and the reported MxA sensitivity sites. Residues are coloured as in the sequence alignment above, except for residue 16, which was not resolved in the crystal structure. (A) Alignment was carried out using ESPript 3.0 (Robert and Gouet, 2014) and sequence similarity was depicted by calculating percentage amino acid equivalence, which is computed considering physico-chemical properties with a global similarity threshold score of 70 %, as standard. The alignment was coloured using the 'flashy' option, where white characters on red background illustrate strictly conserved residues (strict identity), black characters on a yellow background illustrate chemically similar residues (where bold characters in these instances show strict identity, and normal emphasis characters show chemically similar residues). Black text on a white background illustrates physico-chemically divergent amino acid residues, which are less than 70 % similar. The Greek characters α accompanied by a helical line represent alpha helices, β accompanied by a solid black arrow represent beta sheets, and η accompanied by a helical line represent 3_{10} helices. TT represents strict β turns. Secondary structure prediction was carried out and annotated above each line of sequence alignment by importing the PDB file for an NT60 strain nucleoprotein (PDB: 7NT8). RNA binding grooves G1 74-5, 174-5 and G2 150-63 (Fangzheng Wang et al., 2022) are boxed in black (G1) and dark blue (G2); Innate immune factor sensitivity residues for BTN3A3 alone are boxed in blue (52), for both BTN3A3 and MxA are boxed in purple (313), and for MxA alone are boxed in green (16, 100, 283). Residues determining sensitivity to BTN3A3 restriction are taken from recent work (Pinto et al., 2023) and are reported to be Y 52 NHQ, and F 313 YV. Residues determining sensitivity to MxA are taken from earlier work (Götz et al., 2016) and are reported to be G 16 D, R 100 V, L 283 P, and F 313 Y. (Residues reported before the residue number are taken to be avian mutations which are sensitive to restriction, and residues reported after the residue number are taken to be human adaptations which are resistant to restriction). Residues involved in either RNA or immune factor binding are also shown as atoms.

3.1.4. Nanobody-Derived Tools for Structural Biology

Where proteins which require structural study are smaller than the broad 50 kDa cutoff for cryo-EM SPA (described above) or are too flexible to enable protein crystals to be successfully grown or collect cryo-EM SPA datasets, any number of binding partner complexes may be formed with the protein of interest and a specifically binding scaffold proteins to increase the protein size, reduce preferential orientation within vitreous ice (in the case of cryo-EM), and/or decrease the target protein's overall flexibility. Some such examples used in cryo-EM SPA are outlined in the figure below (figure 3.3).

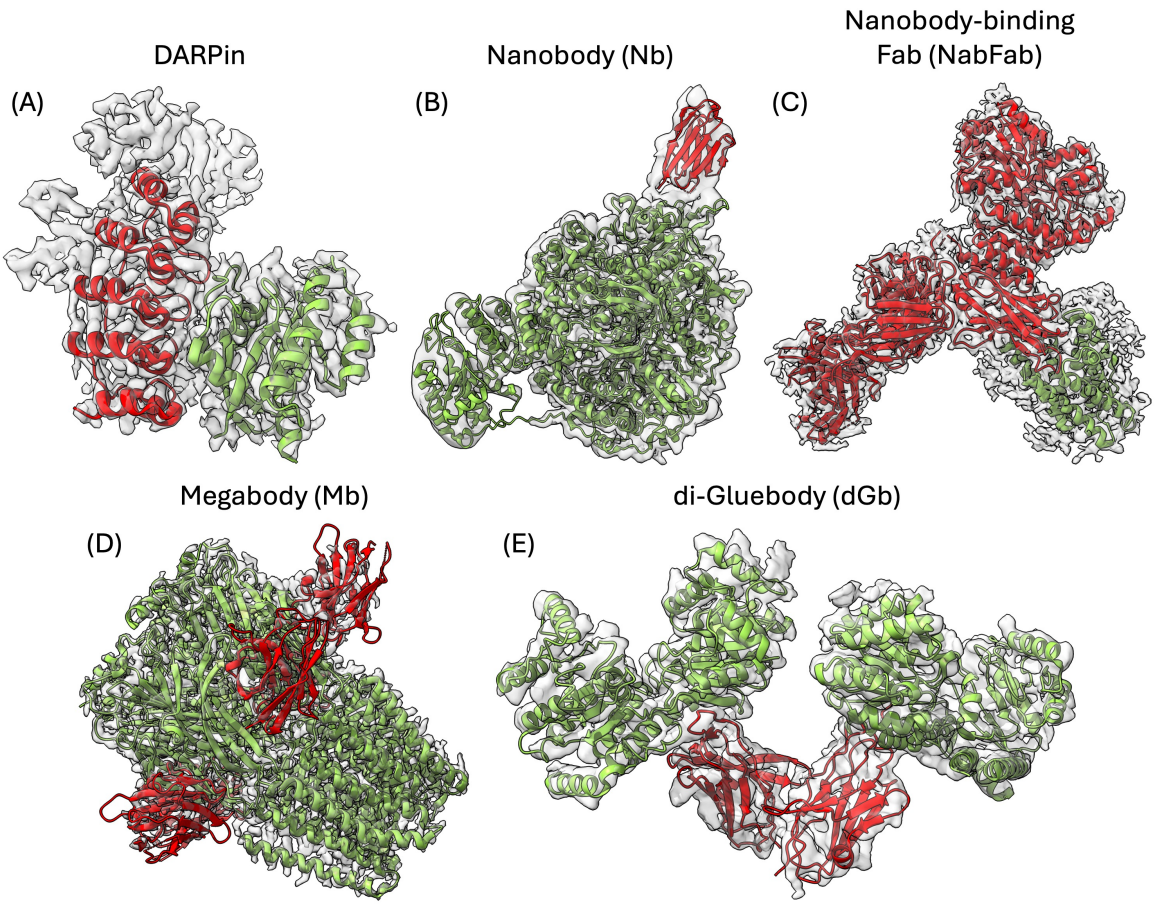


Figure 3.3: Example Protein Scaffold Technologies used for Structural Determination.

This figure illustrates example use cases of small, engineered proteins used as scaffolds to aid structure determination of proteins of interest. The protein scaffolds are coloured in red; the proteins of interest are coloured in green and are shown fit into the cryo-EM map at a transparency of 25 %. (A) The structure of a 19 kDa oncogenic GTPase, KRAS, involved in signal transduction in human cell proliferation pathways, bound to a validated mutant 5-repeat DARPin (18 kDa) forming a rigid scaffold, which is then bound to a rigid protein cage to increase size further (965 kDa complex) (Castells-Graells et al., 2023) (PDB: 8G4H). (B) The structure of the 255 kDa influenza polymerase stabilised by a 15 kDa nanobody (Nb) (Keown et al., 2022) (PDB: 7NKA). (C) The structure of a 23 kDa signal cascade-initiating membrane protein, KDEL, bound to a nanobody-binding Fab (NabFab), a nanobody which is recognised by a nanobody-stabilised Fab (antigen binding fragment from antibody), which additionally binds to maltose binding protein (MBP) to form a 120 kDa so-called 'Legobody' (Wu and Rapoport, 2021) (PDB: 7RXC). (D) The structure of a ≈ 720 kDa complex of homopentameric β3 $GABA_AR$ receptor, with each monomer bound by a ≈ 60 kDa Megabody (Mb) (Uchański et al., 2021) (PDB: 6QFA). (E) The structure of two copies of the 49 kDa DNA helicase, RECQL5, bound to a 30 kDa di-Gluebody (dGb) (Yi et al., 2025) (PDB: 8RL5). Importantly, nanobodies are shown in this figure not only because they aid in structural determination but also because they are the building blocks from which other protein scaffold technologies have been produced, e.g. (D) Megabodies, (E) di-Gluebodies.

3.1.4.1. Nanobodies (Nb)

Conventional human monoclonal antibodies (mAb) are composed of two heavy chains and two light chains (Woof and Burton, 2004; Chiu et al., 2019). The two heavy chains are linked together by disulfide bonds in the hinge region of the protein (Moritz and Stracke, 2017). The association of the light chain to the heavy chain is stabilised by both disulfide bonds between the CH1 and CL constant domains, and by a hydrophobic interface between the VH and VL variable domains. The minimal antigen binding domain of a conventional human mAb is composed of the two variable domains VH and VL, where in each variable domain, three peptide loops named complementarity determining regions (CDRs) form the binding interface between the mAb and its antigen (figure 3.4, panel A). By comparison, camelid-derived antibodies have only two heavy chains, and are thus called heavy chain antibodies (hcAb) (Jin et al., 2023) (figure 3.4, panel B).

Similarly, the two heavy chains are joined by disulfide bonds in the hinge domains. The variable antigen binding domain (VHH) contains the CDRs, and importantly where the binding interface in mAbs are constituted of CDRs from both the VH and the VL domain, in hcAbs, as there is no light chain, the CDR is made only of one copy of the VHH domain. Camelid VHH domains can be isolated and recombinantly produced. These isolated VHH domains are commonly called nanobodies (Nbs). They are small, approximately 15 kDa proteins which are highly soluble and very stable in solution (Hamers-Casterman et al., 1993; Muyldermans, 2013) (figure 3.4, panel C).

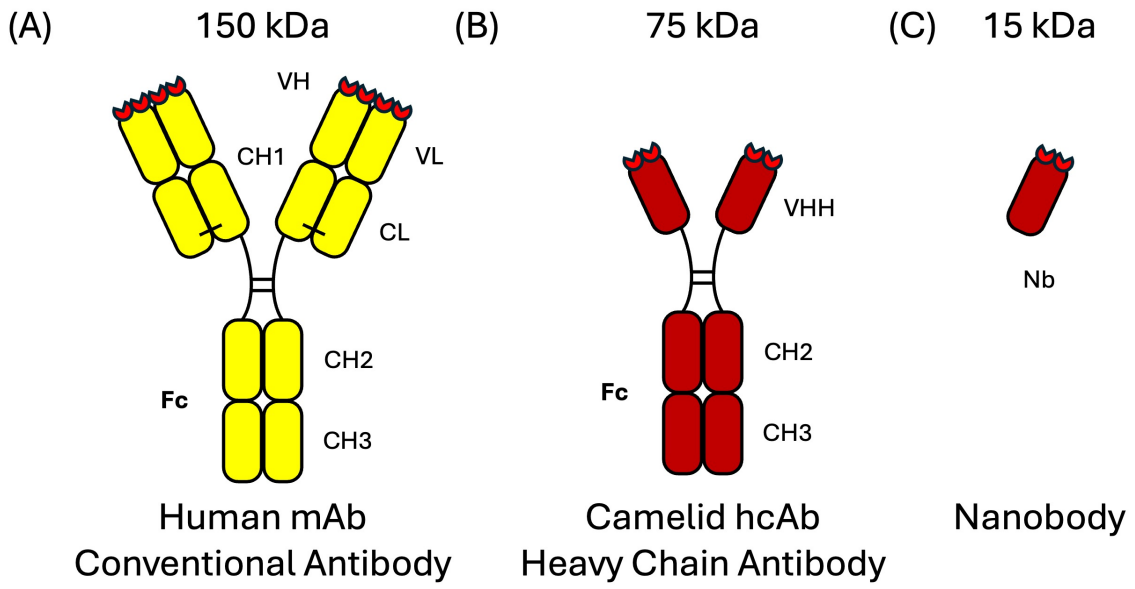


Figure 3.4: Diagrammatic Representation of Antibodies and their Relation to Nanobodies.

Diagrammatic representation of (A) conventional human monoclonal antibodies (mAb), (B) camelid heavy chain antibodies (hcAb), and (C) camelid-derived nanobodies (Nb). (A) The 150 kDa mAb is composed of two copies of two chains. The heavy chain is composed of complementarity determining regions (CDRs) (red) at the tip of the heavy chain variable domain (VH), and three heavy chain constant regions (CH1, CH2, CH3). The two heavy chains are bound by disulfide bonds in the hinge region (denoted by two horizontal lines between the CH1 and CH2 domains). The light chain is composed of CDRs (red) at the tip of the light chain variable domain (VL) and a light chain constant region (CL). The light chain is associated with the heavy chain by disulfide bonds (denoted by one horizontal line between the CH1 and CL domains) and additionally by hydrophobic interactions between the VH and VL domains. (B) The 75 kDa camelid heavy chain antibody (hcAb) is composed of two copies of one chain linked by disulfide bonds in the hinge domain (denoted by horizontal lines between the VHH and CH2 domains). The heavy chain is similar to that of the human mAb, with two constant domains (CH2, CH3), but a simplified variable domain. The variable antigen binding domain (VHH) is the variable domain of the hcAb and contains the CDRs (red). In A and B, the crystallisable fragment (Fc) is also denoted. This region is formed entirely by constant heavy chain domains and mediates downstream immune effector functions. (C) The 15 kDa camelid-derived nanobody (Nb) is composed of the VHH domain of the hcAb and contains the antigen-binding CDR regions of the donor hcAb. Figure adapted from (Bannas et al., 2017).

Nanobodies contain nine beta sheets (A, B, C, C', C'', D, E, F, and G), and three CDR loops: CDR1 (between beta sheets B and C), CDR2 (between beta sheets C' and C'') and CDR3 (between beta sheets F and G) (Uchański et al., 2021) (figure 3.5). The CDR3 loop is known to play the most significant role in target binding, given its longer length compared to the CDR1 and CDR2 loops. Do to its longer length, it is also known to be

able to access otherwise inaccessible cavity-like epitopes which are often more structured and conserved (Tohidi et al., 2024).

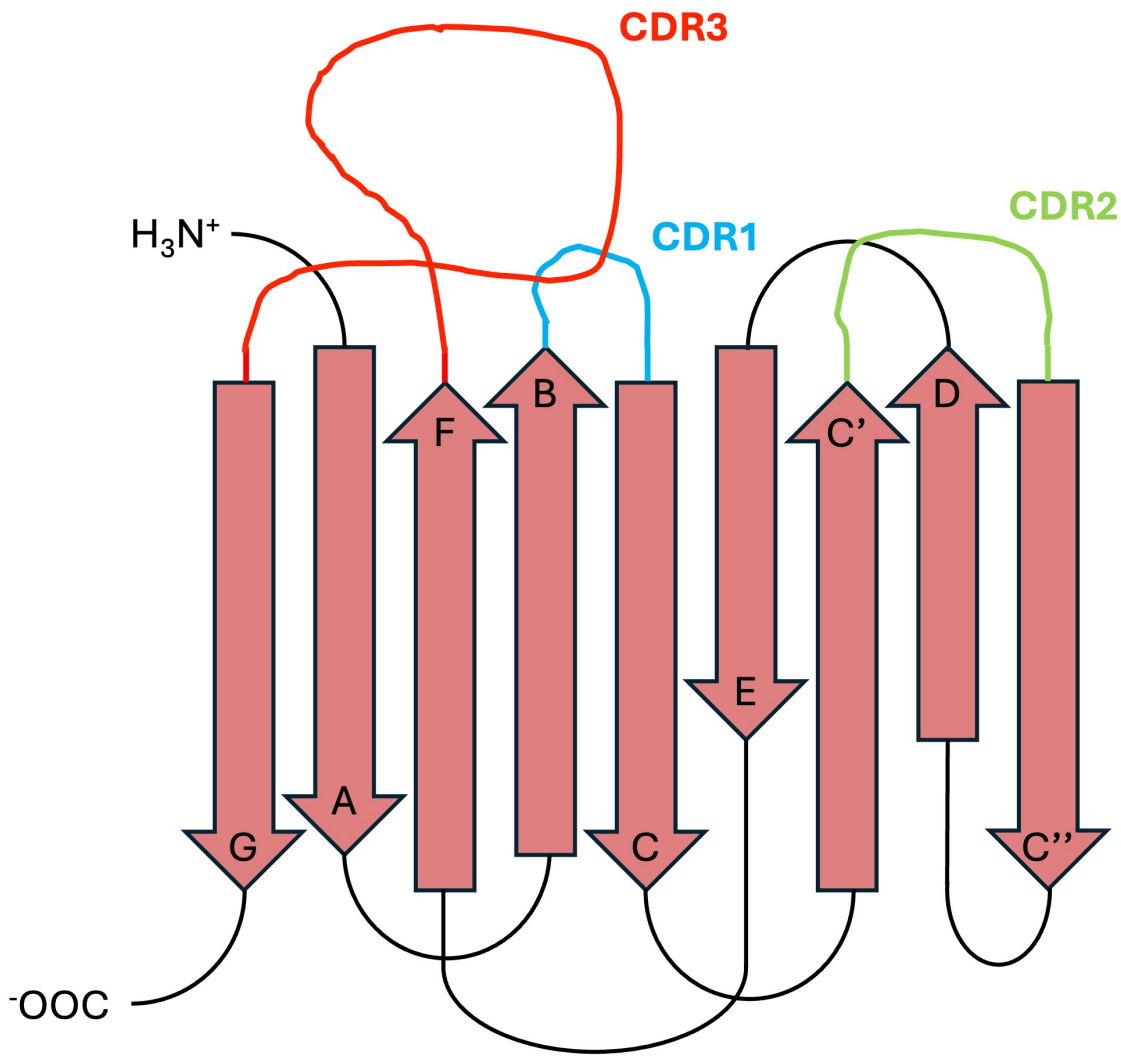


Figure 3.5: Diagram of Nanobody Domain Organisation.

The nanobody is composed of nine beta sheets (red arrows labelled with sheet identifier at the arrowhead) and three CDR loops (coloured and labelled: CDR1 in blue, CDR2 in green, and CDR3 in red).

In order to produce nanobodies, camelids (typically llamas, but also dromedary camels and alpacas) undergo an immunization regime of multiple doses of the purified antigen of choice (Fridy et al., 2014). Camelid serum is taken, and lymphocytes are isolated. In one method of recombinant protective nanobody screening against

influenza virus infection (Schmidt et al., 2016), mRNA was extracted from purified lymphocytes, and VHH coding sequences were each cloned into a lentiviral vector. This vector library was transfected into cells and expression of nanobodies induced prior to a lethal challenge with influenza virus. Genomic DNA libraries from surviving cells were produced and VHH sequences amplified by PCR to determine protective VHH sequences.

Nanobodies have a range of uses including as therapeutics (Bannas et al., 2017), in nanobody mediated protein purification (Stevens et al., 2024), and in multiple structural biology approaches. Nanobodies are able to stabilise flexible domains of proteins, reducing conformational heterogeneity in the protein sample of interest. Nanobodies can also be utilised as crystallisation chaperones, promoting crystal packing conditions which are otherwise not explored by the protein on its own (Rasmussen et al., 2011). In cryo-EM, nanobodies and their derivatives are able to reduce conformational heterogeneity, but additionally, they can be used to change the distribution of particles embedded in vitreous ice. In virology, nanobodies have shown great promise as therapeutics, showing potent neutralisation against viruses including SARS-CoV-2, influenza A, and respiratory syncytial virus, among others (Bhattacharya et al., 2023; Yang et al., 2024). Recent studies using nanobodies against influenza virus have demonstrated their use in the elucidation of viral protein structures and their identification of potential antiviral targets (Schmidt et al., 2016; Hanke et al., 2016; Fan et al., 2019).

The nanobodies used in the research reported in this chapter (and their gluebody and megabody derivatives) originate from the work of Schmidt and colleagues (Schmidt et al., 2016), where a phenotypic lentiviral screen was used to identify nanobodies which target the influenza nucleoprotein.

3.1.4.2. (di-)Gluebodies (dGb)

As demonstrated in figure 3.3, panel E, and in figure 3.7 below, gluebodies represent an engineered nanobody variant incorporating a strategically positioned disulfide bond on the non-CDR surface (specifically before in the A strand, the end of the E strand, and after the G strand of the nanobody) to enable covalent dimerisation without compromising antigen recognition. These engineered disulfide bridges facilitate either homo- or hetero-dimerisation at sites distant from the antigen-binding interface.

The primary advantage of this design is the creation of protein complexes with artificially imposed C2 symmetry, effectively doubling the molecular weight of target proteins while potentially also reducing conformational flexibility. Crucially, unlike the megabody platform which incorporates a large, well-ordered scaffold protein, the di-gluebody does not introduce a large protein, and so, maintains the focus of computational alignment on the protein of interest rather than the auxiliary framework. This design principle preserves native protein features as the primary determinants of structural reconstruction, while providing the molecular weight enhancement necessary for improved cryo-EM analysis.

The versatility of this platform has been validated across multiple applications, demonstrating successful implementation in both X-ray crystallography (Ye et al., 2024), and cryo-EM SPA (Yi et al., 2025). This dual compatibility underscores the broad utility of the di-gluebody approach for structural biology investigations requiring enhanced molecular stability and size augmentation.

3.1.4.3. Megabodies (Mb)

Megabodies are nanobody derivatives which are grafted onto a larger scaffold protein with two short peptide linkers. The nanobody is joined to the larger scaffold protein after the first beta sheet (labelled A in figure 3.6), and the rest of the nanobody is then encoded after the scaffold protein from the second beta sheet (labelled B in figure 3.6). The two scaffold proteins chosen for the development of the megabody system were both large, secreted bacterial proteins which were found to be stable, rigid, efficiently folded in the cell, amenable to genetic manipulation, and contain accessible beta strands for the required grafting. These were the *Helicobacter pylori* HopQ protein (45 kDa, PDB: 5LP2), and the *E. coli* YgjK protein (86 kDa, PDB:3W7T) (Uchański et al., 2021).

The resulting chimeric megabody complex retains the full ability of the nanobody to bind its target, whilst having the benefit of further increased molecular weight and the ability to combat issues of preferred orientation on EM grids at the air-water interface. Megabodies have been used to solve structures of membrane proteins (Uchański et

al., 2021), human serum albumin (De Felice et al., 2024) and have been applied to the study of components of influenza virus replication including the polymerase and the nucleoprotein by a previous DPhil student in this laboratory (Zhu, 2024).

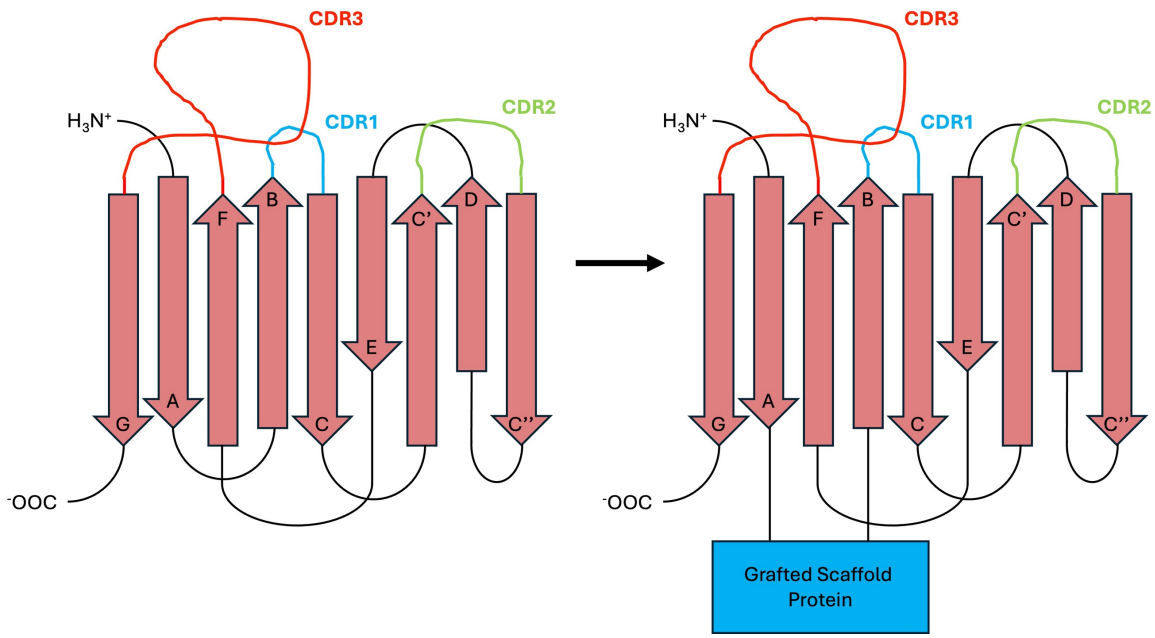


Figure 3.6: Diagram of Scaffold Protein Grafting onto a Nanobody to produce a Megabody.

The nanobody is composed of nine beta sheets (red arrows labelled with sheet identifier at the arrowhead) and three CDR loops (coloured and labelled: CDR1 in blue, CDR2 in green, and CDR3 in red). In the grafting process, the first beta turn (connecting nanobody beta strands A and B) is disrupted, and a mutant bacterial scaffold protein (HopQ or YgjK, here depicted as a blue rectangle, not to size) is linked via an exposed beta hairpin to the end of the nanobody A strand, with the rest of the nanobody (strands B through G) grafted onto the end of the selected scaffold protein through PCR amplification. This cloned construct is inserted into a megabody expression vector for recombinant expression.

Table 3.4: Nanobody Strand Boundaries of Nanobodies compared to Megabodies.

Nanobody strand boundaries from two proteins, 5TJW (Hanke et al., 2016) of an influenza nucleoprotein bound to a nanobody, and 6XUX (Uchański et al., 2021) of a megabody, with the domains of each beta strand (A-G) or CDR loop (CDR1-3) based on predicted protein secondary structure in ChimeraX.

Strand / Loop	Residues (PDB: 5TJW, Nb)	Residues (PDB: 6XUX, Mb)
A	5 – 13	3 – 11
B	19 – 25	789 – 796
CDR1	**26 – 33**	**797 – 803**
C	34 – 40	804 – 810
C’	47 – 53	827 – 823
CDR2	**54 – 58**	**824 – 827**
C’’	59 – 61	828 – 831
D	69 – 73	839 – 844
E	79 – 84	849 – 854
F	93 – 100	863 – 870
CDR3	**101 – 112**	**872 – 885**
G	113 – 122	886 – 894

Nanobody derived protein production methods pertinent to this thesis are displayed in a cartoon below (figure 3.7).

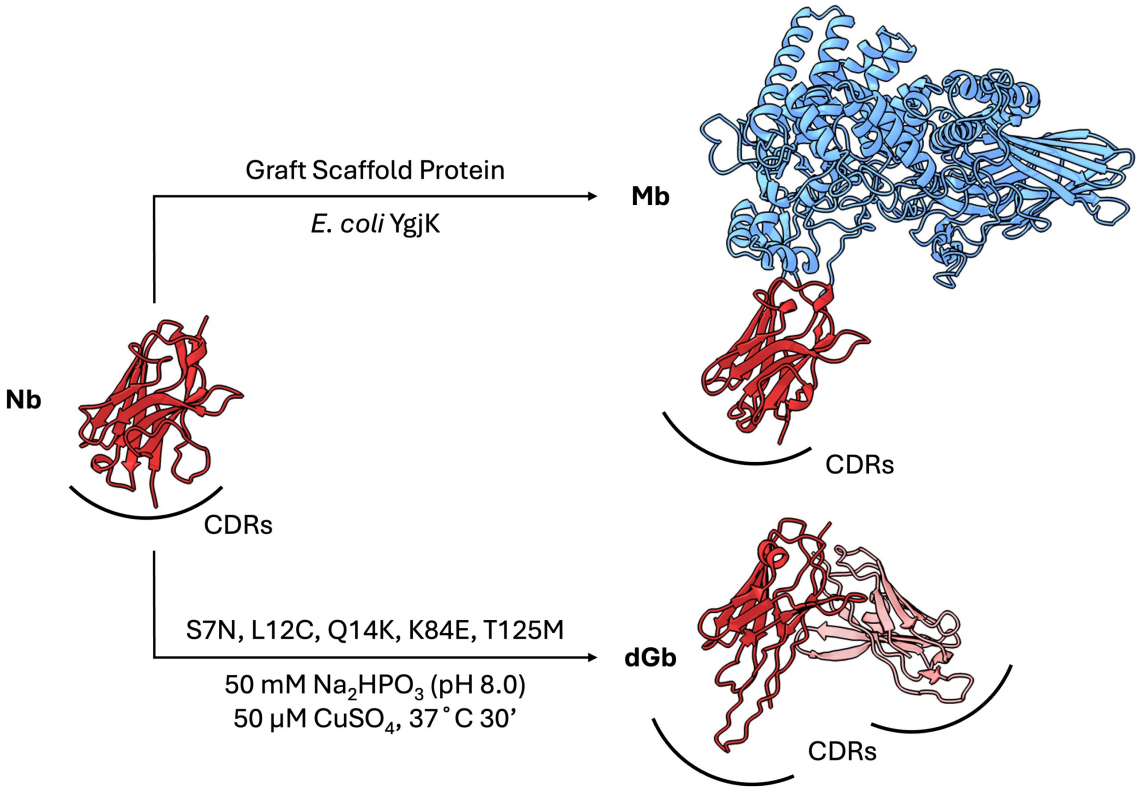


Figure 3.7: Nanobody-Based Protein Scaffolds in this Thesis.

A nanobody (Nb, red, AlphaFold3 prediction of Nb170 used in this study) is converted into a megabody (Mb, Nb, red, Mb scaffold, blue. AlphaFold3 prediction of Mb151 used in this study) by grafting of a larger scaffold protein, in this case the E. coli YgjK protein, via the loop between beta sheets A and B of the nanobody. A nanobody is converted into a (di-)Gluebody (dGb, Nb portions in red with the second copy set to 25 % transparency, taken from (Yi et al., 2025) (PDB: 8RL5) by addition of the five mutations noted in the diagram, and incubation in the conditions noted in the diagram. All Nb-derived protein scaffolds are similarly oriented, with complementarity-determining regions (CDRs) annotated.

3.1.5. Chapter Aims

In the work reported in this chapter, I aimed to express and purify recombinant influenza A nucleoproteins. I also aimed to determine the structure of the S09 avian influenza nucleoprotein bound to RNA using both X-ray crystallography and cryo-EM approaches. I additionally aimed to develop the di-gluebody and megabody systems to study the structure of the influenza A nucleoprotein and the binding of innate immune factors to the nucleoprotein by cryo-EM SPA. A platform to study influenza nucleoprotein and binding partner structure is of particular interest given the potential to be able to study influenza nucleoprotein binding partner structures. This will both improve our understanding of the mode of immune factor binding to monomeric nucleoprotein, and also enable future structure determination of binding partners to improve therapeutics targeting the nucleoprotein.

3.2. Results

3.2.1. Expression and Purification of Influenza Nucleoproteins

In this work, the monomeric nucleoprotein mutant, R416A, was used for structural studies. The behaviour of this mutant has been well documented, with a number of structures of this protein as a monomer (or lower order oligomer than compared to wild type) (Tarus et al., 2012; Chenavas et al., 2013; Knight et al., 2021).

The expression and purification of nucleoprotein constructs are described in the materials and methods chapter (section 2.2.2.). Briefly, the cloned nucleoprotein constructs were affinity purified via an N-terminal GST tag, with elution by cleavage of the GST tag where untagged nucleoprotein was required for structural studies, or elution by L-reduced glutathione where tagged nucleoprotein was required for interaction studies. There was then a size exclusion chromatography (SEC) step to provide a homogeneous protein stock for further structural and biophysical experiments. The purification was carried out until the SEC step in high salt (1 – 1.5 M NaCl) to ensure no residual binding to bacterial RNA. The example chromatogram (fig. 3.8, panel C) illustrates a high level of monodispersity, and the elution corresponds to a 56 kDa protein.

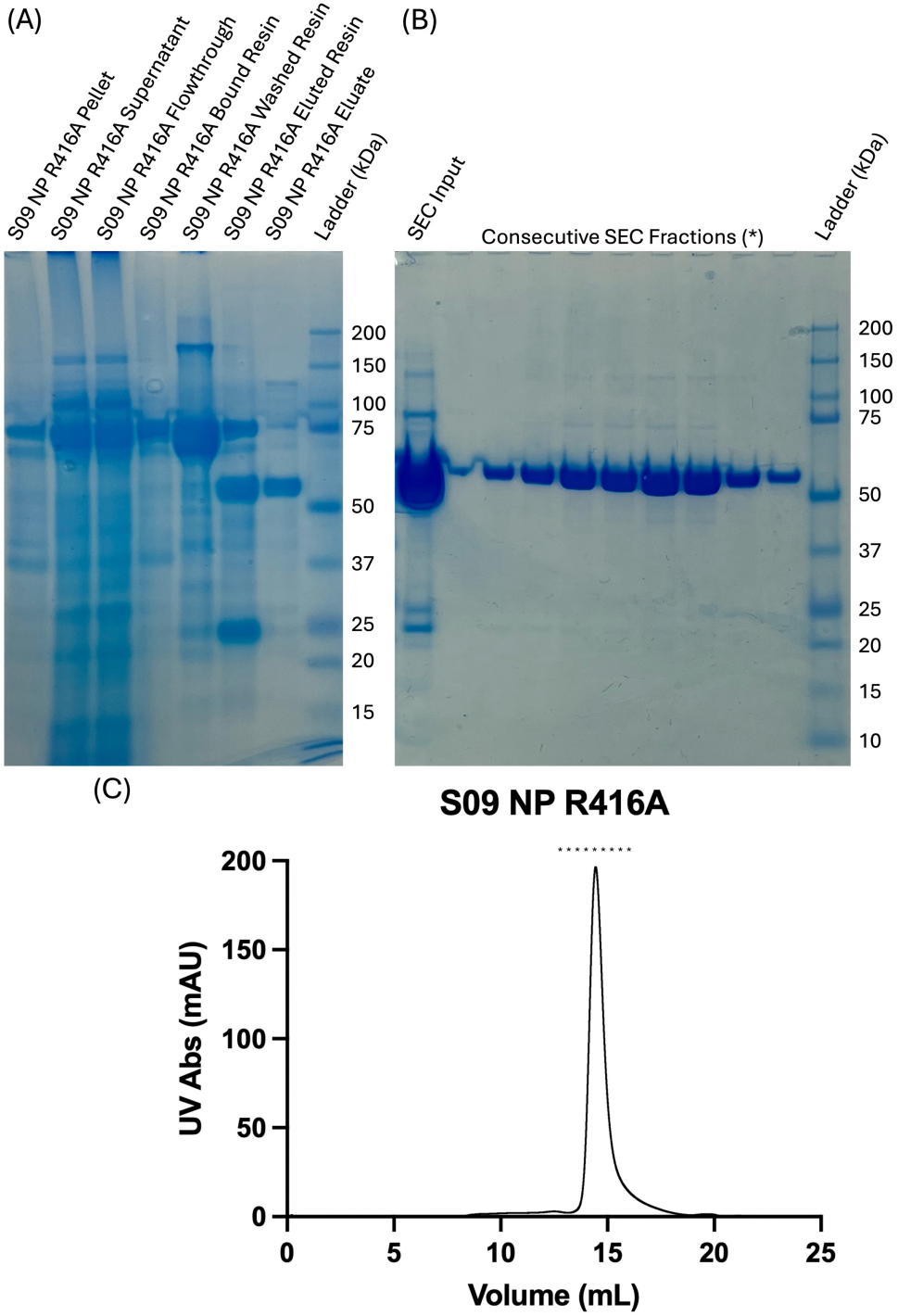


Figure 3.8: Expression and Purification of S09 NP R416A Monomeric Mutant Influenza A Virus Nucleoprotein.

(A) SDS-PAGE gel of expression and purification steps to elution from GST resin. (B) SDS-PAGE of SEC fractions. (C) Chromatography trace of a SEC run of this protein on a Superdex S200 (10/300) increase GL column.

As shown above (figure 3.8, panel C), Influenza nucleoproteins expressed and purified as described eluted in a monodisperse peak at around 14.8 mL on a Superdex S200 10/300 GL increase column, consistent with the expected elution volume of a 56 kDa protein. Fractions with a star (*) above on the chromatogram were run on an SDS-PAGE gel (figure 3.8, panel B), illustrating good separation of SEC input products. During SEC, the salt concentration was reduced to 150 mM NaCl, which would allow oligomerisation of the nucleoprotein.

3.2.2. Crystallography of Nucleoproteins

The nucleoprotein construct was subjected to crystallisation trials. 96 well trays of concentrated nucleoproteins (≈ 10 mg/mL) were set up as described in the materials and methods chapter using a Cartesian Technologies robotics system (Genomic Solutions) (Walter el., 2005) and monitored over time using the standard commercially-available crystallisation screens available – these included Index, PACT, PEGRx, JCSG (all Hampton Research), and other custom screens assembled by Dr. Thomas Walter of the Oxford Protein Production Facility. Both types of 96-well crystal tray available (CrystalQuick Greiner and SWISSCI 3) were used given varying reports of crystal production in types of commercial trays. The plates were imaged by a Formulatrix Rock Imager (Formulatrix) daily by visible light, and the plates were monitored for up to three months. Where precipitate was observed, optimisation screens around these conditions were attempted. Unfortunately, no crystals were found of nucleoproteins either alone or with added RNA (at 2-fold molar excess of either 14-mer RNA (Knight et al., 2021) or 18-mer RNA (Chenavier et al., 2025)).

3.2.3. Nucleoprotein in Complex with Nanobody

In a further attempt to grow nucleoprotein crystals, a nucleoprotein nanobody was used to increase the chance of forming protein crystals. Nb170 (Schmidt et al., 2016) is a nucleoprotein-binding nanobody which is observed not to interfere with vRNP nuclear import, nucleoprotein nuclear import, or vRNP activity (either replication or transcription). Since no structure has been obtained of this nanobody in complex with the nucleoprotein, and the binding interface is unknown, it poses an interesting biological question as it may be helpful in further studies of the influenza genome *in situ*.

3.2.3.1. Expression and Purification of Nanobody Nb170

The nucleoprotein nanobody, Nb170 was expressed and purified as described in the materials and methods chapter (section 2.2.5.). Briefly, the nucleoprotein nanobody was expressed in BL21 strain *E. coli*. As the nanobody construct contains a PelB signal sequence, which is known to direct recombinant proteins to the *E. coli* periplasm after cleavage of the PelB sequence (Power et al., 1992; Su et al., 2015; Shi et al., 2021a), the protein purification was carried out initially with a bacterial osmolysis step, rather than a bacterial lysis step, which leads to the release of pure protein from the bacterial periplasm. After osmolysis, the periplasmically-expressed nanobody was affinity purified via its C-terminal His_6 tag.

Eluted nanobody was then purified further via size-exclusion chromatography. As shown in the SEC trace and accompanying SDS-PAGE gel (figure 3.9, panels C and D), the purified nanobody eluted from the SEC column with a sharp, peak at around 15 mL, and pure protein on the SDS-PAGE gel with a molecular weight of 15 kDa.

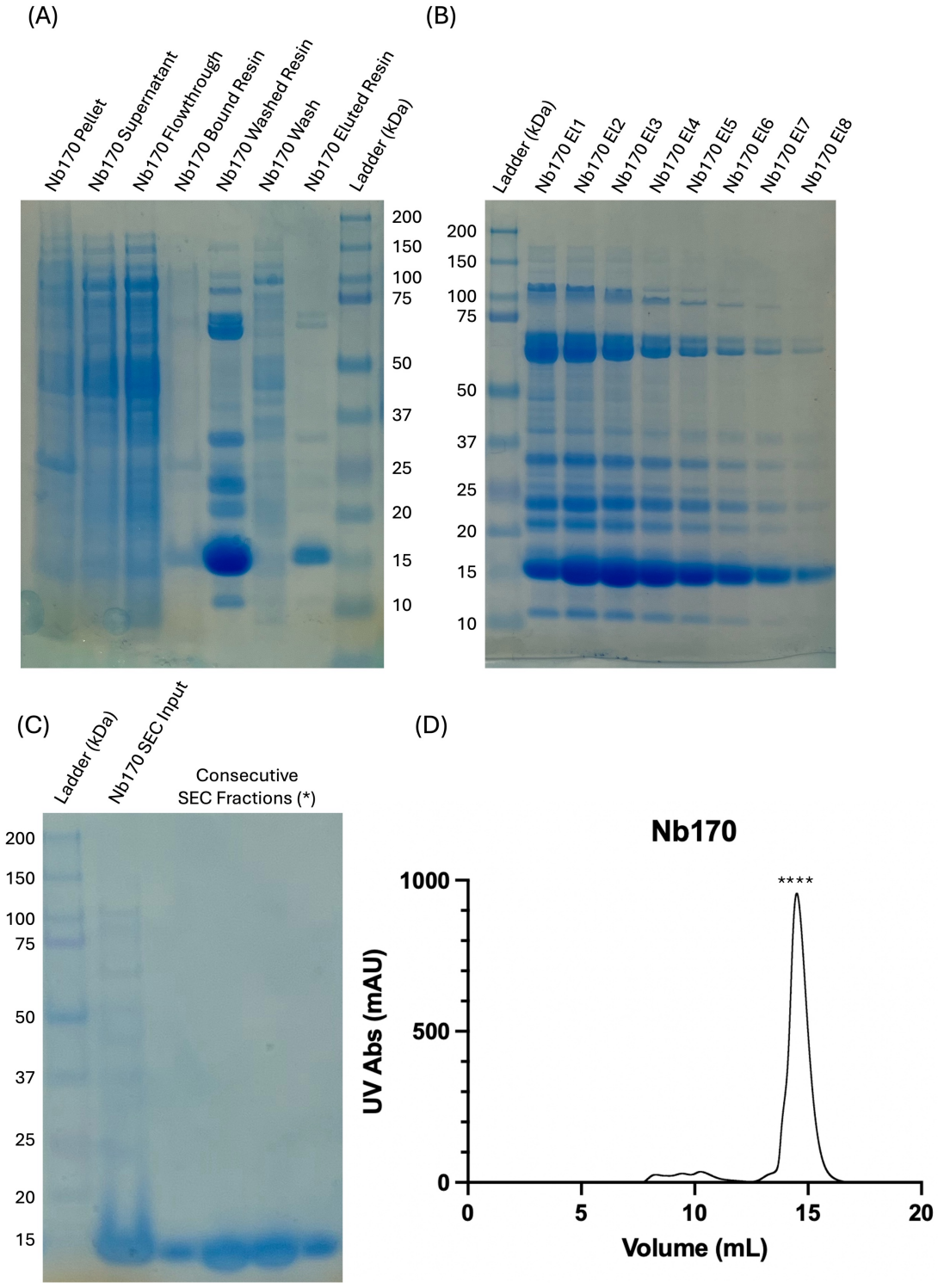


Figure 3.9: Nb170 Purification.

(A) Nb170 purification SDS-PAGE gel from bacterial pellet through to eluted resin. (B) SDS-PAGE gel of elution fractions from Nickel-NTA resin. (C) SDS-PAGE gel of SEC fractions from SEC run of Nb170. (D) Chromatogram from SEC run of Nb170 on S200 10/300 GL Increase column.

3.2.3.2. Complex Formation

A complex was formed between the monomeric nucleoprotein (S09 NP R416A) and an influenza nucleoprotein nanobody which is known not to interfere with the influenza lifecycle (Schmidt et al., 2016), Nb170 by incubating the two proteins in a 1:1 molar ratio on ice for 30 minutes. Control SEC experiments of the nucleoprotein alone and nanobody alone were carried out to determine the positions of these peaks on a SEC chromatogram without complex formation occurring and are shown in the below figure (figure 3.10, panel B). This mixture was then separated by SEC and the resulting peaks analysed with SDS-PAGE.

All complex formation reported in this chapter was analysed using an analytical Superdex S200 5/150 GL Increase column. The nanobody and the nucleoprotein eluted at volumes of 2.7 mL and 2.2 mL respectively. When allowed to incubate together on ice for 30 minutes, the resulting complex that formed eluted at a volume of 2.0 mL. Analysis of peak fractions by SDS-PAGE shows the formation of a broadly equimolar complex between the two proteins. These peak fractions were pooled for further structural studies.

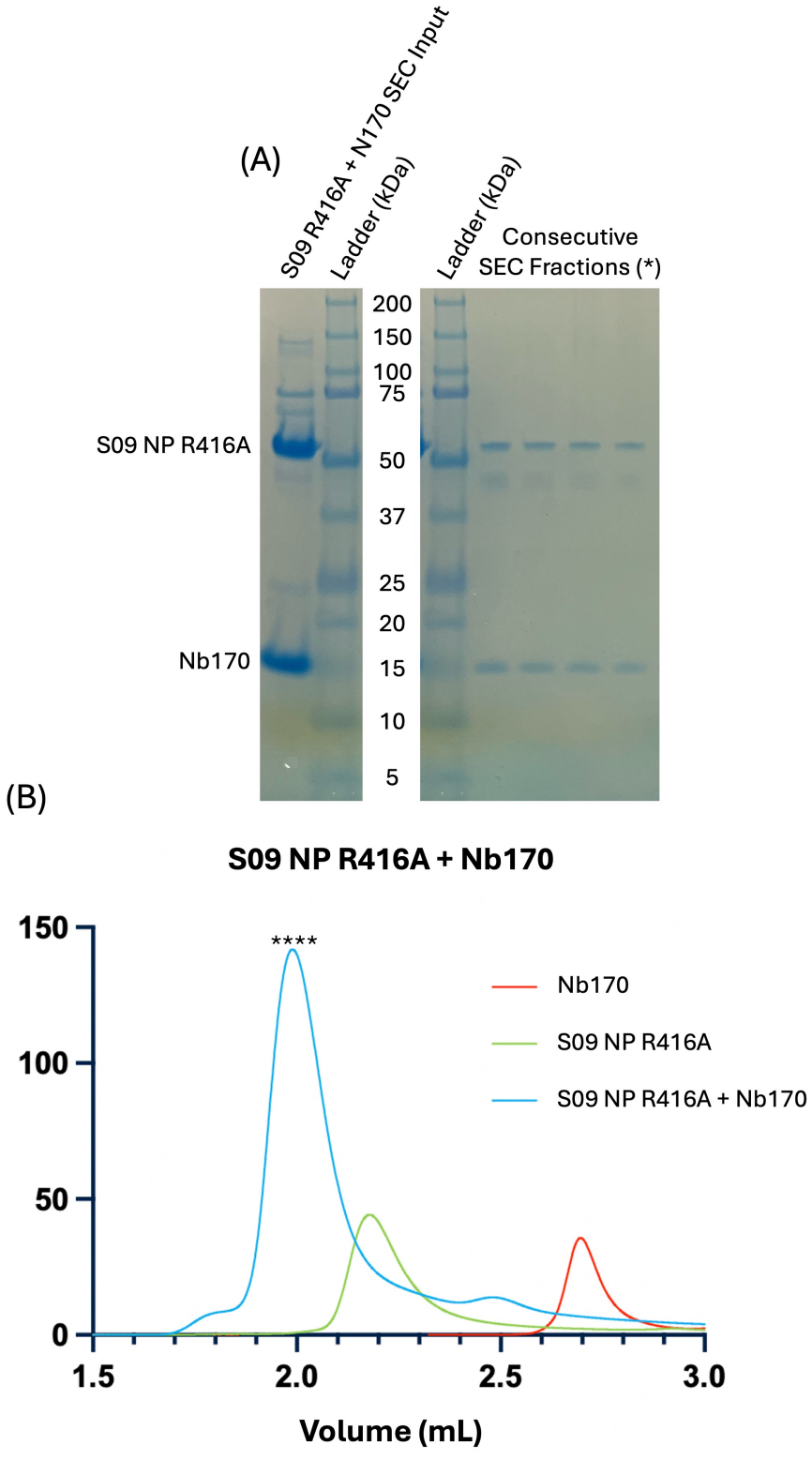


Figure 3.10: Complex Formation of Nucleoprotein and Nanobody.

(A) SDS-PAGE gel of SEC input, with NP and Nb170 annotated, and SEC fractions resulting from the incubation of S09 NP R416A and Nb170. (B) SEC profiles overlayed of Nb170 alone (red), S09 NP R416A alone (green), and the resulting complex (blue).

3.2.3.3. BLI Analysis of S09 NP R416A and Nb170

The affinity of binding between the nucleoprotein and nanobody was investigated with Bio-Layer Interferometry (BLI) (introduced in detail in section 4.1.3.). BLI experiments were carried out as detailed in the materials and methods chapter (section 2.3.5.3.) Briefly, S09 NP R416A was immobilised on anti-GST tips as the ligand, and Nb170 was used as the analyte. A constant concentration of nucleoprotein ligand (20 µg/mL ≈ 0.36 µM) was used in this experiment, with a five point dilution series of Nb170 protein analyte – 60 µg/mL (= 4 µM), 30 µg/mL (= 2 µM), 15 µg/mL (= 1 µM), 7.5 µg/mL (= 0.5 µM), and 3.75 µg/mL (= 0.25 µM). A baseline was run in parallel, where no ligand was immobilised with the highest concentration of analyte flown over the tips to determine a background response due to ligand. This was negligible, suggesting no background binding of Nb170 to the anti-GST tips, but the background reading was still subtracted from the signal.

Kinetic parameters were calculated for the interaction and are displayed below (figure 3.11). The determined K_D value for the interaction between S09 NP R416A and Nb170 was 290 nM, suggesting a moderate to weak binding interaction between the nucleoprotein and the nanobody. The model demonstrates an acceptable fit to the data, given an R^2 value of 0.9296, and a χ^2 value of 0.0013. However, the clustering of low concentration data points in the steady-state analysis plot, and R^2 value of 0.9296, suggest that the data could be improved by repeating this experiment with more data points over a greater concentration range. Taking the data in its current form, this confirms that the nanobody binds the nucleoprotein, with a K_D value of 290 nM, but this should not be considered to be a strong nanobody-target interaction, as the K_D of

ultra-tight nanobody-antigen binding has been reported to be in the low picomolar range (Schedler et al., 2023), and to be in the low nanomolar range for high affinity binders (Vasylieva et al., 2019).

These binding data provide the first biophysical quantification of the binding kinetics of Nb170 to any influenza nucleoprotein, although they suggest that the binding interaction with the nanobody is weak compared to other more tightly-binding nanobodies (Vasylieva et al., 2019; Schedler et al., 2023). This experiment could be repeated with other influenza nucleoproteins and nanobodies to compare the relative affinities of interaction, giving a rational way of selecting nucleoprotein – nanobody combinations for downstream structural work.

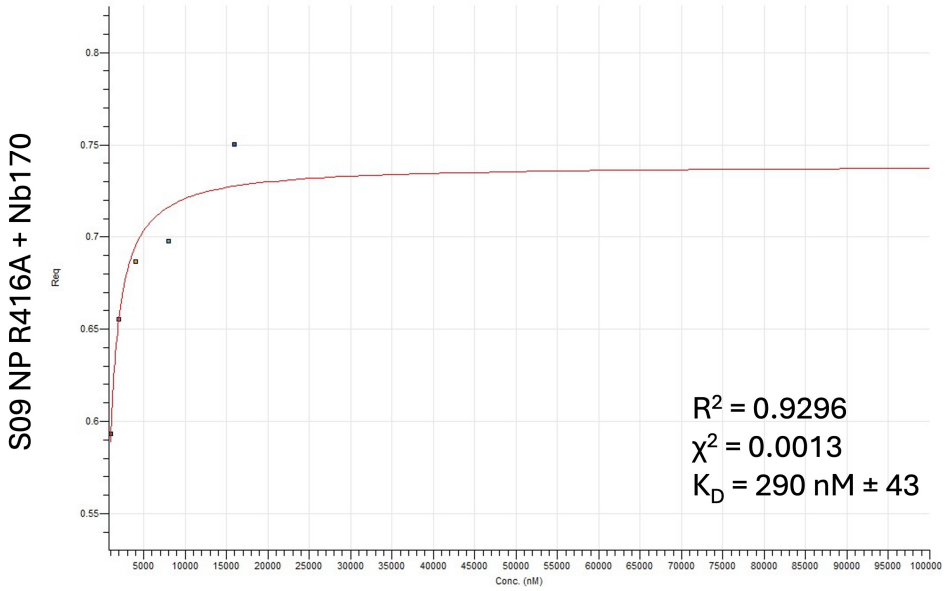


Figure 3.11: Nucleoprotein Nanobody Complex Kinetic Analysis.

Steady-state analysis plot of the BLI experiment with S09 NP R416A immobilised as the ligand and Nb170 used as the analyte. Key kinetic parameters are reported on the steady-state analysis plot.

3.2.3.4. Crystal Trials

The S09 NP R416A + Nb170 complex was concentrated to 10 mg/mL and RNA was added in a 2-fold molar excess (14-mer RNA (Knight et al., 2021)). Ten 96-well crystal trays of different commercially available crystallisation screens were set up, and one crystal was obtained from the Morpheus screen in well E4, corresponding to a condition of 0.12 M di/tri/tetra/penta-ethylene glycol, 0.1 M Imidazole/MES buffer (pH 6.5), 37.5 % (v/v) MPD (2-methyl-2,4-pentanediol), PEG1000, PEG3350). This crystal first appeared 1 day after the crystal tray was set up and stopped growing 4 days after the crystallisation experiment was set up. The crystal grew to an approximate length of 315 µm (figure 3.12). The well was opened, and cryo-protectant (15 % (v/v) ethylene glycol) was added to the drop prior to mounting on an appropriately sized MiTeGen loop and cryo-cooling. The crystals were soaked in the cryoprotectant for one minute prior to mounting on crystal loops. Ethylene glycol was chosen as a cryoprotectant because it was present in the crystallisation condition, and therefore it was rationalised that further addition of ethylene glycol would not cause chemical damage to the crystal through dissolving the protein crystal or osmotic shock.

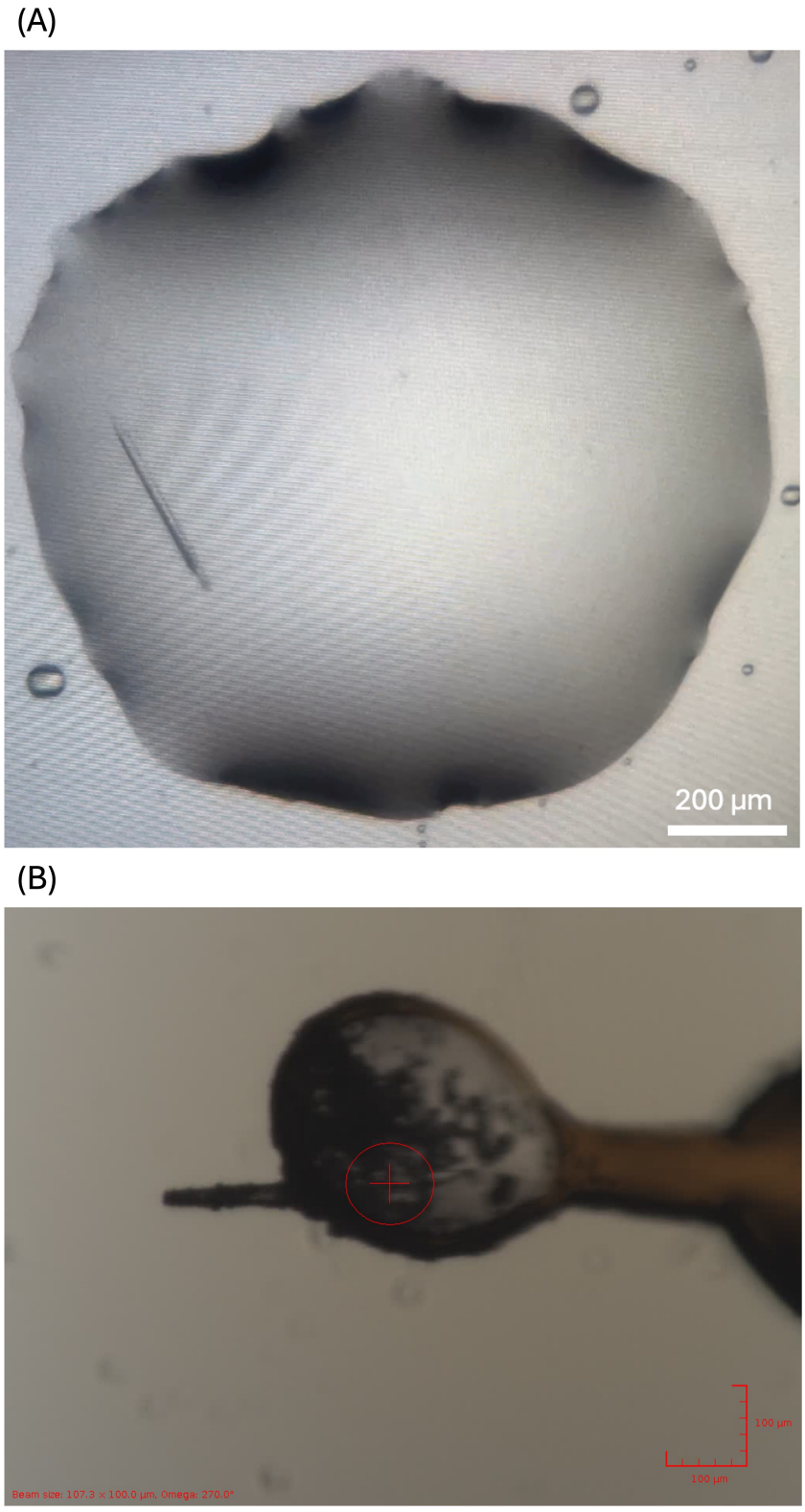


Figure 3.12: Images of S09 NP R416A + Nb170 Crystal.

(A) Image of crystal at its longest in the crystal drop from its source tray. Scale bar = 200 µm. (B) Image of crystal mounted in MiTeGen crystal loop maintained at 100 K at the Diamond Light Source, at beamline I04.

3.2.3.5. X-Ray Diffraction Data Collection and Processing

Diffraction data were obtained at 100 K at the Diamond Light Source, using beamline I04. 360˚ of diffraction data were collected as a series of 3600 images with an oscillation of 0.10˚ per image. The data were auto processed using xia2 dials (Gildea et al., 2022). The crystal belonged to space group $P6_4$ with unit cell measurements of a = b = 133.09 Å, c = 102.64 Å and angles α = β = 90.00˚, γ = 120.00˚. Analysis of intensity statistics in PHENIX XTRIAGE suggested that the data were severely twinned, with a twin operator of h, -h-k, -l, and an estimated twin fraction of 0.490. Additionally, the data were very weak beyond 3.5 Å (table 3.5), and I/σI values were very low beyond 4.08 Å (table 3.6). For this reason, data beyond 3.5 Å were not considered for phasing by molecular replacement (MR) or further refinement. Based on packing considerations for a complex of one nucleoprotein bound to one nanobody in an asymmetric unit, this would give a solvent content of 65 %

Table 3.5: Completeness of X-Ray Diffraction Data Collected.

Resolution Range (Å)	Completeness (%)
115.22 – 18.41	100.0
18.37 – 14.50	100.0
14.50 – 11.43	100.0
11.41 – 8.99	100.0
8.98 – 7.08	100.0
7.08 – 5.58	100.0
5.57 – 4.39	100.0
4.39– 3.46	92.8
3.46 – 2.62	14.6

Table 3.6: Completeness and Intensity for X-Ray Diffraction Data Collected.

Values of I / σI are reported giving a measure of the strength of signal at various signal-to-noise ratios.

Resolution Range (Å)	I / σI > 1	I / σI > 2	I / σI > 3	I / σI > 5	I / σI > 10	I / σI > 15
115.33 – 6.47	99.9	99.6	99.0	98.2	93.6	87.7
6.47 – 5.14	98.5	96.6	93.9	88.0	67.6	50.7
5.14 – 4.49	98.0	95.3	92.0	82.2	60.2	42.0
4.49 – 4.08	96.1	91.2	85.2	69.5	41.8	25.9
4.08 – 3.79	84.6	75.2	62.0	42.1	20.7	12.2
3.79 – 3.56	69.4	52.0	39.1	21.8	9.3	5.2
3.56 – 3.38	42.2	28.3	20.6	11.2	4.3	1.5
3.38 – 3.24	29.1	20.3	13.9	7.9	2.9	1.1
3.24 – 3.11	17.9	12.0	7.8	3.6	0.7	0.6
3.11 – 3.00	11.4	7.3	3.8	1.2	0.1	0.0
3.00 – 2.91	8.9	5.9	3.3	0.7	0.1	0.0
2.91 – 2.83	4.5	2.3	1.1	0.2	0.1	0.0
2.83 – 2.75	2.4	1.3	0.6	0.1	0.0	0.0
2.75 – 2.69	1.1	0.4	0.1	0.0	0.0	0.0

The structure was solved by molecular replacement using PHASER (McCoy et al., 2007) using crystal structures PDB: 7NT8 (NT60 NP R416A nucleoprotein) (Knight et al., 2021) and the nanobody present in crystal structure PDB: 5TJW (Influenza A virus nucleoprotein in complex with inhibitory nanobody) (Hanke et al., 2016) which found one copy of the nucleoprotein and one copy of the nanobody in the asymmetric unit with a log likelihood gain (LLG) score of 200 and a translation function Z score (TFZ) of 10.8, metrics which suggests a correct molecular replacement solution. These final solution scores were reported in the Phaser-MR log file, and the best solution was confirmed to have two components – one copy of the nucleoprotein, and one copy of the nanobody.

Although the packing of the nanobody to the nucleoprotein appears sensible, with the CDR3 loop of the nanobody binding to the nucleoprotein, refinement of the two proteins, initially as rigid bodies followed by atomic positions, grouped *B*-factors, and TLS refinement gave an R-factor and R-free score of 36.92 % and 38.56 % respectively

with correspondingly poor maps. The geometry of the final model was validated using MOLPROBITY (Chen et al., 2010).

To ensure that the good LLG and TFZ scores were not erroneous, molecular replacement was attempted in the same way with an unpublished structure of the approximately 50 kDa polymerase acidic (PA) C-terminal domain of the influenza polymerase (Dr. Cecilia Rocchi, personal communication). The molecular replacement search finished with an LLG value of 14 and a TFZ value of 5.8, which suggests that molecular replacement of the S09 NP R416A nucleoprotein + Nb170 complex was indeed successful using the published structures, as outlined. Data quality was further assessed by considering multiplicity and precision-indicating merging R-factor values (table 3.7).

Given that the overall multiplicity of the data was high and the Rpim values were good, the data are likely reliable at low and medium resolution. Furthermore, as the nanobody is oriented such that the CDR3 loop is the binding interface with the nucleoprotein, knowledge which PHENIX does not consider in its refinement, the initial molecular replacement solution presented here (figure 3.13), is likely to be correct. The severe twinning of the data (49 %) may account for the pathology of the refinement (even though the twin factor was used in the refinement). This twinning may potentially be caused by disordered packing of the complex which cannot be modelled in the refinement, thus causing the poor refinement statistics.

Table 3.7: X-Ray Diffraction Data Collection Parameters.

Data collection parameters for the S09 NP R416A + Nb170 crystal. Values in brackets represent the outer shell of data collected. Data collected using Diamond data collection software (ISPyB) and REFMAC and PHENIX statistics.

Data Collection	
Wavelength (Å)	0.954
Number of Images	3600
Ω Oscillation per Image (˚)	0.10
Exposure Time per Image (s)	0.06
Beam Size (µm)	107 x 100
Transmission (%)	100
Space Group	$P6_4$
Unit Cell Dimensions	
a, b, c (Å)	113.092, 113.092, 102.64
α, β, γ (˚)	90.0, 90.0, 120.0
Total Reflections	319,780
Unique Reflections	15,660
Resolution Range (Å)	115.2 – 2.6 (2.8 – 2.6)
CC (1/2)	1.0 (0.6)
R_{merge}	0.045 (0.242)
Mean I / σI	14.2 (1.9)
Rpim (I)	0.064 (1.070)
Completeness (%)	100 (14.6)
Multiplicity	21.0 (20.7)
Wilson *B*-factor ($Å^2$)	68.82

The putative molecular replacement solution is shown in figure 3.13. In this solution, the nanobody CDR3 loop is positioned at the interface of the nanobody and the nucleoprotein. The interface it interacts with is across two alpha helices from approximate residues 224 to 240. Residues in close enough proximity for interaction are found to be from the CDR3 loop, with additional putative contacts contributed from the A and G strands of the nanobody.

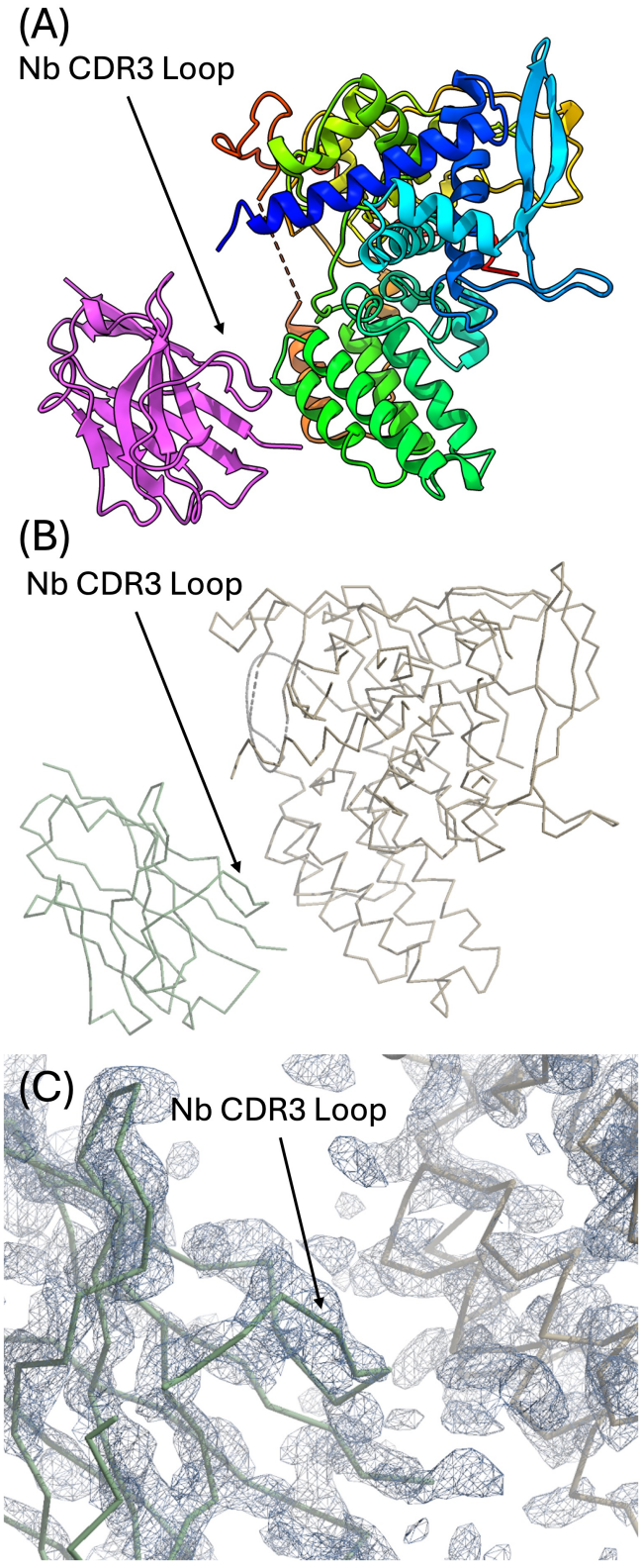


Figure 3.13: Initial Molecular Replacement Solution of S09 NP R416A + Nb170.

(A) The nucleoprotein used in molecular replacement (PDB: 7NT8) (Knight et al., 2021) is coloured using ChimeraX's rainbow colouring scheme, and the nanobody (PDB: 5TJW) (Hanke et al., 2016) is coloured in magenta with its CDR3 loop facing the nucleoprotein, labelled. (B) Image of molecular replacement solution in Coot. Nb170 is coloured by Cα backbone in green and S09 NP R416A is coloured in brown, with the CDR3 loop of the nanobody annotated. (C) Zoomed-in view of molecular replacement solution

in Coot with electron density overlayed. Nb170 is coloured by Cα backbone in green and S09 NP R416A is coloured by Cα backbone in brown, with the CDR3 loop of the nanobody annotated.

3.2.4. Nucleoproteins in Complex with (di-)Gluebodies

A new technology for presenting two copies of the same protein in a two-fold symmetric dimer was developed during this work, the di-Gluebody. The benefit of this technology is that the imposed symmetry, whereby two identical protein molecules are held by a rigid scaffold at a set angle and distance apart, gives a C2 symmetry function between the two monomers. This allows for the imposition of symmetry during cryo-EM data processing: improving signal to noise ratio, aiding particle alignment, and improving the resolution achievable. It was rationalised that complex formation between dimeric nanobodies (di-Gluebodies) and nucleoproteins would form a readily-producible complex sufficiently large and rigid for high-resolution structure determination both of the nucleoprotein of interest alone, and in complex with any number of binding partners.

In order to produce (di-)Gluebodies, five amino acid mutations must be made to the scaffold region of the nanobody (i.e. not affecting the nanobody CDR loops) to facilitate the formation of a disulfide bond under certain conditions (Ye et al., 2022; Yi et al., 2025). The nanobody scaffold is highly conserved, with only the CDR region varying between different nanobodies. The mutations required are S7N, L12C, Q14K, K84E, and T125M, corresponding to mutations in the A strand (residues 7 and 12), the A-B loop (residue 14), the E strand (residue 84) and the C-terminal loop after the G strand (residue 125) (as reported in table 3.4). As the amino acid sequence of

nanobodies is conserved, they can be easily aligned and equivalent residues across varying nanobodies can be identified. This is carried out using the ANARCI web server (Antigen receptor Numbering And Receptor Classification) (Dunbar et al., 2016), and the numbering scheme with which the nanobody is numbered is the IMGT scheme (the international ImMunoGeneTics information system) (Manso et al., 2022). Subsequently, synthetic gene fragments were ordered for Gibson assembly into a nanobody expression vector.

For clarity in nomenclature, the nanobody is the single-domain antibody to which mutations are added. The gluebody is the nanobody with the five required mutations added (as per above). The di-gluebody (dGb) is the dimeric gluebody (post-dimerisation).

3.2.4.1. (di-)Gluebody Expression and Purification

(di-)Gluebodies are expressed and purified as described in the materials and methods chapter. Briefly, synthetic gene fragments of nanobodies produced by Schmidt and colleagues (Schmidt et al., 2016) with the added five mutations required for di-gluebody formation were ordered from Twist Bioscience. Bacterial codon-optimised nucleoprotein-binding nanobody fragments (with the required mutations for di-gluebody formation) were used in Gibson assembly reactions to form bacterial expression plasmids with a range of nucleoprotein-binding gluebodies. The bacterial expression backbone was either pNIC_MBP (containing an MBP-His_6 tag), or pNIC_NHS (containing a streptavidin tag). The choice of expression vector for each

gluebody was arbitrary, with half cloned into the pNIC_MBP vector, and half cloned into the pNIC_NHS vector, with the idea that they would be tested in parallel and the most successful backbone chosen for all further experiments.

In figure 3.14, below, the results of an expression test are shown whereby 1 mL of transformed competent KRX *E. coli* (Promega) was induced to express gluebodies by the addition of 1 mM IPTG, and then at 3 h post-induction, SDS-PAGE fractions were taken of whole cell lysates by boiling 10 µL of the transformed bacteria in 10 µL SDS loading dye, and the relative expression of gluebody construct was compared.

From this expression test, two gluebodies were chosen to carry out further experiments: 50M as it was the best-expressed gluebody, and 54S as it is based on a nanobody which does not interfere with the influenza replication cycle (Nb170), and as such, may be more amenable to use with innate immune factors of interest. (A complete list of nanobodies, gluebodies, and their reported effect on the influenza virus lifecycle is found in appendix 1).

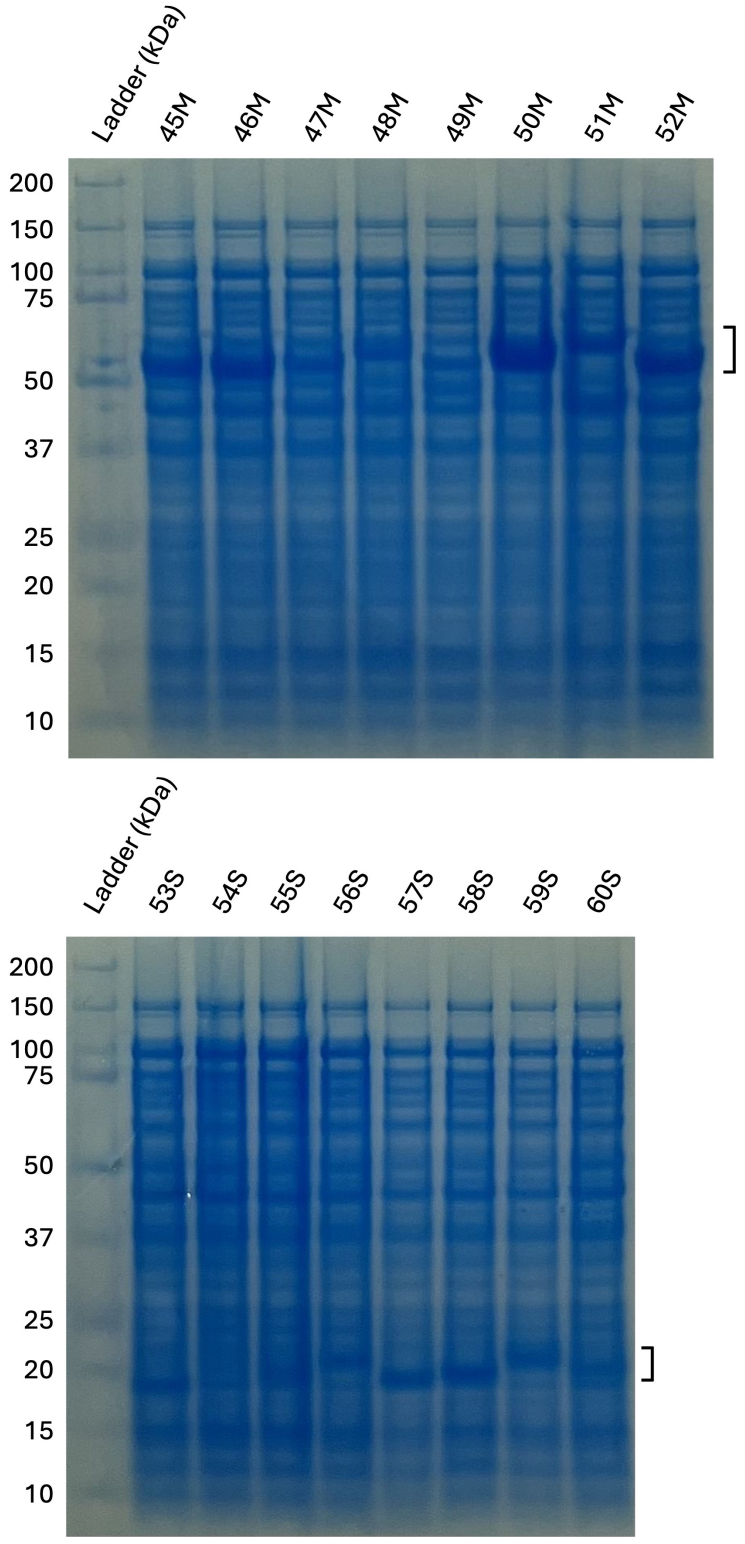


Figure 3.14: SDS-PAGE of Gluebody Expression Test.

Whole cell lysate of bacteria expressing 16 different gluebodies with a black box around bands corresponding to proteins expressed in a pNIC_MBP vector (MBP-His_6-tagged proteins, labelled #M, top) or proteins expressed in a pNIC_NHS vector (streptavidin-tagged proteins, labelled #S, bottom). Details of the nanobody each gluebody was produced from are found in Appendix 1.

Having chosen two candidate gluebodies with which further experiments were to be carried out, both gluebodies were both purified as detailed in the materials and methods chapter (section 2.2.6.). Briefly, MBP-His_6-tagged gluebodies were affinity purified by their N-terminal His_6 tag, thoroughly washed, and eluted in wash buffer supplemented with 300 mM imidazole. The imidazole was removed with the use of a PD-10 desalting column before cleaving the MBP-His_6 tag overnight with the addition of 0.5 mg His_6-tagged TEV protease to the pooled eluate sample.

The following morning, the protein – protease mix was purified by reverse immobilised metal ion affinity chromatography (IMAC), whereby fresh Ni-NTA resin was used to bind the His_6-tagged TEV protease and the MBP-His_6 tag cleaved from the gluebody construct. Streptavidin-tagged gluebodies were affinity purified by their N-terminal twin-strep tag, thoroughly washed, and eluted with three elution washes of 10 mM, 25 mM, and 50 mM biotin. Affinity-purified gluebodies from either tagging system then underwent dimerisation into di-gluebodies (if spontaneous dimerisation had not occurred already). Briefly, gluebodies were diluted to ≤ 0.5 mg/mL to avoid precipitation and dimerised by increasing pH from 7.0 to 9.5 and adding $Cu(II)SO_4$ to catalyse the formation of disulfide bonds. Where required, gluebody samples were also incubated at 37˚C for 30 minutes to improve dimerisation efficiency.

Where a mixture of gluebody and di-gluebody was present in the resulting sample, size exclusion chromatography was carried out to separate the resulting mixture. The example purification of gluebody 50M is shown below (figure 3.15). It is of note that between the SDS-PAGE gel fractions being taken from the TEV cleavage (figure 3.15,

panel B) to the size exclusion chromatography experiment being carried out (figure 3.15, panels C, D), the gluebody underwent spontaneous dimerisation.

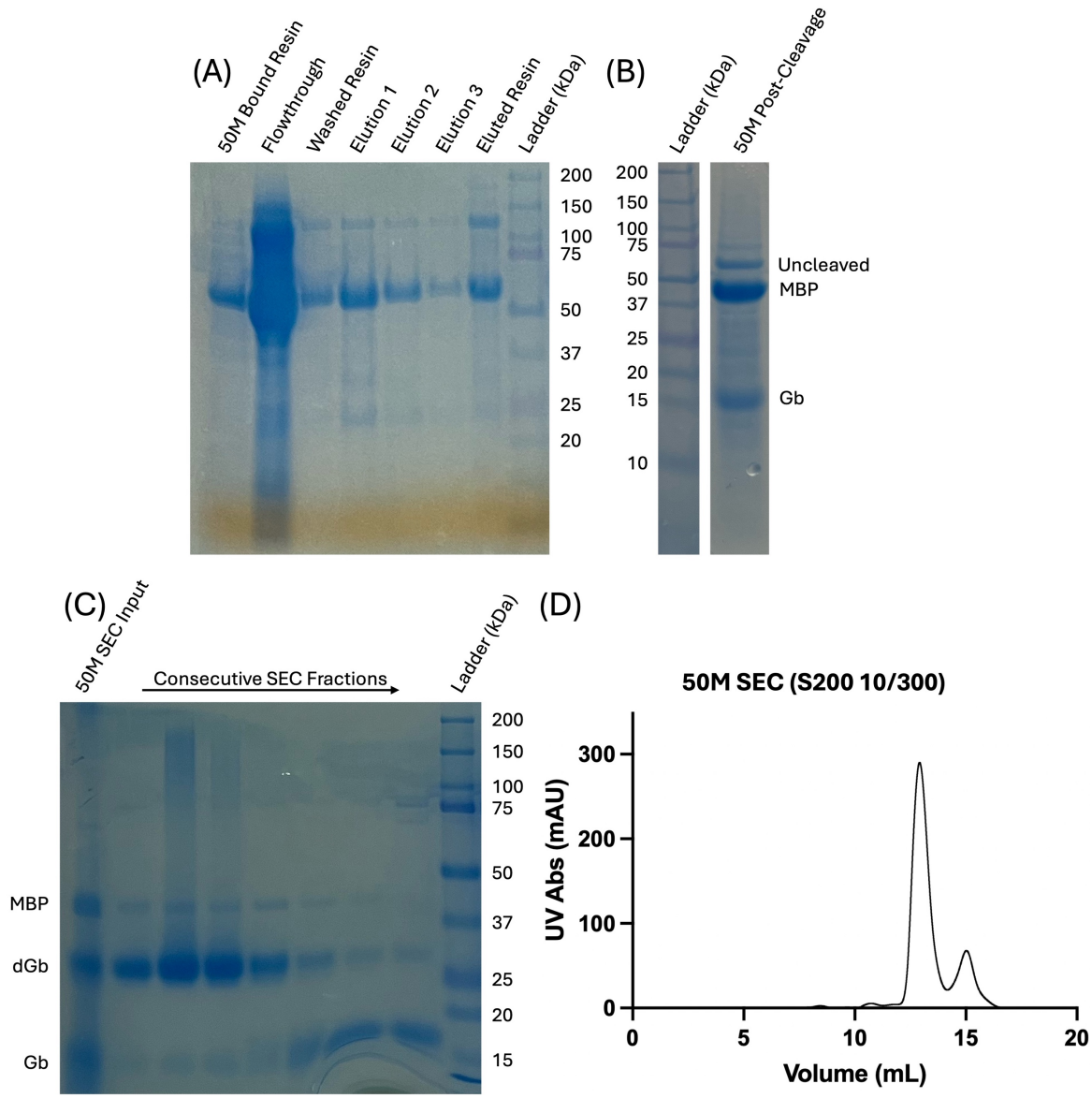


Figure 3.15: 50M Purification.

(A) SDS-PAGE gel of 50M protein through nickel affinity purification to eluted MBP-tagged protein. (B) SDS-PAGE gel lanes of 50M post-cleavage. The elution fractions were pooled and cleaved with TEV protease overnight to produce free MBP and gluebody. (C) Post-reverse IMAC, the cleavage product was concentrated and injected on a size exclusion column which was able to separate MBP from dGb and Gb. (D) Chromatography trace of the 50M SEC run.

Gluebodies and di-gluebodies, where found in a mixed sample, can be separated by size-exclusion chromatography (figure 3.16, panel A). Additionally, the use of 700 mM

beta-mercaptoethanol (the standard concentration of beta-mercaptoethanol used in the reducing SDS-loading dye used in this work) is sufficient to break the disulfide bond formed between the two gluebodies in a di-gluebody, reversing the dimerisation of the gluebody into a di-gluebody, and thus resulting in the formation of a homogeneous population of gluebody (figure 3.16).

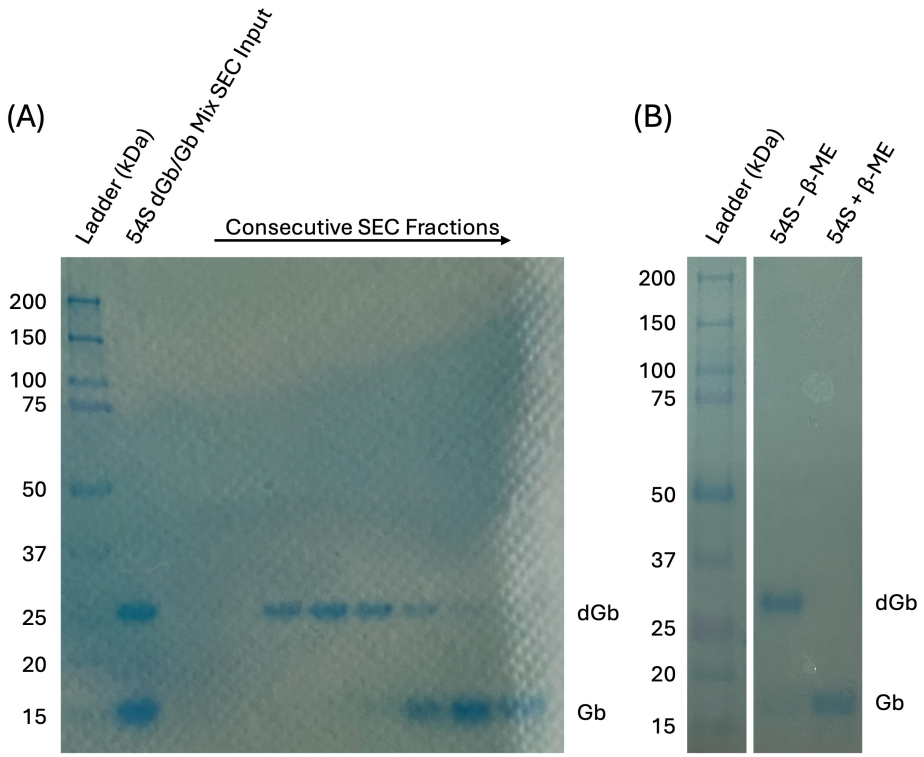


Figure 3.16: Separation and Monomerisation of di-Gluebodies.

(A) SDS-PAGE gel of size exclusion chromatography run of 54S. The input is a mixture of dimer and monomer and is clearly separated into a di-gluebody peak and a gluebody peak. (B) SDS-PAGE gel lanes showing di-gluebodies can be efficiently monomerised by the addition of the reducing agent β-mercaptoethanol.

3.2.4.2. Complex Formation

The two di-gluebodies, 50M and 54S, were complexed with the S09 NP R416A mutant nucleoprotein by equimolar incubation on ice for 30 minutes prior to separation by size-exclusion chromatography. Complexing and separation of nucleoprotein-di-gluebody complexes from the input mixture was carried out by Dr. Gangshun Yi.

Representative SEC traces are shown in figure 3.17 panels (B) and (C). It is notable that there is a band beneath the nucleoprotein band. This is a commonly observed degradation product from long-term storage of influenza A nucleoprotein. Since the vast majority of the complex was with full-length nucleoprotein, this sample was deemed appropriate to be applied to EM grids. A shift in SEC elution volume is observed when complexed with the di-gluebody consistent with complex formation greater than 100 kDa.

Samples of each fraction were taken, and grids were made by Dr. Gangshun Yi. Data collection was carried out by Dr. Loïc Carrique. Downstream data processing was carried out independently.

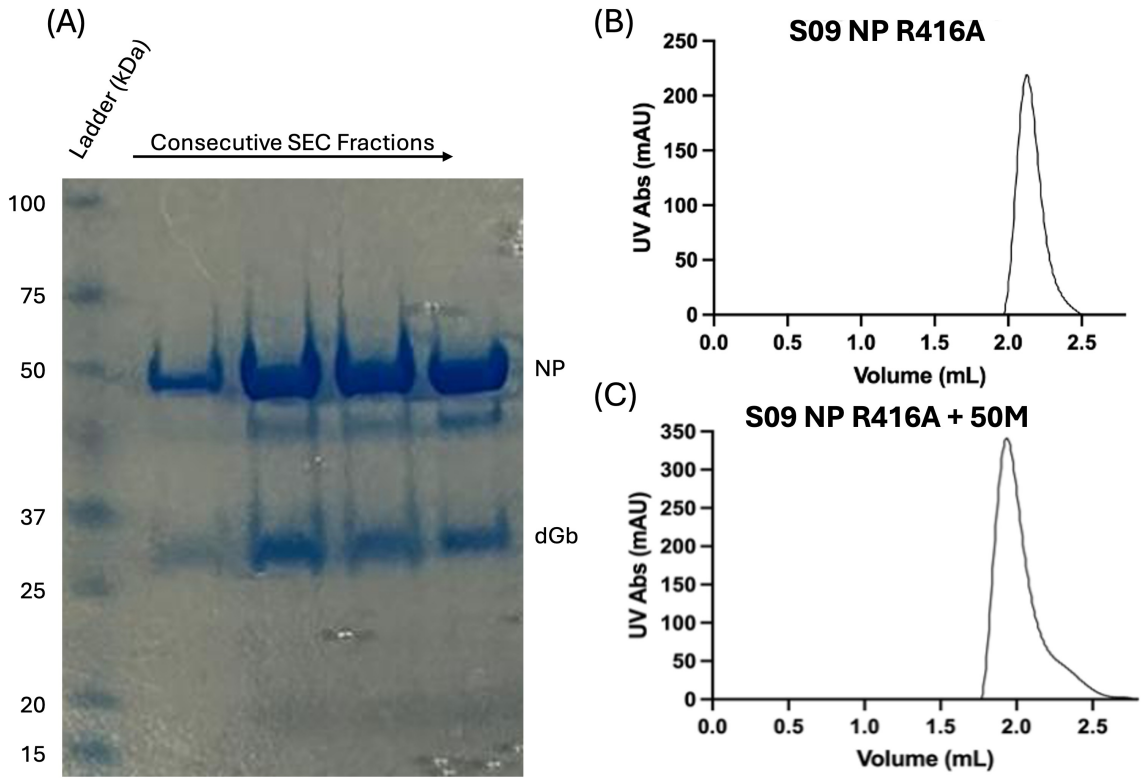


Figure 3.17: Nucleoprotein-di-Gluebody Complex Formation.

Example complex formation between S09 NP R416A mutant nucleoprotein and 50M di-gluebody. (A) SDS-PAGE gel showing consecutive SEC fractions of nucleoprotein – di-gluebody SEC experiment in panel C. (B) Chromatography trace of S09 NP R416A SEC experiment alone on an S200 5/150 column. (C) Comparative chromatography trace of the SEC experiment containing the S09 NP R416A nucleoprotein – 50M di-gluebody complex.

3.2.4.3. Cryo-EM Data Processing

Two datasets were collected of nucleoprotein – di-gluebody complexes: one with 50M bound, and one with 54S bound. The first (with 50M) was a full dataset, with over 12,000 exposures recorded. The second (with 54S) was a smaller dataset during a time of attempted optimisation of complex formation. Data collection and data processing parameters are outlined in table 3.8, below. Data collection was carried out by Dr. Gangshun Yi, but all data processing was carried out independently.

Table 3.8: Data Collection and Data Processing Statistics of Nucleoprotein-di-Gluebody Complexes.

Data collection and data processing statistics of the two nucleoprotein – di-gluebody complexes formed, with processing statistics relevant to the final round of density maps obtained.

Data Collection		
	NP – 50M Dataset	**NP – 54S Dataset**
Microscope	Titan Krios	
Voltage (keV)	300	
Detector	Falcon 4i	
Magnification	165,000	
Pixel Size (Å)	0.7303	
Total Dose ($e^-/Å^2$)	50	
Defocus Range (μm)	-0.5 to -2.5	
Frames / Movie	50	
Data Processing		
Number of Movies	12,373	4,160
Box Size (Pixels)	384	
Final Number of Particles	102,397	32,168
Resolution (FSC 0.143)	3.29	> 7
Symmetry	C1	C1 / C2

A standardised processing pipeline was implemented for both datasets, incorporating multiple rounds of 2D classification coupled with iterative Topaz training for particle

selection (Bepler et al., 2020). This was followed by *ab-initio* reconstruction, heterogeneous refinement, and non-uniform refinement to generate final density maps. This systematic approach was carried out to ensure a large selection of homogeneous particle populations containing both di-gluebody and at least one copy of the nucleoprotein.

3.2.4.3.1. NP – 50M Dataset Analysis

Comprehensive analysis of 2D class averages revealed an unexpected stoichiometry, with only single nucleoprotein copies observed to be bound to the 50M di-gluebody, rather than the anticipated 2:1 nucleoprotein:di-gluebody stoichiometry. Consequently, only 2D classes containing one copy of the nucleoprotein and the di-gluebody were selected to obtain a homogeneous particle stack of 103,000 particles.

Promisingly, 2D class averages exhibited clear secondary structure signatures, indicating a robust particle alignment, and good signal quality with the potential for high-resolution reconstruction. The best nucleoprotein – di-gluebody map was the output of a non-uniform refinement job with 103,000 input particles. The nominal resolution of this reconstruction was 4.25 Å, however, upon examination of the resulting density map, this is likely an overestimate due to the lack of high-resolution detail. Because of this, unambiguous fitting of a nucleoprotein and subsequent model building of the structure were not possible (figure 3.18).

An attempt at further processing was made using particle subtraction. This was attempted as the lowest resolution area of the density map was at the interface

between the nucleoprotein and the di-gluebody. It was therefore rationalised that conformational flexibility at the binding interface was limiting the resolution of the reconstruction. By computationally removing the nanobody density and carrying out focussed refinement using a mask around the nucleoprotein, the effect of low-resolution density may be reduced, and the overall reconstruction may be improved. Particle subtraction of density representing the nanobody was carried out and one final local refinement job was run with the density subtracted particles. This approach improved the overall map quality, resulting in a nominal resolution of 3.21 Å, which facilitated unambiguous fitting of the nucleoprotein into the density, and also a broad understanding of the epitope to which this specific nanobody binds. Further processing did not improve the overall density map, and the final reported map was of insufficient quality to carry out model building due to a lack of clear secondary structure elements and map anisotropy (figure 3.20).

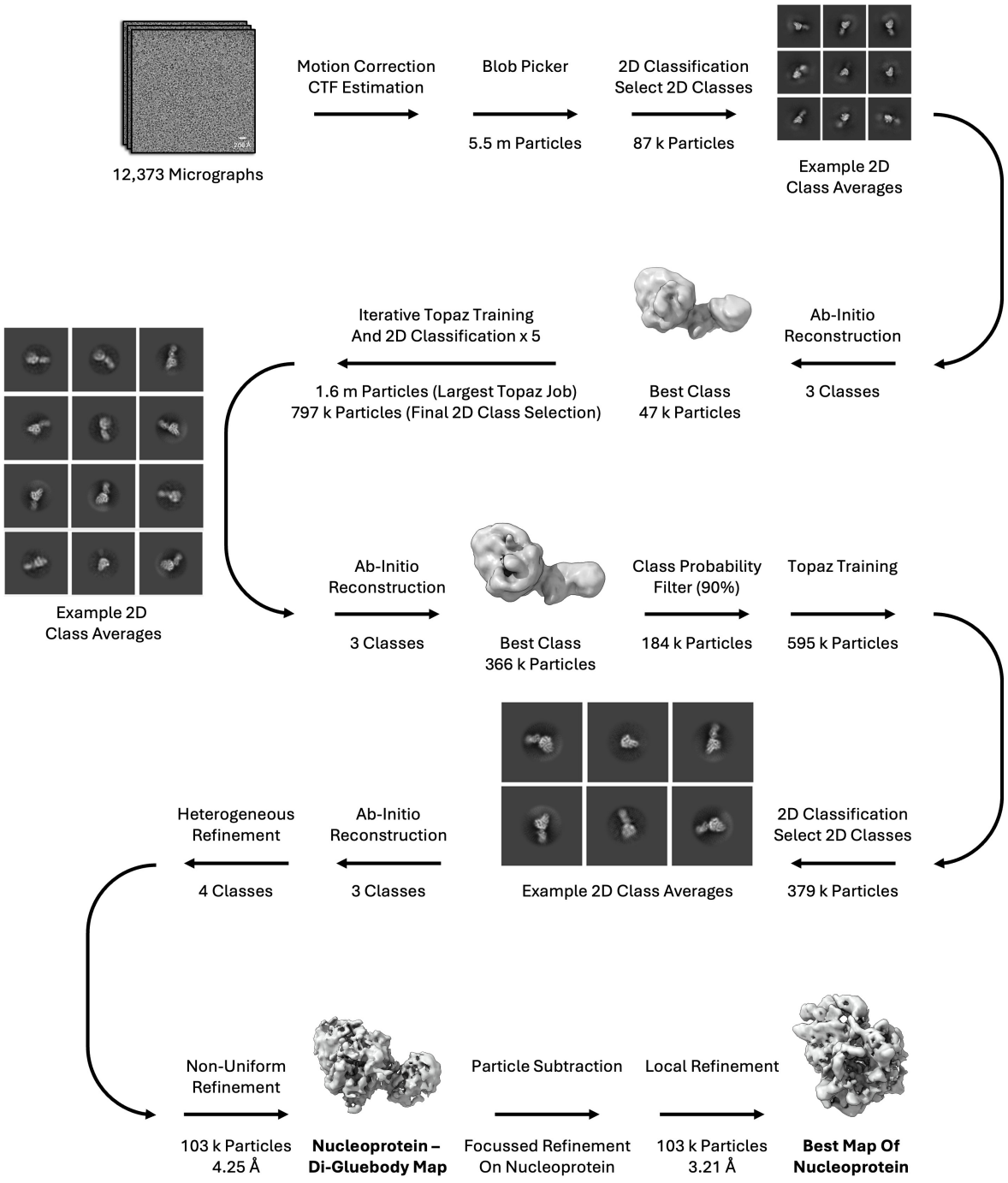


Figure 3.18: Cryo-EM SPA Data Processing Pipeline of 50M dGb + S09 NP R416A Dataset.

3.2.4.3.2. NP – 54S Dataset Analysis

The NP – 54S dataset yielded 2D classes containing the expected 2:1 nucleoprotein:di-gluebody stoichiometry, suggesting differential binding characteristics between the two di-gluebodies tested in this work. The same processing strategy of iterative 2D

classification and Topaz training was maintained to ensure homogeneous particle selection. However, reduced secondary structure visibility in 2D classes suggested a diminished potential for high-resolution reconstruction compared to the 50M dataset.

Despite fewer total exposures, comparable numbers of particles were observed throughout data processing. While some 2D classes retained secondary structure signatures, final density maps lacked sufficient quality for unambiguous nucleoprotein fitting. Multiple computational strategies were explored, including *ab-initio* reconstruction with imposed C2 symmetry to exploit inherent sample symmetry, but failed to yield improved reconstructions (figure 3.19).

It is likely that data processing was hindered by a flexible binding interface between the two gluebodies, and additionally between the gluebodies and nucleoprotein, leading to difficulty in accurate particle alignment.

These results highlight the technical challenges inherent in structural characterisation of nucleoprotein – di-gluebody complexes, likely reflecting conformational heterogeneity and dynamic binding interfaces that limit achievable resolution in single particle analysis.

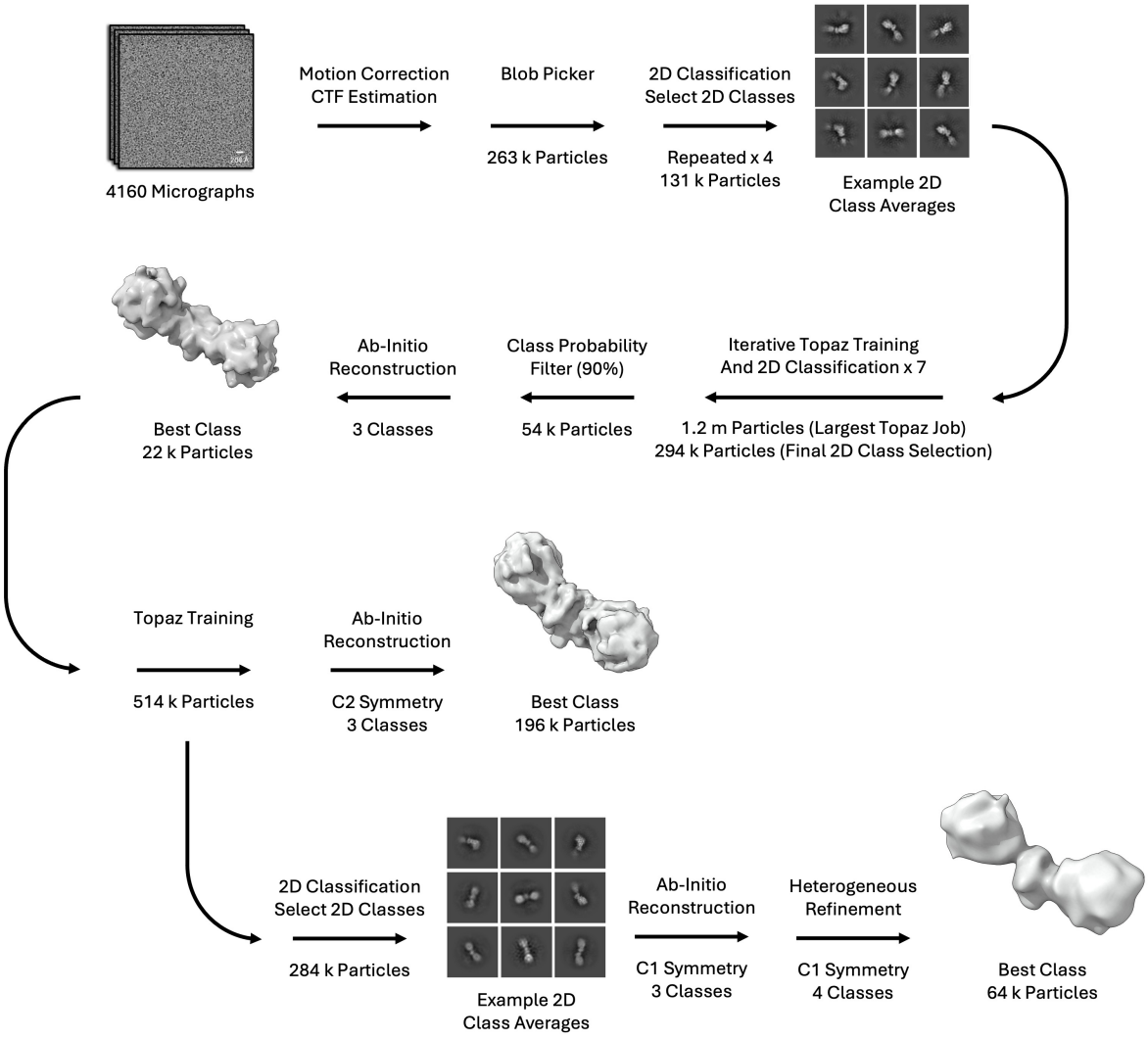


Figure 3.19: Cryo-EM SPA Data Processing Pipeline of 54S dGb + S09 NP R416A Dataset.

3.2.4.3.3. 50M dGb + S09 NP R416A Complex

The final density map obtained of the S09 NP R416A nucleoprotein bound to the 50M di-gluebody allowed for the unambiguous rigid body fitting of the influenza A nucleoprotein into the density. This was carried out with the NT60 NP R416A mutant (PDB: 7NT8). The rigid body fitting of this nucleoprotein into the focussed refinement density map is illustrated below (figure 3.20, panel A). Maintaining the same orientation, the nucleoprotein is additionally illustrated when fitted into the final nucleoprotein – di-gluebody complex map below (figure 3.20, panel B).

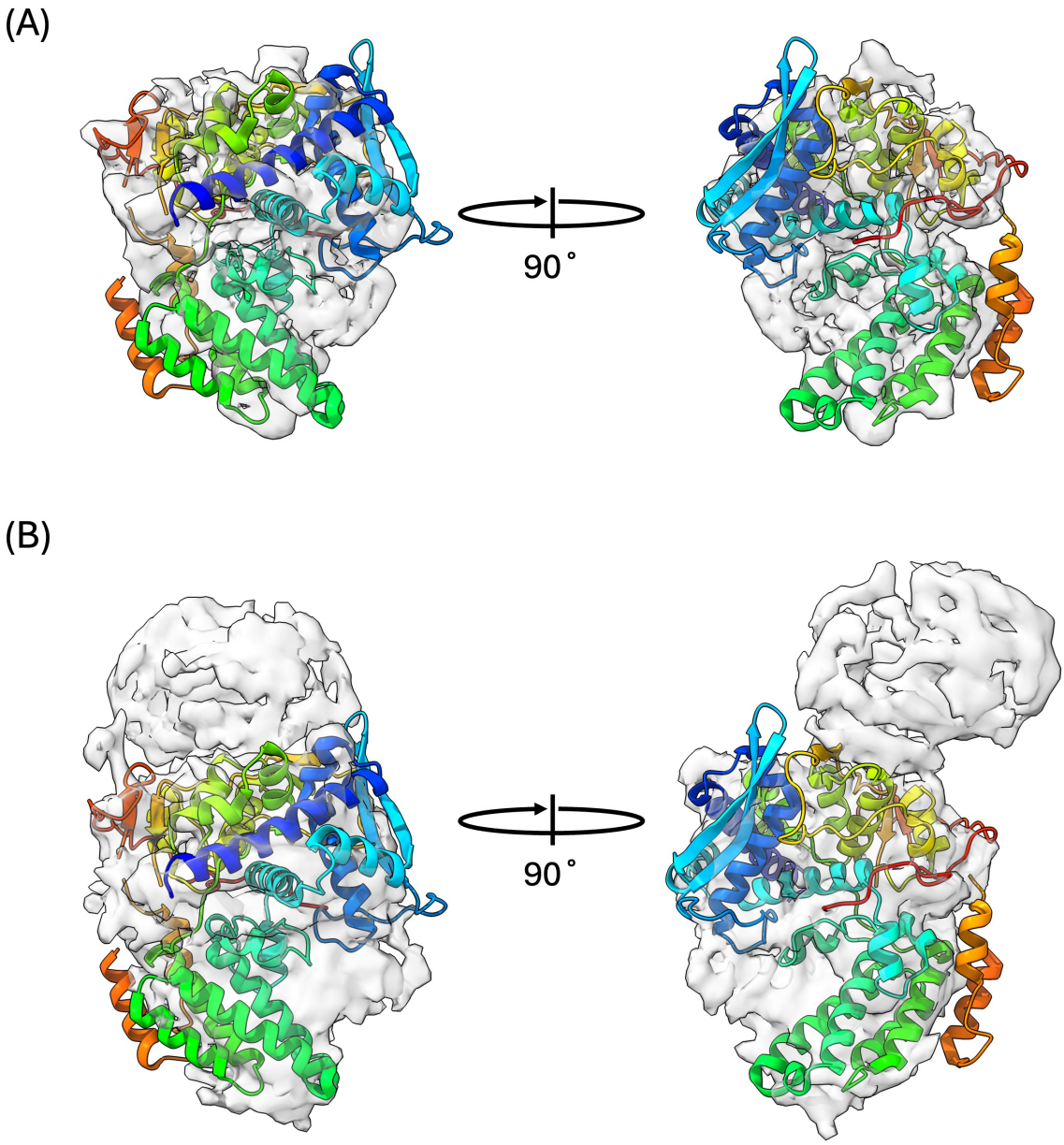


Figure 3.20: Rigid Body Fitting of NT60 NP R416A into Final 50M + S09 NP R416A Density Map.

Rigid body fitting of NT60 NP R416A mutant (PDB: 7NT8) into final density maps. (A) The nucleoprotein rigid body fitted into the final nucleoprotein density map. (B) The nucleoprotein rigid body fitted into the final nucleoprotein – di-gluebody 50M density map. In both panels, the nucleoprotein is coloured using ChimeraX's rainbow colouring scheme.

When the canonical sensitivity residues for MxA and BTN3A3 are considered (residues 16, 52, 100, 283 and 313, as introduced earlier in this chapter), then the 50M di-gluebody is observed to bind to a nearby site (figure 3.21) which likely means that

forming a complex with both 50M and MxA, or 50M and BTN3A3, would not be possible due to steric clashes.

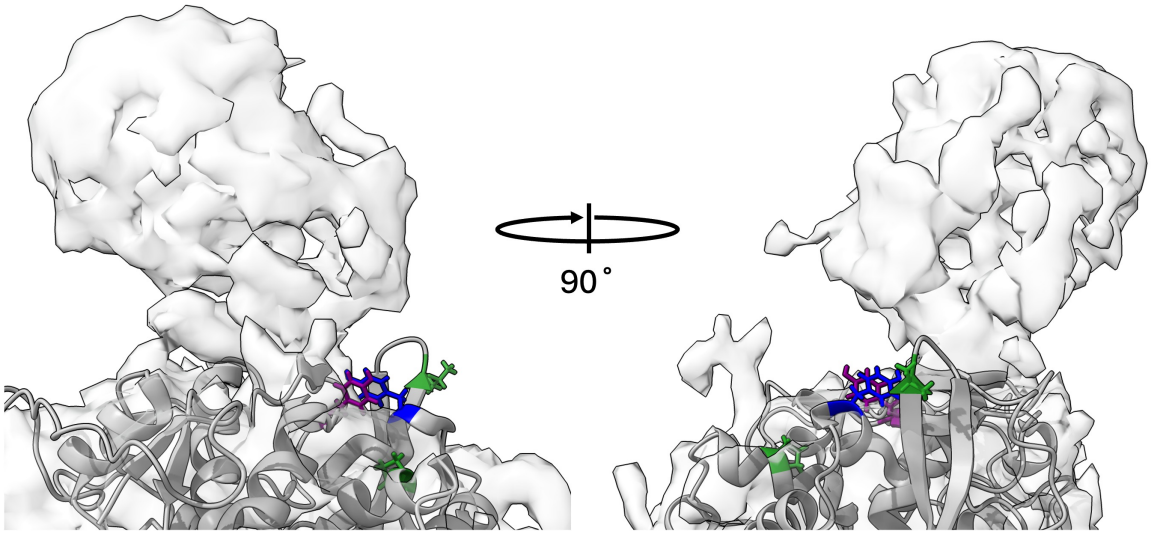


Figure 3.21: 50M dGb Binding Interface.

The 50M di-gluebody is observed to bind to the nucleoprotein at a site close to the understood sensitivity sites for MxA and BTN3A3 restriction. The density map for the nucleoprotein – di-gluebody is shaded in light grey with 25 % opacity. The nucleoprotein structure of NT60 NP R416A (PDB: 7NT8) is rigid body fitted into the density map, as above, and coloured in grey. The understood sensitivity sites for BTN3A3 restriction are coloured in blue (residue 52), for both MxA and BTN3A3 restriction are coloured in purple (residue 313), and for MxA restriction are coloured in green (residues 100 and 283). All sensitivity residues are additionally shown as atoms.

Given that firstly no 2D class averages of two nucleoproteins were found bound to the 50M di-gluebody, that secondly there is likely to be significant flexibility present at the binding interface of 50M and S09 NP R416A, and thirdly that the binding site of this gluebody and the immune factors of interest are very close (leading to potential steric clashes), it is not advisable that work with this gluebody is continued. Thus, other nanobody binding sites should be characterised in pursuit of a doubly-occupied di-gluebody whose nucleoprotein binding site does not occlude potential interactions with immune factors of interest.

3.2.5. Nucleoprotein in Complex with Megabody Mb151

As an alternative method of determining the structure of the nucleoprotein, a larger protein scaffold was used, a megabody (Mb), in an attempt to produce a stable complex of S09 NP R416A nucleoprotein in complex with Mb151, a megabody (Mb) derived from a nucleoprotein nanobody (VHH151). The nanobody on which the megabody was designed was again produced by Schmidt and colleagues (Schmidt et al., 2016), and the megabody derivative was cloned and produced by a previous DPhil student, Dr. Amanda Zhu, and donated for use in this study. This nanobody is known to block vRNP nuclear import and transcription/translation, but not to block nuclear import of the nucleoprotein. Here, I demonstrate the formation of an S09 NP R416A complex, and present cryo-EM SPA data processing of the resulting sample.

3.2.5.1. Complex Formation

A complex was formed between S09 NP R416A and Mb151 by incubating the two proteins in a 1:1 molar ratio on ice for 30 minutes. The resulting mixture was then separated by size-exclusion chromatography and the peak fractions analysed by SDS-PAGE. A control SEC run of the nucleoprotein alone was carried out to determine the position of the nucleoprotein in comparison to the complex.

The resulting SEC chromatogram and SEC fraction SDS-PAGE gel are illustrated below (figure 3.22). The starred fraction (*), which corresponded to the top of the top of the left-most peak illustrated in figure 3.22, panel B, was taken and concentrated to 0.6 mg/mL before cryo-EM grids of the complex were made using freshly glow-discharged

Quantifoil Cu 1.2/1.3 300 mesh grids. This specific fraction was chosen given the similar intensity of staining of both the megabody and nucleoprotein where earlier fractions had more intense staining for the megabody, and later fractions had more intense staining for the nucleoprotein. This suggested that the starred fraction may contain more equal amounts of both component proteins in complex than compared to earlier or later fractions.

Screening was carried out independently on a 200 keV Thermo Fisher Glacios microscope to identify good grids, and data collection was carried out on a 300 keV Thermo Fisher Krios microscope by Dr. Loïc Carrique. Downstream data processing was carried out independently.

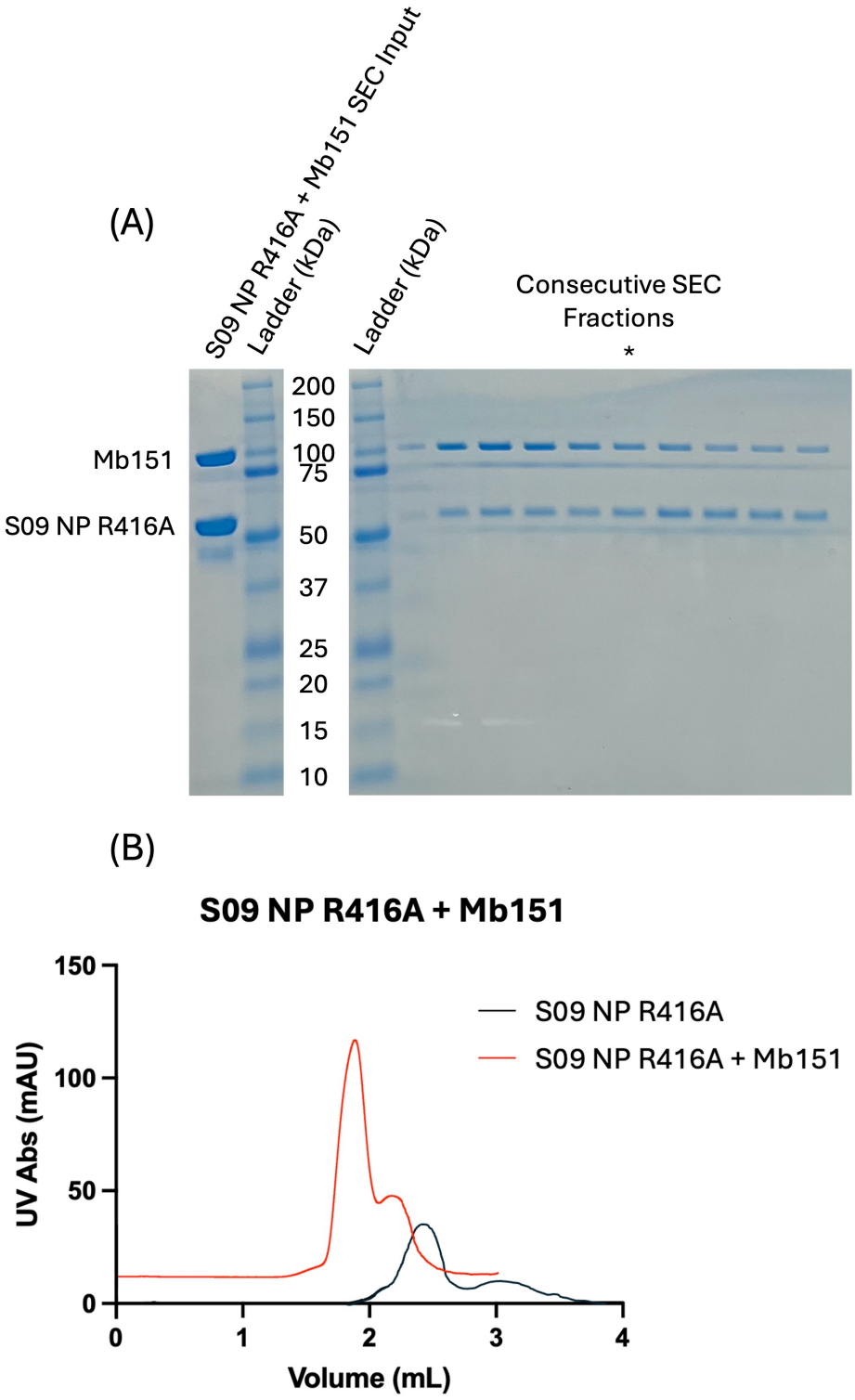


Figure 3.22: S09 NP R416A + Mb151 Complex Formation.

(A) SDS-PAGE gel of SEC input and consecutive SEC fractions from the complex formation experiment. The gel is split into two to allow for clear labelling of the ladder. The starred fraction (*) was used to make cryo-EM grids of the complex and corresponds to the first red peak. (B) SEC chromatogram of experiments with either S09 NP R416A alone (in black) or S09 NP R416A in complex with Mb151 (in red).

3.2.5.2. Cryo-EM Data Processing

Tabular information on data collection, data processing, and model refinement is found in the table below (Table 3.9).

Table 3.9: Cryo-EM Data Collection, Data Processing, and Model Refinement Statistics.

Data collection parameters, key data processing values, and model refinement statistics for the S09 NP R416A + Mb151 composite map structure.

Data Collection – Mb151 + S09 NP R416A	
Microscope	Titan Krios
Voltage (keV)	300
Detector	Falcon 4i
Magnification	165,000 x
Pixel Size (Å)	0.7303
Total Dose ($e^-/Å^2$)	50
Defocus Range (μm)	-0.5 to -2.5
Frames / Movie	60
Data Processing – Mb151 + S09 NP R416A	
Number of Movies	12,864
Box Size (Pixels)	350
Final Number of Particles	38,054
Resolution (FSC 0.143)	Mb Core = 2.85 Å, Nucleoprotein = 3.15 Å
Map Resolution Range (Å)	1.50 – 46.00
Symmetry	C1
Model Refinement – Mb151 + S09 NP R416A	
Initial Model Used	ModelAngelo
Model Resolution (Å) (FSC = 0.143)	3.1
FSC model-map (FSC = 0.5)	0.287
Map Sharpening *B* Factor ($Å^2$)	81.6
Composition	
Non-Hydrogen Atoms	9785
Protein Residues	1229
Protein B Factor ($Å^2$) (max/mean)	131.27 / 85.74
RMSD from Ideal	
Bond Lengths (Å)	0.006
Bond Angles (°)	0.889
Validation	
MolProbity Score	1.83
Clashscore	10.10
Rotamer Outliers (%)	1.27
Ramachandran	
Favoured (%)	96.55
Allowed (%)	3.45
Outliers (%)	0.00

Following motion correction, CTF correction, and blob picking, 6.8 million (m) particles

were extracted from 12,408 curated micrographs with a box size of 350 pixels (px) (binned to 128 px, ≈ 2.7 x bin factor). Specific box sizes were chosen in considering the maximum dimensions of the particle of interest, with added room around this to ensure little data were lost to allow for accurate CTF (Contrast Transfer Function) estimation and correction. Specific box sizes were chosen in accordance with advice in the EMAN2 documentation (https://blake.bcm.edu/emanwiki/doku.php?id=eman2:boxsize) to minimise processing power required in initial *ab-initio* processing.

An initial round of 2D classification was run to select particles in classes resembling the Mb-nucleoprotein complex, and these 2D class averages were fed into a template picker job to find more, similar particles. 18,000,000 (18 m) such particles were found. Inspect picks was used to ensure *bona fide* particles were being selected, not ice and not samples on carbon, after which the particles were extracted from their micrographs and put through 2D classification again. 246,000 (246 k) particles were selected in the select 2D class job, and the same round of template picker, inspect particle picks, extract from micrographs, 2D classification, and select 2D classes workflow was run to select 246 k genuine particles in 27 promising classes (as shown in figure 3.23, below).

Ab-initio reconstruction was then carried out with 5 classes (due to the flexibility of both the Mb and nucleoprotein). The best class had 89 k particles, and clear regions of Mb and nucleoprotein. Iterative Topaz training was used three times until a point where less than a 50 % increase in particle number was observed. After this, 9 rounds of

iterative heterogeneous refinement were used on the binned particles (Haas et al., 2025). Each round had one ‘good’ class from the previous refinement job and three ‘junk’ classes from an earlier refinement job of a bad particle set.

The particles from the one good class were then fed into the next heterogeneous refinement job and this process repeated until fewer than 10 % of the particle set was being removed at each step. A local refinement step was then carried out prior to un-binning of the particles in an extract from micrographs job. A final non-uniform refinement job was carried out on the resulting 38 k particles, with the resulting map corresponding to the Mb core (and some flexible information from the nanobody and nucleoprotein) being reported to 2.85 Å (see figure 3.24, below). As there was little to no genuine signal for the nucleoprotein, a strategy of particle subtraction and local refinement was carried out in order to focus the refinement on the nucleoprotein molecule. After the final local refinement step, the nucleoprotein volume was reported to 3.15 Å (see figure 3.24, below).

A composite map of the best Mb core and best nucleoprotein maps was created in PHENIX for model building and refinement.

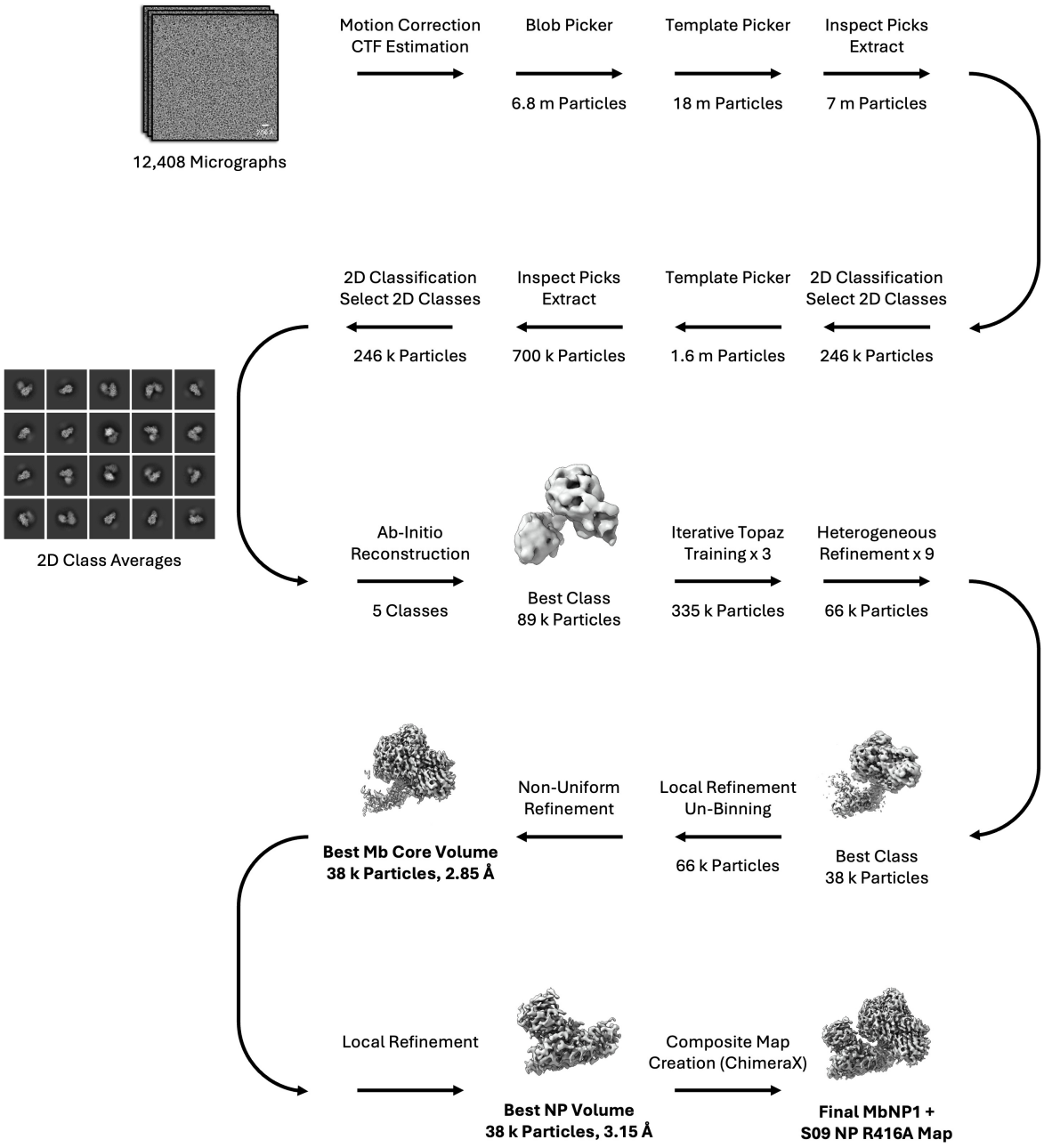
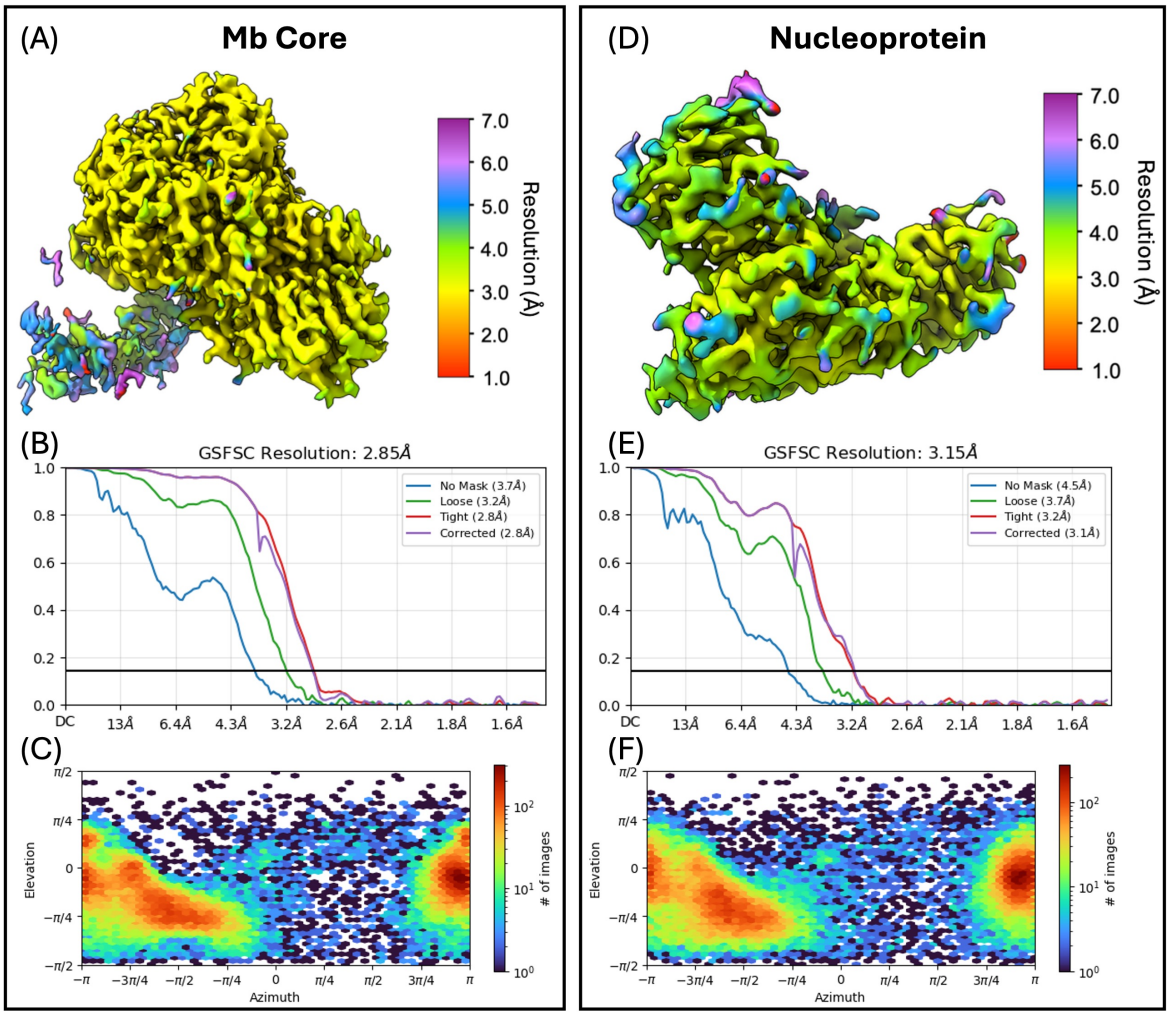


Figure 3.23: Cryo-EM Processing Pipeline for S09 NP R416A + Mb151 Complex.

Pipeline for EM processing of S09 NP R416A + Mb151 data from micrographs through to final composite map production. All data processing steps were carried out in cryoSPARC unless otherwise stated.


(A)
Mb Core
7.0
6.0
5.0
4.0
3.0
2.0
1.0
Resolution (Å)
(B)
GSFSC Resolution: 2.85Å
No Mask (3.7Å)
Loose (3.2Å)
Tight (2.8Å)
Corrected (2.8Å)
DC
13Å
6.4Å
4.3Å
3.2Å
2.6Å
2.1Å
1.8Å
1.6Å
(C)
Elevation
Azimuth
# of images
(D)
Nucleoprotein
(E)
GSFSC Resolution: 3.15Å
No Mask (4.5Å)
Loose (3.7Å)
Tight (3.2Å)
Corrected (3.1Å)
(F)

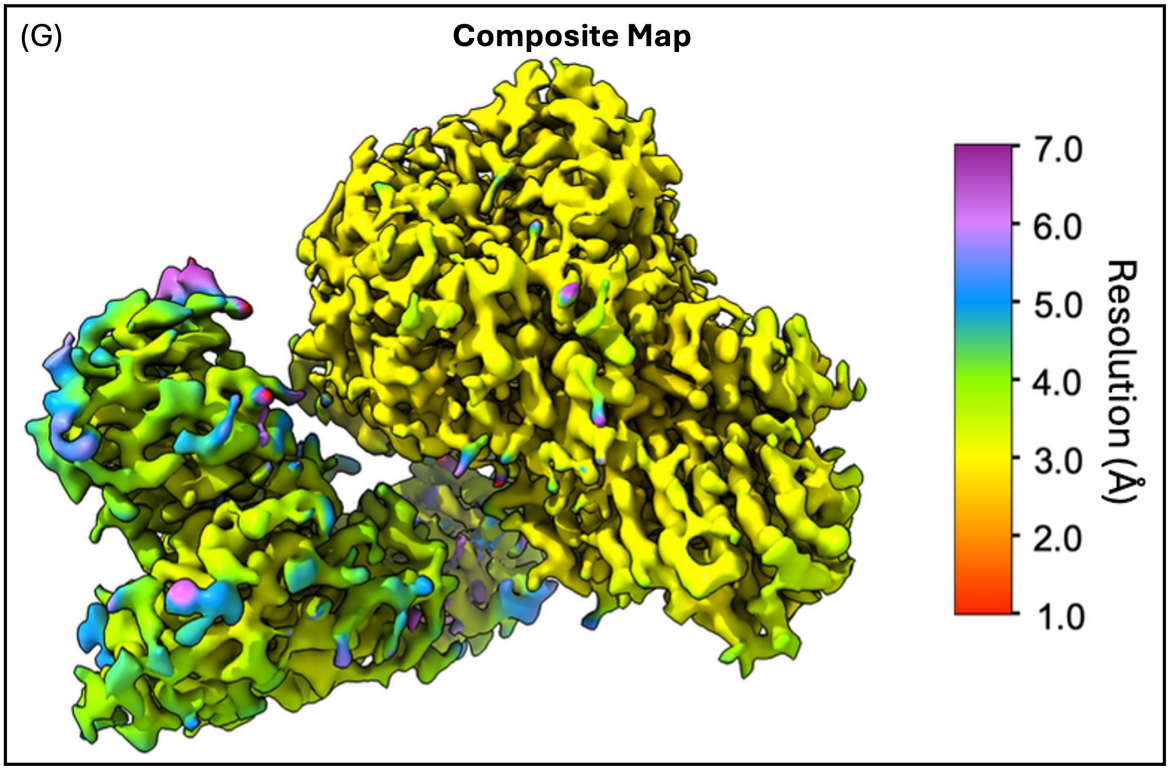

(G)
Composite Map
7.0
6.0
5.0
4.0
3.0
2.0
1.0
Resolution (Å)

Figure 3.24: Local Resolution Maps and Accompanying Resolution and View Distribution Information for Final S09 NP R416A + Mb151 Complex Density Map.

(A) Mb core local resolution map. (B) Mb core gold-standard Fourier Shell Correlation curve. Nominal resolution (FSC = 0.143) = 2.85 Å. (C) Mb core particle Angular Distribution Plot. (D) Nucleoprotein local resolution map. (E) Nucleoprotein gold-standard Fourier Shell Correlation curve. Nominal resolution (FSC = 0.143) = 3.15 Å. (F) Nucleoprotein particle Angular Distribution Plot. (G) Composite map of Mb core and nucleoprotein at the same threshold value coloured with the same local resolution scale.

3.2.5.3. Structure of S09 NP R416A + Mb151 Complex

With the processing and refinement strategy outlined above, ModelAngelo (Jamali et al., 2024) was used to build initial models of both the S09 NP R416A nucleoprotein and Mb151 by providing the composite map and sequences of the two component proteins to a centrally installed version of ModelAngelo on the Biomedical Research Computing (BMRC) cluster.

Initial ModelAngelo model production did not include the first and final beta strands of the nanobody domain of the megabody (strands A and G, residues 3-13 and 880-885 based on the predicted secondary structure of the Mb151 AlphaFold3 prediction in ChimeraX). Rather, the model was built from residue 15 (after the first beta strand) until residue 871 (before the final beta strand). This is likely due to relatively poor density of these strands compared to others in the nanobody domain, as well as the flexible nature of the linker regions connecting the nanobody domain to the megabody core. Being flexible, these regions were poorly resolved in the experimental density map and this likely lowered the ModelAngelo confidence score of these regions, leading to their absence in the initial model (figure 3.25). As a result, the nanobody domain of the megabody was rebuilt by fitting reference structures and predictions to assess optimal nanobody placement within the visible density.

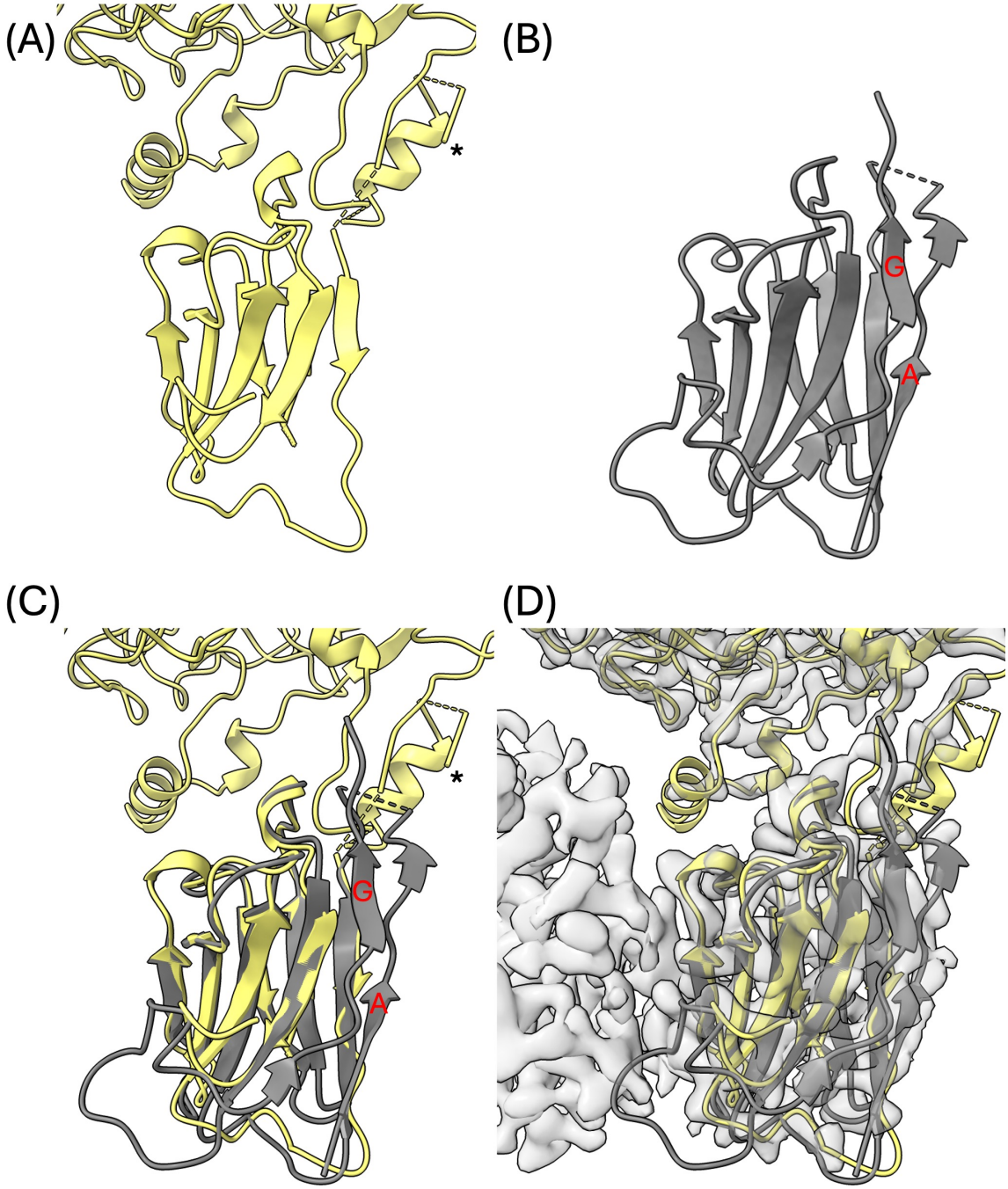


Figure 3.25: Missing Chains from Initial ModelAngelo Model of Mb151 Nanobody.

*(A) Initial ModelAngelo model of constituent nanobody of Mb151 (yellow). * Indicates the start of the ModelAngelo model, at residue number 15. (B) Constituent nanobody domain of a solved megabody structure (PDB: 6XUX) (grey) aligned with the ModelAngelo model in panel (A) using the matchmaker command in ChimeraX. The start of the first beta strand, A, and the end of the final beta strand, G, are annotated. (C) Superimposition of the ModelAngelo model and the matchmaker aligned nanobody domain of the megabody structure 6XUX. The missing chains from the ModelAngelo model are labelled as in panel (B), and the start of the ModelAngelo at residue 15 is again annotated by *. (D) The superimposition displayed in panel (C) fitted into the experimentally determined density map, displaying*

the fit of strands A and G of the solved structure 6XUX into the density observed in this experimental density map.

A series of rigid body fitting experiments with various predictions was carried out in order to attempt to determine the quality of the initial ModelAngelo model, and to assign density to the first and last beta strands which were missing. Structural models of the nanobody based on which the megabody was produced again using ModelAngelo with a map which had been auto-sharpened in PHENIX.

A model given the sequence of the nanobody from which the megabody was designed alone, and a model of the megabody alone were observed to contain the last beta strand but not the first beta strand. AlphaFold3 predictions of the original nanobody and megabody either alone or in complex with the nucleoprotein predicted all beta sheets and were highly similar to the solved structure of the megabody with the same cYgjK backbone (Uchański et al., 2021) (PDB: 6XUX). The structural models, AlphaFold3 predictions, and solved megabody structure were all superimposed onto the initial ModelAngelo model nanobody domain and nanobody structure was compared.

Cross-validation against the ModelAngelo models, AlphaFold3 predictions and PDB structure confirmed the accuracy of the initial ModelAngelo model and provided both missing strands. Both missing beta strand were consistently positioned within a gap which corresponded to the solved megabody nanobody domain structure. This rigid body fitting was well-supported by the density in the experimental map, as shown (figure 3.26).

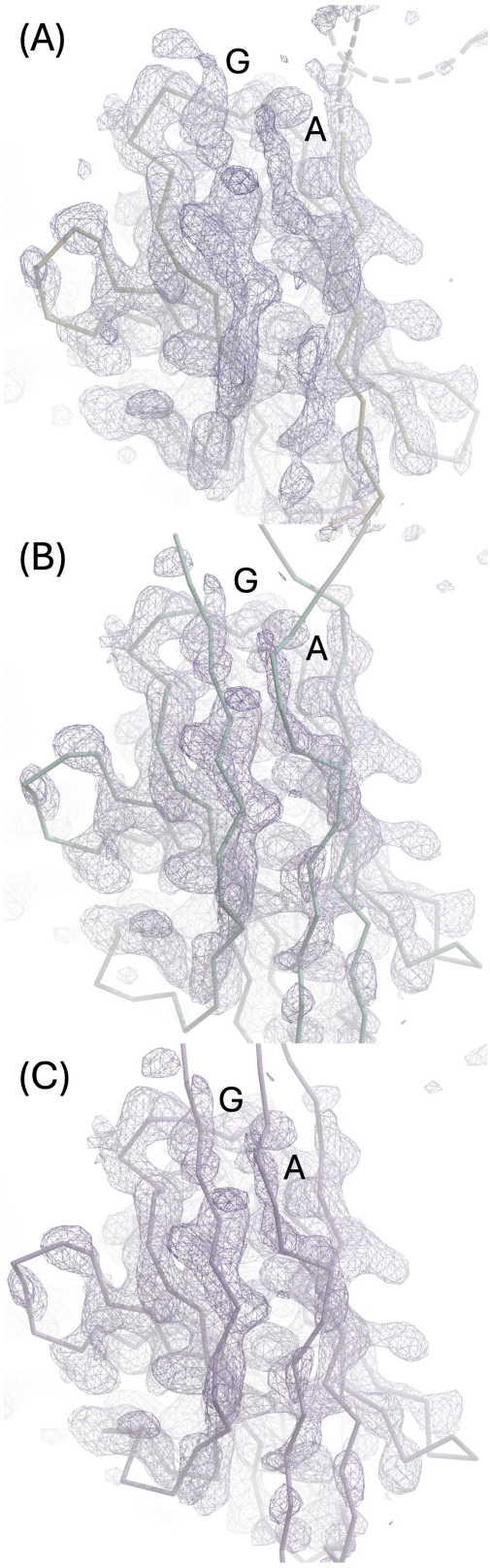


Figure 3.26: Nanobody Missing Chain Replacement.

(A) The initial ModelAngelo model of the megabody displayed in the initial map. (B) The rigid body fitted solved structure of the megabody (PDB: 6XUX) superimposed onto the auto-sharpened map. (C) The

second ModelAngelo model of the megabody displayed in the auto-sharpened map. In each panel, the density for strands A and G are annotated.

As a result, the AlphaFold3 prediction of the constituent nanobody domain of the megabody was used to rebuild the nanobody section of the initial ModelAngelo model of the megabody. The final structure was refined with iterative rounds of PHENIX (Adams et al., 2010) real-space refinement and manual model inspection and improvement in Coot (Emsley et al., 2010). The final model was validated with MOLPROBITY (Chen et al., 2010) and refinement statistics are additionally reported in table 3.9.

The overall structure of the megabody – nucleoprotein complex is displayed below (figure 3.27).

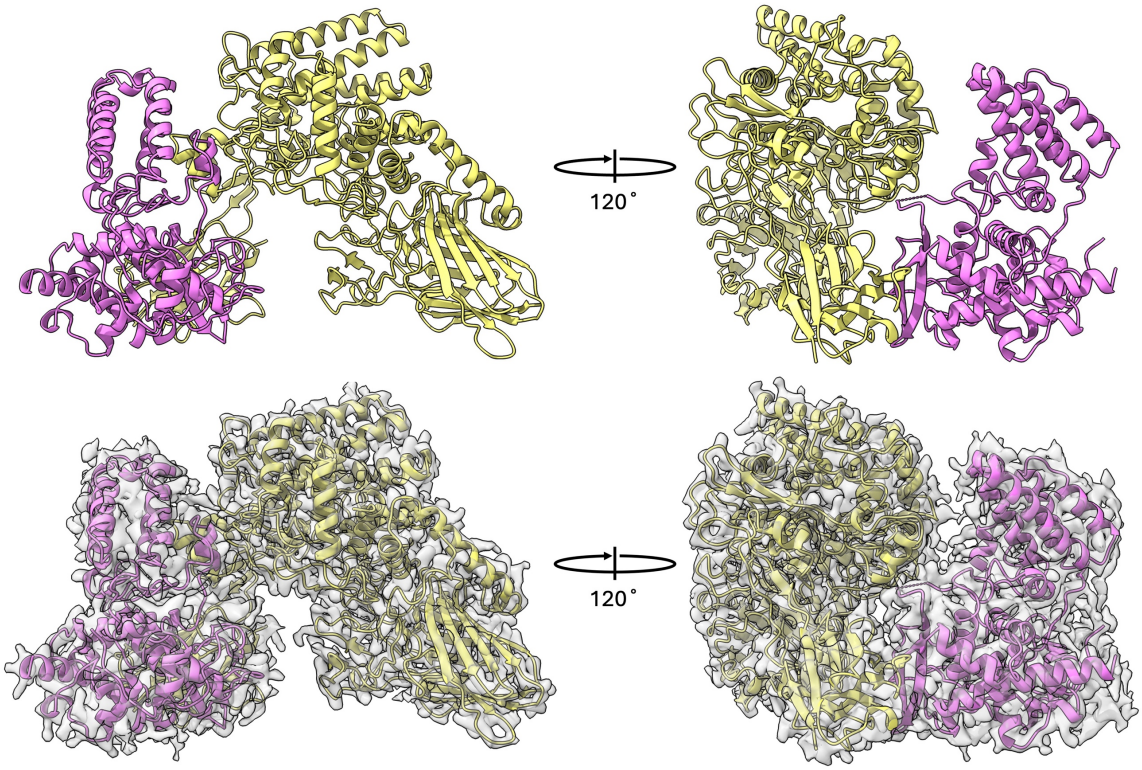


Figure 3.27: Overall Cryo-EM Structure of the Mb151 Megabody in Complex with the S09 NP R416A Nucleoprotein.

The megabody is displayed in yellow and the nucleoprotein is displayed in pink. The structure is fit into the density map at a transparency of 30 % in the bottom half of the figure.

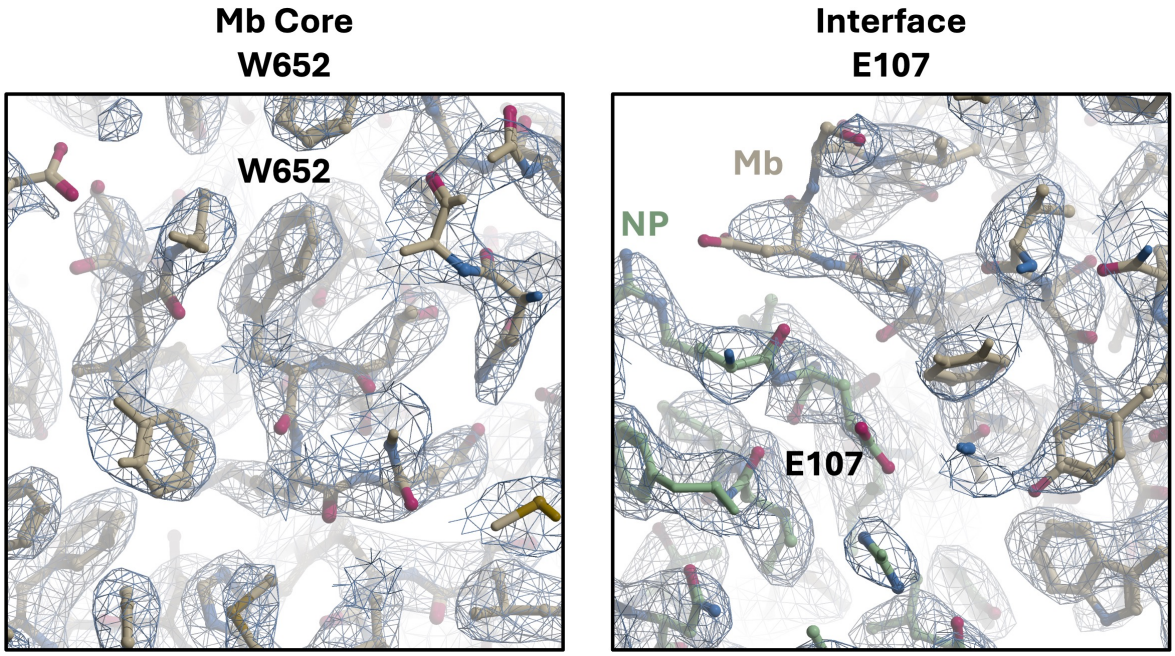


Figure 3.28: Example Density Fitting of the Mb151 Megabody in Complex with the S09 NP R416A Nucleoprotein.

(Left) Example density fitting within the megabody core centred around residue W652 (labelled). (Right) Example density fitting at the nucleoprotein – megabody interface centred around nucleoprotein residue E107 (labelled). The proteins modelled are coloured by heteroatom, with the backbone coloured by chain (nucleoprotein – green, megabody – brown).

The observed density of the S09 NP R416A + Mb151 complex does not support a significant CDR3 loop interaction. Instead, the interaction is mediated through residues involved in two beta sheets. The beta sheet from the nucleoprotein implicated in the interaction from the nucleoprotein is composed of residues 106-112. The beta sheet from the megabody implicated in the interaction is composed of residues 828-832. When comparing predicted secondary structure elements, this corresponds to residues 57-61 of a nanobody, which relates to the C'' strand of the nanobody (with some potential overlap with the end of the CDR2 loop).

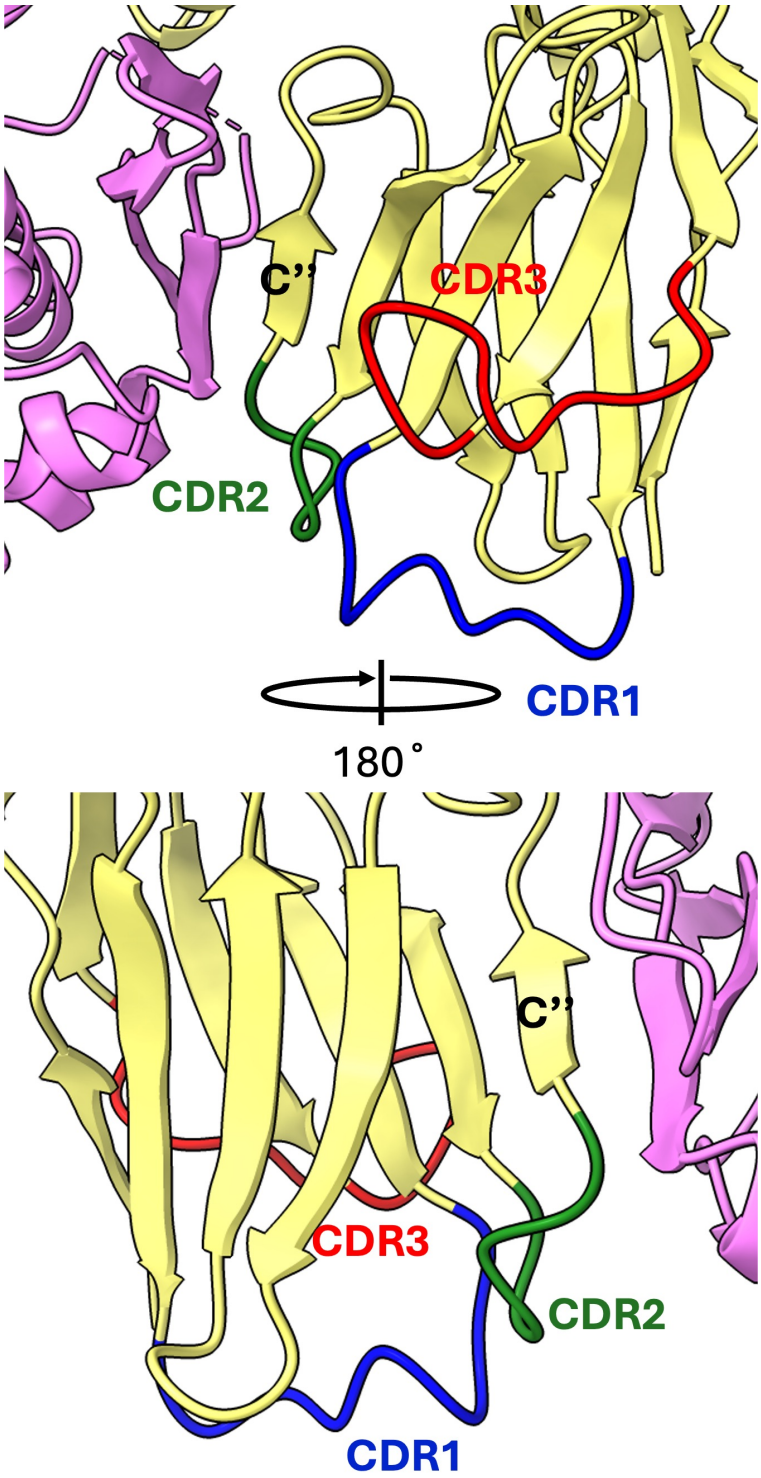


Figure 3.29: Interaction Interface between Mb151 and S09 NP R416A.

The interface between the Mb151 megabody (yellow) and the S09 NP R416A nucleoprotein (pink) is made predominantly of the C'' strand of the constituent nanobody of Mb151, with some potential contribution from the CDR2 and CDR3 loops.

The binding interface of the resulting structure was analysed using PISA (Krissinel and Henrick, 2007). This analysis showed that six hydrogen bonds and two salt bridges were present stabilising the interaction between S09 NP R416A and Mb151 (figure 3.30; table 3.10) . The residues implicated in the interaction interface on the nucleoprotein were all within the residue range 106 – 115, which corresponds mostly to a predicted beta sheet. The residues implicated in the interaction interface on the megabody were mostly localised to the residue range 827 – 833, which corresponds to the end of the CDR2 loop, the length of the C'' strand, and one residue of the subsequent loop of the constituent nanobody. The salt bridges were both contributed by R871, which belongs to the CDR3 loop – the only residue from this hypervariable loop detected to form a stabilising interaction at the interface.

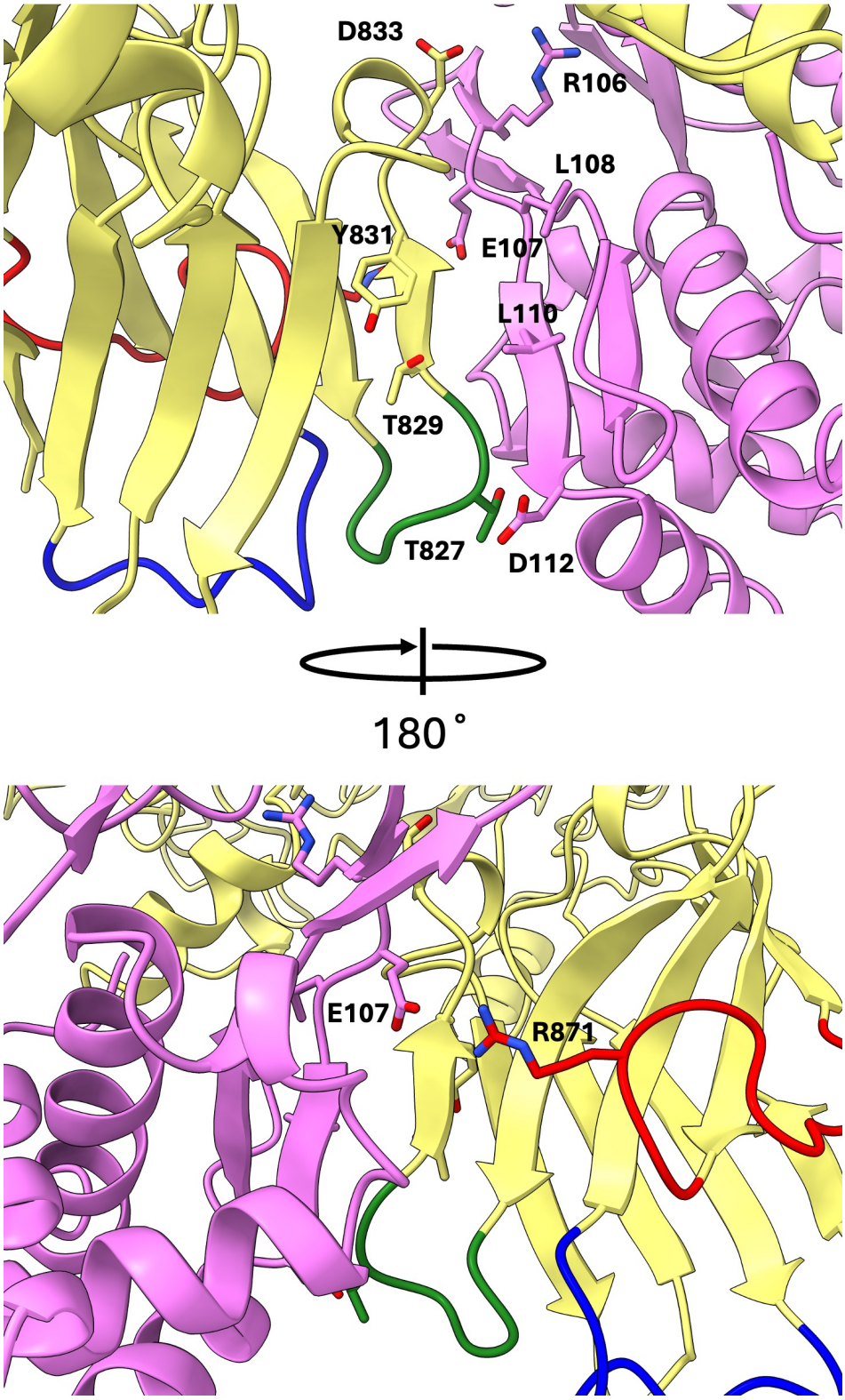


Figure 3.30: Residues Contributing to the Binding Interface of S09 NP R416A and Mb151.

Residues identified by PISA analysis to contribute to the binding interface of S09 NP R416A and Mb151. Mb151 displayed in yellow, S09 NP R416A displayed in pink. Residues determined to be contributing to the binding interface are shown as atoms, coloured by heteroatom, and labelled. Residues from the nucleoprotein are always coloured in pink. Residues from the Mb151 nanobody core are always coloured in yellow. Residues from the Mb151 CDR loops are coloured blue (CDR1), green (CDR2), or red (CDR3).

Table 3.10: Residues Contributing to the Binding Interface of S09 NP R416A and Mb151.

Table indicating residues identified by PISA analysis to contribute to the binding interface of S09 NP R416A and Mb151 by either forming hydrogen bonds or salt bridges. No disulfide or covalent bonds were present at the interaction interface.

Hydrogen Bonds		
Nucleoprotein	**Megabody**	**Distance (Å)**
R106 (O)	R833 (N)	3.89
E107 (OE2)	R871 (NH2)	3.30
E107 (OE2)	R871 (NH1)	3.37
L108 (O)	Y831 (N)	3.78
L110 (O)	T829 (N)	3.28
D112 (N)	T827 (O)	2.71
Salt Bridges		
E107 (OE2)	R871 (NH2)	3.30
E107 (OE2)	R871 (NH1)	3.37

Whilst CDR3 is understood to be the main contributor to binding partner interactions in nanobody binding, whereas CDR1 and CDR2 are reported to assist in the binding strength (Mitchell and Colwell, 2018b; Mitchell and Colwell, 2018a; Zavrtanik et al., 2018; Jovčevska and Muyldermans, 2020), it is also known that antigen binding can be assisted by residues form the nanobody core (i.e. non-CDR regions) (Zavrtanik et al., 2018). In this study, Zavrtanik and colleagues determined that, from 105 nanobody – antigen complexes deposited on the protein data bank, although 50.6 % of interaction interfaces were mediated by the CDR3 loop, 16 % of all nanobody – antigen contacts were mediated by non-CDR loop residues. Although uncommon, there is clear density for this interaction interface, and the mode of binding is supported by multiple AlphaFold3 predictions of this megabody, or the original nanobody, binding to the S09 NP R416A nucleoprotein in this manner.

The buried surface area between the nucleoprotein and nanobody is approximately 540 $Å^2$. This is less than the standard interface of binding between nanobodies and

their targets, commonly around 600 – 800 Å^2 (Beghein and Gettemans, 2017). However, the buried surface area of nanobodies (and sybodies, synthetic nanobodies) and their targets are also observed between a range of 585 Å^2 and 1663 Å^2 in one study (Ahmad et al., 2021), and the only solved structure of an influenza nucleoprotein bound to a nanobody also found a buried surface area of 542 Å^2 (Hanke et al., 2016) (PDB: 5TJW). The interface in the published structure however is seen to use residues from all three CDRs, whereas in the megabody-bound structure presented in this chapter, the main interactions between the nanobody and nucleoprotein originate from the C'' sheet of the nanobody, supplemented by one additional residue from each the CDR2 and CDR3 loops.

The low buried surface area observed in this megabody – nucleoprotein complex is likely explained due to the relatively short CDR3 loop, 10 amino acids (residues 869 – 879). This is within the expected range of the CDR3 loop, between 3 and 28 amino acids (Bannas et al., 2017), but shorter than others observed.

In figure 3.31, the nanobody domain of Mb151 (yellow) is shown overlaid with the nanobody domain of the megabody in the PDB deposition 6XUX (grey). Here, the CDR3 loops are both coloured in red. The CDR3 loop length of the nanobody domain of the deposited structure 6XUX is 14 residues (871 – 885) which offers a greater range of flexibility to the CDR3 loop to form more interaction contacts with its binding partner.

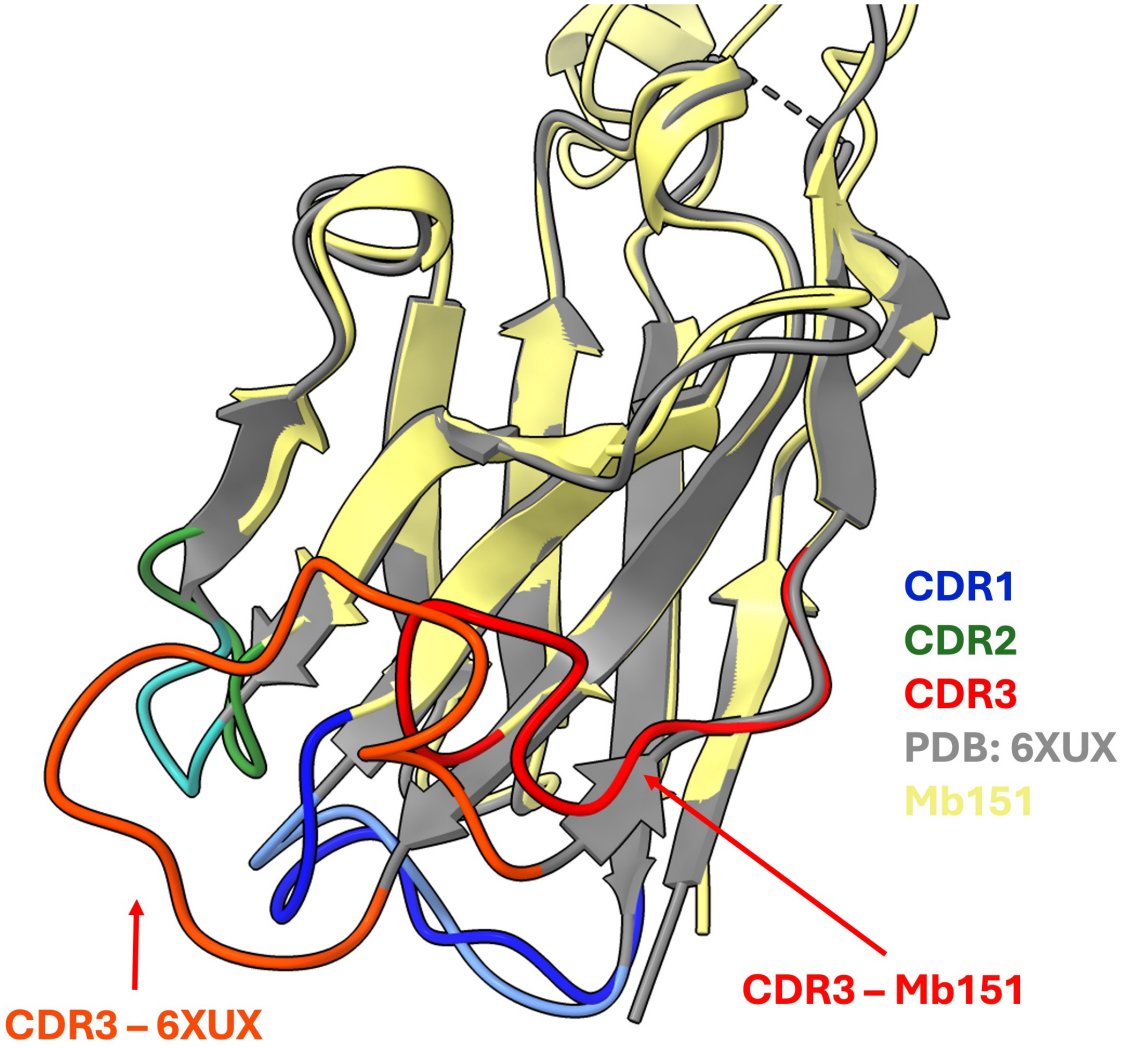


Figure 3.31: Comparison of CDR3 Loops from Nanobody Domains of Megabodies.

The nanobody core of the deposited structure (PDB: 6XUX) is displayed in grey, and the nanobody core of the solved structure in this chapter, Mb151, is displayed in yellow. The CDR loops are coloured by chain, with the CDR loops from the deposited structure being lighter than that of the determined structure: CDR1 is blue (Mb151) or light blue (PDB: 6XUX), CDR2 is green (Mb151) or light green (PDB: 6XUX), and CDR3 is red (Mb151) or orange (PDB: 6XUX). The CDR3 loops corresponding to both structures are labelled for clarity.

Where no density was present to clearly model peptide backbone, protein residues were deleted. Therefore, the nucleoprotein was only modelled from residue 22. Residues 83 – 86 were not resolved, and thus, were not modelled. Additionally, the C-terminal of the protein (residues 391 to 498) was not resolved and also not modelled. As such, the flexible C-terminal 402 – 428 loop was not present in the solved structure.

The mutated residue, A416, was therefore not modelled, but the other required residue for oligomerisation, E339 (in the well-ordered body domain of the nucleoprotein) was resolved.

The overall structure of the solved nucleoprotein is very similar to that of the crystal structure of the NT60 NP R416A mutant (PDB: 7NT8), with only marginal differences in secondary structure elements. The RNA binding grooves between the two structures adopt a similar conformation, and the major difference between the two proteins is observed to be the unresolved C-terminal tail loop (figure 3.32).

No density for RNA was observed, as expected with the R416A monomeric nucleoprotein (Knight et al., 2021). This is likely to be due to the fold of two flexible loops – earlier identified as the D72-K90 and G490-N498 loops. Here, only the former is resolved (figure 3.32), but it adopts a very similar conformation to that observed in the previously solved crystal structure (PDB: 7NT8).

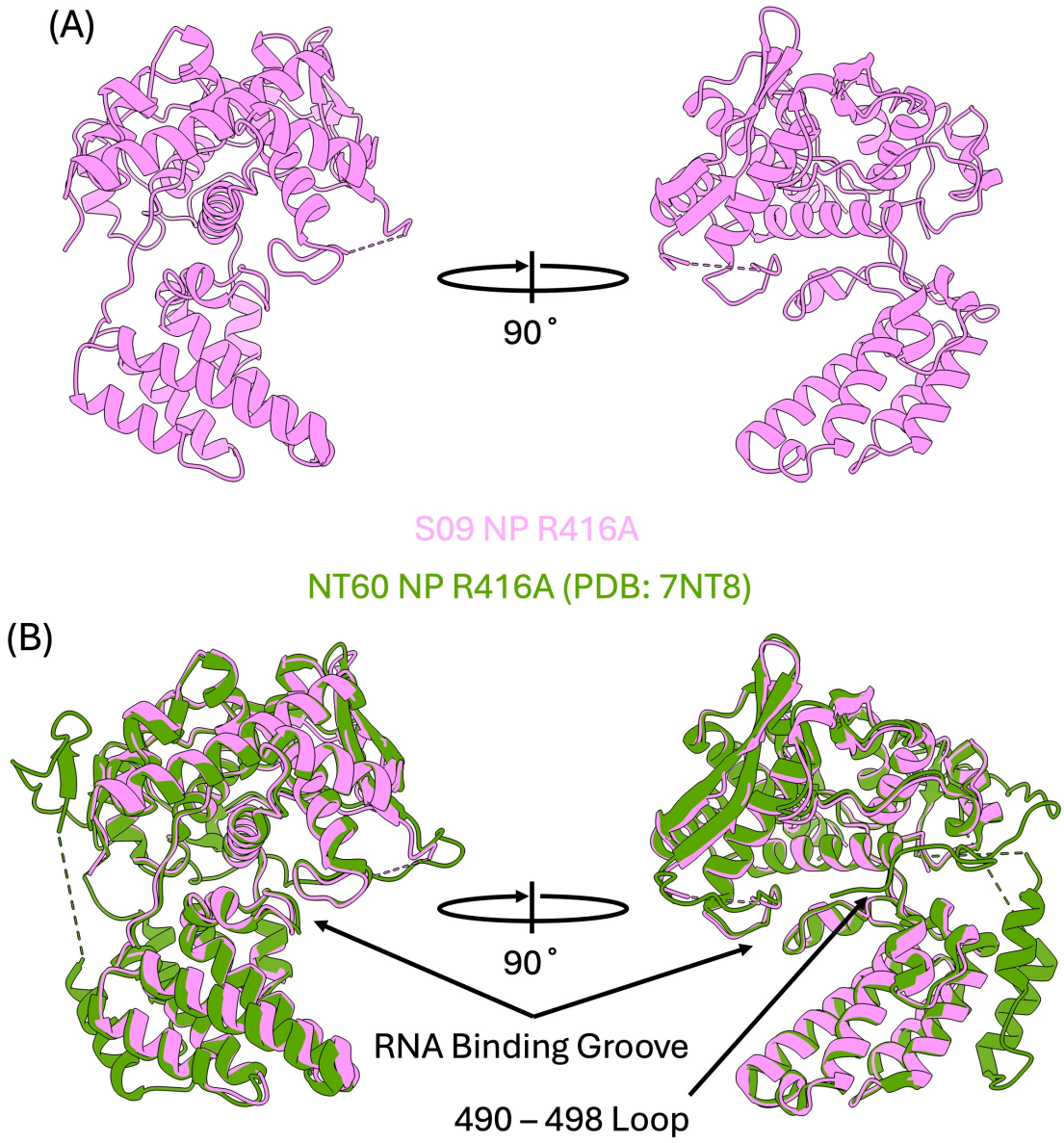


Figure 3.32: Comparison of S09 NP R416A and NT60 NP R416A.

Comparison of the S09 NP R416A structure described in this chapter with the solved NT60 NP R416A structure (PDB: 7NT8). (A) The solved S09 NP R416A structure in pink. (B) The previously solved NT60 NP R416A structure overlayed in green, with the RNA binding groove and 490 – 498 loop annotated. The reported RMSD when aligning the two structures in ChimeraX is 0.844.

3.3. Discussion

The motivation of this work was first to solve the structure of a H2N1 avian influenza nucleoprotein, which had not been done prior to this work. This nucleoprotein would then be used to develop a scaffold system, such that routine high resolution structural

determination of the nucleoprotein in complex with binding partners of interest would be possible. This would improve our understanding of the molecular basis of nucleoprotein binding partner interactions, including partners required for the viral lifecycle, immune factors of interest, and potential therapeutics.

This chapter presents a comprehensive structural investigation of S09 NP R416A influenza virus nucleoprotein, employing multiple complementary approaches: X-ray crystallography, cryo-EM, and nanobody-assisted structural determination strategies. The work encompasses the expression and purification of S09 NP R416A nucleoprotein, nucleoprotein-binding nanobody Nb170, engineered gluebodies, and a megabody construct, all integrated into systematic structural biology workflows.

3.3.1. X-Ray Crystallography of Influenza Nucleoproteins

X-ray crystallography is a powerful technique to study the high-resolution structure of macromolecular proteins. It has been successfully used for the determination of multiple influenza nucleoprotein structures since 2006 at resolutions ranging from 2.2 to 3.3 Å, as reviewed in the introduction to this chapter. While nucleoprotein crystallisation on its own proved intractable in this work, the use of a nucleoprotein nanobody (Nb170) facilitated the reproducible formation of protein crystals. This crystal diffracted to a reported 2.6 Å resolution, but due to crystal pathologies outlined in this chapter, the structure was not able to be refined to produce an accurate model. Nonetheless, a convincing molecular replacement solution was obtained from the nucleoprotein – nanobody complex, validated by the positioning of the nanobody in

such a position where the CDR3 loop provides the largest contribution to the predicted binding interface.

The nanobody used (Nb170) is known not to interfere with influenza virus replication or transcription (Schmidt et al., 2016), suggesting that the observed nanobody binding site is on an exposed surface in the biological context of the vRNP, which will be further explored in chapter 5.

3.3.2. Di-Gluebodies as a Tool to Study Influenza Nucleoprotein Structure

The application of the novel di-gluebody platform represents an innovative approach to nucleoprotein structural determination. At 56 kDa, the nucleoprotein represents an ideal test case for this technology, positioned at the current practical limit of routine cryo-EM SPA (Herzik et al., 2019).

Size-exclusion chromatography confirmed stable nucleoprotein – di-gluebody complex formation; however, cryo-EM analysis revealed significant heterogeneity. Only one dataset exhibited fully occupied di-gluebodies, suggesting differential binding characteristics between the two different di-gluebodies tested. This observation warrants further investigation of affinity-driven effects on binding to the di-gluebody scaffold.

While the di-gluebody platform has demonstrated utility in both X-ray crystallography and cryo-EM applications (Ye et al., 2024; Yi et al., 2025), the conformational flexibility observed at binding interfaces described in this study prevented successful structure determination even with fully occupied complexes. Combined with the reduced expression yields compared to conventional nanobodies, these technical challenges represent significant hurdles that must be addressed for broader platform adaptation.

Future optimisation should focus on reducing interface flexibility and improving stoichiometric occupancy to fully realise the platform's potential for small protein structure determination.

3.3.3. Megabodies as a Tool to Study Influenza Nucleoprotein Structure

Megabody Mb151 provided the most successful approach for nucleoprotein structural characterisation in this study. Stable nucleoprotein-megabody complexes were readily formed via size-exclusion chromatography and maintained structural integrity during cryo-EM sample preparation.

Single particle analysis yielded high-quality reconstructions suitable for atomic modelling through focussed refinement strategies. A composite map approach was employed, generating two complementary reconstructions: the megabody core region at 2.85 Å resolution and the nucleoprotein – nanobody interface at 3.15 Å resolution.

Initial model building was facilitated using ModelAngelo (Jamali et al., 2024), followed by iterative refinement to produce the final structure.

This structure represents a significant achievement as the first monomeric nucleoprotein structure solved by cryo-EM. Remarkably, the reconstruction reveals an unconventional nanobody – target binding mode, with interactions predominantly mediated through the C'' strand, rather than the canonical CDR loops. Only two additional CDR contacts were observed – one each from the CDR2 and CDR3 loops – highlighting the unique binding mechanism employed by this nanobody.

This atypical binding mode expands our understanding of nanobody – target interactions and demonstrates the structural diversity achievable through nanobody engineering approaches.

3.3.4. Resolved Nanobody Binding Interfaces on the Influenza Nucleoprotein

Prior to this work, only one nanobody-bound influenza nucleoprotein structure had been solved and reported in the literature. This structure employed nanobody VHH1, which was shown to inhibit replication and transcription of long RNA segments (Hanke et al., 2016) (PDB: 5TJW). In the figure below, this structure is overlayed with the Mb151 – S09 NP R416A nucleoprotein structure solved in this chapter, and the putative molecular replacement solution of the Nb170 + S09 NP R416A crystal (figure 3.33).

This chapter presents two additional nanobody-nucleoprotein complexes: the high-resolution structure of the Mb151 megabody bound to S09 NP R416A, and the putative molecular replacement solution of Nb170 bound to S09 NP R416A. Figure 3.33 shows these three structures superimposed to reveal their distinct binding sites on the nucleoprotein.

The nanobodies exhibit markedly different functional properties that correlate with their binding sites. Mb151 is derived from nanobody VHH151. This nanobody inhibits nuclear import of vRNPs and blocks transcription and replication of influenza A virus (Schmidt et al., 2016). The solved structure displays a binding interface close to the top of the head domain of the nucleoprotein.

VHH1, as reported (Hanke et al., 2016) binds to the head domain and inhibits replication and transcription of long RNA segments, similar to Nb151 but at a distinct site on the head domain.

Nb170 is uniquely placed among the nanobodies tested by Schmidt and colleagues, showing no detectable effects on influenza A virus vRNP nuclear import, nucleoprotein nuclear import, or on vRNP activity (replication and transcription) (Schmidt et al., 2016). The molecular replacement solution positions the nanobody at the base of the body domain of the nucleoprotein.

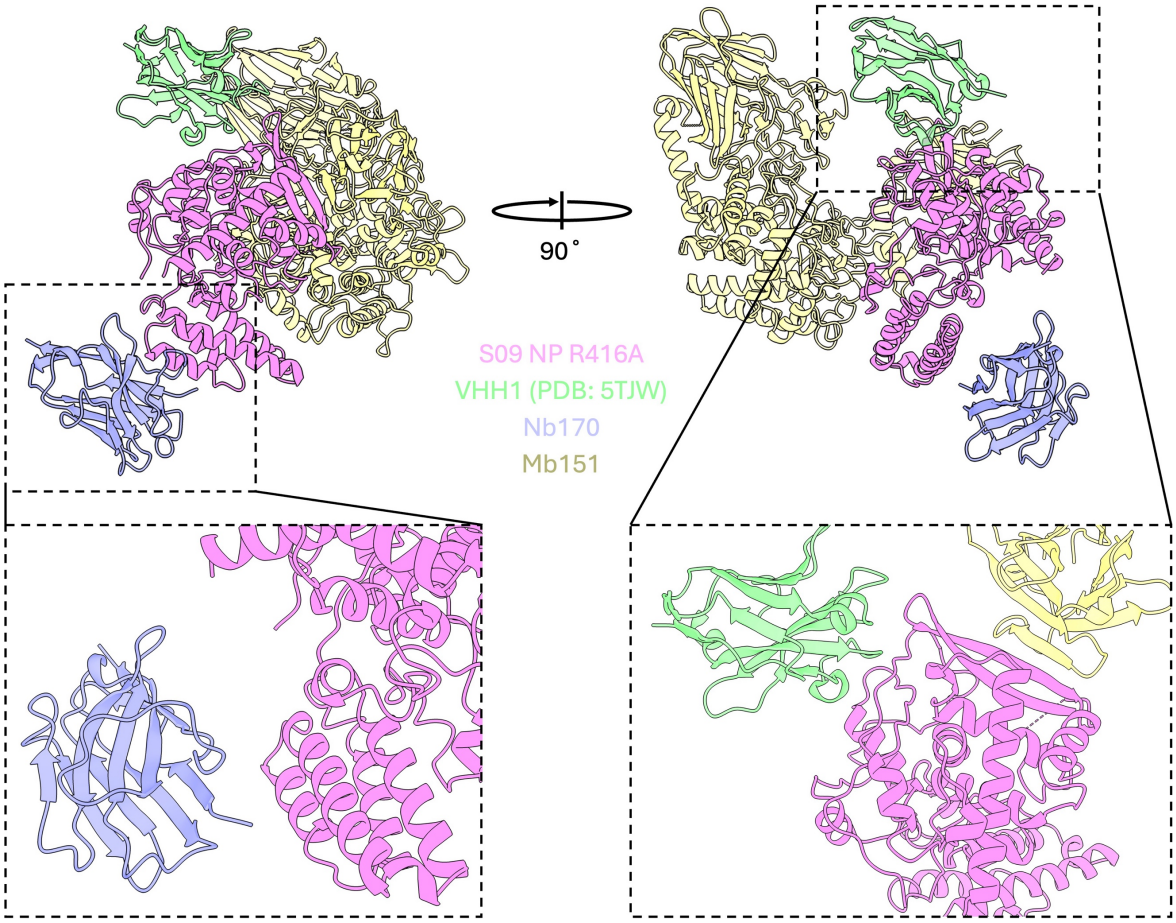


Figure 3.33: Superimposition of Nanobody and Nanobody-Derived Structures and Predictions on one Central Nucleoprotein.

The published structure of WSN nucleoprotein (PDB: 5TJW) bound by VHH1 (green), the molecular replacement structure of S09 NP R416A + Nb170 presented in this chapter (lilac), and the solved structure of S09 NP R416A + Mb151 presented in this chapter (yellow) all coloured with 50 % transparency with one central nucleoprotein, S09 NP R416A (pink) presented in this chapter solved in complex with Mb151 displaying the binding sites of the one previously solved nucleoprotein nanobody in comparison with the structure solved and molecular replacement solution presented in this chapter. Sections are additionally displayed zoomed in to demonstrate the interaction interfaces of the nanobodies with the nucleoprotein.

As discussed by Hanke and colleagues (Hanke et al., 2016), VHH1 used in their study of the structure of the nucleoprotein overlaps significantly with regions that determine sensitivity to MxA. When these residues are additionally highlighted (figure 3.34), this overlap is clear. In contrast, both Nb151 and Nb170 do not directly overlap with this binding site, suggesting both may be suitable candidates for the study of MxA and BTN3A3 interactions with the nucleoprotein without interference.

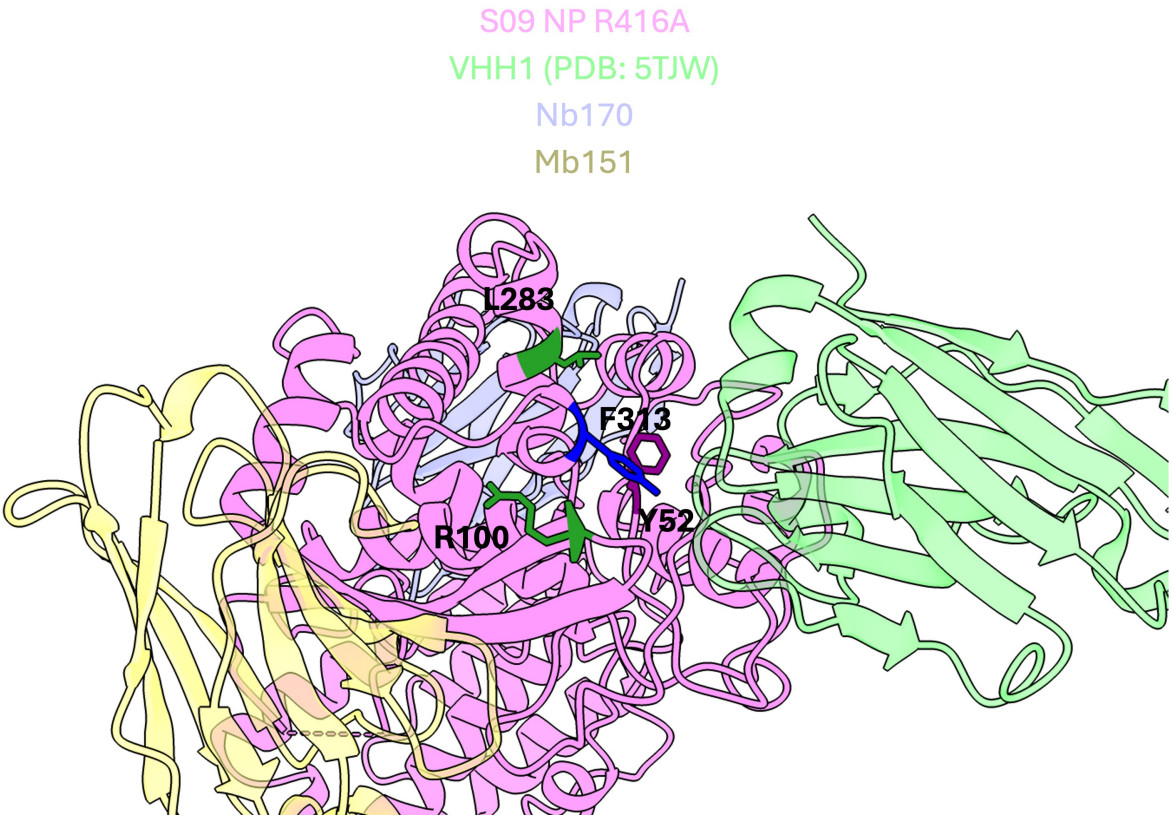


Figure 3.34: Immune Factor Sensitivity Residues and Nanobodies Bound to one Central Nucleoprotein.

S09 NP R416A structure solved in this chapter (pink) with sensitivity residues for MxA (100, 283), BTN3A3 (52), and both MxA and BTN3A3 (313) drawn as atoms in dark green, blue, and purple respectively. The mode of binding of Mb151 (yellow), VHH1 (green, PDB: 5TJW), and Nb170 (light purple) are overlayed.

By superimposing the nanobody onto low-resolution structures of the vRNP, Hanke and colleagues also reason that in one model (Moeller et al., 2012), the VHH molecule would block the major groove of the vRNP complex. This would reduce accessible surface area of the vRNP and hence limit interactions of the vRNP with other host and viral proteins. In another model (Arranz et al., 2012), the VHH molecule would clash with the head domain of an adjoining nucleoprotein molecule and additionally occlude the NLS2 (nuclear localisation signal 2) of the nucleoprotein, explaining the observed inhibition of vRNP nuclear import.

The NLS's of the nucleoprotein are illustrated in the figure below (figure 3.35). It is observed that NLS2 is at the base of the body domain of the nucleoprotein and given the top-to-bottom packing of nucleoproteins in early vRNP models (Arranz et al., 2012; Moeller et al., 2012 – both introduced fully in section 5.1.2), this would indeed restrict access to NLS2 in the context of the vRNP and likely disrupt physiological vRNP packaging.

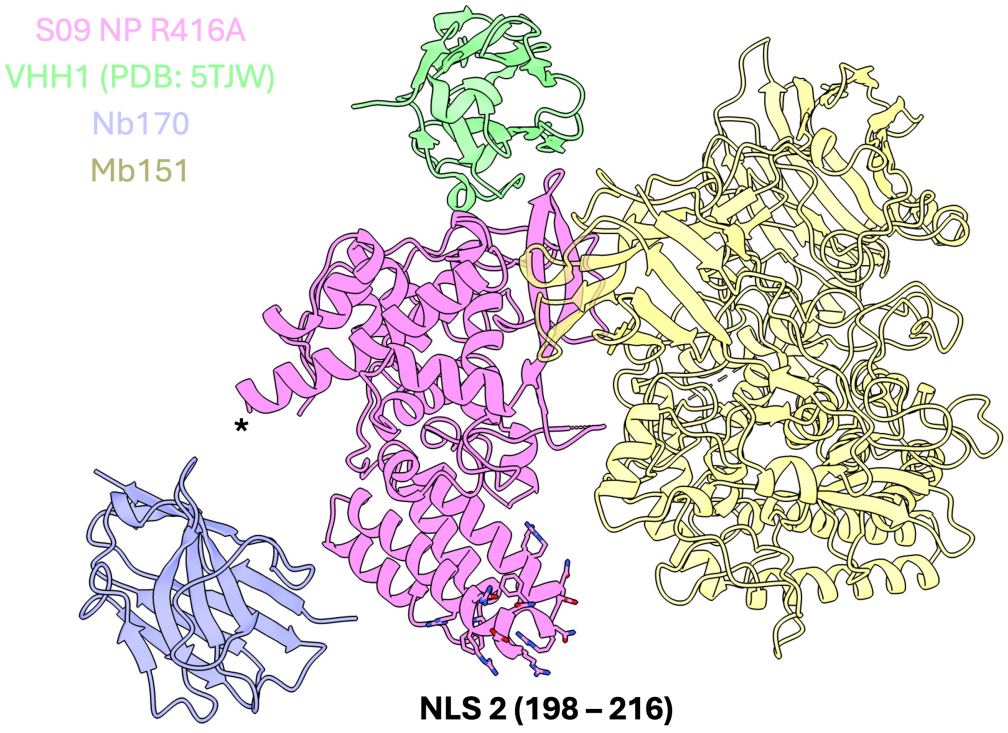


Figure 3.35: Nanobodies bound to one Central Nucleoprotein with NLS2 Displayed.

S09 NP R416A structure solved in this chapter (pink) with NLS2 (residues 198 – 216) shown as atoms, coloured as heteroatoms, and labelled at the base of the body domain of the nucleoprotein. The start of the resolved structure, beginning at residue 22 is shown by a star (*). This is the closest location on the resolved protein of NLS1 (residues 3 – 13). The mode of binding of Mb151 (yellow), VHH1 (green, PDB: 5TJW), and Nb170 (light purple) are overlayed.

Since the conceptualisation of this work, two recent studies have been published which yield interesting data on vRNP and vRNP-like complexes (Chenavier et al., 2025; Peng et al., 2025), which are both further discussed in chapters 5 and 6.

In the prior paper, Chenavier and colleagues employ N-terminal nucleoprotein truncation mutants (Δ1-14) and the use of a synthetic piece of RNA ($(UC)_{18}$-fluorescin$^{3'}$ (FAM)) to form a pseudo-vRNP made of recombinant truncated nucleoprotein and synthetic RNA (Chenavier et al., 2025). The resulting complex was observed to form an antiparallel double helix with the canonical major and minor grooves observed of influenza vRNPs.

Chenavier and colleagues were able to obtain a 3.04 Å reconstruction of the nucleoprotein by employing a focussed refinement strategy on a nucleoprotein dimer. However, the physiological relevance of this structure must be called into question, as varying nucleoprotein truncations or synthetic RNA lengths were observed to lead to the production of highly variable helical assemblies – from single helices to parallel double helical assemblies with no major or minor grooves.

In the latter paper, Peng and colleagues use an influenza D reconstitution system to purify the shortest segment of the influenza D genome, NS. They carried out cryo-EM SPA and cryo-ET on the reconstituted influenza D RNPs, and additionally, cryo-ET on native influenza A vRNP segments.

In their cryo-EM SPA pipeline, they were able to obtain a 5.1 Å reconstruction of a 4-mer of influenza D nucleoproteins, and further lower resolution information for the influenza D helical RNP.

Both studies converge on a consistent architectural model of the influenza RNA: an antiparallel double helix with RNA encapsidated in the minor groove, where individual nucleoprotein subunits are connected by flexible tail loops. This conservation across the two experimental approaches provides strong support for this being the physiologically relevant vRNP architecture.

Because a consistent architectural model of nucleoprotein packaging was agreed upon across both studies, and the helical reconstruction of the influenza D RNP offered significantly more information across one whole turn (compared to much less in other reconstructions), comparisons were made between the helical packaging of the influenza D RNP and the nanobody interaction interfaces presented in earlier figures in order to analyse the mode of binding of each of the nanobodies presented to the native vRNP (figure 3.36).

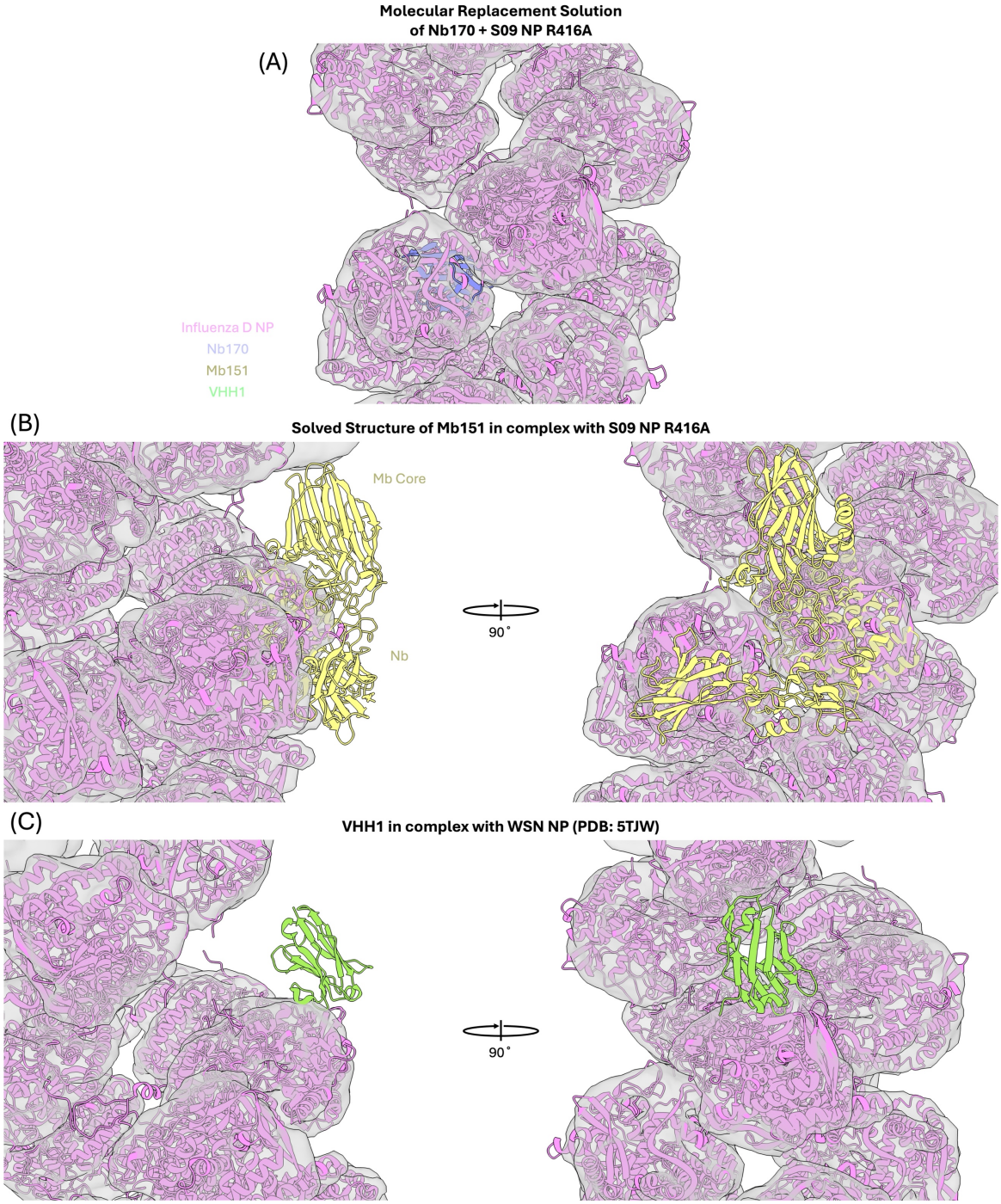


Figure 3.36: Comparison of Nanobody Interaction Sites with the Structure of the Helical Influenza D Virus RNP.

Comparison of nanobody interaction sites with the influenza D helical RNP structure recently published (Peng et al., 2025). Briefly, the density map of the influenza D virus helical RNP (EMDB: 44980) was displayed in grey at 40 % transparency and the associated influenza D helical RNP structure (PDB: 9C4H) is displayed in this density in magenta. The three structures presented earlier: the molecular replacement solution for the S09 NP R416A + Nb170 crystal (A) was displayed in magenta with the bound nanobody displayed in lilac, the structure solved for the Mb151 megabody in complex with the S09 NP R416A nucleoprotein (B) was displayed in magenta with the bound megabody displayed in yellow, and the deposited VHH1 nanobody bound to the WSN nucleoprotein (C) was displayed in magenta with the bound nucleoprotein in green (broadly maintaining the colour palate from the preceding figure).

As expected for a megabody, Mb151 (figure 3.36, panel B) would cause significant steric clashes if the vRNP were fully coated. However, the nanobody domain itself is not observed to clash with neighbouring nucleoproteins. The inhibition of vRNP nuclear import and activity are likely results from steric interference with host and viral protein interactions rather than with neighbouring nucleoproteins or NLS occlusion.

In the case of VHH1 (figure 3.36, panel C), the overlay suggests that this nanobody does not directly clash with neighbouring nucleoproteins, calling into question the proposed mechanism of vRNP nuclear import via the occlusion of NLS2 on neighbouring nucleoprotein molecules. Instead, the functional effects may be the result of a similar global vRNP conformational perturbation or reduced accessibility for binding partner interaction when multiple copies of the nanobody bind to the vRNP.

Finally, the molecular replacement solution of Nb170 bound to S09 NP R416A (figure 3.36, panel A) places the nanobody directly in density occupied by neighbouring nucleoproteins, which leads to a direct contradiction of the functional data showing no effect on vRNP activity.

Given this apparent contradiction, an AlphaFold3 prediction was performed for the S09 NP R416A + Nb170 complex. In 4 out of 5 predictions, the binding interface was positioned on the back / head domain of the nucleoprotein, rather than the body suggested by the molecular replacement solution (figure 3.37, panel A).

When this predicted binding site is similarly mapped onto the solved structure of the influenza D helical RNP published, Nb170 would bind to a highly exposed portion of the head domain that would allow complete coating of the vRNP without structural perturbation. This positioning is consistent with the observed lack of functional effects on vRNP nuclear import and activity. However, when the mode of binding of the AlphaFold3 prediction of Nb170 to the nucleoprotein is examined, it is observed that there are two alpha helices present in the AlphaFold3 prediction which do not exist in experimentally-solved structures of nanobodies, and that the residues corresponding to the CDR3 loop are not implicated in binding to the nucleoprotein, but rather, residues in the C strand, the C-C' loop, and the E-F loop.

These findings suggest that the molecular replacement solution may be incorrect despite favourable data processing statistics (good LLG and TFZ scores) and biologically reasonable placement of the CDR3 loop. Because of this, further molecular replacement and data processing is necessary to correctly solve the structure of the nucleoprotein in complex with the nanobody in this crystal.

To investigate other potential modes of binding of Nb170 to the S09 NP R416A nucleoprotein, AlphaFold3 predications were made. Both AlphaFold2-Multimer and AlphaFold3 can be used as helpful screening tools for nanobody binding, and have been able to predict binding interfaces which have been shown to be correct given experimental data. However, the overall success rate of such predictions can be lower than 50% (Eshak and Goupil-Lamy, 2025; Sachdev et al., 2025).

AlphaFold3 predictions demonstrate a binding site which would fit with the observation that influenza virus replication and transcription aren't affected by this nucleoprotein, but highlight issues in a dependence on structural predictions, given the incorrect fold of the nucleoprotein observed. This highlights the importance of validating structural solutions experimentally (figure 3.37, panel B).

The reported pTM score of 0.75 suggests an overall accurate structure of the component proteins in the prediction, whereas the low ipTM score of 0.13 suggests that this prediction, too, may be wrong. Therefore more work must be carried out to verify the mode of Nb170 binding to the nucleoprotein.

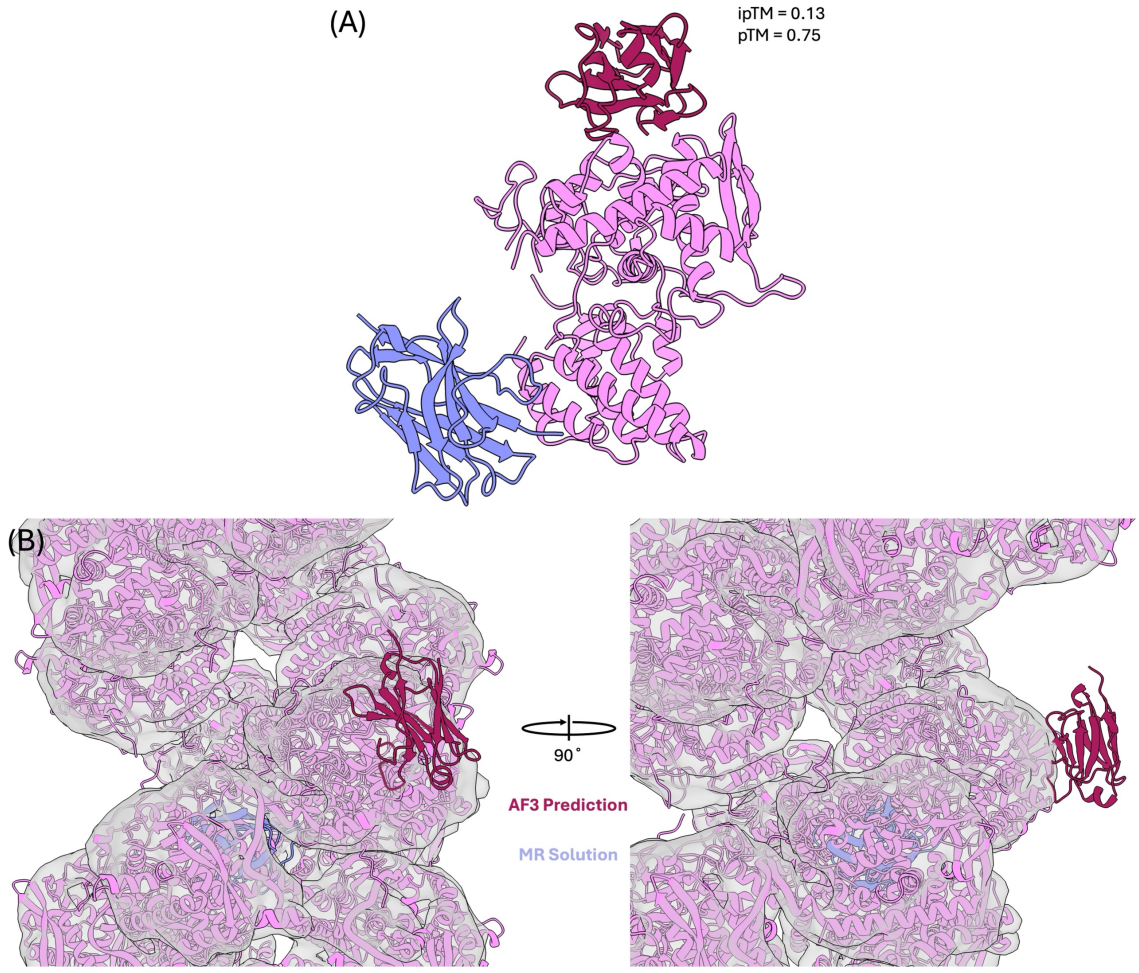


Figure 3.37: Comparison of Molecular Replacement Solution and AlphaFold3 Prediction of Nb170 Binding Site.

Comparison of nanobody interaction sites with the influenza D helical RNP structure recently published (Peng et al., 2025). Briefly, (A) an AlphaFold3 prediction of Nb170 bound to the S09 NP R416A nucleoprotein was carried out and this was aligned to the molecular replacement solution introduced earlier to compare predicted nanobody binding sites. Then, the density map of the influenza D virus helical RNP (EMDB: 44980) was displayed in grey at 40 % transparency and the associated influenza D helical RNP structure (PDB: 9C4H) is displayed in this density in magenta. The molecular replacement solution for the S09 NP R416A + Nb170 crystal was displayed, the nucleoprotein coloured in magenta and the bound nanobody coloured in lilac (MR Solution). Additionally, the AlphaFold3 prediction of Nb170 bound to the S09 NP R416A nucleoprotein was displayed, the nucleoprotein coloured in magenta and the bound nanobody coloured in maroon (AF3 Prediction). The molecular replacement solution and AlphaFold3 prediction were aligned in ChimeraX to one copy of the nucleoprotein in the published influenza D helical RNP structure (PDB: 9C4H).

Whilst Nb170 was chosen as it does not interfere with the replication cycle of influenza virus, and Mb151 was used as it was a well-expressing megabody available in the laboratory, these nanobody and nanobody-derived proteins may not be ideal. This is due either to weak binding affinity, or to the epitope targeted on the influenza virus nucleoprotein. It may be possible to develop tighter binding nanobodies which target either similar exposed epitopes, or epitopes which have not yet been targeted by our current repertoire of nanobodies.

Future work in this area of research should consider the possibility of going back to the camelid source of nanobodies, and producing more nanobodies which may perform better in structural and functional experiments. By producing a new, broader panel of nanobodies, other epitopes may be exploited either to study the virus *in situ*, or used as a therapeutic (e.g. to block nucleoprotein oligomerisation). Such nanobodies may be able to be developed for both research and therapeutic purposes.

4. Nucleoprotein Interaction Partners of the Human Innate Immune System

4.1. Introduction

In order for the successful replication of influenza virus to occur in the host cell, the virus must be able to evade host immune restriction, as detailed in chapter 1. The nucleoprotein is both critical in the context of the viral lifecycle – protecting the viral genome from recognition by the host immune response and acting as a processivity factor for viral transcription and replication – and critical in the context of the immune response. Being so abundant in viral infection, the nucleoprotein is also necessarily exposed to the host cell, forming a common target for the host immune system. For this reason, there are a number of antiviral targeting mechanisms of both the innate and adaptive immune response to the nucleoprotein, including MxA and BTN3A3.

Due to this targeting by the host immune system, there is an evolutionary pressure to evade the host immune response. Regions of the nucleoprotein targeted by the host immune response necessarily mutate to evade host immune factor restriction, and to lead to more productive infection cycles. These mutations, observed across different influenza virus strains, provide insights into mechanisms of host factor restriction and viral immune evasion strategies. Therefore, comparative studies of the binding of human innate immune factors to nucleoproteins from different influenza virus strains are informative on the effects of specific nucleoprotein mutations on the ability of the immune factors to bind to the nucleoprotein.

Two innate immune factors known to interact with the influenza virus nucleoprotein are MxA and BTN3A3, as introduced in chapter 1.

MxA is a well-studied protein of the innate immune system, and the specific residues involved in the MxA – nucleoprotein interaction are well-understood (Götz et al., 2016), with clear evidence for the MxA L4 loop being the determinant of antiviral specificity (von der Malsburg et al., 2011; Patzina et al., 2014; Verhelst et al., 2015).

Given this wealth of understanding, a reasonable model has been proposed of MxA binding to vRNPs by forming higher order oligomeric complexes including rings around vRNPs in order to restrict their transport into the host cell nucleus (Gao et al., 2010; Haller et al., 2010; Gao, von der Malsburg, et al., 2011), and a similar binding to nascently-produced vRNPs inducing their clustering followed by transport and sequestration at the MTOC (McKellar et al., 2025). Additionally, it is suggested that MxA dimers bind nascently produced influenza nucleoprotein in the host cell cytoplasm, and inhibit their transport into the host cell nucleus (Haller et al., 2015; Nigg and Pavlovic, 2015). However, there is a lack of structural data to support the proposed model of restriction.

BTN3A3 has been recently identified as an innate immune factor important in the restriction of avian influenza A virus infection (Pinto et al., 2023). This work is fundamental in the emerging understanding of the butyrophilin family of proteins and their roles in the innate immune response. In their work, Pinto and colleagues describe

the specific restriction of avian influenza A infection as a result of BTN3A3, showing the localisation of BTN3A3 to the nucleus by immunoblotting. The mechanism underlying this nuclear localisation remains unclear. Furthermore, the direct biochemical interaction between avian influenza A virus nucleoprotein and the expected domain of interaction, B30.2, has not been demonstrated, nor has the mode of binding of BTN3A3 to the nucleoprotein been explained.

The widespread resistance of human-adapted influenza virus to restriction by the innate immune factors outlined underscores the critical evolutionary importance of immune factor evasion in influenza pathogenesis and host adaptation. While interactions between these immune factors and susceptible avian-origin influenza A virus strain nucleoproteins have been suggested, in both cases, no direct binding interactions have been biochemically demonstrated. Systematic biochemical characterisation across multiple sensitive and resistant influenza A virus strains using quantitative and biochemical techniques are therefore lacking.

Improved knowledge of these interactions on the molecular level is essential for fully understanding the mechanisms of immune factor action and additionally understanding the importance of specific residues in immune evasion of these proteins.

This chapter presents the production and biophysical characterisation of MxA and the BTN3A3 B30.2 domain. Interactions with multiple influenza A nucleoprotein strains are characterised by complementary biochemical methods. This chapter attempts to

address gaps in our understanding of host innate immune factor binding to the avian influenza nucleoprotein to establish direct interactions and quantitative binding parameters.

4.1.1. Protein Choice Rationale

4.1.1.1. Influenza A Virus Nucleoproteins

Three influenza A virus strain nucleoproteins were selected to examine immune factor binding across strains which are predicted to be sensitive and resistant to restriction by MxA and BTN3A3. Two avian-adapted nucleoproteins were chosen based on their predicted sensitivity to both MxA and BTN3A3 restriction.

The S09 strain (A/Green-Winged Teal/Ohio/175/1966 (H2N1)) and Tk strain (A/turkey/Turkey/1/2005 (H5N1)) both contain the canonical sensitivity residues for these immune factors and share 97.6 % sequence identity, with identical RNA binding and immune factor sensitivity sites (see chapter 3, table 3.3, and figure 3.2).

The Tk strain was specifically selected as a representative H5N1 highly pathogenic avian influenza (HPAI) virus, which are of particular global concern due to their enhanced pandemic potential. H5N1 viruses are classified as HPAI based on their multibasic cleavage site (MBCS) in the HA protein, which enables cleavage by ubiquitous proteases across tissue types, enhancing their pathogenicity (Joseph et al., 2017; Mostafa et al., 2024; Neumann et al., 2025). The recent widespread zoonosis of H5N1 in cattle (Neumann and Kawaoka, 2024; Mostafa et al., 2024), increased

pandemic potential, and currently limited human spillover (Garg et al., 2024; Uyeki et al., 2024) has heightened surveillance interest in this strain. It was studied in this work to compare the differences in restriction between two avian influenza strain nucleoproteins.

The NT 60 strain nucleoprotein (A/Northern Territory/60/1968 (H3N2)) was chosen as a negative control, representing a human-adapted influenza A virus strain expected to be resistant to restriction by both innate immune factors. This nucleoprotein contains the canonical resistance mutations to MxA and BTN3A3 and shares 92.4 % and 92.0 % sequence identity with the S09 and Tk strain nucleoproteins respectively, providing an ideal comparative framework to assess the functional impact of specific resistance mutations.

Table 4.1: Sequence Identity of Nucleoprotein Strains Used.

Amino acid sequence identity of S09, Tk, and NT60 influenza A virus nucleoproteins. The two avian-adapted influenza A virus strain nucleoproteins differ in 11 surface-exposed residues: (33, 34, 77, 105, 371, 373, 377, 408, 417, 450, and 482) which are not clustered in any one location, and are instead spread around the head and body domains of the nucleoprotein. Aside from the sensitivity sites to MxA and BTN3A3, differences between the two avian influenza A virus strain nucleoproteins and the human-adapted NT60 influenza A virus nucleoprotein are predominantly surface-exposed (with some in the interior of the protein), and located in the head and body domain of the nucleoprotein, with some residues additionally located in the G1 and G2 RNA binding regions: (16, 31, 61, 98, 100, 109, 114, 127, 136, 146, 214, 217, 283, 293, 305, 313, 334, 344, 351, 353, 357, 372, 375, 411, 421, 422, 423, 442, 452, 455, 456, 472).

S09	100 %		
Tk	97.6 %	100 %	
NT60	92.4 %	92.0 %	100 %
	S09	**Tk**	**NT60**

4.1.1.2. Innate Immune Factors

4.1.1.2.1. MxA

The interferon-stimulated innate immune factor MxA is known to restrict influenza virus through binding and inhibiting the nuclear import of vRNP complexes, as well as by sequestering nascent nucleoprotein molecules produced in the host cell cytoplasm. MxA is thought to function primarily as a dimer capable of forming higher-order oligomeric structures around vRNPs to prevent their nuclear transport (Gao et al., 2010; Haller et al., 2010; Gao et al., 2011; von der Malsburg et al., 2011; Haller et al., 2015; Nigg and Pavlovic, 2015). Within this mechanism, the MxA L4 loop serves as the key determinant of antiviral specificity across multiple viral targets (von der Malsburg et al., 2011; Patzina et al., 2014; Verhelst et al., 2015).

As introduced in section 1.5.2. and in figure 1.9, the MxA L4 loop is located at the tip of the stalk domain of MxA. Its function is highly conserved across species, acting as the primary determinant of antiviral action (Verhelst et al., 2015), however it is not well-conserved at the amino acid sequence level – especially among primates (Mitchell et al., 2012).

To characterise MxA – nucleoprotein interactions, an L4 loop peptide was used in an initial pull-down experiment to verify that the L4 region alone is sufficient for specific nucleoprotein recognition. Subsequently, full-length wild-type MxA protein was employed in a pull-down assay and size exclusion chromatography with a susceptible

avian influenza strain nucleoprotein to study physiologically relevant interactions of MxA with the nucleoprotein.

Quantitative binding analysis was performed using BLI, where GST-tagged influenza A nucleoproteins were immobilised on sensor tips and MxA was used as the analyte across a wide concentration range. This experimental design allowed for MxA oligomerisation to occur in solution while measuring direct binding kinetics with immobilised nucleoproteins, allowing the determination of kinetic parameters for MxA – nucleoprotein binding events across three nucleoprotein strains.

4.1.1.2.2. BTN3A3

Also recently identified as an interferon-stimulated innate immune factor against influenza A virus (Pinto et al., 2023), BTN3A3 has been reported to selectively restrict avian origin influenza A virus infection in humans. However, the molecular mechanism underlying restriction of avian influenza A virus infection by BTN3A3 remains poorly understood.

As introduced in chapter 1, BTN3A3 contains a cytoplasmic B30.2 domain, which shares structural homology with other innate immune proteins known to interact with viral components (Ozato et al., 2008; Pertel et al., 2011; D'Cruz et al., 2013; Tomar and Singh, 2015). Given this domain architecture and its cytoplasmic localisation, the B30.2 domain represents the most likely candidate for direct interaction with the influenza A nucleoprotein.

To test this hypothesis and provide evidence for direct interaction of the B30.2 domain with influenza A nucleoprotein, the isolated B30.2 domain was used in pull-down experiments with an example susceptible avian influenza nucleoprotein and was subsequently studied by means of size exclusion chromatography. This approach allows for the specific assessment of the B30.2 domain interaction with the nucleoprotein without potential complications resulting from the use of the full-length transmembrane protein.

Nucleoprotein – B30.2 domain interactions were studied as for MxA. GST-tagged influenza A nucleoproteins were immobilised on sensor tips, and the B30.2 domain was used as the analyte across multiple concentrations. This configuration enables direct comparison between B30.2 interactions and the three influenza A virus nucleoprotein strains used, providing quantitative insights into the affinity of BTN3A3-mediated restriction.

4.1.2. Methodological Approach

In this chapter, three complementary techniques were used to identify binding interactions between nucleoproteins and immune factors. First, resin-based pull-down assays were carried out to investigate the hypothesised interaction between the S09 avian influenza nucleoprotein and three interaction partners: the MxA L4 peptide, the full-length wild-type MxA protein, and the BTN3A3 B30.2 domain. Subsequently, the interaction was investigated between the S09 avian influenza nucleoprotein and the full-length wild-type MxA and the BTN3A3 B30.2 domain by means of size-

exclusion chromatography. Finally, the interaction between the S09 and Tk avian influenza nucleoproteins and the NT60 human-adapted influenza virus nucleoprotein, and the full-length wild-type MxA or the BTN3A3 B30.2 domain was investigated by means of Bio-Layer Interferometry.

In all cases, the monomeric mutant influenza nucleoprotein, containing the R416A mutation, was chosen to investigate interactions, as this simplified the model system. Experimentally, this comes at the expense of studying higher order oligomeric nucleoprotein interactions with their binding partners, but when using BLI or similar techniques, one component must be immobilised in an assumed low-order oligomeric state.

Resin-based pull-down assays serve as the foundational technique for the detection of protein-protein interactions, allowing initial identification of binding partners through specific binding interactions. In resin-based pull-down assays, one protein (the bait) is immobilised by a specific affinity tag to resin which binds specifically tagged protein – in this chapter, the immobilisation is carried out using a GST tag which specifically binds glutathione sepharose resin – while the expected binding partner (or prey) is added in solution and binding is allowed to occur. Following this incubation and sufficient washing steps, proteins that remain bound to the resin are visualised by means of SDS-PAGE, providing initial, qualitative, evidence for a direct binding interaction.

Size-exclusion chromatography enables characterisation of protein complexes in solution in their native state by separating proteins and protein complexes based on their hydrodynamic radii. Packed resin columns made of crosslinked polymers with strictly controlled pore sizes allow smaller molecules to enter pores formed by the polymer, causing smaller molecules to be retained for a longer time in the packed resin. Larger molecules by comparison cannot enter the polymer-formed pores and are retained for less time in the column, eluting earlier. When two proteins interact in solution, they elute from the size exclusion column earlier than either individual component, owing to their increased hydrodynamic radius. This technique is useful for studying the oligomeric state of proteins and protein complexes in solution and facilitates their separation into distinct homogeneous fractions. It is however a limitation that if binding partner interaction is weak, the complex may not be able to be observed by size-exclusion chromatography.

Bio-Layer Interferometry (BLI) is a biophysical method used to measure protein-binding partner affinity and kinetics by immobilising a ligand (or bait) molecule on the tip of a glass biosensor by a specific affinity tag or peptide linker. In this chapter, all nucleoprotein bait molecules were immobilised via their GST tag to anti-GST-coated biosensor tips. Once bait molecules have been successfully captured, a baseline measurement is taken. Following this, the expected binding partner (also called the analyte or prey), in this case the immune factors MxA or the BTN3A3 B30.2 domain, are exposed to the ligand at varying concentrations and binding allowed to occur.

White light is transmitted through the glass biosensor, which consists of an optical layer and a biocompatible matrix where target molecules are immobilised (see figure 4.1, below). When light reflects from the sensor surface, interference occurs between light reflected from the optical layer and light reflected from the biocompatible matrix. Binding events alter the optical thickness of the biolayer, causing a shift in the interference pattern that is proportional to the mass of the bound molecules. The recorded difference in interference provides a direct readout of protein binding to the biocompatible matrix layer and is recorded as the response signal (in nm).

The tips are then placed in fresh experimental buffer to allow dissociation of the bound analyte from the ligand. By monitoring the response in both the binding and dissociation stages of this experiment, binding kinetic parameters of the ligand and analyte can be obtained (Shah and Duncan, 2014; Sultana and Lee, 2015; Petersen, 2017) (figure 4.1, below).

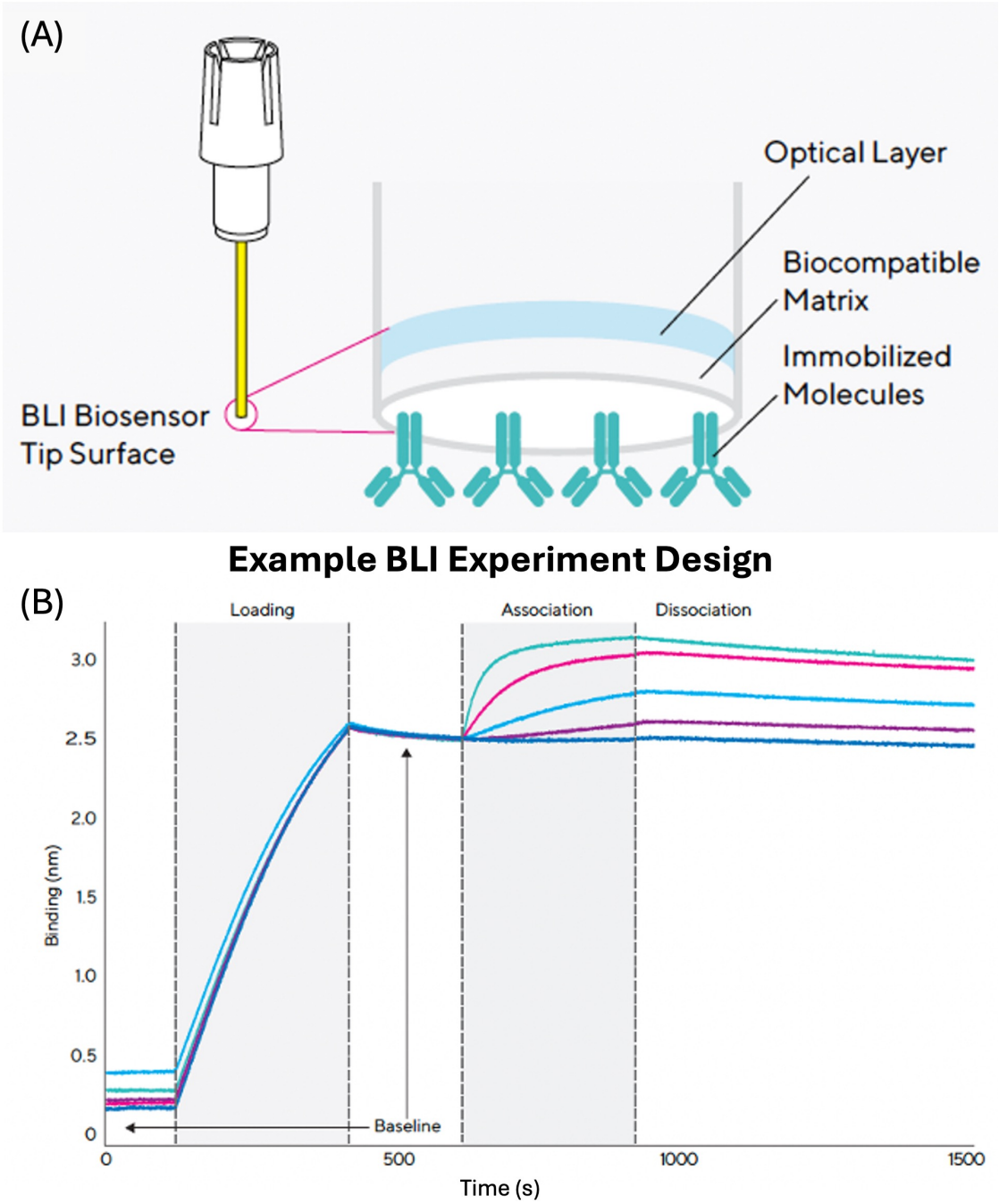


Figure 4.1: Basic BLI Experimental Design

(A) The glass biosensor (or 'tip') is displayed in diagrammatic form with both the optical layer and biocompatible matrix labelled. At the end of the tip (conjugated onto the biocompatible matrix) are immobilised molecules (in this case, anti-GST antibodies). (B) An example BLI experiment sensogram with the steps of the experiment labelled above. The first baseline step equilibrates the tips in the reaction buffer. The ligand loading step is where a controlled amount of ligand is specifically added to each tip by the conjugated biomolecule (in this case, GST-tagged proteins are immobilised by an anti-GST antibody). The analyte association step is the point at which the tips are incubated in a dilution series of analyte (in this case, MxA or BTN3A3 B30.2 domain).In this step, k_a is calculated. The analyte dissociation step is the point at which the tips are incubated in fresh reaction buffer and analyte dissociation is allowed to occur. In this step, k_d is calculated. The different coloured curves represent different concentrations of analyte used in this example experiment. Figures adapted from Sartorius Octet® documentation.

Due to background binding of the immune factors to the tips, all experiments were accompanied by ligand-free replicates, where the dilution series of analyte was flowed over GST tips with no ligand bound. These data were then subtracted from each concentration of experimental runs. Additionally, a 0 μM analyte sample run was carried out in each experiment as a reference and subtracted from the analyte runs. In the data processing, all curves were aligned to the average of the final 5 s of baseline.

Data analysis was carried out with the associated Octet® Data Analysis HT software (Sartorius, Germany). Once kinetic data have been acquired, they are fit to a binding model in order to derive kinetic parameters. In order to fit data to a kinetic model in BLI, multiple binding models can be used, where the aim is always to best fit the data observed. The simplest model is the Langmuir model of binding, which describes a simple 1:1 stoichiometric interaction, assuming a single binding site and no binding cooperativity.

Other models of binding may also be considered if the Langmuir model does not describe the data well. They include binding models which consider multiple states of reactions, for example when binding induces a structural change in the resulting complex, which leads to multiple kinetic parameters to consider, and also binding models which consider higher order binding – whereby binding involves a non-stoichiometric ratio of the two components being examined in the BLI experiment.

In all cases presented in this chapter, the 1:1 binding model was the most accurate in terms of fit to model and was always used. A global 1:1 binding model was used to fit the association and dissociation phases of all curves, assuming a 1:1 mode of binding between nucleoprotein and immune factor.

In an assumed 1:1 mode of binding between nucleoprotein (here, ligand, L), and immune factor (here, analyte, A):

$$A + L \underset{k_d}{\overset{k_a}{\rightleftharpoons}} AL$$

The equilibrium dissociation constant, K_D (M), is calculated using the kinetic parameters of the 1:1 binding model using the equation below:

$$K_D = \frac{[A]\,[L]}{[AL]} = \frac{k_d}{k_a}$$

Where the parameters [A], [L], and [AL] correspond to the concentration of analyte, ligand, and analyte:ligand complex respectively (M), and the parameters k_d and k_a correspond to the dissociation rate constant (s^{-1}) and association rate constant ($M^{-1}s^{-1}$) respectively.

Notably, no kinetic parameters are reported for the experiments where NT60 NP was used as the ligand, as no accurate values were able to be calculated due to the minimal binding of the immune factors to the resistant nucleoprotein.

4.2. Results

4.2.1. MxA and Nucleoprotein Interactions

In this section of the chapter, I demonstrate a direct interaction between the S09 avian influenza A nucleoprotein and a synthetically-produced MxA L4 loop peptide. I then demonstrate the bacterial expression and purification of recombinant full-length, wild-type MxA protein, before characterising its oligomerisation in solution with Size Exclusion Chromatographs – Multi Angle Light Scattering (SEC-MALS) and Dynamic Light Scattering (DLS). I then demonstrate the interaction of this MxA construct with the S09 avian influenza A nucleoprotein by means of a resin-based pull-down assay, and size-exclusion chromatography, before characterising the binding of two avian influenza A nucleoproteins and one human-adapted influenza A nucleoprotein to MxA by BLI.

4.2.1.1. Interaction Studies of the MxA L4 Loop Peptide with Influenza A Virus Nucleoprotein

A synthetic peptide of the MxA L4 loop was ordered and used in resin-based pull-down assays to detect binding to S09 strain nucleoprotein, an influenza A nucleoprotein which was hypothesised to bind to MxA. This was carried out to validate the L4 loop as the domain required for direct interaction with the nucleoprotein in a minimal biochemical system.

4.2.1.1.1. Resin-Based Pull-Down Assay

In order to confirm the interaction between the MxA L4 loop and influenza nucleoprotein, a resin-based pull-down was carried out, as described in the materials and methods chapter. Briefly, GST-tagged monomeric S09 NP R416A was immobilised by its GST tag using glutathione sepharose resin (Cytiva). Approximately 0.35 mg of S09 NP R416A was thawed and incubated with 50 µL of glutathione sepharose resin which was pre-equilibrated with binding buffer (25 mM HEPES-NaOH (pH 7.5), 150 mM NaCl) and allowed to incubate on a rotating platform at 4˚C for 1 hour. After binding, the resin was washed three times with 1 mL of binding buffer. After washing, an approximate 2-fold molar excess of MxA L4 loop peptide was added (10 µM of MxA L4 peptide in a final volume of 1 mL) and incubated at 4˚C on a rotating platform for 1 hour. After binding, the resin was washed three times with 1 mL of binding buffer. After the final wash, the resin was resuspended in 500 µL of binding buffer. SDS-PAGE samples were taken at each stage of the pull-down assay and analysed.

In the pull-down assay carried out (figure 4.2, below) the S09 NP R416A input was seen to bind the glutathione sepharose resin, with excess nucleoprotein being washed away in the wash steps. Enriched impurities were washed away in subsequent wash steps. Upon the addition of the approximate 2-fold molar excess of MxA L4 peptide loop, much of the L4 peptide is observed to be removed in the wash steps. However, a weak band of MxA L4 peptide is still present after thorough washing. Given the presence of a weak band of L4 peptide and the difference in the intensity of the S09 NP R416A band and MxA L4 peptide band, this suggests that a direct interaction between the

nucleoprotein and the L4 peptide has occurred, but there is not complete occupancy of the nucleoprotein.

In order to confirm that this interaction was not due to a non-specific interaction of the L4 peptide and the glutathione sepharose resin used, the L4 peptide was incubated with the glutathione sepharose resin alone. The same binding and wash protocol as detailed above was carried out, and the peptide was detected in both input and flowthrough lanes, but crucially not in the bound resin lane, suggesting that the L4 peptide alone does not bind to the glutathione sepharose resin, and that the interaction detected was indeed genuine (figure 4.2, panel B).

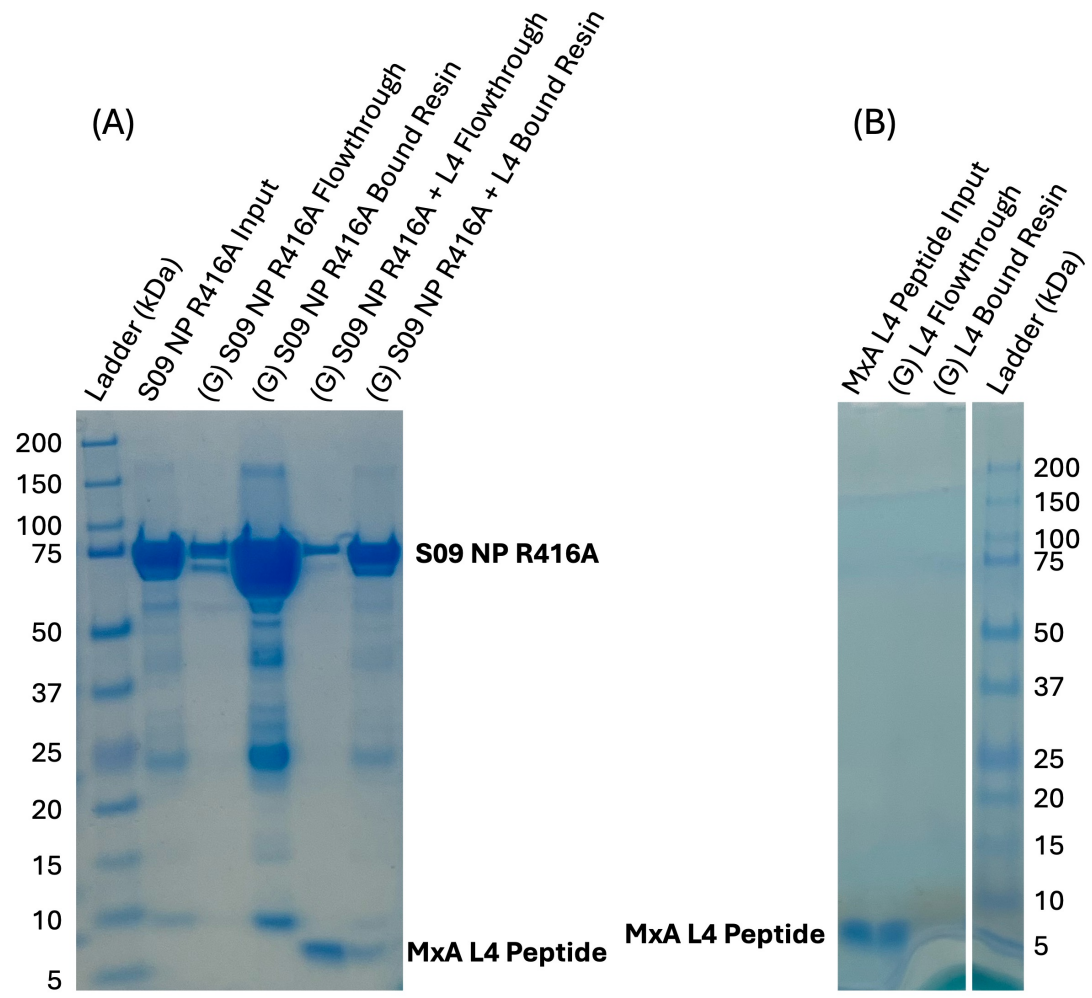


Figure 4.2: SDS-PAGE Gel of Resin-Based Pull Down Experiment of S09 NP R416A Nucleoprotein with MxA L4 Loop Peptide.

SDS-PAGE gel of resin-based pull-down experiment of S09 NP R416A with MxA L4 loop peptide demonstrating interaction between the S09 NP R416A nucleoprotein and the synthesised MxA L4 peptide. (A) S09 NP R416A is first immobilised to glutathione sepharose resin (G), and the MxA L4 peptide is added and the resin washed. Flowthrough and bound resin fractions of S09 NP R416A alone, and S09 NP R416A plus MxA L4 peptide are shown, with the approximate molecular weights fo the bands for MxA L4 peptide (L4 Peptide) and S09 NP R416A (NP) annotated. (B) A control experiment to confirm that the MxA L4 peptide does not bind to the glutathione sepharose resin.

4.2.1.2. Expression and Purification of MxA

Bacterial expression systems are commonly used to produce recombinant MxA which has been used to functionally characterise MxA, as it is observed to form varied orders of oligomer in solution (Kochs et al., 2002; Gao et al., 2010; Gao, von der Malsburg, et

al., 2011; von der Malsburg et al., 2011; Chen et al., 2017; McKellar et al., 2023), establishing bacterial expression as the standard for MxA structural and biophysical studies.

Given that MxA can be expressed in bacteria, the pET151 bacterial expression plasmid, with an N-terminal His_6 tag and TEV cleavage site was chosen to express soluble MxA. Expression and purification were carried out as described in the materials and methods chapter. Briefly, MxA was affinity purified with the N-terminal His_6 tag using Ni-NTA resin and eluted with imidazole before a final SEC step.

As shown in the SDS-PAGE gels below (figure 4.3, panels A and B), a significant amount of protein still remains in the bacterial pellet and could be optimised further for higher protein yields. MxA (75 kDa) is seen to bind to the resin, but some protein is observed to be removed in the subsequent wash steps. A high concentration of protein is observed in the elution steps, and these elutions are pooled and applied to an S200 10/300 column.

The SEC run is observed to separate most impurities from the sample (figure 4.3, panel B, 'Non-Consecutive SEC Fractions' lanes), and the resulting chromatography trace of the SEC run (figure 4.3, panel C) demonstrates a major peak corresponding to the first three SEC fraction lanes in figure 4.3, panel B. This peak is around the expected molecular weight of 440 – 669 kDa (Cytiva documents). Given the one major protein product in the SDS-PAGE, this suggests oligomers of MxA from around 6-mers to approximately 9-mers are forming in solution.

Given a lack of observed issues of aggregation, and a significant loss of protein when cleavage and reverse-immobilised metal affinity chromatography was attempted in order to purify cleaved MxA, the His_6 tag was decided not to be removed.

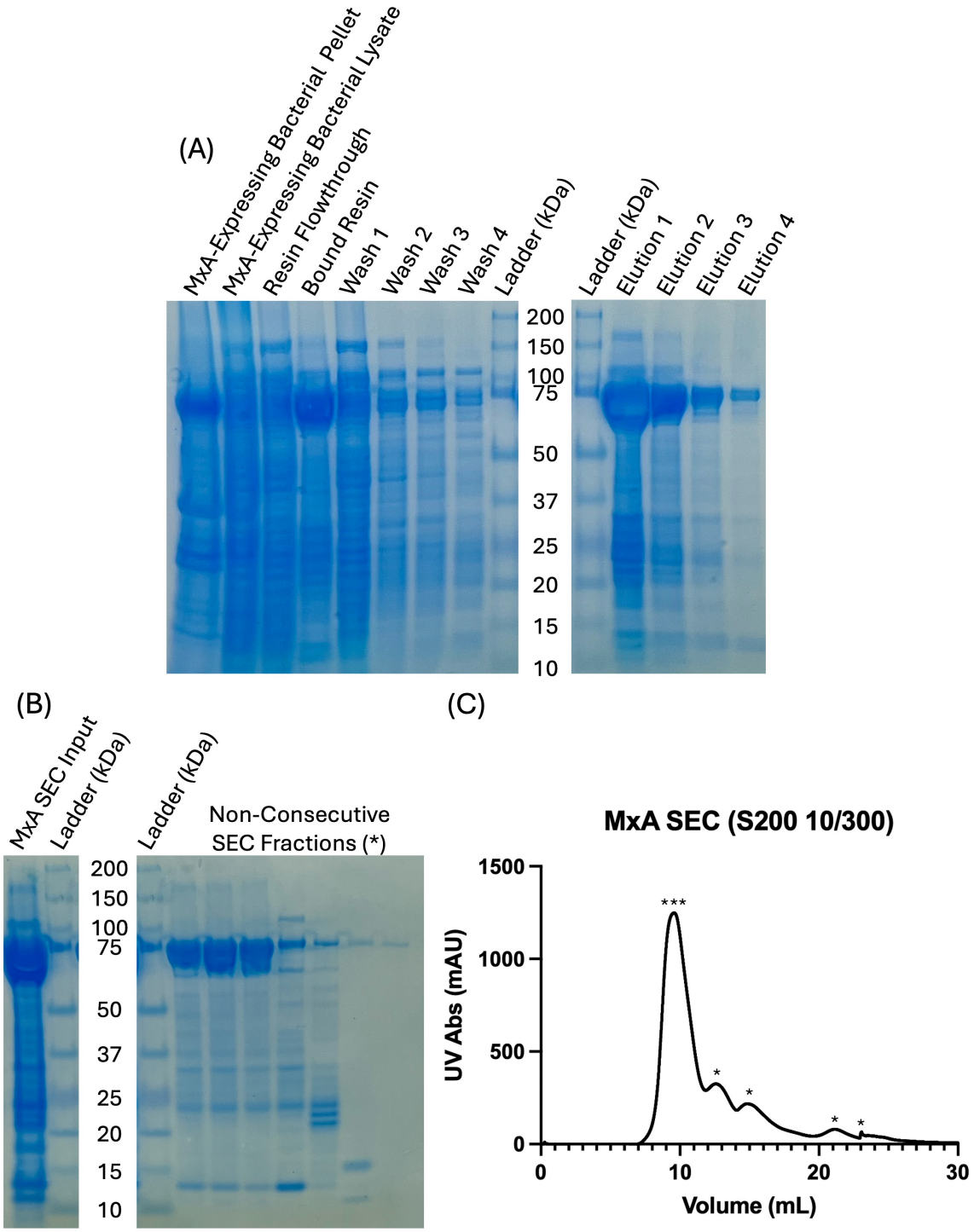


Figure 4.3: Purification of MxA.

(A) SDS-PAGE gel of expression and purification steps to elution from Ni-NTA resin. (B) SDS-PAGE of SEC fractions. (C) Chromatography trace of SEC run of this protein on a Superdex S200 (10/300) increase GL column. The SDS-PAGE gels in each panel are one gel, but because of the lane the ladder was run on, are split and the ladder lane repeated to allow for ease in interpretation of molecular weights according to the ladder. Note: It may have been more appropriate to choose a column with a greater fractionation range, e.g. a Superose 6 column, as MxA is present very early in the elution volume.

4.2.1.3. Characterisation of MxA

In an attempt to better understand the oligomeric state of the MxA sample produced, a number of biophysical methods were chosen to characterise the protein.

4.2.1.3.1. Size-Exclusion Chromatography – Multi-Angle Light Scattering

SEC-MALS experiments were designed and carried out together with Dr. David Staunton of the Molecular Biophysics Suite at the Department of Biochemistry, University of Oxford.

In order to determine the characteristics of MxA oligomerisation, SEC-MALS of MxA was performed at 0.6 mg/mL in a buffer of 25 mM HEPES-NaOH (pH 7.5) and either 150 mM NaCl – known to promote higher order MxA oligomerisation – or 500 mM NaCl – known to abolish MxA oligomerisation (Kochs et al., 2002). Additionally, the dependence of concentration on oligomerisation was investigated by carrying out these experiments at three concentrations: 0.6 mg/mL, 0.3 mg/mL, and 0.15 mg/mL. UV measurements (green) were used to determine protein concentration, rather than light scattering (LS), or refractive index (RI), and in doing so, the peaks shown in figure 4.4, below, are obtained. This data is tabulated in table 4.2.

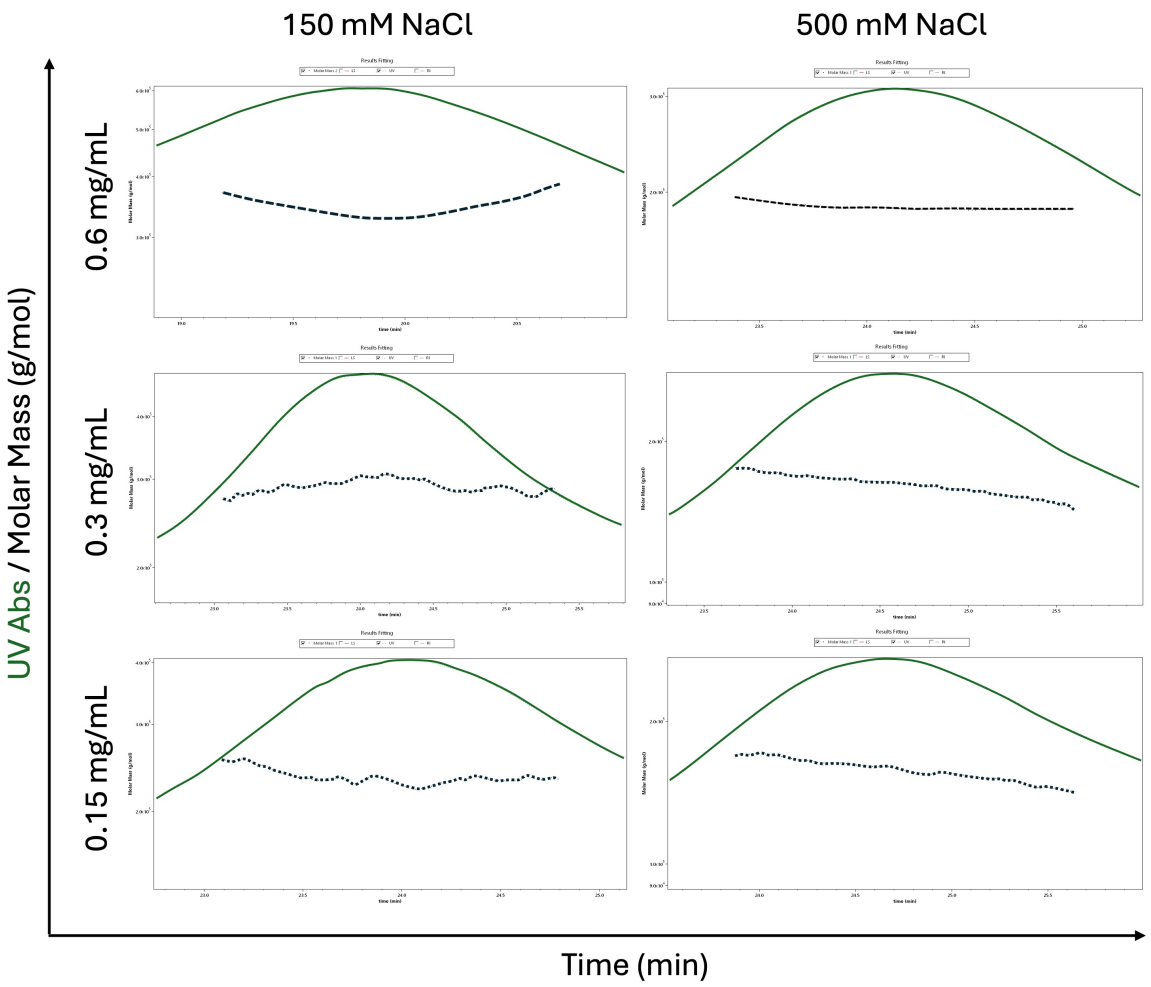


Figure 4.4: SEC-MALS Experiments of MxA.

MxA was studied by SEC-MALS to determine molecular weight of the MxA protein sample in two different salt concentrations, 150 mM NaCl (left), and 500 mM NaCl (right), and in three different protein concentrations, 0.6 mg/mL (top), 0.3 mg/mL (middle), and 0.15 mg/mL (bottom). UV absorbance peaks (green) and molar mass traces (dotted black) are plotted.

Table 4.2: Observed Molecular Weight of MxA by SEC-MALS.

Observed protein molecular weights of MxA samples in 500 mM and 150 mM NaCl measured at 0.6, 0.3 and 0.15 mg/mL protein concentration. Observed protein molecular weight is reported in kDa with corresponding oligomeric state reported in backets by dividing the observed molecular weight by the molecular weight of MxA (75 kDa).

Observed Protein Molecular Weight (kDa)	**NaCl Concentration**	
MxA Concentration	**150 mM**	**500 mM**
0.60 mg/mL	346.0 (4.61)	187.3 (2.50)
0.30 mg/mL	291.1 (3.88)	161.5 (2.15)
0.15 mg/mL	233.4 (3.11)	157.9 (2.11)

These data demonstrate differing predominant states of oligomericity in differing protein and salt concentrations. For any given concentration of MxA, the observed molecular weight was greater in 150 mM NaCl than in 500 mM NaCl, and in greater concentrations of MxA up to 0.6 mg/mL, the observed molecular weight of MxA is greater.

The starkest comparison of shift in oligomeric state being between 0.6 mg/mL MxA in 500 mM NaCl and in 150 mM NaCl, with the reported molecular weight in the high salt condition being 187.3 kDa compared to 346.0 kDa, thus suggesting that MxA oligomerises to higher orders in a concentration-dependent manner. This oligomerisation is inhibited in higher salt concentrations.

The expected oligomeric state is reported in brackets next to the observed molecular weight and suggests that the highest detectable oligomeric order observed in the data were approximately a 4.6-mer. This is likely due to the contributions of other order oligomers in solution. It is likely that higher order oligomers also exist, but their relative abundance is less than the populations reported above. The predominance of an approximate dimer in high salt and across the concentration range tested gives further support that the likely building block for MxA oligomerisation is the MxA dimer.

These data demonstrate the successful production of MxA which is able to dynamically oligomerise to varying orders dependent on protein and salt concentration in solution.

4.2.1.3.2. Dynamic Light Scattering

Dynamic light scattering was used to assess homogeneity and hydrodynamic diameter of MxA protein solution at 0.6 mg/mL in a buffer of 25 mM HEPES-NaOH (pH 7.5) but different salt concentrations: 150 mM, and 500 mM NaCl (figure 4.5, below) and tabulated (below, table 4.3). DLS measurements were captured using an UNcle machine with all DLS measurements being carried out at room temperature (24˚C).

In the DLS size distribution plots (figure 4.5, below), it is clear that there is a change in hydrodynamic diameter between the two salt conditions. In the higher salt condition (500 mM NaCl), the protein has a smaller hydrodynamic diameter than in the lower salt condition (150 mM NaCl), over an average of two reported datasets, a 191.7 nm hydrodynamic diameter in the high salt condition compared to 298.0 nm hydrodynamic diameter in the low salt condition. These data further support the above SEC-MALS observations that MxA forms different sized oligomeric assemblies dependent on salt concentration.

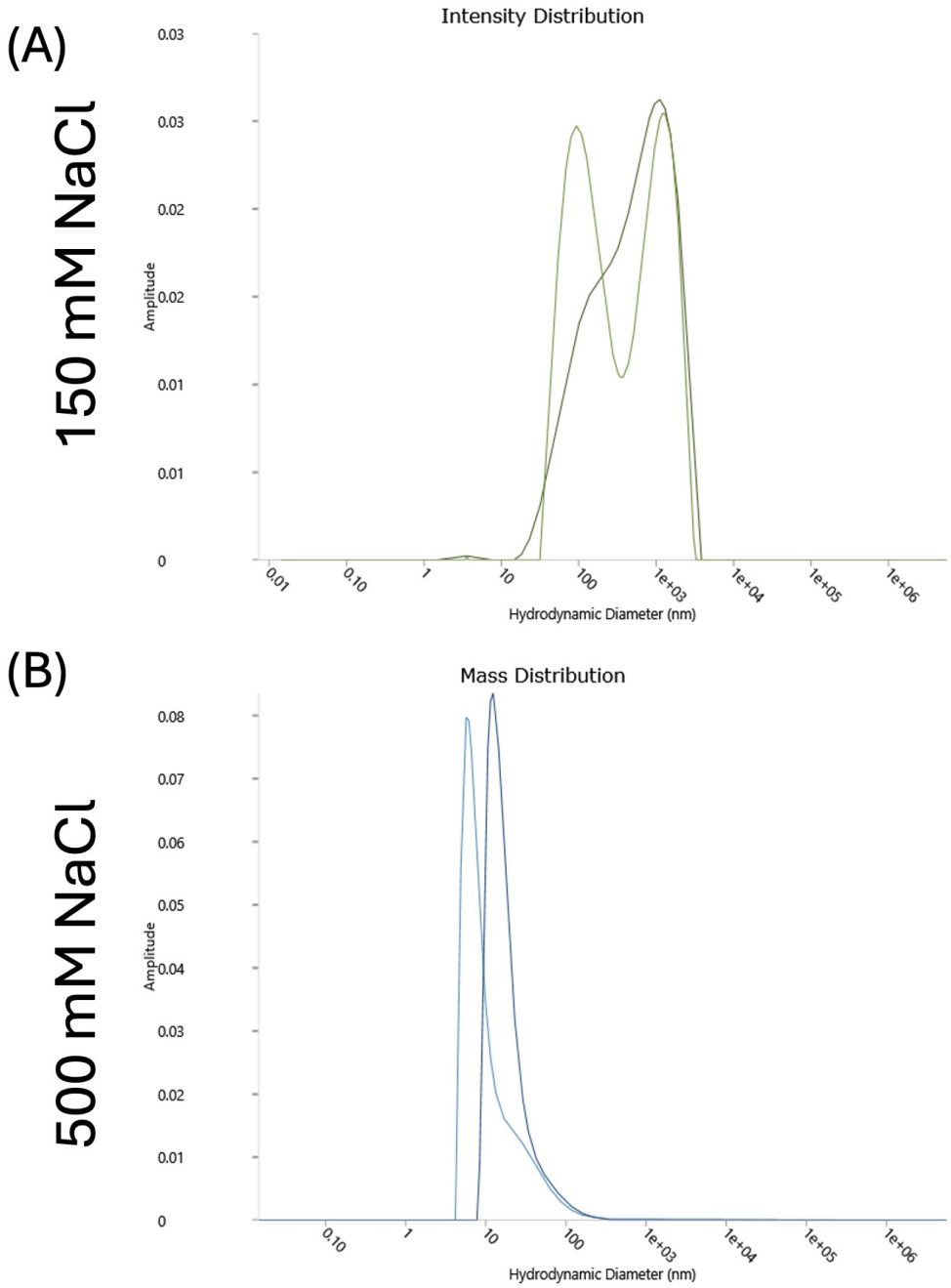


Figure 4.5: DLS of MxA.

DLS size distribution plot of MxA protein at 0.6 mg/mL concentration in two different salt concentrations (A) 150 mM NaCl, and (B) 500 mM NaCl. Different colours demonstrate individual repeats, as presented in the table below. Note: In future, it would be useful to carry out a control experiment using BSA to determine whether the UNcle machine can distinguish between different oligomeric states of protein.

Table 4.3: Tabular Data of MxA DLS Experiment.

Tabular data of DLS experiment in figure 4.5, above. Duplicate experiments were run, where the results of each experiment are displayed as different shades of green (150 mM NaCl condition) and blue (500 mM NaCl condition). Z diameter is a measure of hydrodynamic radius. SD diameter is the standard deviation of the measured diameter. Polydispersity Index (PDI) is calculated as the square of the SD diameter divided by the observed Z diameter, where a monodisperse population of molecules is taken

to have a PDI of between 0.0 and 0.1, a moderate polydisperse population of molecules is taken to have a PDI of between 0.1 and 0.4, and a broad polydisperse population of molecules is taken to have a PDI of above 0.4.

Sample	Z Diameter (nm)	SD Diameter (nm)	PDI
MxA 150 mM NaCl 1	251.7	135.1	0.288
MxA 150 mM NaCl 2	344.2	198.6	0.333
MxA 150 mM NaCl Average	**298.0**	**166.9**	**0.311**
MxA 500 mM NaCl 1	172.0	95.4	0.291
MxA 500 mM NaCl 2	211.3	114.0	0.308
MxA 500 mM NaCl Average	**191.7**	**104.7**	**0.300**

4.2.1.4. Interaction Studies of MxA with Influenza A Virus Nucleoprotein

The interaction of MxA with S09 NP R416A nucleoprotein was investigated by means of resin-based pull-down assay, SEC-based interaction studies, and finally quantitively with BLI.

4.2.1.4.1. Resin-Based Pull-Down Assay

In order to confirm the interaction between full-length wild-type MxA and S09 strain avian influenza nucleoprotein, a resin-based pull-down was carried out, as described in the materials and methods chapter. In this pull-down system, because GST-tagged influenza nucleoproteins are approximately the same molecular weight as the recombinantly produced MxA (≈ 75 kDa), MxA was immobilised to Ni-NTA resin (Qiagen) by its His_6-tag as a ligand, and cleaved recombinant monomeric S09 NP R416A nucleoprotein (≈ 55 kDa) was used as an analyte. Briefly, 0.35 mg of MxA was thawed and incubated with 50 µL of Ni-NTA resin which was pre-equilibrated with binding buffer (25 mM HEPES-NaOH (pH 7.5), 150 mM NaCl) and was allowed to incubate on a rotating platform at 4˚C for 1 hour. After binding, the resin was washed three times with 1 mL of binding buffer. After washing, an approximate 2-fold molar

excess of S09 NP R416A nucleoprotein (cleaved, without its GST tag) was added (55 µL of cleaved S09 NP R416A at a concentration of 10 mg/mL in a final volume of 1 mL) and incubated at 4 ˚C on a rotating platform for 1 hour. After binding, the resin was washed three times with 1 mL of binding buffer. After the final wash, the resin was resuspended in 500 µL of binding buffer. SDS-PAGE samples were taken at each stage of the pull-down assay and analysed.

In the pull-down assay carried out (figure 4.6, below) the MxA input was seen to bind to the Ni-NTA resin, with excess MxA being washed away in the wash steps. Other unidentified bands were seen to be enriched and could not be removed in repeat experiments. Upon the addition of the approximate 2-fold molar excess of cleaved S09 NP R416A, much of the added nucleoprotein sample is observed to be removed in the wash steps. However, a protein band approximately equal in intensity to the protein band for MxA was observed to be enriched in the final bound resin sample (figure 4.6). This suggests a direct interaction between the full-length wild-type MxA protein and the cleaved S09 NP R416A nucleoprotein.

In order to confirm that this interaction was not due to a non-specific interaction of the cleaved nucleoprotein with the Ni-NTA resin used, the cleaved nucleoprotein was incubated with the Ni-NTA resin alone. The same binding and wash protocol as detailed above was carried out, and the nucleoprotein was detected in both the input and flowthrough lanes, but crucially not significantly in the bound resin lane, which suggests that the nucleoprotein alone does not bind to the Ni-NTA resin, and that the interaction detected was indeed genuine (figure 4.6, panel B).

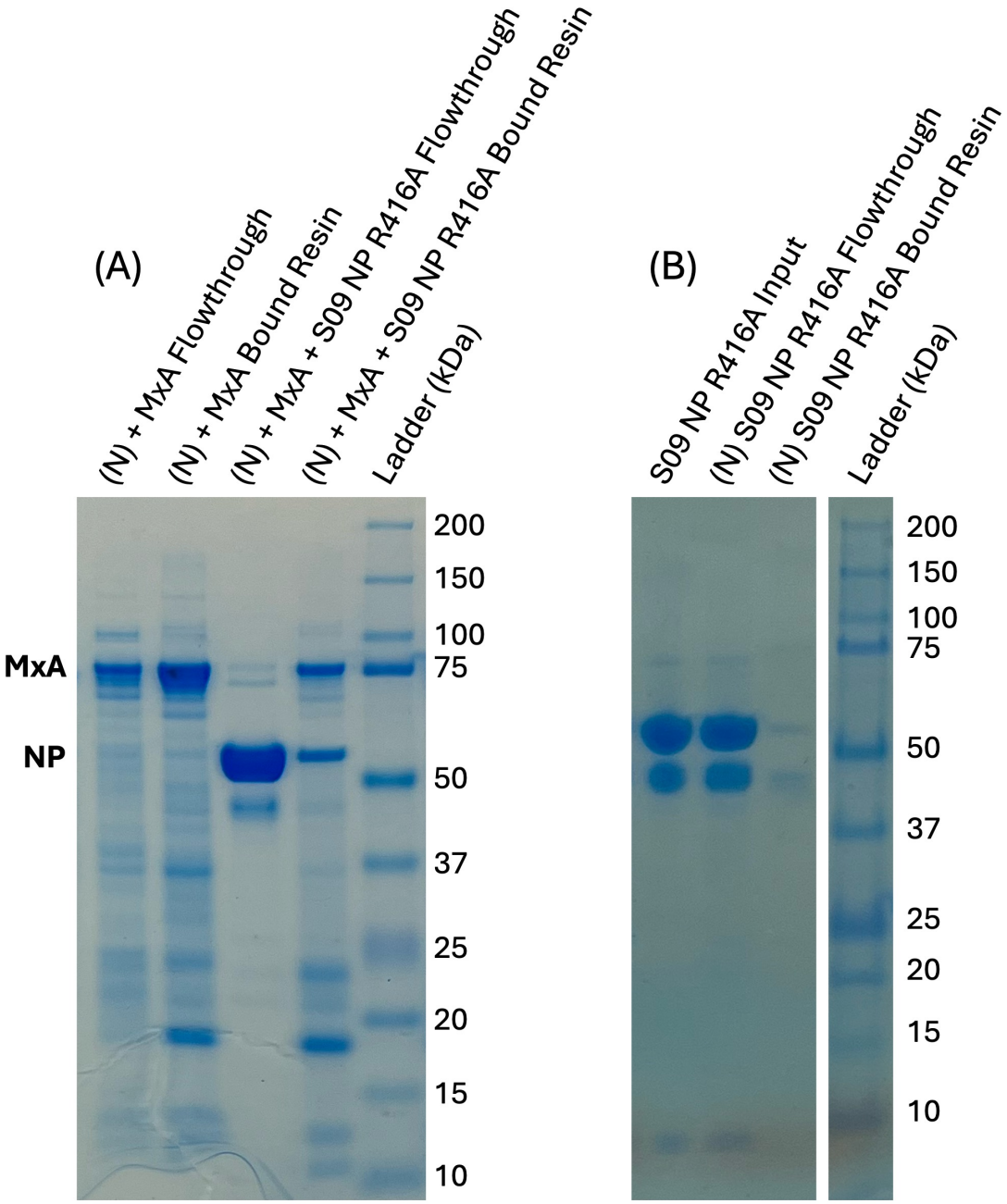


Figure 4.6: SDS-PAGE Gel of Resin-Based Pull Down Experiment of S09 NP R416A Nucleoprotein with MxA.

SDS-PAGE gel of resin-based pull-down of MxA and S09 NP R416A demonstrating interaction between the S09 NP R416A nucleoprotein and full-length bacterially expressed MxA. (A) MxA is first immobilised to Ni-NTA (N) resin, and then S09 NP R416A is added and the resin washed. Flowthrough and bound resin fractions of MxA alone, and MxA plus S09 NP R416A are shown, with the approximate molecular weights of the bands for MxA and S09 NP R416A (NP) annotated. (B) A control experiment to confirm that the cleaved nucleoprotein does not bind to the Ni-NTA resin.

4.2.1.4.2. Size-Exclusion Chromatography

Following the demonstration of direct S09 nucleoprotein – MxA binding in pull-down assays, size-exclusion chromatography was employed to reconstitute and analyse the complex in solution. Insufficient amounts of sample from the pull-down assay precluded downstream structural analysis, necessitating complex re-formation through controlled stoichiometric mixing.

Equimolar amounts of His_6-tagged MxA and untagged S09 NP R416A were incubated together in 150 mM HEPES-NaOH (pH 7.5), 150 mM NaCl buffer for 30 minutes on ice and injected in a 250 µL injection loop on an S200 5/150 increase GL analytical SEC column for hydrodynamic radius-based separation. A single peak was obtained (figure 4.7, panel B) and fractions of this peak (starred) were taken and analysed by means of SDS-PAGE (figure 4.7, panel A).

SDS-PAGE analysis revealed both nucleoprotein and MxA bands across all four peak fractions. MxA distribution was non-uniform and sub-stoichiometric, with an enrichment of MxA in the centre of the peak. This pattern is consistent with non-equimolar interaction stoichiometry rather than stable 1:1 complex formation. Repeated attempts to reproduce the results of this experiment were not successful, yielding insufficient material for crystallisation trials or cryo-EM analysis.

The elution profile indicated a protein complex of between 440 kDa and 690 kDa. Given that the nucleoprotein exists as a monomer in solution and MxA forms higher-order oligomers, this size range likely reflects heterogeneous complexes with a variable MxA :

nucleoprotein stoichiometry. The stronger nucleoprotein staining observed suggests both incomplete MxA occupancy of nucleoprotein, and the presence of more nucleoprotein than MxA in these fractions.

Attempts to reproduce these SEC results were unsuccessful. Combined with the non-uniform protein distribution across the peak fractions, this suggests that the MxA – nucleoprotein interaction represents a dynamic equilibrium rather than the formation of a discrete, stable complex in solution.

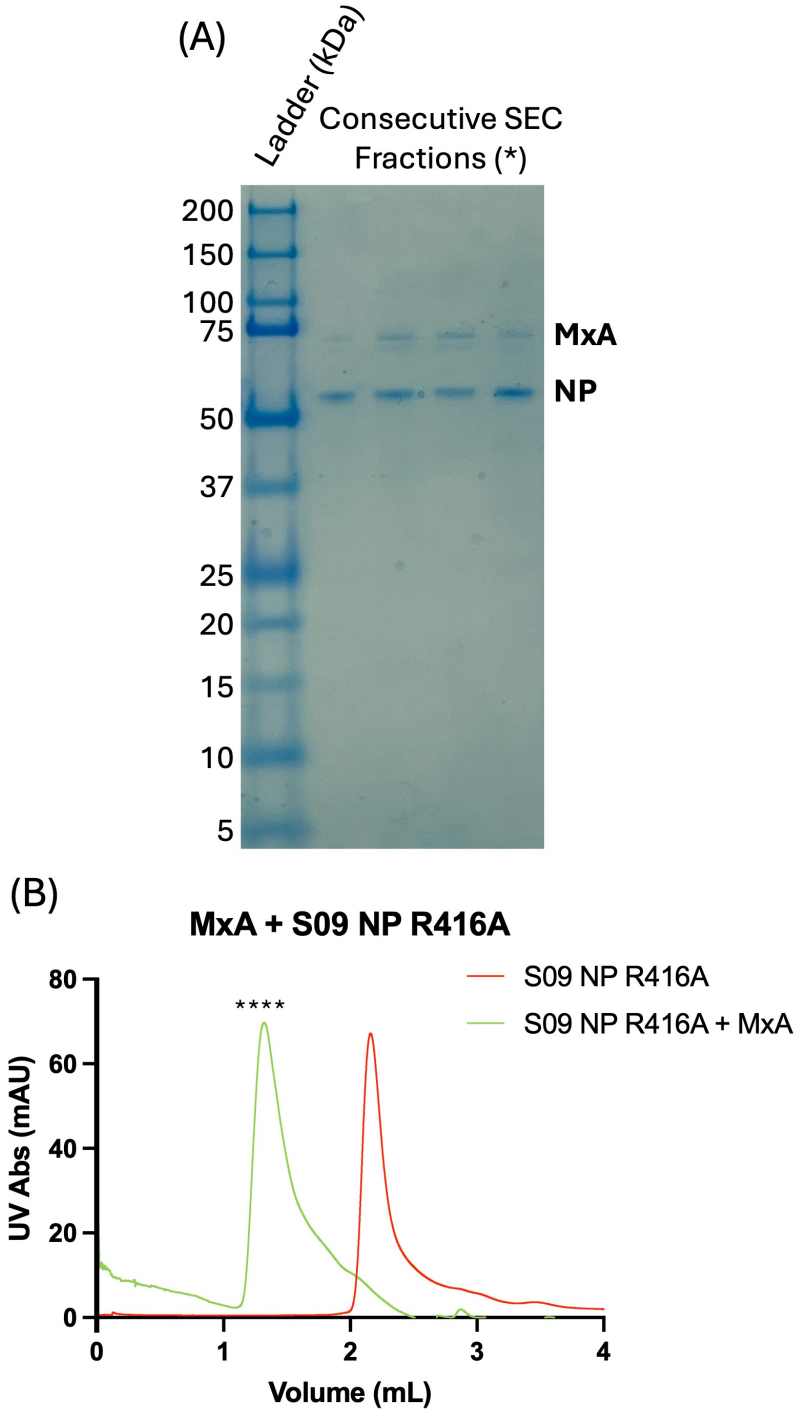


Figure 4.7: Analytical SEC of MxA and S09 NP R416A Complex.

(A) SDS-PAGE gel of peak fractions from analytical SEC run. (B) Chromatogram of analytical SEC run using a Superdex S200 5/150 increase GL column, with S09 NP R416A run alone (red) and in complex with MxA (green). Starred peak fractions were run on the SDS-PAGE gel accompanying.

4.2.1.4.3. Bio-Layer Interferometry

In order to further investigate the interaction between MxA and nucleoproteins, a set of BLI experiments were carried out where S09 NP R416A, Tk NP R416A, and NT60 NP R416A were immobilised on anti-GST tips as the ligand and MxA was used as the analyte. S09 and Tk nucleoproteins were used as test avian influenza nucleoproteins, assumed to be sensitive to MxA restriction, and thus, MxA binding. NT60 nucleoprotein was used as a negative control. As NT60 is a human-adapted influenza virus, it is known to have resistance mutations to both MxA and BTN3A3, and as such, no response is hypothesised to be observed upon the addition of MxA.

In this experiment, recombinant GST-tagged nucleoprotein was immobilised on glass anti-GST tips by the GST tag at a constant concentration of 20 µg / mL (≈ 0.36 µM), as suggested by the manufacturer. A five point dilution series of MxA analyte was incubated with independent nucleoprotein – bound tips of 60 µg / mL (0.8 µM), 30 µg / mL (0.4 µM), 15 µg / mL (0.2 µM), 7.5 µg / mL (0.1 µM), and 3.75 µg / mL (0.05 µM).

Since background binding of MxA to the anti-GST tips was detected, each experiment was run in parallel with an equivalent concentration gradient of analyte, with no ligand bound to the tips. The resulting response of the control was subtracted from the observed experimental values to account for the background binding.

In the experiment carried out, the three higher concentrations of MxA were not reported due to errors with the tips, which lead to erroneous curves not fitting with the concentration gradient used. As these experiments were carried out immediately prior

to writing this thesis, there was not time to repeat these experiments. For this reason, kinetic parameters were not able to be derived from the observed data. However, comparable experiments of the two lower concentrations of MxA used as an analyte are reported for all three nucleoproteins.

In the processed sensograms below, over the period of 0 – 1000 s, varying concentrations of MxA analyte were incubated with the nucleoprotein – bound tips and binding is allowed to occur. Between 1000 – 2000 s, the tips are moved into fresh experimental buffer (with no analyte), and the complex begins to dissociate. In each case reported below, the top curve corresponds to a MxA concentration of 7.5 µg / mL (0.1 µM MxA) and the bottom curve corresponds to a MxA concentration of 3.75 µg / mL (0.05 µM MxA).

The reported Rmax value corresponds to the maximum measured interference between incident and reflected light transmitted through the biosensor tip, representing the total amount of analyte bound at equilibrium.

While technical issues with tips corresponding to the three highest MxA concentrations in all three experiments prevented the determination of kinetic parameters of this interaction, the Rmax values of S09 = 1.15 nm, Tk = 0.39 nm, and NT60 = 0.029 nm obtained using a constant concentration of immobilised nucleoprotein and a constant MxA concentration of 20 µg / mL (≈ 0.36 µM) suggest differing binding affinities to different influenza A nucleoproteins. This is broadly expected given both S09 and Tk influenza A virus nucleoproteins contain residues

expected to make this virus sensitive to MxA restriction, whereas NT60 contains residues expected to make this virus resistant to MxA restriction.

Future repeat experiments will be useful to determine full kinetic parameters of the interaction between avian influenza nucleoproteins and MxA.

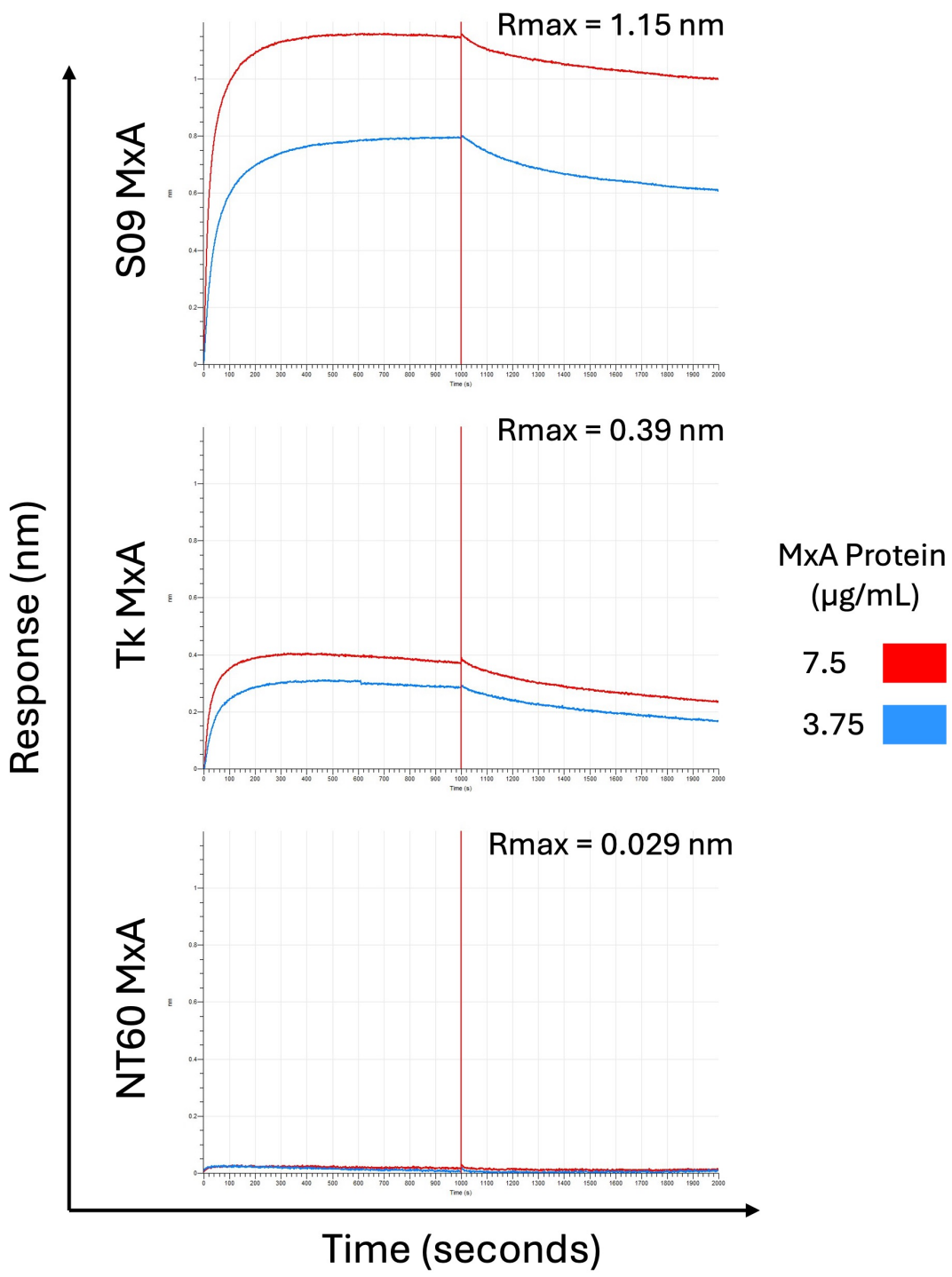


Figure 4.8: Processed Sensograms of MxA + Nucleoprotein BLI Experiments.

Processed sensograms of BLI experiments with a constant concentration of nucleoprotein ligand (20 µg/mL) immobilised to each GST tip of S09 NP R416A (top), Tk NP R416A (middle), and NT60 NP R416A nucleoproteins (bottom), and a constant dilution series of MxA of 60 µg/mL, 30 µg/mL, 15 µg/mL, 7.5 µg/mL, and 3.75 µg/mL. In these sensograms, only the traces for 7.5 µg/mL (red) and 3.75 µg/mL (blue) are shown. Analyte association is displayed on the left hand side of the red line (0 – 1000 s), and analyte dissociation is displayed on the right hand side of the red line (1000 – 2000 s). Recorded Rmax values are annotated on each sensogram.

4.2.2. BTN3A3 B30.2 Domain and Nucleoprotein Interactions

In this section of the chapter, I demonstrate the bacterial expression and purification of recombinant BTN3A3 B30.2 domain protein. I characterise the produced protein by SEC-MALS and DLS before demonstrating the interaction of this protein with the S09 avian influenza A nucleoprotein by means of a resin-based pull-down assay and size-exclusion chromatography. I then characterise the binding of two avian influenza A nucleoproteins and one human-adapted influenza A nucleoprotein to the B30.2 domain by BLI, extracting kinetic parameters of interaction for the two avian origin influenza A virus nucleoproteins.

4.2.2.1. Expression and Purification of BTN3A3 B30.2 Domain

The B30.2 domain of human BTN3A3 was ordered as a synthetic gene in a bacterial expression plasmid, pET-29b, with an N-terminal His_8 tag and TEV cleavage site (Twist Bioscience, USA). This expression plasmid was chosen due to its commercial availability and common use as a bacterial expression plasmid system. Expression and purification were carried out as described in the materials and methods chapter. Briefly, B30.2 was affinity purified with the N-terminal His_8 tag using a 5 mL HisTrap column and eluted with imidazole (figure 4.9) before a final SEC step (figure 4.10).

The expression and purification figures below show that after bacterial lysis, there is still some B30.2 protein insoluble in the bacterial pellet, and there is a significant amount of protein which remains in the flowthrough after 2 column volumes of HisTrap column loading. This reflects the high levels of expression of the protein in bacteria.

The 5 mL HisTrap column used in the purification of the protein was subject to thorough washing with the protein was buffer (50 column volumes) and then bound protein was eluted with a linear gradient of imidazole in wash buffer up to a concentration of 250 mM imidazole. The resulting elution fractions containing B30.2 domain protein were pooled and applied to an S200 10/300 SEC column and separated. The major peak at 16.2 mL corresponds to a protein of around 27 kDa. The fractions from the major peak (figure 4.10, panel B, starred) were run on an SDS-PAGE gel (figure 4.10, panel A) and a protein of the expected molecular weight (27 kDa) was obtained, with a minor product below which was present in each purification despite more thorough washing attempted.

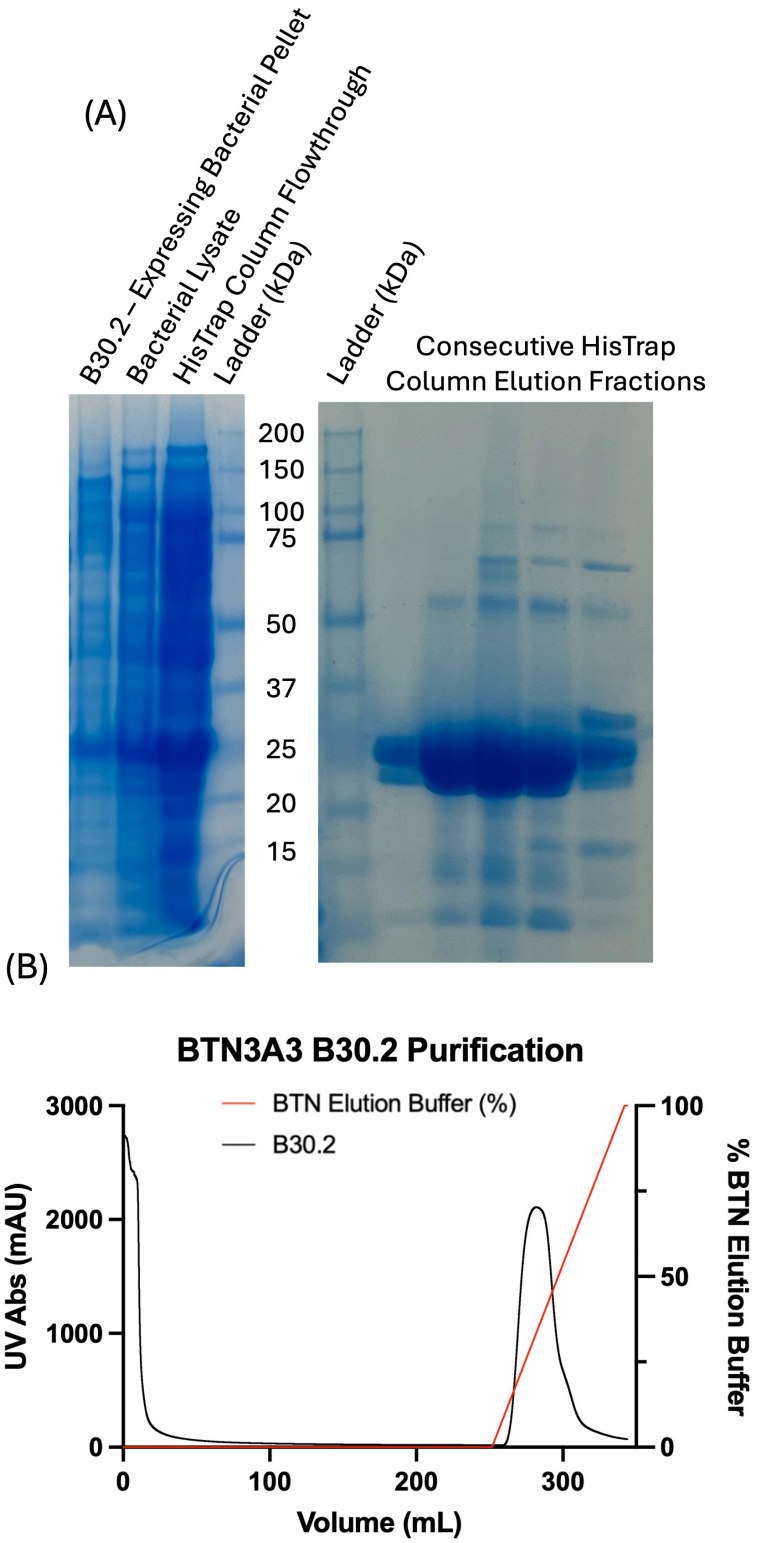


Figure 4.9: Purification of BTN3A3 B30.2 Domain.

(A) SDS-PAGE gels of purification of B30.2 domain using a 5 mL HisTrap column including elution fractions from HisTrap column. (B) SEC Chromatogram of B30.2 domain purification showing thorough wash step and elution with imidazole.

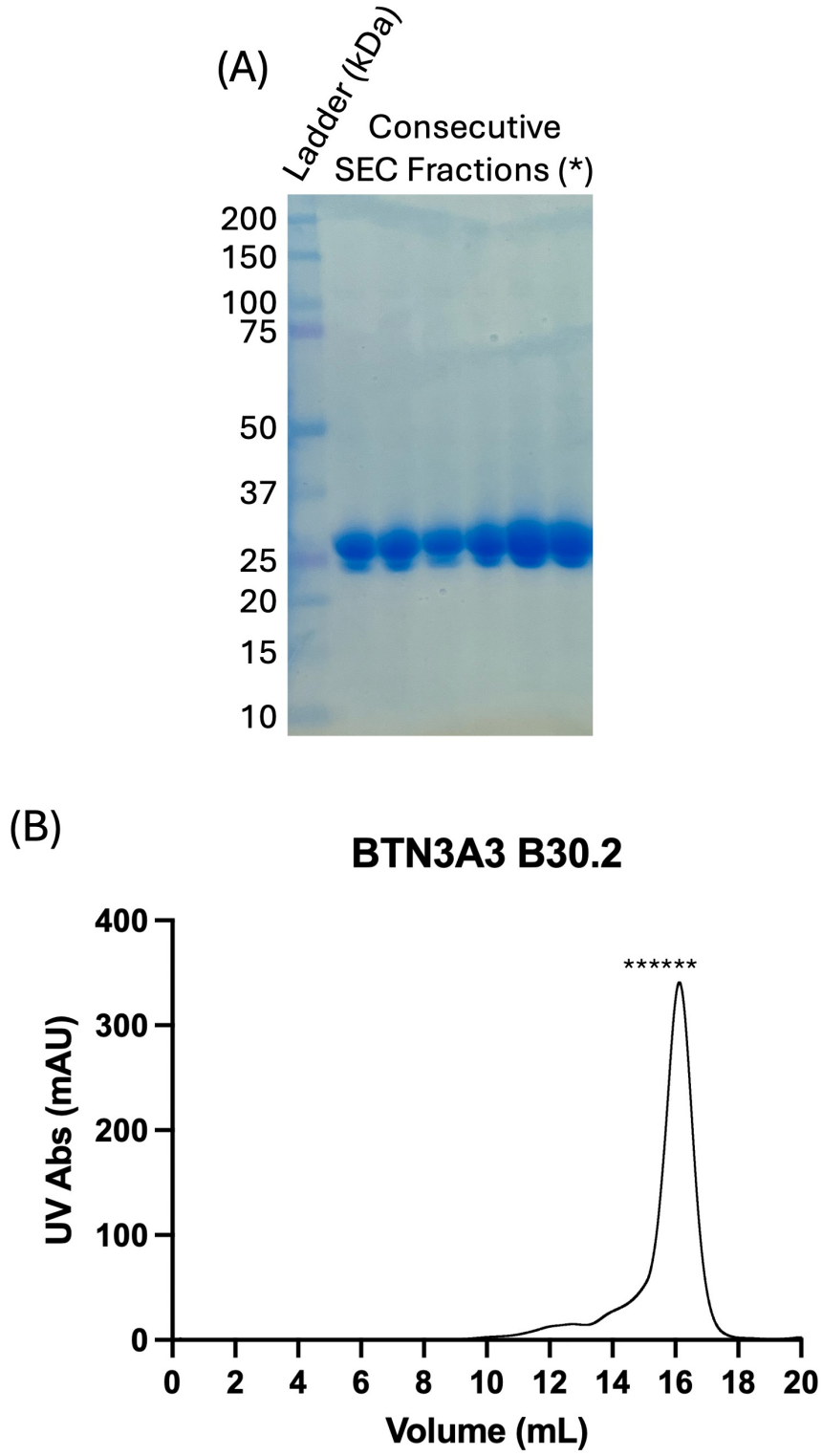


Figure 4.10: SEC of BTN3A3 B30.2 Domain.

SDS-PAGE gel (A) and SEC chromatograph (B) of BTN3A3 B30.2 domain purification.

4.2.2.2. Characterisation of BTN3A3 B30.2 Domain

The B30.2 domain protein was, as with MxA above, characterised with a number of biophysical methods to better understand the sample obtained.

4.2.2.2.1. Size-Exclusion Chromatography – Multi-Angle Light Scattering

SEC-MALS experiments were carried out together with Dr. David Staunton of the Molecular Biophysics Suite at the Department of Biochemistry, University of Oxford.

SEC-MALS analysis of the B30.2 domain was performed at 5 mg/mL in a buffer containing 25 mM HEPES-NaOH (pH 7.5) and 150 mM NaCl (figure 4.11). The overlapping light scattering (LS, red), UV absorbance (UV, green), and refractive index (RI, blue) traces demonstrate excellent sample homogeneity, while molar mass determination across the elution peak confirmed a molecular weight of 28 kDa. This experimental value corresponds closely to the theoretical monomeric mass, indicating that the B30.2 domain exists as a stable, non-aggregated monomer under these conditions in solution. These results confirm the production of monomeric B30.2 domain protein suitable for subsequent biophysical characterisation and structural studies.

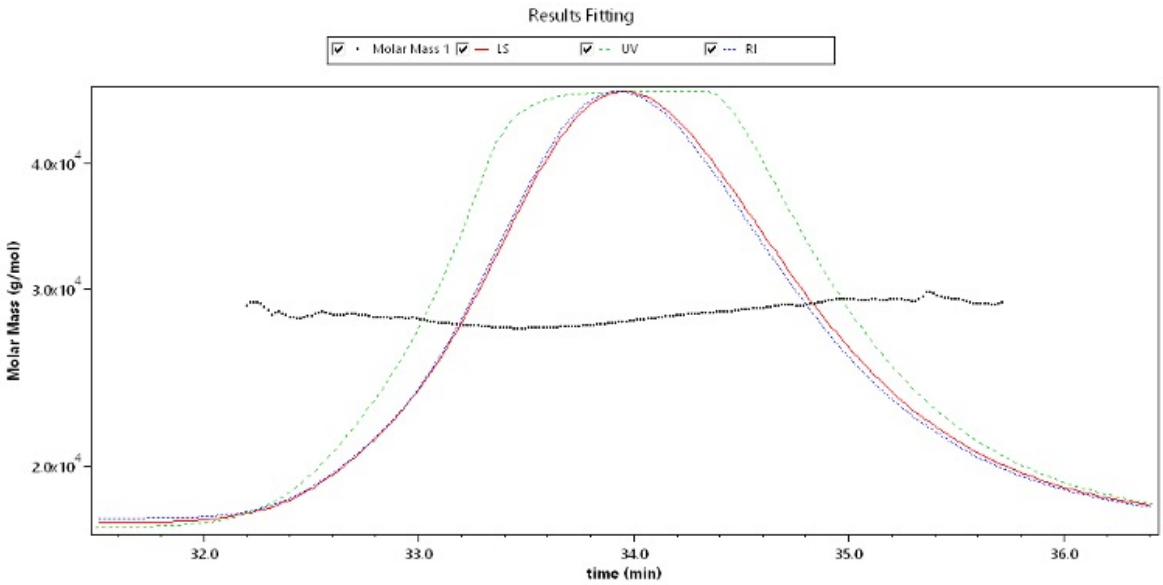


Figure 4.11: SEC-MALS of BTN3A3 B30.2 Domain Protein.

SEC-MALS peak trace of BTN3A3 B30.2 domain protein with light scattering (LS, red), UV absorbance (UV, dashed green), and refractive index (RI, blue), and molar mass traces (dotted black) plotted.

4.2.2.2.2. Dynamic Light Scattering

Dynamic light scattering experiments were carried out in triplicate using an UNcle machine to assess sample homogeneity and hydrodynamic diameter. Here, BTN3A3 B30.2 domain protein was diluted to 5 mg/mL and DLS measurements were taken at 24˚C.

The intensity distribution of triplicate B30.2 domain protein is displayed from a DLS experiment (figure 4.12). The reported traces have multiple peaks, one major peak at around 6 nm hydrodynamic diameter, similar to the expected particle diameter. In the tabular data (table 4.4) it is clear that one of the three samples has biased the overall data with a Z diameter of 56.15 nm compared to 6.14 nm and 6.22 nm. This data point is not included in the calculated means.

The majority of the intensity distribution is centred around the approximately 6 nm hydrodynamic diameter peak, and the second measurement is therefore likely erroneous, but reported for completeness.

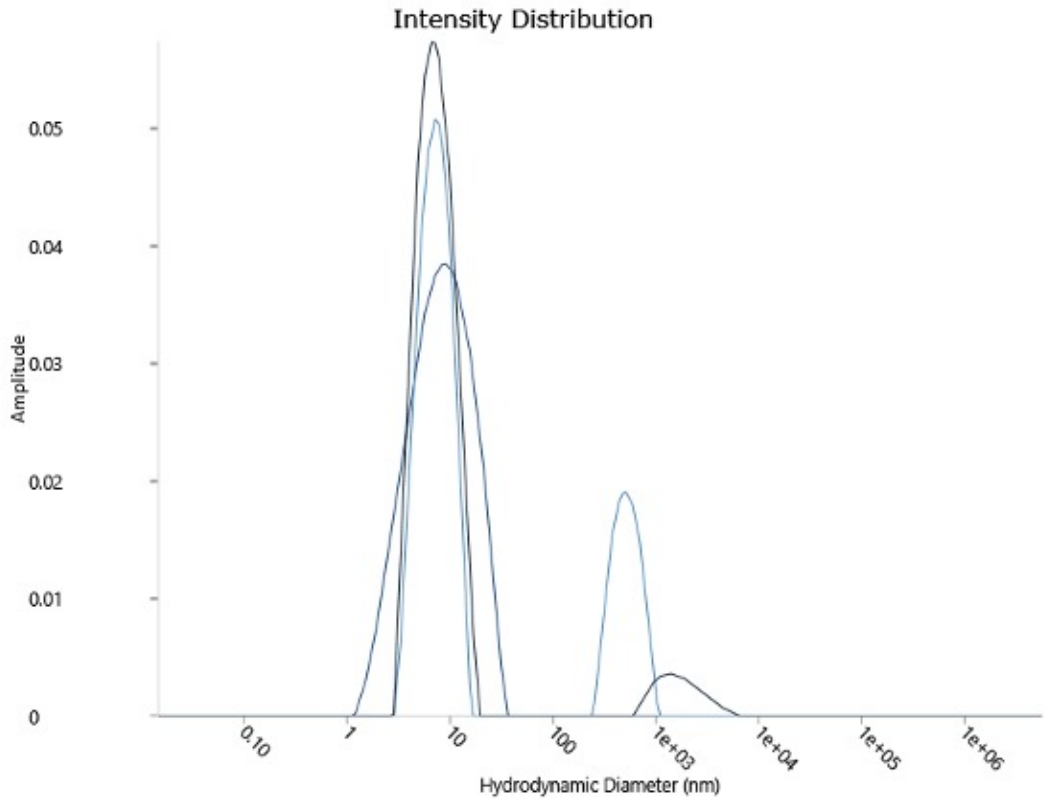


Figure 4.12: DLS of BTN3A3 B30.2 Domain Protein.

DLS size distribution plot of B30.2 domain protein at 5 mg/mL concentration in 150 mM NaCl. Different shades of blue represent individual repeats, as presented in the table below. Note: In future, it would be useful to carry out a control experiment using BSA to determine whether the UNcle machine can distinguish between different oligomeric states of protein.

Table 4.4: Tabular Data of BTN3A3 B30.2 Domain Protein DLS Experiment.

Tabular data of DLS experiment in figure 4.12, above. Duplicate experiments were run, where the results of each experiment are displayed in different shades of blue. Z diameter is a measure of hydrodynamic radius. SD diameter is the standard deviation of the measured diameter. Polydispersity Index (PDI) is the square of the SD diameter divided by the observed Z diameter, as defined in table 4.3 above. Note: the second repeat is included in the reported data for completeness, but not included in the calculated mean values. This data point is likely erroneous. In future, a calibration of the UNcle machine should be used to ensure correct measurement of DLS data.

Sample	Z Diameter (nm)	SD Diameter (nm)	PDI
B30.2 1	6.14	3.35	0.298
B30.2 2	56.15	22.30	0.158
B30.2 3	6.22	3.34	0.311
B30.2 Mean	**6.18**	**3.35**	**0.305**

4.2.2.3. Interaction Studies of BTN3A3 B30.2 Domain Protein with Influenza A Virus Nucleoprotein

The interaction of BTN3A3 B30.2 domain with S09 NP R416A nucleoprotein was investigated with a series of resin-based pull-down assays, SEC-based interaction studies, and finally quantitively with BLI.

4.2.2.3.1. Resin-Based Pull-Down Assay

In order to investigate the hypothesized interaction between the BTN3A3 B30.2 domain and an influenza A nucleoprotein expected to be sensitive to BTN3A3 restriction, a resin-based pull-down assay was carried out, as described in the materials and methods chapter.

Here, GST-tagged S09 NP R416A influenza A nucleoprotein was immobilised to glutathione sepharose resin (Cytiva) via its GST tag and recombinantly produced B30.2 domain protein was investigated as an analyte. Approximately 0.35 mg of S09 NP R416A nucleoprotein was thawed and incubated with 50 µL of glutathione sepharose resin which was pre-equilibrated with binding buffer (25 mM HEPES-NaOH (pH 7.5), 150 mM NaCl) and incubated on a rotating platform at 4˚C for 1 hour. After binding, the resin was thoroughly washed with 1 mL of binding buffer three times. After washing, an approximate 2-fold molar excess of purified B30.2 domain protein was added (36

µL of 10 mg/mL stock B30.2 domain protein at 10 mg/mL) and incubated at 4˚C on a rotating platform for 1 hour. After binding, the resin was thoroughly washed three times with 1 mL of binding buffer. After the final wash, the resin was resuspended in 500 µL binding buffer and SDS-PAGE samples were taken and analysed.

In the pull-down assay (figure 4.13), the S09 NP R416A input nucleoprotein was observed to bind to the glutathione sepharose resin, with excess nucleoprotein seen to be removed in the flowthrough from the wash steps. Upon the addition of the approximate 2-fold molar excess of B30.2 domain protein, much of the B30.2 domain protein is seen to be removed in the wash steps. However, a thin band of B30.2 domain protein is still present after thorough washing.

Given the presence of this thin band corresponding to B30.2 domain protein, and the significantly more intense band for S09 NP R416A nucleoprotein, a direct interaction between the nucleoprotein and the B30.2 domain protein has occurred, but the nucleoprotein sample is not completely occupied by B30.2 domain protein.

To confirm that this interaction was not due to a possible interaction between B30.2 domain protein and the glutathione sepharose resin used, the B30.2 domain protein was incubated with the glutathione sepharose resin alone. The same binding and wash protocol as detailed above was carried out. The B30.2 domain protein was present in both the input and flowthrough lanes, but not in the bound resin lane, suggesting that the B30.2 domain protein does not bind to the glutathione sepharose resin, and that the interaction observed was genuine (figure 4.13, panel B).

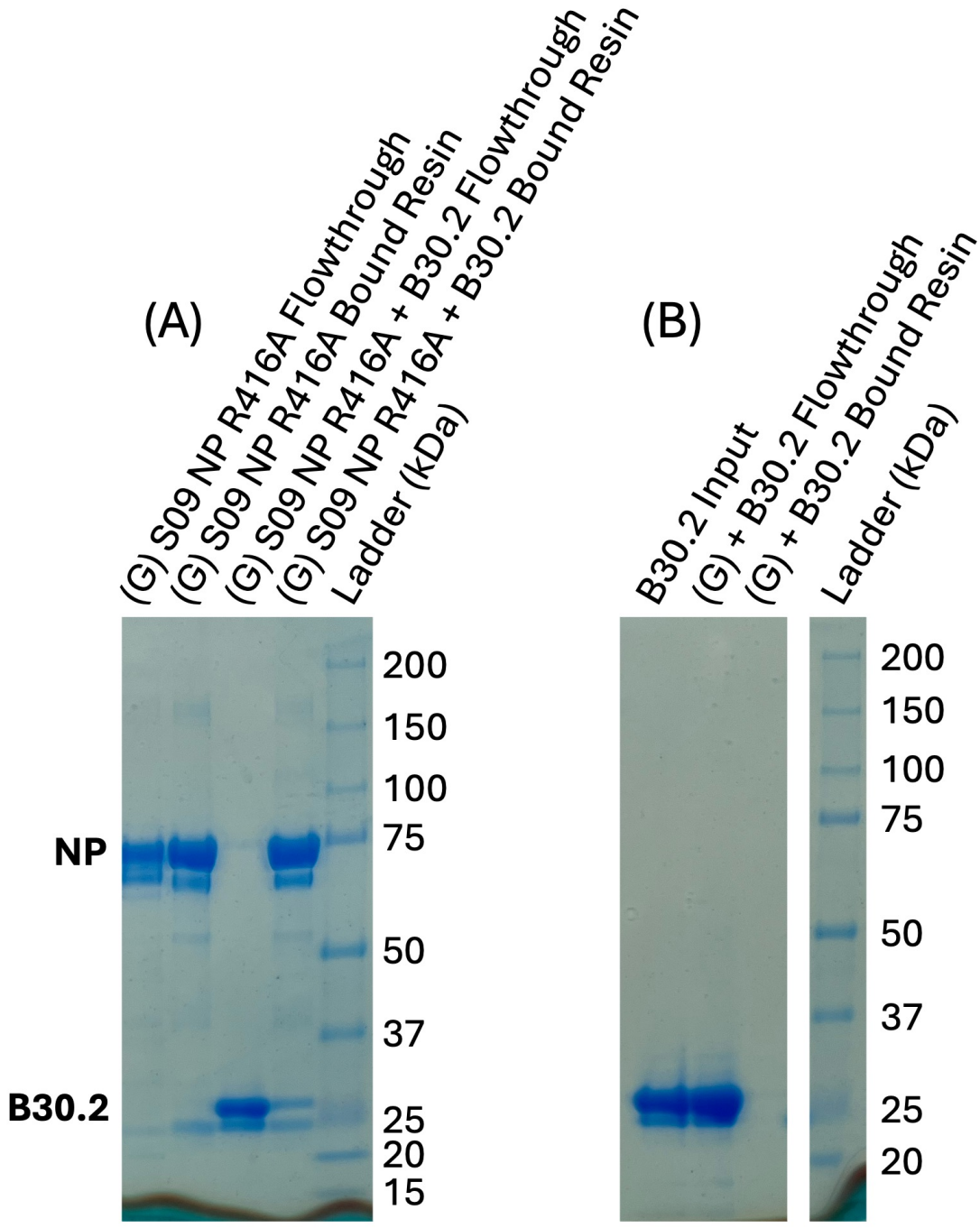


Figure 4.13: SDS-PAGE Gel of Resin-Based Pull Down Experiment of S09 NP R416A Nucleoprotein with BTN3A3 B30.2 Domain Protein.

SDS-PAGE gel of resin-based pull-down of B30.2 domain protein and GST-tagged S09 NP R416A. (A) GST-tagged S09 NP R416A nucleoprotein is first immobilised to glutathione sepharose (G) resin, and then B30.2 domain protein is added and the resin washed. Flowthrough and bound resin fractions of S09 NP R416A alone, and S09 NP R416A plus B30.2 domain protein are shown, with the approximate molecular weights of the bands for S09 NP R416A (NP) and B30.2 domain protein (B30.2) annotated. (B) A control experiment to confirm that B30.2 domain protein does not bind to the glutathione sepharose resin.

4.2.2.3.2. Size-Exclusion Chromatography

In the pull-down experiment above, direct binding of the B30.2 domain protein to the S09 NP R416A nucleoprotein was observed. However, insufficient amounts of complex were able to be eluted from the pull-down experiment for further structural experiments despite multiple attempts.

Complex formation was attempted in solution by the equimolar incubation of S09 NP R416A and B30.2 domain protein for its study by means of size-exclusion chromatography. Briefly, equimolar amounts of B30.2 domain protein and S09 NP R416A were incubated together on ice for 30 minutes in a buffer containing 150 mM HEPES-NaOH (pH 7.5), 150 mM NaOH and injected in a 500 µL injection loop on an S200 10/300 increase GL SEC column and separated by hydrodynamic radius (figure 4.14).

It is seen that no major peak corresponding to a complex with both components is formed, and therefore no complex was detected using SEC. This is consistent with the low-affinity interaction demonstrated by the pull-down assay. If this experiment were to be repeated, a higher molar excess of B30.2 domain protein should be used to attempt to form a complex with the S09 NP R416A nucleoprotein

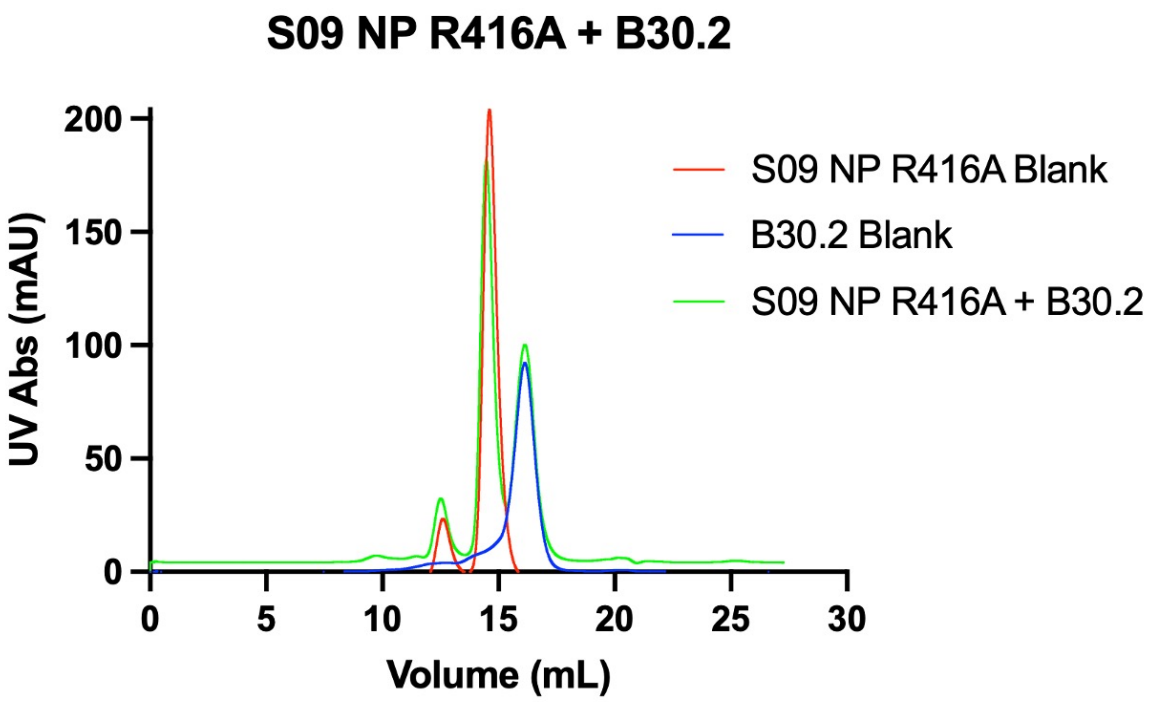


Figure 4.14: SEC of BTN3A3 B30.2 Domain Protein and S09 NP R416A.

Chromatogram of SEC run using an S200 10/300 GL Increase column with S09 NP R416A and BTN3A3 B30.2 domain proteins. S09 NP R416A run alone is depicted in red, BTN3A3 B30.2 domain run alone is depicted in blue, and the two components run together after a 30 minute incubation on ice is depicted in green.

4.2.2.3.3. Bio-Layer Interferometry

In order to further investigate the interaction between BTN3A3 B30.2 domain protein and nucleoproteins, a set of BLI experiments were devised where S09 NP R416A, Tk NP R416A, and NT60 NP R416A were immobilised on anti-GST tips as the ligand, and B30.2 domain protein was used as the analyte. Again, S09 and Tk nucleoproteins were used as test avian influenza nucleoproteins, as both are assumed to be sensitive to BTN3A3 restriction, and thus, be measured by BLI. NT60 nucleoprotein was used as a biological control, as it contains resistance mutations against BTN3A3 restriction. A constant concentration of nucleoprotein ligand (20 µg/mL ≈ 0.36 µM) was used across the three experiments, with a five point dilution series of B30.2 domain protein analyte

– 60 µg/mL (2.2 µM), 30 µg/mL (1.1 µM), 15 µg/mL (0.55 µM), 7.5 µg/mL (0.275 µM), and 3.75 µg/mL (0.1375 µM) used to test binding.

Low-level background binding to B30.2 domain protein to the tips was detected. Therefore, each experiment was run in parallel with an equivalent concentration gradient of analyte, but no-ligand bound, and the difference between the two response curves subtracted to account for background binding of analyte to the tips.

Kinetic parameters were able to be calculated from both S09 NP R416A nucleoprotein and Tk NP R416A nucleoprotein binding to B30.2 domain protein and are displayed below (figure 4.15). Both experiments reported below demonstrate a good steady-state kinetics model fit, with high R^2 values (0.9709 for the S09 NP R416A and B30.2 experiment, and 0.9694 for the Tk NP R416A and B30.2 experiment) corresponding to a kinetic model which fits the data very well, and low χ^2 values (0.0201 for the S09 NP R416A and B30.2 experiment, and 0.0059 for the Tk NP R416A and B30.2 experiment, both less than 10 % of the Rmax value of the experiment) which suggest a reliable fit.

The reported K_D of binding for the S09 NP R416A and B30.2 domain protein interaction is 1.20 µM, and the reported K_D of binding for the Tk NP R416A and B30.2 domain protein interaction is 0.42 µM, which indicates that the interaction of the B30.2 domain protein to the Tk NP R416A nucleoprotein is stronger than compared to that of the S09 NP R416A nucleoprotein.

The binding kinetics of NT60 NP R416A and B30.2 were not able to be calculated due to be calculated due to a poor spread of Req/Rmax values and poorly fitting data, which is due to the lack of sufficient binding response.

The processed sensograms of the S09 NP R416A, Tk NP R416A, and NT60 NP R416A interaction experiments with B30.2 domain protein are shown below (figure 4.16). In the processed sensograms below, over the period of 0 – 1000 s, varying concentrations of B30.2 analyte are incubated with the immobilised nucleoprotein and binding can occur. Between 1000 – 2000 s, the tips are moved into fresh experimental buffer (without analyte) and the complex formed in the binding step is able to dissociate. In each case below, the concentration of B30.2 domain protein decreases from the top trace, 60 µg/mL, to the bottom trace, 3.75 µg/mL, as detailed in the figure legend.

As in section 4.2.1.5.3., the Rmax value corresponds to the maximum measured interference between incident and reflected light transmitted through the biosensor and is used as a comparative measure of maximum binding response at equilibrium.

The recorded Rmax values of S09 = 1.61 nm, Tk = 2.52 nm, and NT60 = 0.24 nm obtained using a constant concentration of immobilised nucleoprotein and of B30.2 domain analyte (60 µg/mL = 2.2 µM) suggest different binding affinities to the nucleoproteins tested. The result that the two avian influenza nucleoproteins had a much higher Rmax value is expected given that they contain residues expected to make the virus sensitive to BTN3A3 restriction, whereas the human-adapted influenza

A virus nucleoprotein contains residues expected to make the virus resistant to BTN3A3 restriction.

These quantitative data provide the first biophysical evidence for direct binding interaction between the BTN3A3 B30.2 domain and BTN3A3-sensitive influenza A virus nucleoproteins, while further supporting the resistance phenotype of human-adapted influenza A virus strains. The micromolar binding affinities observed are consistent with the transient nature of the innate immune factor interaction and support the biological relevance of BTN3A3-mediated restriction.

Furthermore, the different binding affinities observed between the two avian influenza virus strains tested, Tk and S09, suggest that BTN3A3 sensitivity may vary among circulating influenza viruses, with additional residues potentially mediating sensitivity to BTN3A3 restriction which have yet to be identified.

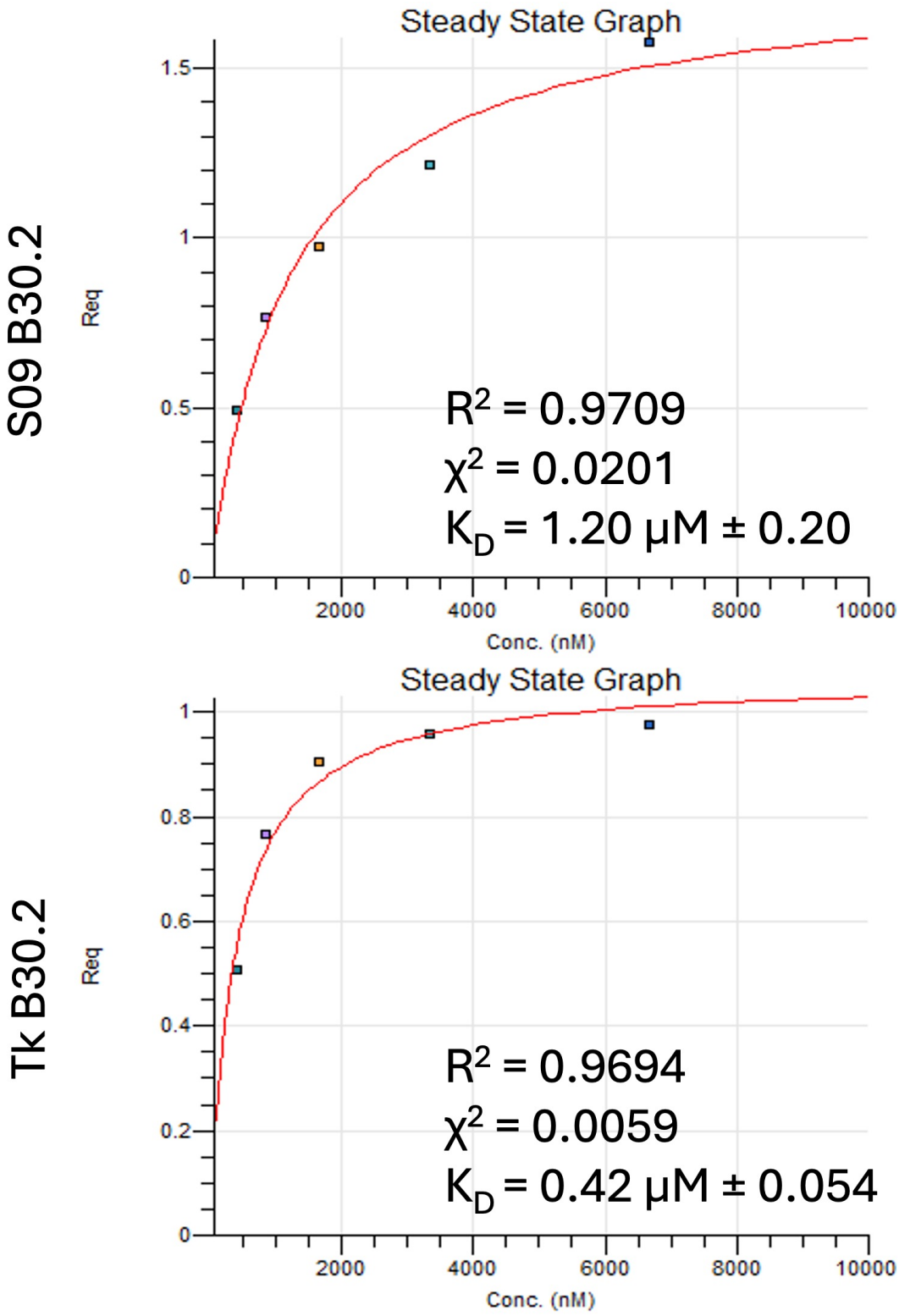


Figure 4.15: Steady-State Analysis Plots of BTN3A3 B30.2 Domain Protein BLI Experiments.

Steady-state analysis plots of BLI experiments with either S09 NP R416A (top) as the immobilised ligand, or Tk NP R416A (bottom) as the immobilised ligand.

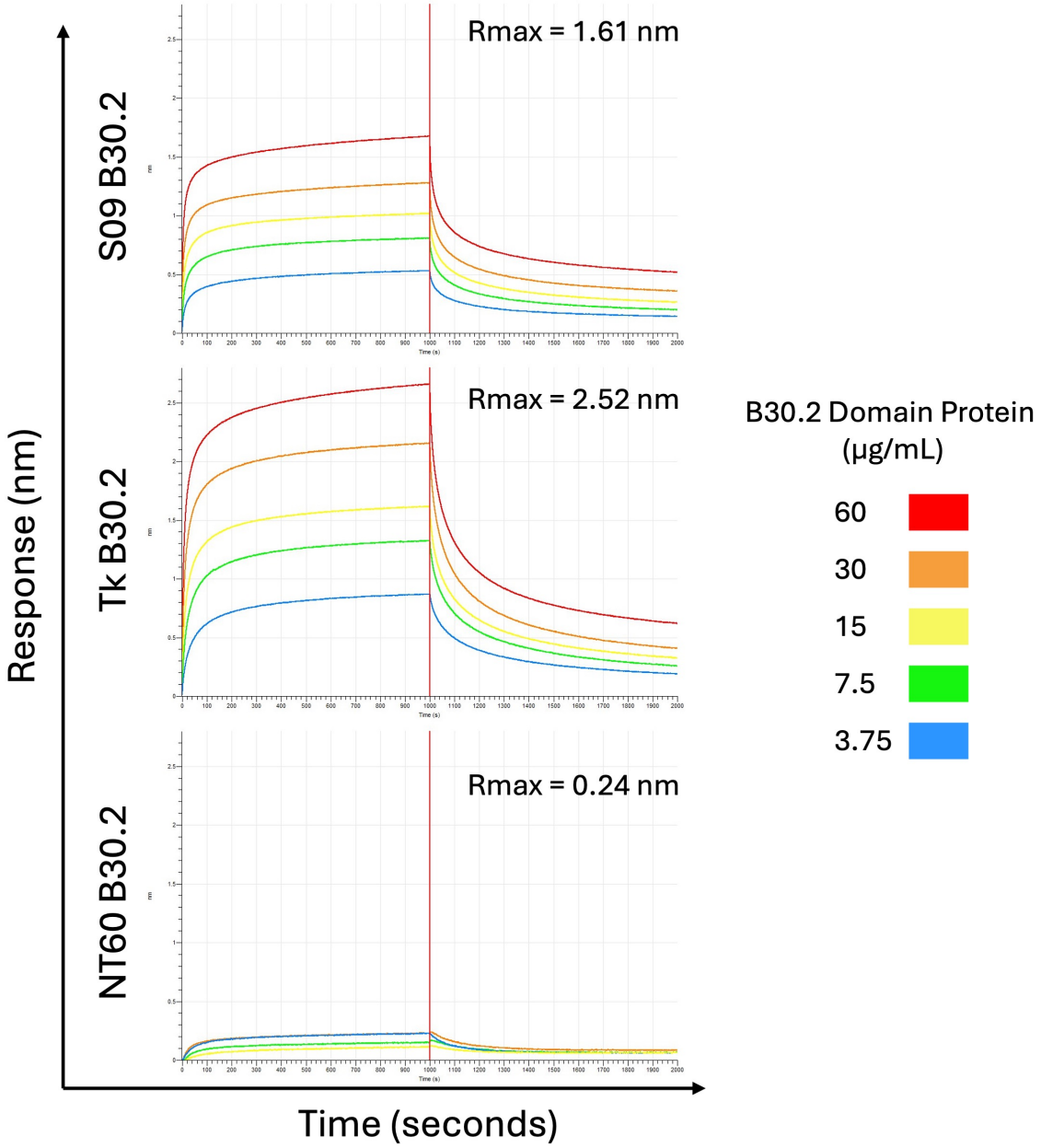


Figure 4.16: Processed Sensograms of BTN3A3 B30.2 Domain Protein + Nucleoprotein BLI Experiments.

Processed sensograms of BLI experiments with a constant concentration of nucleoprotein ligand (20 µg/mL) immobilised to each GST tip of S09 NP R416A (top), Tk NP R416A (middle), and NT60 NP R416A nucleoproteins (bottom), and a constant dilution series of B30.2 domain protein of 60 µg/mL (red), 30 µg/mL (orange), 15 µg/mL (yellow), 7.5 µg/mL (green), and 3.75 µg/mL (blue). Analyte association is displayed on the left hand side of the red line (0 – 1000 s), and analyte dissociation is displayed on the right hand side of the red line (1000 – 2000 s). Recorded Rmax values are annotated on each sensogram.

4.3. Discussion

This chapter presents the biochemical and biophysical characterisation of direct interactions of the innate immune factors MxA and BTN3A3 with the influenza A virus nucleoprotein. This chapter demonstrates strain-specific interaction patterns which correlate with predicted antiviral restriction profiles based on nucleoprotein sequence analysis.

4.3.1. MxA – Nucleoprotein Interactions

The MxA L4 loop peptide in isolation was shown to be able to bind to the nucleoprotein, demonstrating weak but detectable binding to the S09 NP R416A avian origin nucleoprotein in pull-down assays. Recombinantly produced full-length wild-type MxA exhibited concentration- and salt-dependent oligomerisation behaviour which was consistent with established literature (Kochs et al., 2002; Gao, von der Malsburg, et al., 2011; von der Malsburg et al., 2011). All MxA experiments were conducted in 150 mM NaCl in order to maintain physiological relevant oligomerisation states of MxA.

A significant experimental limitation was the absence of GTP in binding assays. GTP is known to promote the disassembly of MxA oligomers into dimers which are capable of binding to the influenza A nucleoprotein (Nigg and Pavlovic, 2015). In future work, it would be valuable to investigate MxA – nucleoprotein interactions in the presence of GTP to determine whether nucleotide-induced conformational changes affect the binding affinity of MxA interaction with influenza A nucleoproteins.

While resin-based pull-down assays demonstrated MxA binding to S09 NP R416A nucleoprotein, size-exclusion chromatography results remained irreproducible, likely reflecting the dynamic nature of MxA oligomerisation in solution. Despite employing a monomeric nucleoprotein construct, MxA oligomeric heterogeneity generates complex mixtures of potential binding partners. Given that current models propose MxA antiviral activity to involve dynamic re-oligomerisation around viral targets (Nigg and Pavlovic, 2015), this oligomeric heterogeneity may fundamentally preclude stable complex formation under standard biochemical conditions.

In the future, these experiments should be re-attempted with a vast excess of MxA protein in order to attempt to isolate sufficient amounts of complex for further structural experiments.

BLI experiments, while unable to yield quantifiable parameters of binding affinity due to inconsistent biosensor performance, provided valuable comparative binding data through Rmax analysis. These measurements revealed preferential binding of MxA to avian influenza nucleoproteins versus human-adapted influenza nucleoproteins. This differential binding pattern correlates with the observed pattern of MxA's restriction of avian origin influenza viruses.

The results in this chapter have confirmed the MxA L4 loop as the minimal region for binding to the influenza A nucleoprotein. However, it is likely that the affinity of the MxA L4 loop peptide is less than that of the full-length wild-type protein as the full-length protein is thought to be antivirally active in its dimeric form (Dick et al., 2015; Nigg and

Pavlovic, 2015). This interpretation is supported by the similar intensity of Coomassie staining of the nucleoprotein and full length MxA by SDS-PAGE in the pull-down presented above.

4.3.2. BTN3A3 B30.2 Domain Protein – Nucleoprotein Interactions

SEC-MALS analysis confirmed that the BTN3A3 B30.2 domain exists as a stable, monomeric species in solution, providing an optimal system for biophysical characterisation. Resin-based pull-down assays demonstrated weak but reproducible binding to S09 NP R416A, though attempts to purify stable complexes via SEC were unsuccessful. In the future, these experiments should be re-attempted with a vast excess of B30.2 domain protein in order to attempt to isolate sufficient amounts of complex for further structural experiments.

Quantitative BLI analysis successfully determined binding kinetics for both avian nucleoprotein variants, revealing differential binding affinities: Tk NP R416A (K_D = 0.42 μM) bound more strongly than S09 NP R416A (K_D = 1.20 μM). Comparative Rmax analysis across nucleoprotein variants confirmed preferential binding to avian strains over human-adapted NT60 NP R416A, providing quantitative validation for the mechanism proposed by Pinto and colleagues for the direct restriction of avian, but not human-adapted influenza A virus infection by BTN3A3.

4.3.3. Structural and Functional Implications

The micromolar binding affinities observed for both restriction factors were insufficient for stable complex isolation required for structural studies, suggesting that these factors operate through transient, dynamic interactions rather than stable complex formation. Future experiments should employ vast molar excesses of immune factors to favour complex formation and enable structural characterisation.

The avian influenza-specific binding preferences of both restriction factors correlate with known sensitivity of these viruses to human innate immune factors. Higher Rmax values obtained with avian nucleoproteins underscore the importance of species-specific residues in determining restriction factor efficiency.

Critical mechanistic questions remain unresolved, particularly regarding BTN3A3 subcellular localisation. While Pinto and colleagues demonstrated nuclear BTN3A3-nucleoprotein interactions through co-immunoprecipitation, the mechanism enabling nuclear import of this transmembrane protein remains unclear. The possibility of proteolytic cleavage and subsequent translocation of the cytoplasmic domain of BTN3A3 cannot be excluded based on current data.

Previous studies reported perinuclear localisation of HA-tagged butyrophilin family members (Karunakaran et al., 2023), suggesting potential subcellular localisation of the butyrophilin family of proteins beyond solely at the plasma membrane. However, nuclear localisation has yet to be reported, highlighting the need for targeted

experiments to resolve BTN3A3 trafficking mechanisms and confirm the proposed nuclear restriction pathway.

4.3.4. Computational Structure Prediction Analysis

To complement experimental binding data, AlphaFold3 predictions were generated for nucleoprotein complexes with B30.2 domain, MxA L4 peptide, and full-length MxA (figure 4.17). The nucleoprotein – B30.2 domain prediction positioned the B30.2 domain such that reported sensitivity residues (52 and 313) were directly implicated in the binding interface, consistent with the reported importance of these residues for BTN3A3 binding to the influenza A nucleoprotein.

However, neither the MxA L4 peptide nor the full-length MxA predictions in complex with the nucleoprotein implicated known MxA sensitivity residues in the binding interface. Additional predictions using dimeric MxA yielded similar stalk domain-mediated binding modes. The absence of conversion on high-confidence predictions for MxA interactions demonstrates the critical need for experimental structural determination of stable nucleoprotein – immune factor complexes.

In all predictions displayed in figure 4.17, both the interface predicted template modelling (ipTM) and predicted template modelling (pTM) scores are reported. The pTM score is a computed measure of how well AlphaFold-Multimer has predicted the overall structure. The ipTM score is a measure of the accuracy of the predicted relative positions of subunits forming the protein-protein interaction.

In figure 4.17, the pTM scores of the B30.2 and L4 peptide complexes are above a threshold of 0.5, which gives confidence to the structural predictions, however the MxA complex is below 0.5, suggesting that this prediction may not accurately reflect the overall structure of the proteins in the complex. In all three cases however, the ipTM is very low. In high-quality predictions, the pTM score should be above 0.5 and the ipTM score is expected to be higher than 0.8.

When a similar approach was taken with the same three protein complexes, but the substitution of the human-adapted NT60 NP R416A, very similar modes of binding were observed and confidence scores obtained, with only a slight difference in the structure prediction of the nucleoprotein in complex with B30.2 domain protein. This suggests that these predictions may not be correct given their low confidence scores and similar binding when compared to a human-adapted, restriction factor-resistant nucleoprotein, and further emphasises the requirement for high resolutions structure solutions to better understand the nature of these interactions.

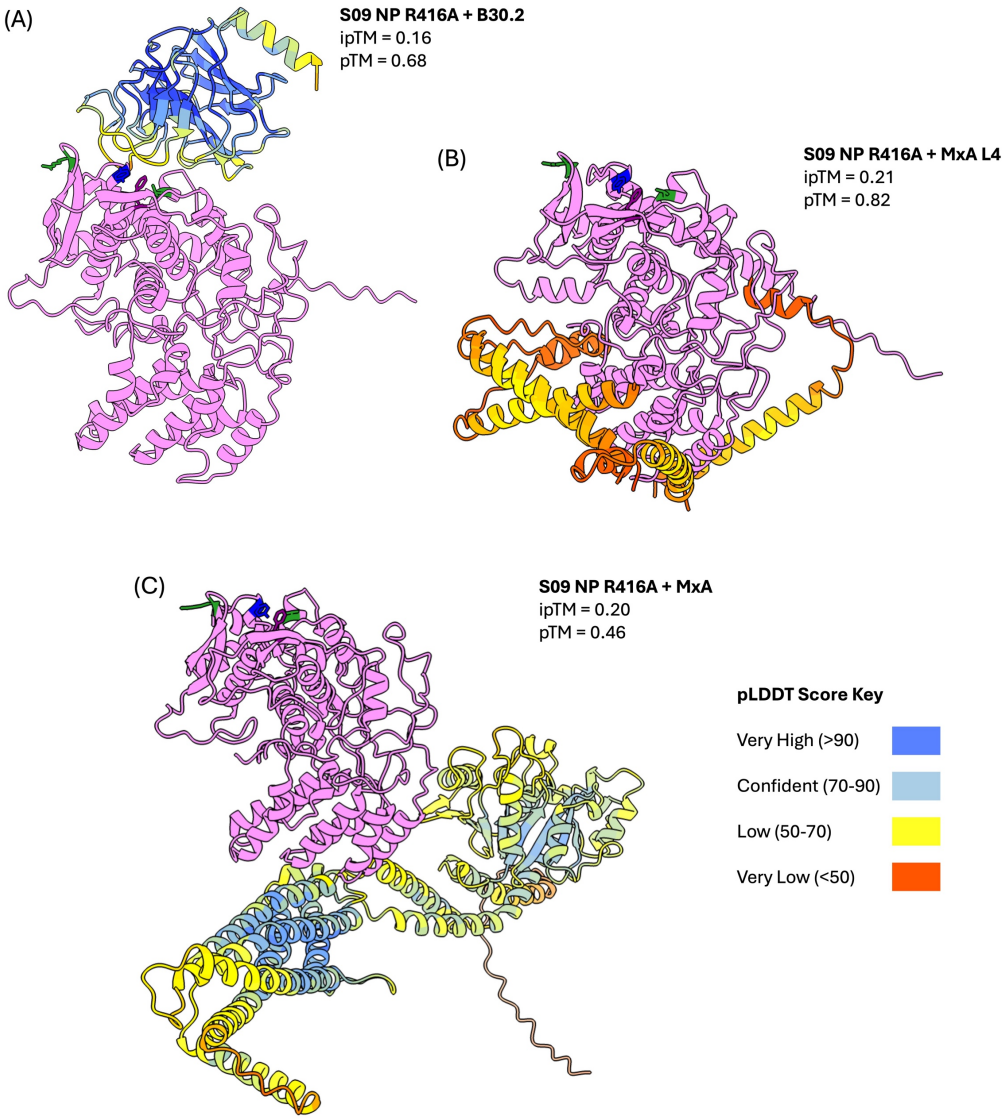


Figure 4.17: AlphaFold3 Predictions of Immune Factor Binding.

AlphaFold3 predictions of the S09 NP R416A nucleoprotein (pink) and BTN3A3 B30.2 domain (panel A), MxA L4 peptide (panel B), and full-length MxA protein (panel C). The immune factors of interest are displayed coloured by pLDDT score. Sensitivity residues on S09 NP R416A for MxA (100, 283), BTN3A3 (52), and both MxA and BTN3A3 (313) drawn as atoms in dark green, blue, and purple respectively. (A) One representative AlphaFold3 prediction of the B30.2 domain binding to S09 NP R416A. The predicted binding site overlaps with the sensitivity residues reported to be involved in BTN3A3 restriction. (B) Five predictions of the position of the MxA L4 peptide are illustrated, none of which overlap with the well-known sensitivity residues for MxA restriction. (C) One representative AlphaFold3 prediction of the full-length MxA protein binding to S09 NP R416A. Neither the MxA L4 loop nor the known sensitivity residues of the nucleoprotein are predicted in binding.

4.3.5. Experimental Limitations and Future Directions

In this chapter, I have demonstrated binding of the MxA L4 loop peptide, full length wildtype MxA protein, and of the BTN3A3 B30.2 domain protein to susceptible avian-origin influenza virus nucleoproteins, but not to resistant human-adapted influenza virus nucleoproteins. These results are impactful in the field, as they demonstrate the first biophysical support for B30.2 domain – nucleoprotein interactions (with a statistically-rigorous K_D calculated for two avian-origin nucleoproteins), and further, they emphasise the importance of the reported sensitivity mutations on restriction by these two innate immune factors.

Several technical improvements would enhance future investigations of these antiviral interactions. For MxA studies, tagged L4 loop peptides would enable direct comparison with full-length MxA by means of BLI.

A key limitation was maintaining the nucleoprotein in its monomeric form, precluding analysis of MxA binding to physiologically relevant oligomeric states. While this approach enabled controlled binding studies, it limits biological relevance. Future work should explore binding to oligomeric nucleoprotein assemblies, despite the increased experimental complexity.

These findings provide the first quantitative framework for understanding innate immune factor – nucleoprotein interactions while highlighting the dynamic nature of

these antiviral mechanisms and the need for continued structural and mechanistic investigation.

5. Novel Approaches to study Influenza vRNP Structure and its Interaction with Nucleoprotein Binding Partners

5.1. Introduction

Two principal methods exist for generating influenza virus genome segments suitable for biochemical and structural investigation. The first approach involves *in vitro* reconstitution of individual viral ribonucleoprotein (vRNP) segments by cellular transfection with plasmids encoding the nucleoprotein, the three influenza polymerase subunits, and RNA sequences corresponding to segments of interest (Peng et al., 2019). Products of this methodology are designated as RNPs (without the 'v' prefix) to distinguish their non-viral origin from native viral particles.

Alternatively, an *in situ* approach enables vRNP isolation either through direct purification of lysed viral particles or enrichment from infected cellular lysates (Zhu et al., 2023). This latter strategy can be implemented using live viral infection or through transfection-based reverse genetics systems that encode for the complete influenza viral genome and produce live viral particles (Fodor et al., 1999; Hoffmann et al., 2000).

A third, recently developed approach produces vRNP-like particles through completely recombinant methods. This system uses a specific nucleoprotein truncation (Δ1-14) in the N-terminus and synthetic $(UC)_{18}$-fluorescein$^{3'}$ (FAM) RNA to

produce completely recombinant vRNP-like particles, also called pseudo-vRNPs (Chenavier et al., 2023; Chenavier et al., 2025). While pseudo vRNPs provide valuable insights into nucleoprotein structure within vRNP-like conformations, their utility is constrained by sensitivity to experimental parameters. Minor variations in protein construct design and RNA composition can dramatically alter pseudo-vRNP architecture – shifting from physiologically relevant antiparallel double helical conformations (exhibiting major and minor grooves) to parallel double helices or single helical arrangements. Furthermore, the absence of a functional polymerase and a continuous RNA segment in pseudo-vRNP systems limits their use for mechanistic studies of viral replication and transcription.

5.1.1. Challenges in Structural Elucidation of Influenza Genome Segments

High-resolution structure determination by cryo-EM requires relatively homogeneous protein samples and datasets with accurate particle alignment. Conformational flexibility can hamper high resolution protein structure determination by distributing particles across multiple conformational classes, thereby leaving too few particles which align to a sufficiently high resolution for interpretation of finer biological features (Arranz et al., 2012; Moeller et al., 2012; Gallagher et al., 2017; Coloma et al., 2020). While processing approaches such as 3D classification and continuous flexibility modelling can partially address these limitations, they cannot fully compensate for samples with extensive structural heterogeneity.

Inherent local nucleoprotein flexibility and global flexibility of the vRNP complex are likely required features for the vRNP to function physiologically. In order to facilitate the study of native vRNP structure, strategies that rigidify the vRNP – either physically on the electron microscopy grid, or by employing vRNP binding partners (e.g. nanobodies, immune factors) – may be useful approaches to improve 3D reconstructions by reducing overall sample flexibility and heterogeneity.

Added to the issue of sample flexibility, biologically-relevant vRNPs must be produced using cell-based techniques, as it is not yet possible to produce true vRNP-like complexes associated with both RNA and the influenza polymerase using recombinantly expressed and purified material.

Consequently, biologically-relevant vRNPs must be purified from infected or transfected cells, or enriched directly from viral particles. This cellular production requirement introduces practical limitations: extended production times and reduced material yields compared to bacterial expression systems, and consequently diminished experimental throughput. The combination of inherent sample flexibility and limited material availability renders vRNP structural studies incompatible with X-ray crystallography approaches. Instead, structural investigation of viral genome segments predominantly employs cryo-electron microscopy single particle analysis, and increasingly, cryo-electron tomography coupled with sub-tomogram averaging techniques, better suited to handling flexible, heterogeneous macromolecular complexes (Graham and Zhang, 2023).

5.1.2. The Current Understanding of Influenza Genome Segment Structure

Despite extensive research on influenza vRNP complexes, high-resolution structural information remained elusive prior to this work. Initial ultrastructural characterisation in the 1970s revealed vRNPs as double helical rod-shaped structures of variable length under negative-stain electron microscopy (Compans et al., 1972).

To circumvent the inherent structural flexibility of full-length vRNPs, early three-dimensional electron microscopy studies focussed on reconstituted mini-RNPs. These simplified systems were generated through mammalian cell transfection with plasmids encoding the three polymerase subunits (PB2, PB1, PA), nucleoprotein, and a construct producing a short 248-nucleotide vRNA transcript that promoted nucleoprotein oligomerisation (Ortega et al., 2000).

Structural analysis by negative-stain electron microscopy (Martín-Benito et al., 2001; Area et al., 2004) and subsequent cryo-electron microscopy (Coloma et al., 2009) revealed mini-RNPs as nine-membered nucleoprotein rings with a single polymerase copy attached via interactions with paired RNP termini (figure 5.1). While this work provided crucial insights into influenza genome organisation, the artificial 248-nucleotide RNA constraint forced mini-RNPs into compact loop conformations that contrasted with the physiologically relevant double helical architecture observed in native vRNPs.

This discrepancy between mini-RNP and native vRNP ultrastructure highlighted the need for structural approaches capable of handling full-length, biologically relevant ribonucleoprotein complexes.

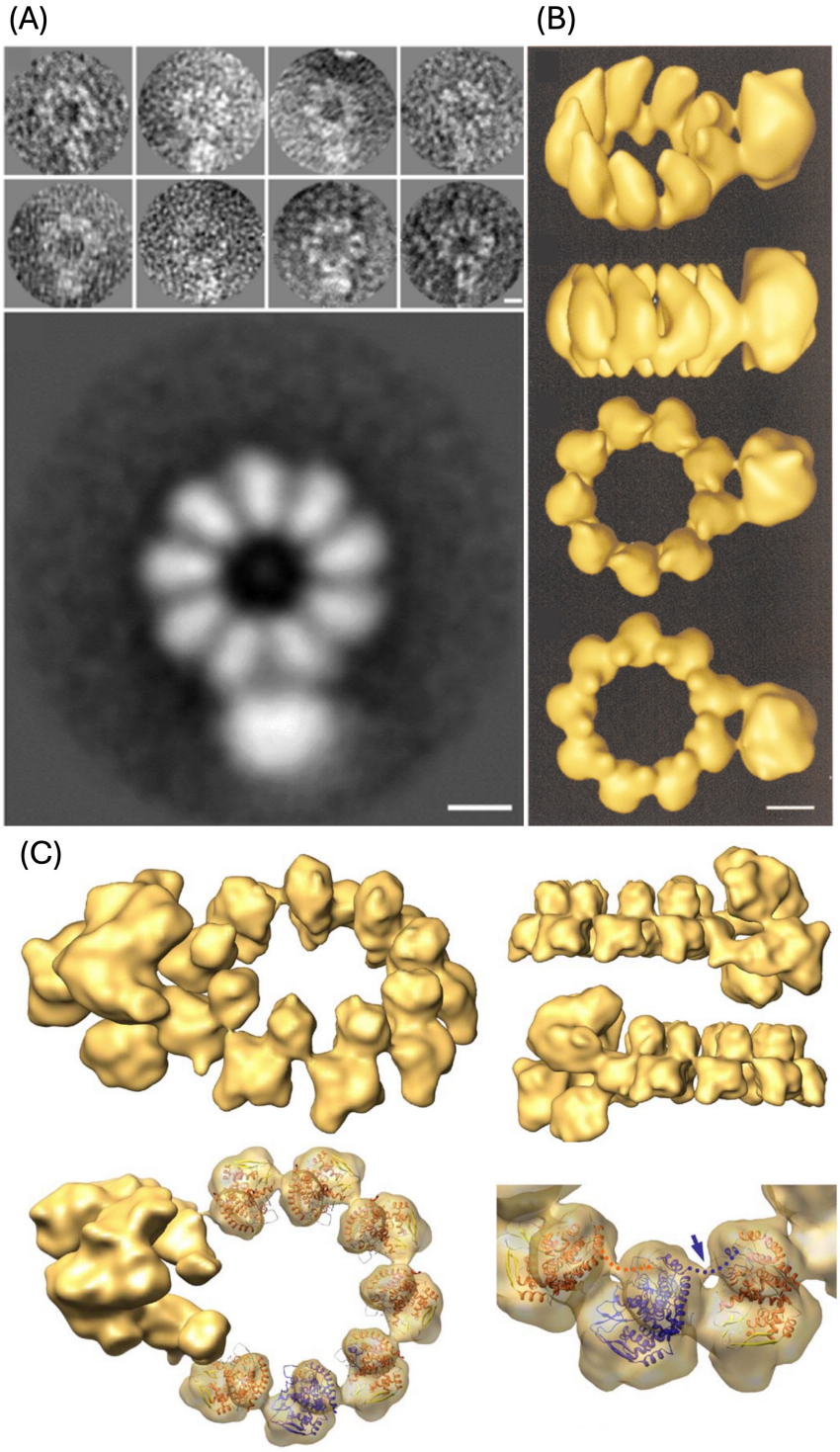


Figure 5.1: Mini-vRNP Structure.

Low-resolution structural data of influenza mini-RNPs. (A) 2D averages of negative stain electron microscopy data of influenza mini-RNPs. Scale bar = 50 Å. (B) 3D density map of influenza mini-RNPs determined by negative stain electron microscopy. Scale bar = 50 Å. Both panels A and B are taken from (Martín-Benito et al., 2001). (C) 3D density map (top) and rigid body fitting of influenza A nucleoprotein into the obtained density (bottom) of influenza mini-RNPs determined by cryo-EM SPA. (Coloma et al., 2019).

Structural determination of native-length RNPs, either as virion-derived vRNPs (Arranz et al., 2012), or by overexpression in mammalian cell lines (Moeller et al., 2012) was later pursued. This was highly advantageous, as full-length RNPs are more representative of the biological context of the RNP. These data confirmed the antiparallel organisation of the nucleoproteins of the RNP into a double helix with a canonical major and minor groove as a result of nucleoprotein oligomerisation. Also observed in both systems were the canonical loop ends, and opposite, a putative influenza polymerase-bound end (figure 5.2).

Key in these two papers is that the handedness of the reconstructed vRNP is different. Arranz and colleagues determined a low-resolution reconstruction of the vRNP with a left-handed double helix (figure 5.2, panels A-G), whereas Moeller and colleagues determined a right-handed double helix (figure 5.2, panels H-K). The handedness of the helix remains ambiguous at this resolution.

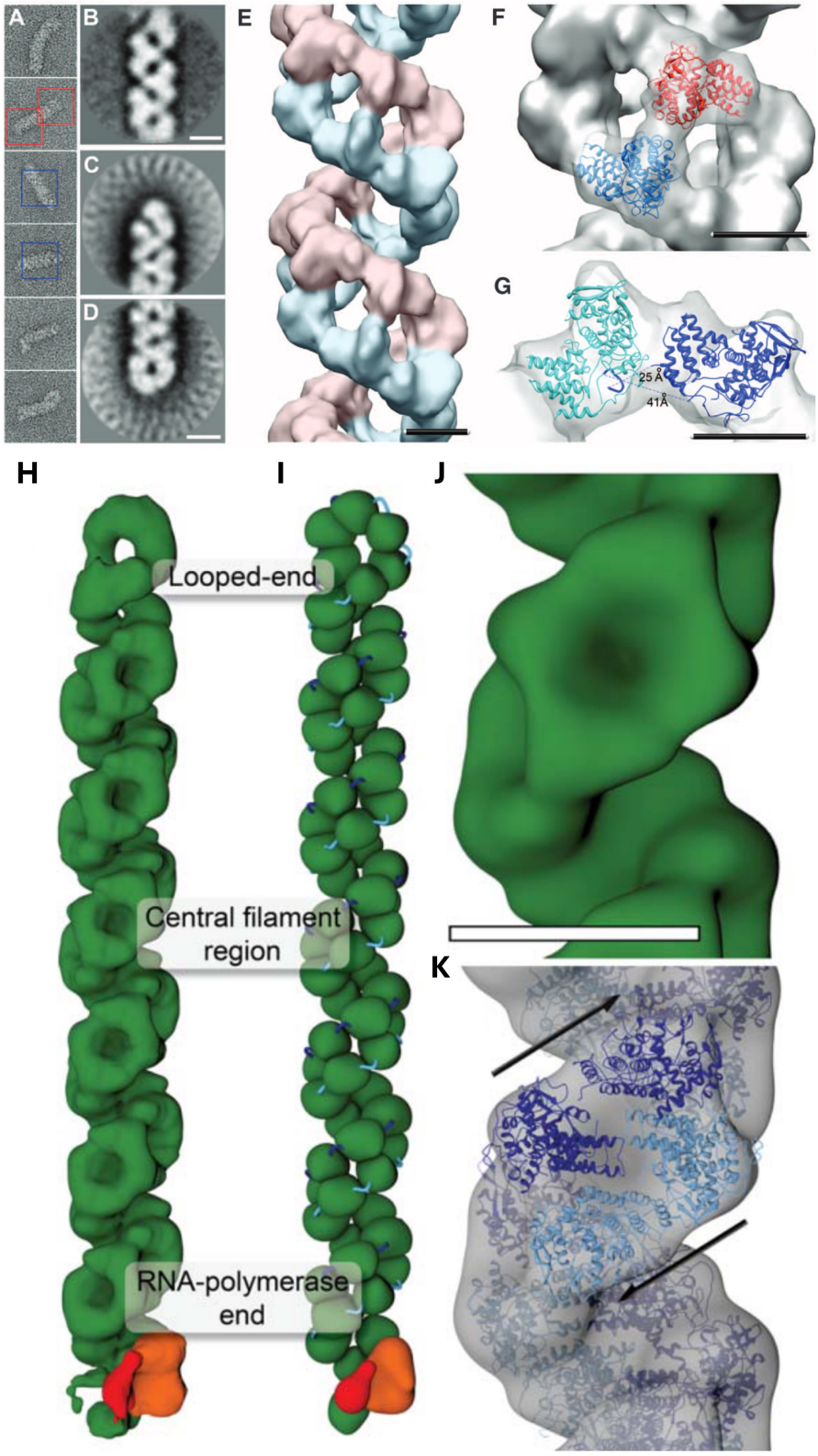


Figure 5.2: vRNP Structures as determined in 2012.

vRNP structures as determined by Arranz and colleagues (Arranz et al., 2012) (panels A – G) and by Moeller and colleagues (Moeller et al., 2012) (panels H – K). (A) Gallery of vRNP images obtained by negative stain electron microscopy. Regions boxed in red correspond to terminal vRNP regions used for terminal vRNP structure processing. Regions boxed in blue correspond to central, helical vRNP regions used to process the central helical section of the vRNP. (B) 2D average obtained from picked central

regions. (C) 2D average obtained from polymerase-containing vRNP termini. (D) 2D average obtained from loop-containing vRNP termini. (E) 3D cryo-EM reconstruction of the central portion of a vRNP. Opposite polarity strands are illustrated in pink and blue. (F) Rigid body docking of two nucleoprotein monomers into opposite strands. (G) Rigid body docking of two nucleoprotein monomers at an interstrand connection. (H) Composite model of cryo-EM reconstruction of three regions of a vRNP – the looped end, the central filament, and the RNA polymerase end. (I) Cartoon representation of corresponding vRNP organisation. (J) Reconstruction of the central filament region imposing helical symmetry. (K) Nucleoprotein monomer structures rigid body fitted into the cryo-EM density. Two opposing arrows represent differing strand polarity. Scale bar represents 100 Å in panels B – D and J, and 50 Å in panels E – G. Figures adapted from Arranz et al., 2012 and Moeller et al., 2012.

Moeller and colleagues additionally employed a C-terminal His_6 tag on the influenza polymerase PB2 domain in order to identify the position of the polymerase on the vRNP using a Ni-NTA nanogold particle. In doing so, they observed that the influenza polymerase was present near to, but not exclusively at the end of the RNP (Moeller et al., 2012).

Cryo-electron tomography was employed in the study by Arranz and colleagues to show that vRNPs also exist in the influenza virion with an approximately parallel arrangement of neighbouring vRNPs (figure 5.3), and the same double helical structure determined as virion-derived vRNPs *ex virio* (figure 5.2, panels A – G). The best nominal resolution from this study employing cryo-EM SPA was around 18 Å. This low-resolution was explained by the conformational heterogeneity of the vRNPs (Arranz et al., 2012).

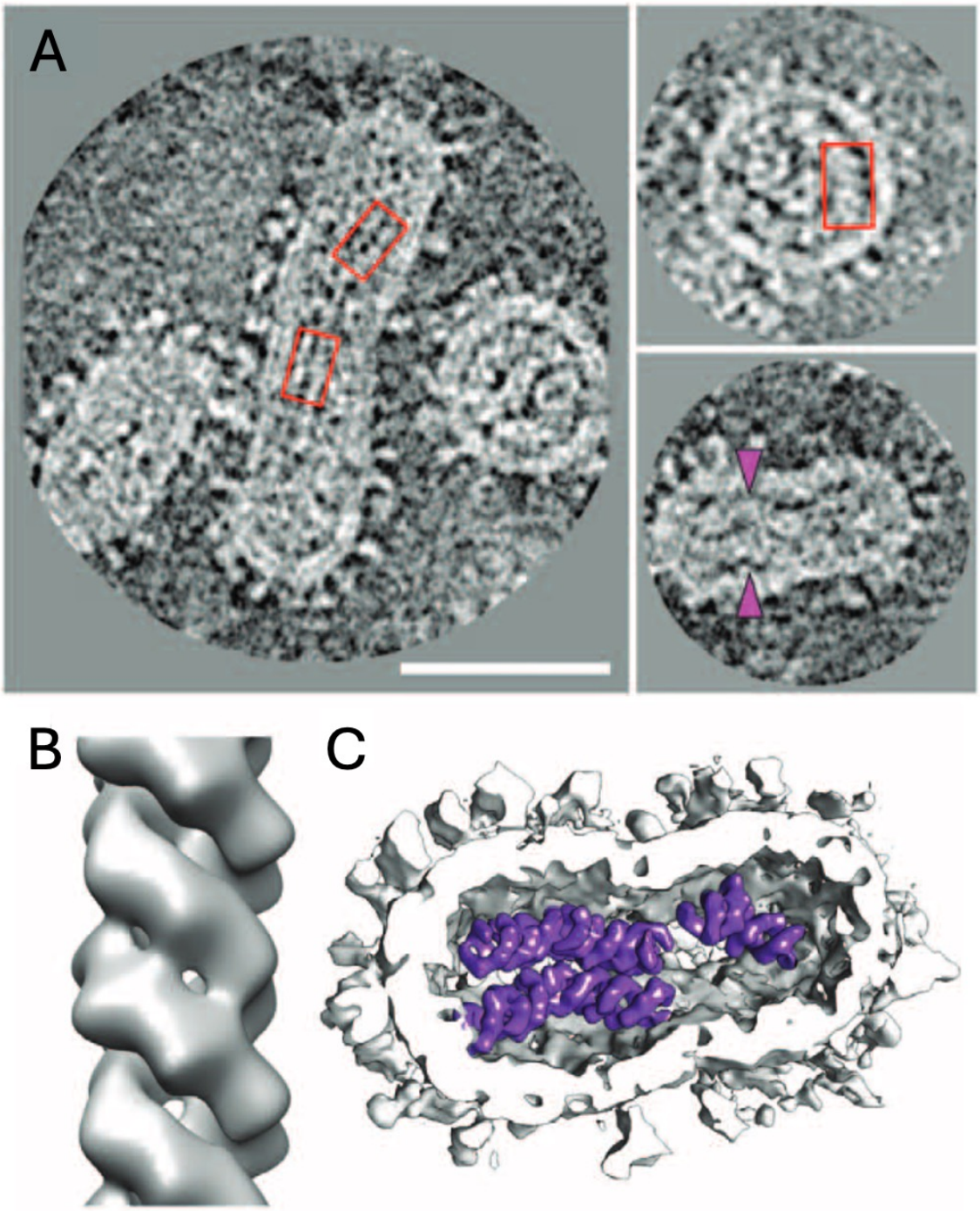


Figure 5.3: vRNPs *in virio*.

The structure of influenza vRNPs in virio as studied by Arranz and colleagues using cryo-electron tomography and sub-volume averaging. (A) Central sections of tomograms with select vRNPs highlighted in red rectangles. (B) Sub-volume average after processing of 288 volumes extracted from tomograms. (C) Volumetric representation of an influenza virus particle segmented with vRNPs coloured in purple. Scale bar = 1000 Å. Figure adapted from Arranz et al., 2012.

The observation of highly flexible vRNPs (underpinning both vRNP function and limitations for structural studies) was further supported by the characterisation of vRNP size, curvature, and organisation by Gallagher and colleagues studying vRNPs isolated from live virus using cryo-electron tomography (Gallagher et al., 2017) (figure 5.4).

Coloma and colleagues later carried out 3D classification of vRNP data using cryo-EM SPA and found that the relative position of the nucleoprotein on one strand was not fixed with respect to the position of the nucleoprotein on the adjacent strand, but could vary extensively, limiting the feasibility of high-resolution reconstructions with such heterogeneous data, but highlighting the required flexibility of the vRNP. Analysis of the reported variability of vRNP conformations during transcription revealed a sliding motion of one strand over the other to facilitate polymerase movement through the RNA template without disrupting the overall double helix structure (Coloma et al., 2020).

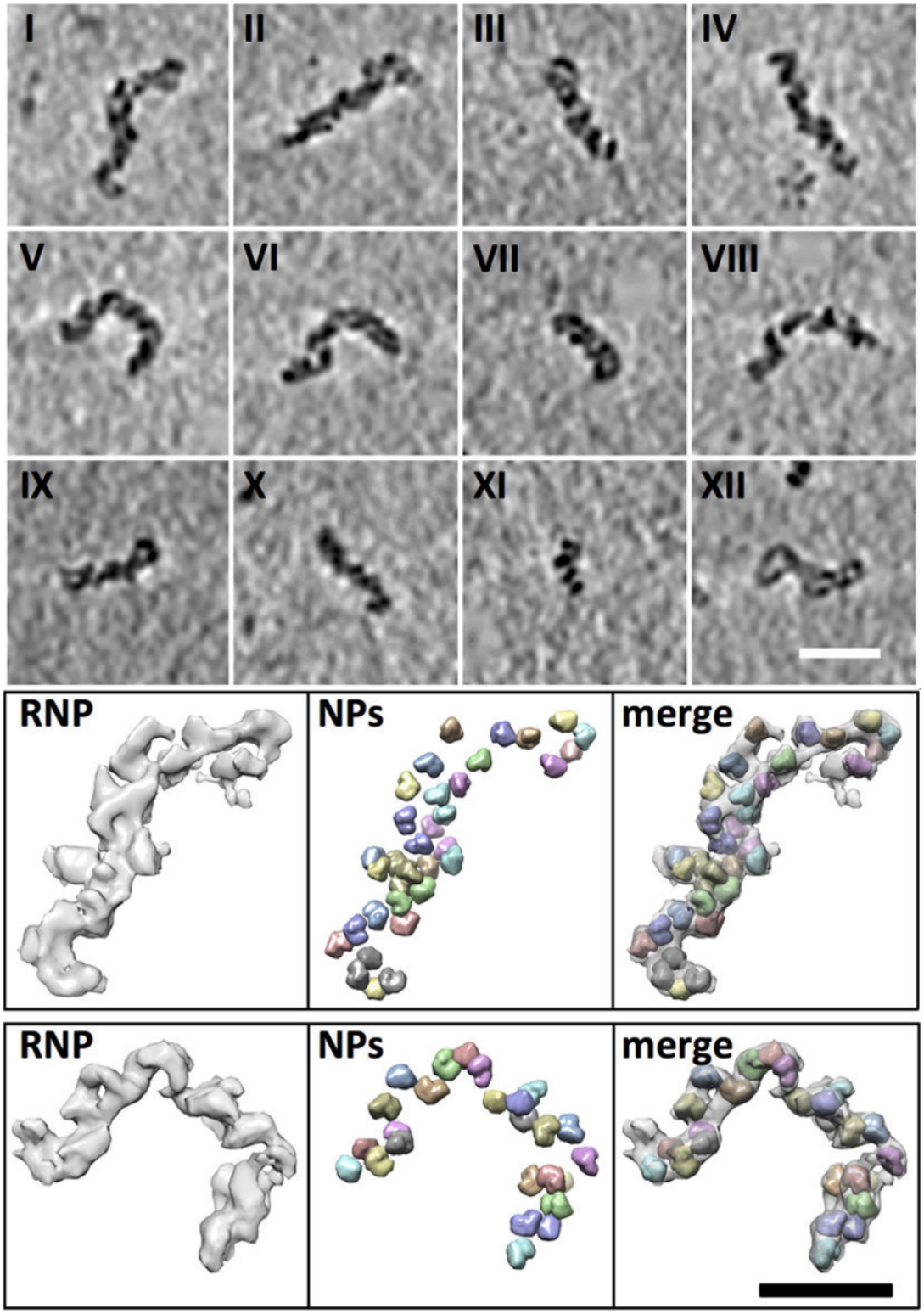


Figure 5.4: vRNP Structural Heterogeneity.

(Top) Gallery (I – XII) of x-y sections through subtomograms of individual vRNP filaments, illustrating their varied morphologies. (Bottom) Calculated molecular envelopes of two vRNPs (RNP) with individual nucleoprotein monomers fitted to the density (NPs) and then overlayed (merge) to illustrate the relative size and occupancy of individual vRNPs. Adapted from Gallagher et al., 2017.

While significant progress on the structure of the influenza vRNP has been made within the last 20 years, with the local resolution of the nucleoprotein in the context of the vRNP now being reported to a nominal resolution of 7 Å (Coloma et al., 2020), there are still gaps in our knowledge. At the outset of this work, these included the high-resolution structure of the influenza nucleoprotein in the context of the vRNP, the molecular mode of binding of the influenza polymerase to the vRNP, and the molecular mode of interaction of the innate immune factor MxA to the vRNP.

5.1.2.1. Previous Laboratory Work on vRNP Structure

Our laboratory has previously attempted to elucidate nucleoprotein organisation within vRNP complexes, and to determine the positioning of the influenza polymerase in the context of the vRNP.

vRNPs were enriched from viral particles using established protocols (Vreede and Brownlee, 2007), described in the present thesis materials and methods chapter (section 2.4.3. Virion-Derived Influenza Virus vRNP Isolation). Three different His_6-tagged influenza polymerase megabodies were added to the vRNP sample to bind any present influenza polymerase. Being His_6-tagged, the megabodies could also be labelled with Ni-NTA gold nanobeads in order to indirectly visualise the polymerase by negative stain electron microscopy (figure 5.5).

Quantitative analysis of negative-stain micrographs revealed clear organisational patterns: 89.9 % of nanogold particles co-localised with vRNP structures (figure 5.5,

panel B), while stoichiometric analysis demonstrated that 85.2 % of vRNPs associated with single nanogold particles, 10.7 % lacked detectable nanogold association, and 4.6 % had two nanogold particles (figure 5.5, panel C). Notably, no vRNP exhibited more than two associated nanogold particles, suggesting that most vRNP complexes have one constituent polymerase associated, and only a small population of vRNPs have two copies of the polymerase associated. Additionally, it was observed that 90.3 % of nanogold particles localised to vRNP termini (figure 5.5, panel D) providing strong evidence for preferential polymerase positioning at ribonucleoprotein ends in this sample, rather than distributed along the helical structure.

Subsequent cryo-EM data were collected without the nanogold particle bound to the megabody to eliminate electron scattering artifacts that compromise high-resolution imaging (figure 5.5). While vRNP filaments remained clearly visible in the resulting dataset, conventional particle picking approaches failed to identify discrete polymerase or megabody-polymerase 2D classes (Zhu, 2024) (figure 5.5, panel F).

This limitation highlights the technical challenges inherent in structural characterisation of polymerase components within the context of the flexible vRNP architecture, necessitating alternative strategies for polymerase localisation and downstream data processing.

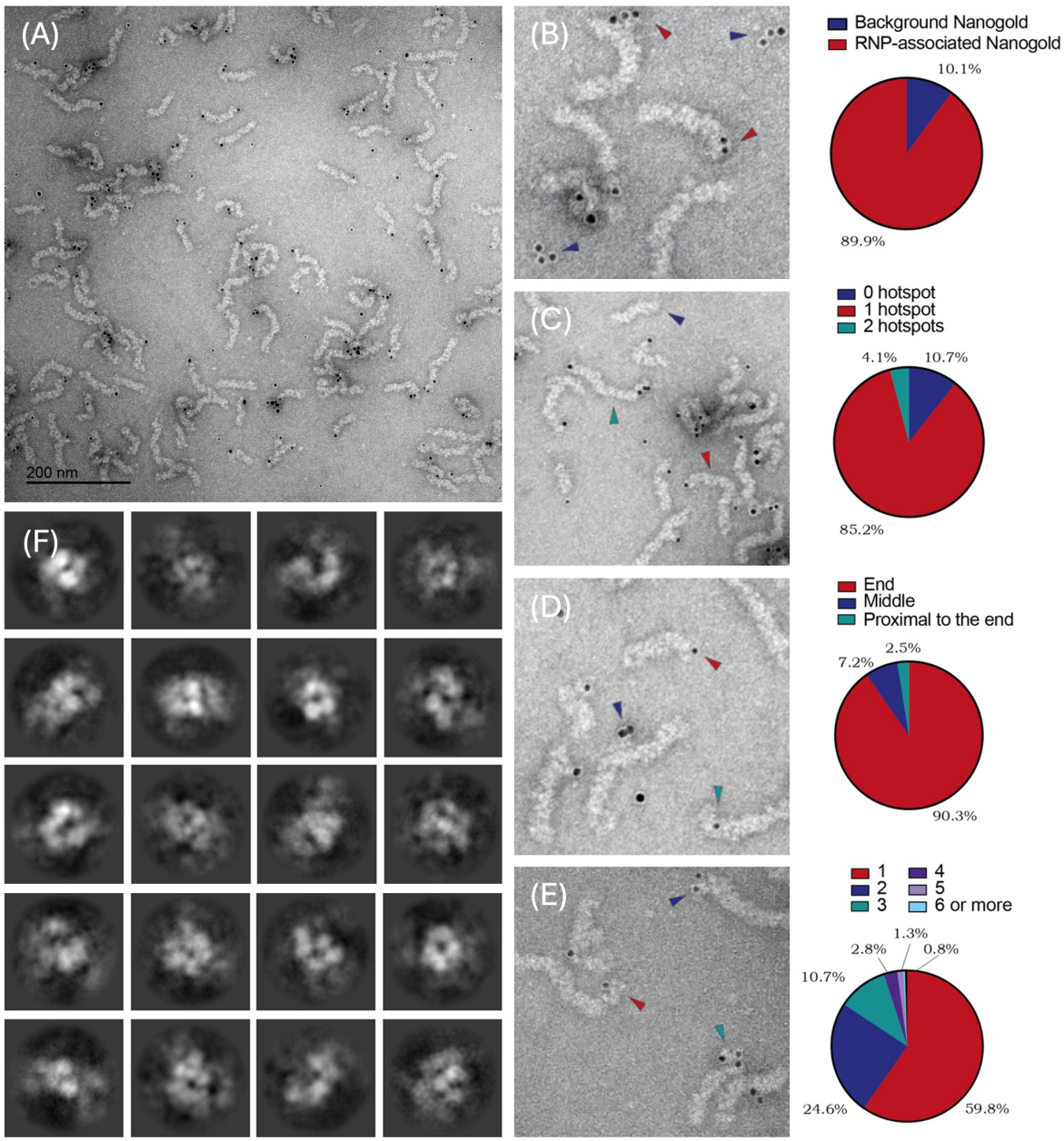


Figure 5.5: vRNP-Megabody Dataset Overview.

(A) Negative stain electron microscopy dataset of influenza A virus-enriched vRNPs with three influenza polymerase megabodies bound, and the megabodies additionally bound with nanogold particles (black dots). (B – E) Representative areas of negative stain electron microscopy micrographs showing vRNPs and nanogold particles, with respective quantification on the right side of each panel. Arrows of different colours represent similarly-coloured areas of pie charts to quantify different characteristics of the data: (B) background nanogold particles vs vRNP-associated nanogold particles (C) number of nanogold particle hotspots on a given vRNP (D) positions of nanogold particles along the vRNP, and (E) quantification of numbers of nanogold particles in a given nanoparticle hotspot. (F) Representative 2D class averages of cryo-EM data collected on this sample. Figure adapted from Zhu, 2024.

5.1.3. Helical Theory

Helical filaments are ubiquitous structural elements across all domains of life. Prokaryotic cells require helical assemblies of actin and tubulin homologues, including FtsA (Lara et al., 2005) and MreB (Busiek and Margolin, 2015) for essential cell cycle processes. Eukaryotic systems extend this organisational principle beyond the canonical cytoskeletal proteins such as actin and tubulin (Kueh and Mitchison, 2009) to encompass diverse protein families including immune system components (Lu et al., 2014) and pathological assemblies such as the misfolded protein aggregates associated with Alzheimer's disease (Lövestam et al., 2024).

Viruses have provided particularly valuable models for understanding helical assembly principles. The first structural data of proteinaceous helical filaments were derived from Tobacco Mosaic Virus by means of X-ray fiber diffraction (Watson, 1954; Franklin, 1955; Franklin, 1956), and the first three-dimensional reconstruction of a helical protein filament by electron microscopy studied a bacteriophage tail (De Rosier and Klug, 1968).

Negative-sense RNA viruses demonstrate sophisticated helical packaging strategies, encapsidating their genomes within ordered nucleocapsid structures (Luo et al., 2020). The degree of helical symmetry varies among different negative-sense RNA viruses, with well-characterised examples including the helical nucleocapsids of Toscana virus (Olal et al., 2014), measles virus (Gutsche et al., 2015), and Hantaan virus (Arragain et al., 2019).

The intrinsic symmetry of these assemblies provides significant advantages for cryo-EM SPA and cryo-ET STA by enhancing signal-to-noise ratios and enabling the imposition of structural helical constraints during 3D reconstruction, contingent upon the accurate determination of helical parameters, which is often extremely challenging.

5.1.3.1. Helical Symmetry Parameters

The mathematical description of biological helical filaments defines the geometric relationship between adjacent subunits that collectively determine the three-dimensional architecture of the complete assembly. Helically symmetric filaments comprise physical subunits (protein monomers) related through rotation and translation along a helical or screw axis (figure 5.6, denoted as z). These subunits can be conceptualised as points distributed on a cylindrical surface with their spatial relationships defined by two fundamental parameters: helical twist, $\Delta\phi$, and helical rise, Δz, between subunits (figure 5.6).

Twist ($\Delta\phi$) (˚) represents the angular rotation (in degrees) between adjacent subunits around the helical axis while rise per subunit (Δz) (Å) describes the axial translation (in Å) between consecutive subunits along the helical axis. Together, these two parameters define the local geometry of subunit – subunit interactions within the helical filament.

Global helical parameters provide complementary architectural descriptors. Pitch (*P*) represents the axial distance required for one complete 360˚ rotation around the helical axis. The number of subunits per turn (*n*) defines the number of individual subunits comprising one complete helical turn. Handedness (*h*) describes the fundamental topology of the helical filament. This property is determined by the direction of rotation when moving along the helical axis: right-handed (+1) helices exhibit clockwise rotation, while left-handed (-1) helices exhibit anti-clockwise rotation. Together, pitch, number of subunits per turn, and handedness describe the global parameters of the helical filament.

Additionally, the radius (*r*) defines the distance from the helical axis to the center of mass of each subunit, determining the overall diameter of the helical filament.

These local and global parameters are mathematically related by the following equations:

$$P = \frac{360^{\circ}}{\Delta\varphi}\ x\ \Delta z$$

The pitch provides a measure of helical compactness relative to twist magnitude.

From *n* and *h* values, $\Delta\phi$ can be calculated as:

$$\Delta\varphi = \frac{360^{\circ}\ h}{n}$$

From *P* and *n* values, *Δz* can be calculated as:

$$\Delta z = \frac{P}{n}$$

And similarly, from *Δz* and *Δϕ*, *P*, *n*, and handedness can be calculated as:

$$n = \frac{360^{\circ}}{|\Delta\varphi|}; P = \frac{360^{\circ}\Delta z}{|\Delta\varphi|}; h = sign(\Delta\varphi)$$

Both parameter sets – (*Δϕ*, *Δz*) and (*P, n, h*) – serve distinct computational purposes in helical analysis. Flexible helical assemblies often show ambiguity in the number of subunits per turn while maintaining consistent pitch values, making the (*P*, *n*, *h*) set more suitable for exploring symmetry parameters. Conversely, the (*Δϕ*, *Δz*) set more directly represents the information needed for symmetry imposition during refinement. CryoSPARC utilises both parameter sets at different stages of helical processing (Wang et al., 2022) (Helical symmetry in CryoSPARC).

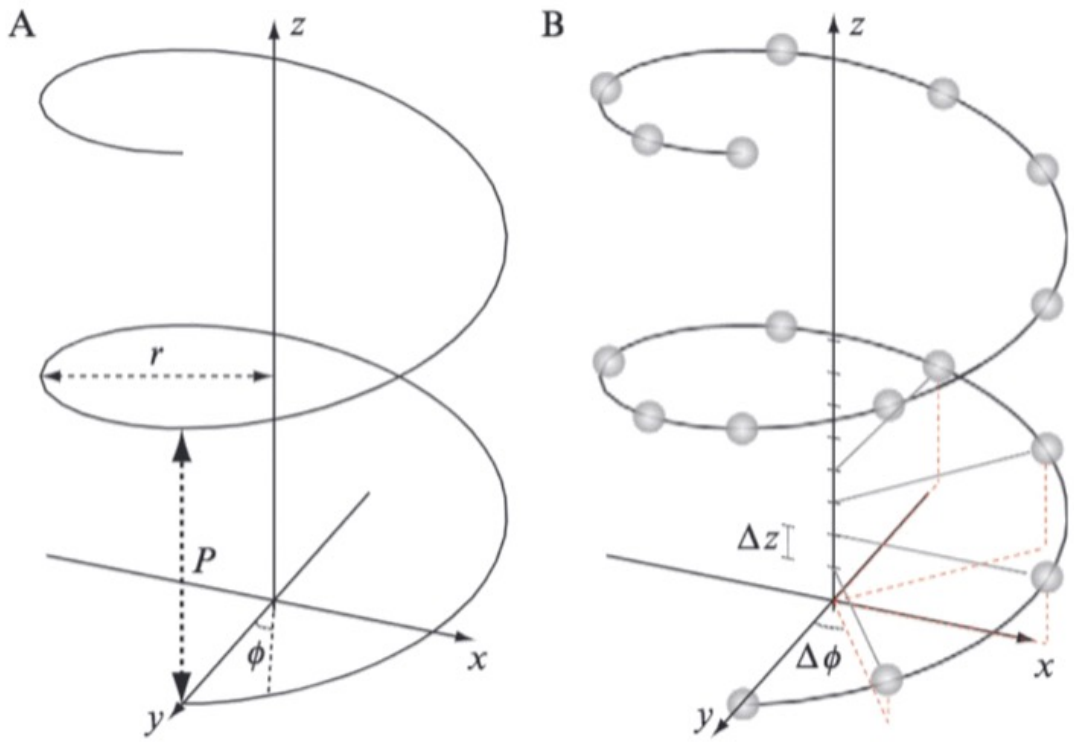


Figure 5.6: Basic Helical Symmetry Parameters.

(A) A continuous helix with ϕ, twist, r, radius, and P, pitch annotated. (B) A discrete helical lattice with Δϕ, the twist between two subunits (coloured as grey spheres), and Δz, the helical rise between two subunits annotated. Figure taken from Bubeck lecture 'Making the most of symmetry' 2019 (http://www.cryst.bbk.ac.uk/embo2019/ppts/EMBO_2019.pdf).

Helical structural classification requires additional consideration of strand multiplicity – specifically, whether assemblies are comprised of single- or multi-start helical architecture. The number of helical starts determines whether the helical lattice follows one continuous helical path or consists of multiple independent, intertwined helical strands. Single-start helices trace one continuous path around the helical axis, while multi-start helices incorporate multiple discrete helical threads that wind together to form the complete filament.

Influenza vRNPs are an example of a 2-start, antiparallel helical filament, that enables efficient packaging of complete genome segments while accommodating a single

polymerase heterotrimer at one end, and a non-helical nucleoprotein loop at the opposite end. This 2-start organisation confers enhanced structural stability through inter-strand interactions while preserving the conformational flexibility required for transcription and replication.

The antiparallel double helical arrangement generates distinct major and minor grooves, which is crucial for the understanding of both structural organisation and biological processes carried out on the vRNP. These grooves likely provide specific binding sites for regulatory factors and may facilitate the directional movement of the polymerase machinery during RNA synthesis, making their structural characterisation crucial for understanding viral replication mechanisms.

5.1.3.2. Experimental Approaches to Helical Parameter Determination

Cryo-electron microscopy has revolutionised helical parameter determination by preserving helical assemblies in their native, hydrated states within vitreous ice. This methodology offers particular advantages for helical filaments, which often resist crystallisation due to their extended, flexible architecture. Cryo-EM enables examination of physiologically relevant conformational states that remain inaccessible through conventional crystallographic methods.

The image processing pipeline for helical structures in cryoSPARC begins with identification of helical segments from micrographs through manual and/or

automated particle picking, combined with filament tracing. Short helical segments are subsequently extracted at regular intervals along filament length to generate particle sets suitable for helical reconstruction.

Box size selection for segment extraction represents a critical optimisation parameter. Larger box sizes capture extended information spanning multiple helical turns, thereby providing better constraints for accurate helical parameter determination. However, the inherent flexibility and polymorphism of biological filaments means that longer segments may exhibit greater conformational heterogeneity (Wang et al., 2022), particularly at the box edges, potentially degrading the quality of averaged structures. The choice of an optimal box size therefore requires a balance between sufficient helical information content against conformational homogeneity within individual segments.

CryoSPARC enables filament selection based on curvature and sinuosity metrics. Filament curvature represents a local geometric property that varies along a given filament that is defined as the degree of local bending of the filament at each point across its length, measuring the deviation of the filament from a straight line at a specific location. Conversely, filament sinuosity is a global measure of overall deviation from linearity, calculated as the ratio of actual path length to straight-line distance between filament end points.

For high-resolution structure determination, filament selection should favour segments with low curvature (< 0.01 Å^{-1}) and low sinuosity (< 1.5). Resulting straighter

filaments minimise conformational heterogeneity and provide more reliable helical parameter estimates, while highly curved or sinuous filaments can introduce noise and result in lower resolution 2D class averages and 3D reconstructions. These selection criteria are essential for achieving the structural homogeneity required for high-resolution helical reconstruction.

5.1.3.2.1. Two-Dimensional Classification and Initial Helical Parameter Estimation

Following filament tracing and segment extraction, 2D classification separates particles into homogeneous classes, each potentially representing distinct helical conformations or polymorphic states. Reliable parameter estimation requires classes containing sufficient helical content (typically > 2 helical turns) and high structural homogeneity within each class.

For each 2D class, segments undergo verticalisation (alignment along the helical axis) prior to power spectra calculation through Fourier transformation. The resulting power spectra exhibit layer lines with observed helical parameters in 2D class averages (Egelman, 2015; Kreutzberger et al., 2024). Well-defined layer lines enable direct measurement of the helical pitch from the meridional spacing, while Bessel order analysis provides information about the radial mass distribution. Computational tools such as Helixplorer (https://rico.ibs.fr/helixplorer/helixplorer/) facilitate matching between layer line positions and helical parameters for use in 3D refinement.

When power spectra lack sufficient signal due to small box sizes or heterogeneity within 2D classes, parameters can instead be estimated from visual inspection of 2D class averages. Helical pitch is often able to be measured directly from the repeat pattern, while the number of subunits per turn can be counted to estimate twist angles. Multiple candidate parameter sets can then be evaluated through preliminary 3D reconstructions to further refine optimal helical parameters.

Alternative approaches become necessary when power spectra lack interpretable signal. A systematic approach involves measuring visible helical features directly from 2D class averages, estimating subunit periodicity, and testing multiple candidate parameter sets in initial 3D reconstruction. Alternatively, an iterative refinement strategy can be employed, where parameters are systematically adjusted until a satisfactory 3D density is achieved. It is critical that high-resolution features, such as secondary structure elements and side chains emerge in order to be certain that the applied symmetry is correct, since applying different helical symmetries can drastically alter the outcome of 3D refinement jobs.

5.1.3.2.2. Three-Dimensional Reconstruction and Parameter Refinement

Once initial helical parameters have been estimated, 3D reconstruction of the data can be attempted through multiple strategies. Non-symmetrised *ab-initio* reconstruction allows parameter extraction through density map inspection, while direct helical reconstruction can be initiated using estimated helical parameters.

CryoSPARC implements advanced helical reconstruction algorithms capable of simultaneously refining both the helical parameters and the 3D structure. This process typically involves a search across a range of potential twist and rise values with the optimal parameters identified based on reconstruction consistency and resolution metrics. The software can accommodate both global parameter refinement – applying uniform parameters across the entire dataset – and local parameter refinement, which detects and accommodates helical polymorphism or conformational flexibility.

5.1.4. Carbon Support Cryo-Electron Microscopy Grids

Protein samples prepared for cryo-EM can suffer from detrimental interactions with the air-water interface, which can lead to preferential orientation in the vitreous ice, denaturation, and aggregation. To mitigate these effects, continuous thin carbon support films (Passmore and Russo, 2016; Thompson et al., 2016, p.201; Drulyte et al., 2018) and graphene-based support films (Pantelic et al., 2010; Palovcak et al., 2018) can be used to minimise protein-interface interactions.

These support films offer several advantages: they can improve protein distribution and reduce preferential orientation by providing a uniform surface for particle adsorption, enable better control of particle concentration across grid holes by facilitating interactions with the support film, and help maintain protein native state during the freezing process (Glaeser, 2018).

However, the use of support films introduces a trade-off, as the additional material layer increases background signal in micrographs due to electron scattering, which reduces image contrast. The choice to use support films must therefore be balanced between the benefits of improved particle distribution against the costs of increased background signal.

5.1.5. Functionalised Cryo-Electron Microscopy Grids

Functionalisation of electron microscopy grids involves the addition of chemical groups that interact selectively with tagged proteins (e.g. Ni-NTA with His_6-tagged proteins, Strep-Tactin with Strep-tag II-tagged proteins). This approach enables concentration and immobilisation of specifically tagged protein complexes directly onto the grid surface.

Functionalised grids serve dual experimental purposes: they can function as a direct affinity purification method, whereby proteins are captured from cell lysate and non-specifically bound material is washed away, or as a final concentration step to increase particle density for samples that resist conventional concentration methods.

This technique has been successfully applied to increase sample concentration on electron microscopy grids, enabling high-resolution structure determination by cryo-EM SPA for samples that would otherwise be too dilute for effective data collection (Kelly et al., 2010; Yu et al., 2016; Benjamin et al., 2016; Ramlaul et al., 2023).

An important consideration when using functionalised grids is the optimisation of protein linker length connecting the target protein to its affinity tag. The system used in this chapter is illustrated below (figure 5.7). Short protein linkers between the nanobody and its His_6 tag would constrain vRNP complexes to limited orientations relative to the grid surface, while extended linkers provide greater conformational flexibility.

The linker length directly impacts both the diversity of particle views available for 3D reconstruction and the potential for preferred orientation artifacts – a common limitation in cryo-EM SPA. Optimal linker design must therefore balance sufficient conformational freedom to ensure comprehensive orientational sampling against the risk of introducing flexibility that could compromise structural homogeneity.

The strategic implementation of functionalised grids represents a powerful approach for overcoming challenges of sample preparation inherent in structural analysis of large, flexible complexes such as viral ribonucleoproteins.

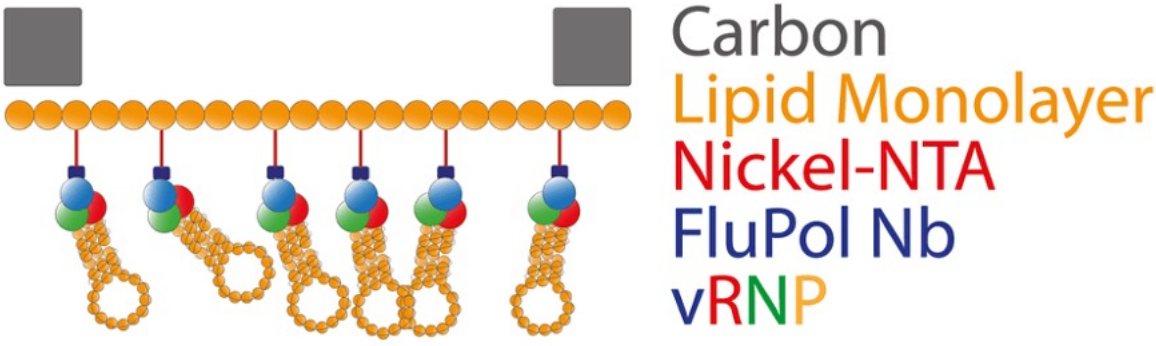


Figure 5.7: Diagram of Functionalised Electron Microscopy Grid.

In the system illustrated, the grid hole is illustrated in grey and annotated 'Carbon'. There is a lipid monolayer grown across the grid, which is charged with Nickel-NTA for the immobilisation of His_6-tagged

proteins. In this system, an influenza polymerase nanobody (FluPol Nb, blue) is immobilised to the Ni-NTA layer, and influenza vRNPs are immobilised in different orientations.

5.1.6. Chapter Aims

Prior to the commencement of this work, several fundamental knowledge gaps existed in our understanding of influenza vRNP structural biology. First, no high-resolution structure of the nucleoprotein in its physiologically relevant RNA-bound state within the context of the vRNP had been determined, constraining our understanding of viral genome packaging and organisation mechanisms.

Second, while previous work indicated that individual vRNPs typically associate with single polymerase complexes, the structural basis underlying polymerase – vRNP interactions remained uncharacterised.

Finally, no structural or biochemical data existed describing the molecular mechanism by which the innate immune factor MxA (introduced in chapters 1 and 4) associates with influenza A nucleoprotein within the context of intact vRNPs.

My first aim was to identify and structurally characterise the influenza polymerase complex within the context of the vRNP. I planned to use convolutional neural network-based particle picking software followed by conventional data processing pipelines in cryoSPARC to reliably pick vRNP ends, and classify out the loop end, selecting for and further processing the influenza polymerase end of the vRNP. This approach would enable classification and isolation of loop-containing termini from polymerase-

associated ends, facilitating focussed analysis on polymerase-vRNP association mechanisms.

My second aim was to determine the high-resolution structure of the influenza A nucleoprotein within virion-derived vRNPs. This required the development and implementation of novel cryo-EM approaches to overcome inherent challenges posed by vRNP flexibility and low on-grid concentrations.

I employed three main strategies: carbon support grids to adsorb vRNPs and reduce filament flexibility, nanobody-mediated rigidification using a specific nucleoprotein-binding nanobody, and functionalised cryo-EM grids as an on-grid concentration method to enrich vRNPs with bound interaction partners.

My final aim was to establish an experimental system to study the molecular mechanism of MxA binding to influenza virus vRNPs. This involved the generation of a chimeric influenza A virus sensitive to MxA restriction. I then planned to demonstrate a direct binding interaction between vRNPs and MxA through biochemical analysis. Finally, I aimed to determine the mode of MxA binding to vRNPs using integrated cryo-electron microscopy and cryo-electron tomography approaches.

These objectives collectively address critical gaps in influenza vRNP structural biology while developing methodological advances applicable to other flexible ribonucleoprotein complexes.

5.2. Results

5.2.1. Enrichment of vRNPs from Live Influenza Virus

vRNPs utilised in this chapter were isolated from live WSN virus (Influenza A/WSN/1933 (H1N1)) following a method established in literature (Vreede and Brownlee, 2007), and as in this thesis materials and methods chapter (section 2.4.3. Virion-Derived Influenza Virus vRNP Isolation) (figure 5.8). This method is well-established and has been shown to reliably isolate vRNPs produced by viral infection. This process generates a mixture of all eight vRNP segments, which are difficult to further separate into individual segment lengths. As a result, the vRNP fractions are pooled, introducing an additional layer of sample heterogeneity, namely the variation in segment length. However, this variability was considered surmountable given the available computational resources for both 2D and 3D classification.

Briefly, six T-175 flasks of MDBK cells were infected with influenza virus at a multiplicity of infection (MOI) of 0.01 and incubated for 48 h. Following this incubation period, the supernatant was harvested, and cell debris removed. The viral particles were then concentrated by means of ultracentrifugation through a sucrose cushion. The resulting pelleted virus was sealed and resuspended overnight in chilled resuspension buffer (10 mM Tris-HCl (pH 7.4), 100 mM NaCl, 1 mM EDTA). The viral suspension was then transferred to a microfuge tube and viral particles lysed using a disruption buffer including 2 % Triton X-100. The suspension was incubated at 31˚C in a shaking incubator for 30 minutes to ensure complete virion lysis. In some cases, recombinantly-produced vRNP interaction partners (Nb170, MxA) were added in

excess to the lysate after the virion lysis step. The resulting lysate was subsequently separated by means of ultracentrifugation through a glycerol gradient. After the separation, a hole was pierced in the base of the ultracentrifuge tube and 250 µL fractions of separated virion lysate were collected. They were vortexed and samples taken for analysis by SDS-PAGE, silver staining, and western blot as appropriate. Fractions containing isolated vRNPs were identified and stored.

In cryo-EM analysis, it was observed that vRNPs obtained directly from the vRNP separation step were insufficiently concentrated to visualise a sufficient number of particles in individual cryo-EM grid holes. Consequently, an additional ultracentrifugation step was necessary to increase the concentration of vRNPs for cryo-EM analysis. Following this step, the vRNPs were pelleted and resuspended overnight at 4˚C in 25 µL of resuspension buffer (20 mM Tris-HCl (pH 8.0), 150 mM NaCl).

All vRNP enrichment procedures were carried out by Dr. Kuang-Yu Chen.

The protocol for the enrichment of virus-derived vRNPs is illustrated in the figure below details the virus-derived vRNP enrichment protocol (figure 5.8).

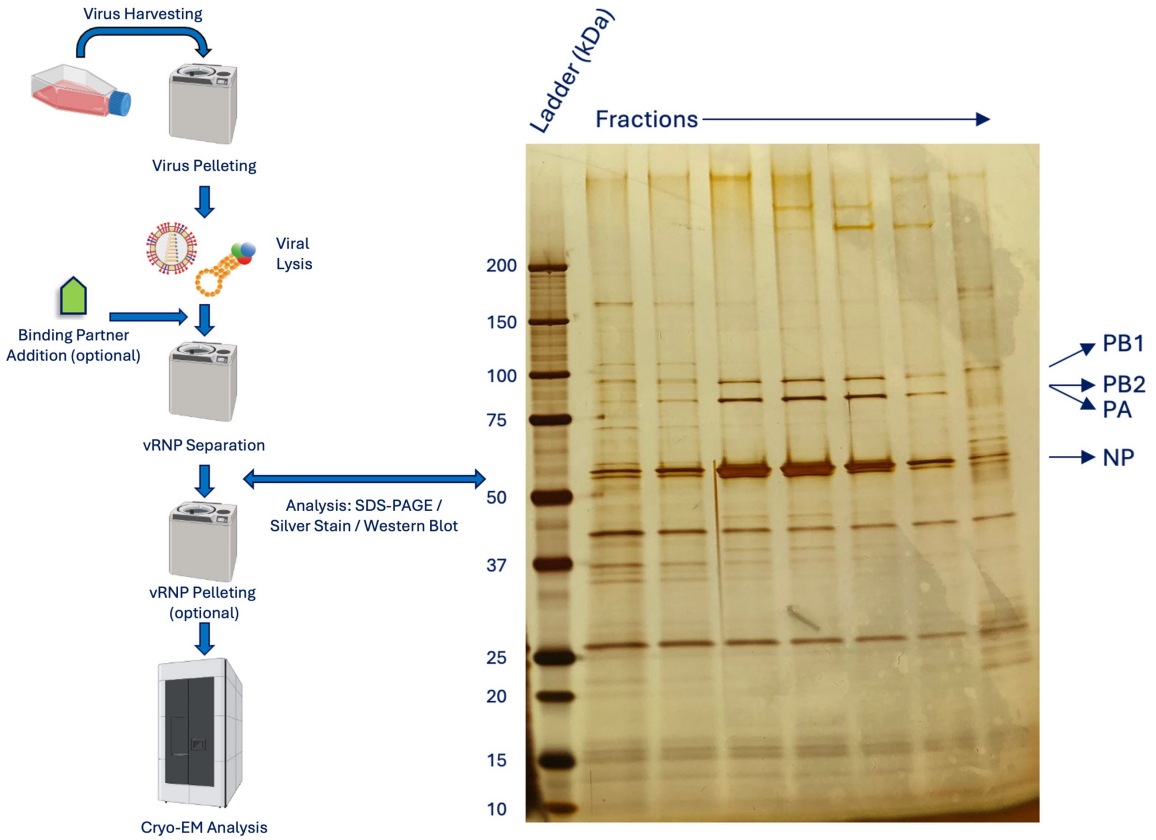


Figure 5.8: vRNP Enrichment Protocol.

(Left) Cartoon diagram of steps in the vRNP enrichment protocol described. Virus harvesting occurs 48 hpi. Viral particles are pelleted by means of ultracentrifugation. Virus particles are lysed, and binding partners are added in excess (optional). vRNPs are separated from the viral lysate and analysed by means of SDS-PAGE, silver stain, or western blot, as appropriate. An optional vRNP pelleting step can be carried out if higher concentrations of vRNPs are required. This is followed by analysis by cryo-EM. (Right) After the vRNP separation step, the fractions isolated are analysed by means of SDS-PAGE, silver stain, or western blot, as appropriate. Here, a standard silver stain gel of separated viral lysate fractions is shown. Fractions are taken from the base of the ultracentrifuge tube, and earlier fractions relate to pelleted and aggregated material, and heavier protein complexes. As fractions are consecutively collected, the molecular weight of the protein complexes observed decreases. The vRNP fractions are boxed in blue and are characterised by a thick protein band of 56 kDa corresponding to the nucleoprotein, a double band which appears to run as one band of corresponding to PB2 (approximately 85 kDa) and PA (approximately 85.5 kDa), and one band which runs at a higher observed molecular weight corresponding to PB1 (approximately 86.5 kDa).

The silver stain gel presented demonstrates the ability of vRNPs to be successfully enriched form virus. The enriched sample can then be analysed by cryo-EM for structure determination or used for biochemical analysis (Zhu et al., 2023).

5.2.2. Automated Particle Picking of Megabody-Bound Influenza Polymerase in the Context of Influenza vRNPs using crYOLO

Traditional particle picking methods, whether manual or template based, are time consuming and prone to inherent human biases. Manual picking can introduce operator bias, as individuals may subconsciously select particles based on preferred orientations, higher contrast, or specific morphological features, while potentially overlooking particles in less favourable conditions or orientations. Such biases can lead to incomplete angular sampling during reconstruction, limiting the achievable resolution and map quality while also potentially introducing artifacts in the final structure.

Template-based particle picking approaches, although more objective than manual picking, still present several limitations. It requires *a priori* knowledge of particle appearance and may fail to detect particles that deviate from the template provided. Additionally, these methods often rely on initial manual particle selections to generate templates, which can propagate human bias throughout the picking process.

To overcome these limitations, crYOLO (cryo-electron microscopy You Only Look Once) has been implemented to use a deep learning approach based on convolutional neural networks (CNNs) for automated particle detection and picking (Wagner et al., 2019; Wagner and Raunser, 2020). Briefly, CNNs learn to identify hierarchical features in cryo-EM micrographs. During training, the network analyses annotated micrographs

to distinguish patterns characteristic of particles, background noise, ice contamination, and other artefacts present in cryo-EM data.

In this study, crYOLO was applied to a cryo-EM SPA dataset collected previously, introduced in chapter 5.1.2.1. (Zhu, 2024) where influenza A vRNPs were bound by three different influenza polymerase-binding megabodies. Conventional particle picking methods had been unsuccessful in identifying megabody- or influenza polymerase-containing 2D classes, despite picking strategies targeting solely the ends of vRNPs. Given that negative stain data that demonstrated the prevalence of influenza polymerase complexes were predominantly located at the ends of vRNPs in this sample (as shown in figure 5.5), crYOLO was specifically trained to pick the ends of vRNPs. Importantly, the loop ends of vRNPs were intentionally excluded from the picking process, as the polymerase complex was mainly observed at the end opposite the loop. Where it was ambiguous which end contained the loop, and which was thought to contain the polymerase, both ends were picked.

A box size of 405 Å x 405 Å (1.35 Å/px, 300 px x 300 px) was selected to ensure sufficient coverage to identify both the influenza polymerase and any potentially bound megabodies, as well as adjacent nucleoprotein molecules. The polymerase has a maximum dimension of approximately 130 Å (Keown et al., 2022) (PDB: 7NK4), while a bound megabody could add up to a theoretical maximum of 100 Å if bound on the edge of the longest dimension of the polymerase (Uchański et al., 2021) (PDB: 6XUX). To ensure adequate padding around the particle, and for signal from nucleoprotein molecules interacting with the polymerase, the box size was chosen to provide

approximately 1.5 x the expected maximum particle dimension. This padding prevents truncation of polymerase signal during particle extraction and allows for particle shifts during alignment without losing signal at the box edges.

The particle picks generated by crYOLO were imported into cryoSPARC, where particles were extracted from their original micrographs. Across a dataset of 7,337 exposures, this led to 207,830 particles. These particles were subjected to two rounds of 2D classification. In both rounds, the resulting 2D class averages were assessed for signal which could be attributed to megabody and polymerase.

An example micrograph with crYOLO particle picks is shown in figure 5.9, panel A, where the picks are marked by white circles. These picks were globally good, being in the majority of cases solely at the ends of vRNPs. Figure 5.9, panel B presents an example 2D class selection job, where clear nucleoprotein was evident across all classes, with the nucleoprotein loop clearly visible in some classes. However, no signal attributable to either megabody or polymerase was observed across the dataset (figure 5.9, panel B).

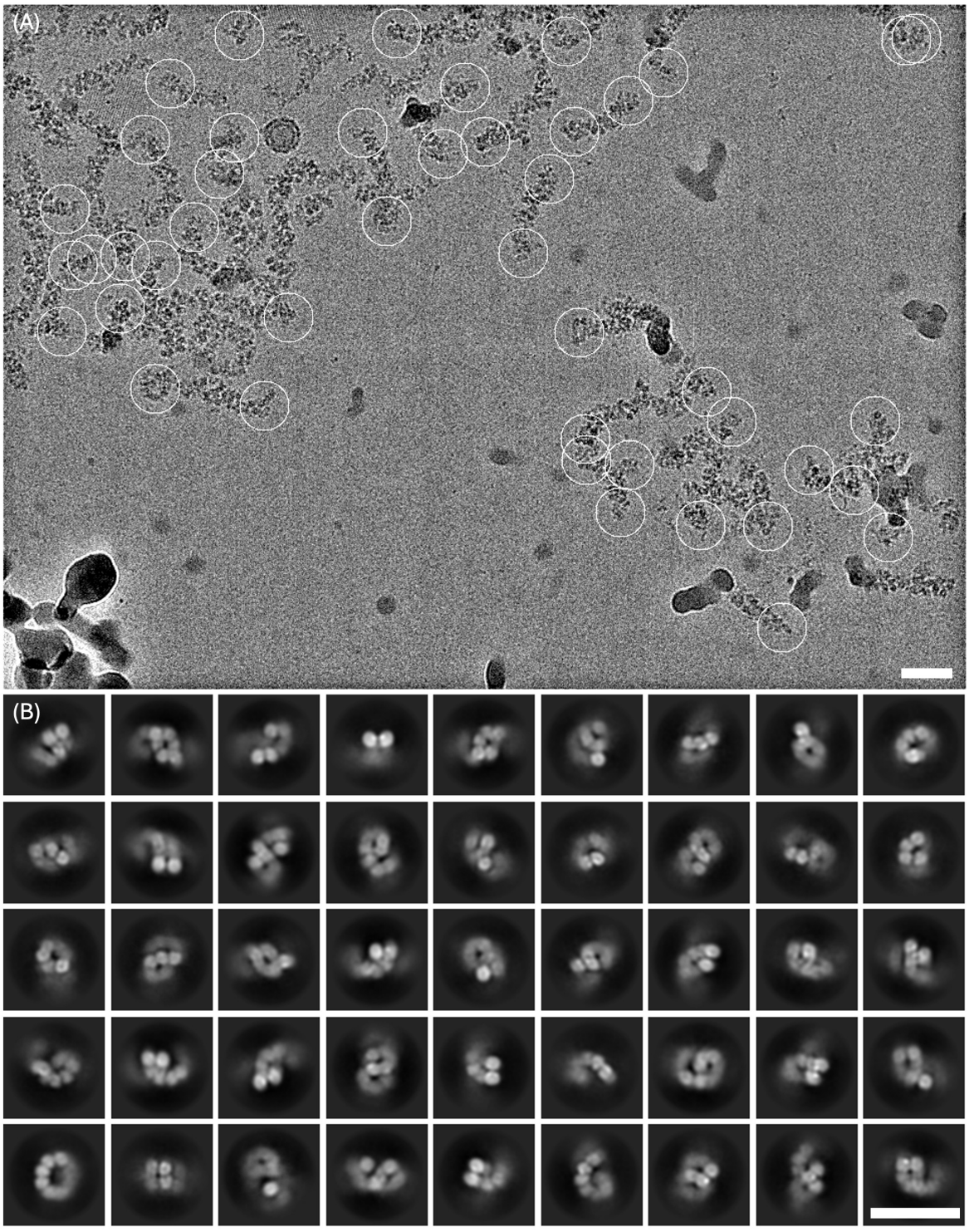


Figure 5.9: crYOLO Picking of vRNP Ends.

crYOLO was employed to pick the ends of vRNPs to obtain a particle stack of influenza polymerase particles. (A) Micrograph with locations of crYOLO picks circled in white. (B) Examples of 2D classes from the above particle selection. In both panels, the scale bar corresponds to 400 Å.

Despite the globally successful identification of influenza vRNP ends by crYOLO, the absence of detectable polymerase or megabody suggests that the picked particles contained few, if any, polymerase complexes.

Several factors may account for this observation. First, bias during the training process of crYOLO could have resulted in the preferential selection of particles that did not contain polymerase complexes. Alternatively, erroneous picking of helical nucleoprotein structures, which may have appeared similar to the desired vRNP ends, could have contributed to the lack of polymerase detection. Another possibility is insufficient 2D classification, which may not have been thorough enough to extract clear polymerase and megabody signal.

The absence of megabody signal further supports the hypothesis of limited polymerase presence but also suggests that the megabody complexes may either be too flexible or have been prone to disassembly during the vitrification process. This observation is consistent with findings from other studies (Uchański et al., 2021; Ackle et al., 2025) where megabody complexes are not observed in final reconstructions due to flexibility between the nanobody domain and the scaffold domain (of HopQ-based megabodies, but this observation could be applicable to YgjK megabodies also). These reports highlight the challenges associated with maintaining the structural integrity of megabody complexes during cryo-EM data collection.

5.2.3. Carbon Support Cryo-EM Grids for the Reduction of Influenza vRNP Flexibility

Carbon support grids have been demonstrated to mitigate several challenges associated with conventional holey carbon grids in cryo-EM sample preparation. These include the prevention of unfavourable protein – air – water interface interactions, reduction of preferential particle orientations, and minimisation of protein denaturation and aggregation during vitrification (Passmore and Russo, 2016; Thompson et al., 2016; Drulyte et al., 2018). The continuous carbon support film provides a uniform adsorption surface that potentially improves particle distribution within grid holes while simultaneously minimising complex flexibility through interactions with the carbon support layer.

Given the inherent global flexibility of influenza vRNPs, it was hypothesised that the use of carbon support grids would enhance sample preparation for cryo-EM SPA. The carbon support layer was expected to improve particle concentration within the grid holes and reduce global vRNP flexibility through interaction with the carbon support film. This conformational restraint could facilitate downstream image processing and potentially enable higher resolution reconstructions.

Briefly, vRNPs were isolated as previously described (chapter 5.2.1.). The observed concentration by nanodrop was 0.6 mg/mL. 2.5 µL of concentrated vRNPs were applied to freshly glow-discharged Quantifoil R 2/1 Cu 200 + 2 nm C grids, blotted, and plunge frozen using a Vitrobot mark IV at 100 % humidity.

The vitrified grids were screened on a ThermoFisher Glacios Transmission Electron Microscope, and grids with a suitable ice thickness and good particle distribution were stored under liquid nitrogen for later data collection, if required. An example micrograph of vRNPs on such a carbon support grid is shown below (figure 5.10).

Comparative analysis of vRNPs on conventional holey carbon grids versus carbon support grids reveals distinct imaging characteristics (figure 5.10, panels A and B). The most significant difference is the reduced image contrast observed with carbon support grids, attributable to increased background signal from the additional 2 nm carbon layer – an inherent limitation that must be considered in data quality evaluation.

Qualitative assessment suggests potential improvement in vRNP concentration within grid holes on carbon support surfaces. However, definitive conclusions regarding the efficacy of carbon support grids in either enhancing particle concentration or reducing vRNP flexibility cannot be drawn from examination of individual micrographs alone. Rigorous comparison would require comprehensive dataset collection from both grid types, followed by quantitative analysis of particle density and direct comparison of reconstruction quality.

This investigation demonstrated the technical amenability of the use of carbon support cryo-EM grids for influenza virus vRNP sample preparation. The vitrification protocol successfully produced grids with adequate ice quality and particle number suitable for data collection. While initial observations suggest potential benefits in

terms of particle distribution, comprehensive evaluation should be carried out in further studies to determine the true benefit of the use of carbon support grids in this capacity. During the course of this DPhil, these grids were kept in cryo-storage as a contingency project should alternative approaches attempted not yield adequate data .

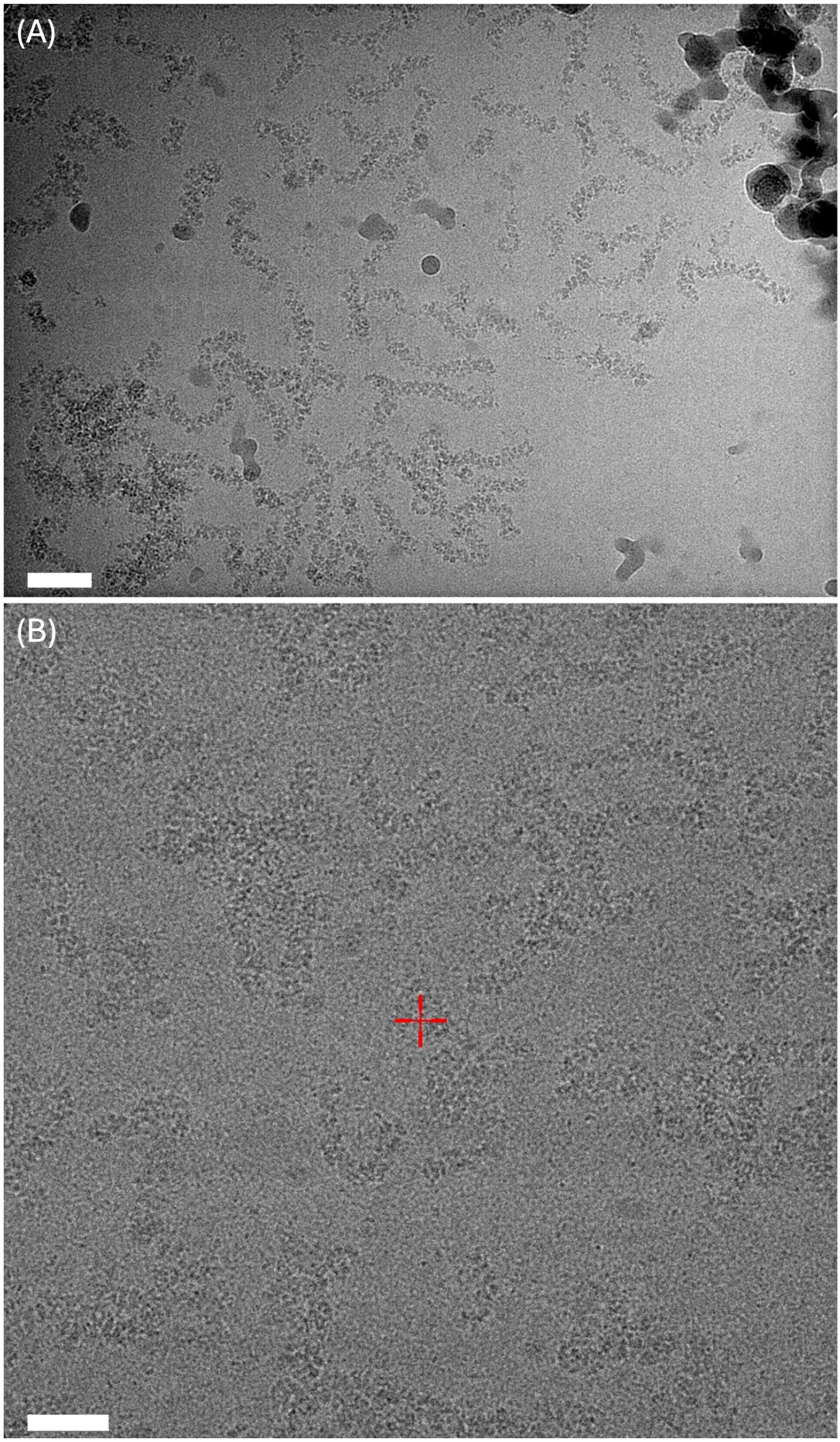

(A)
(B)

Figure 5.10: Carbon Support Grid Comparison.

Cryo-EM micrographs of vRNPs on a holey carbon copper grid (A) compared to a holey carbon copper grid with a 2 nm layer of ultra-thin carbon (B). Scale bars correspond to approximately 50 nm.

5.2.4. Nucleoprotein Nanobody-Mediated vRNP Rigidification

As described in section 3.2.4., Nb170 is a camelid-derived nanobody that exhibits the unique property of binding to the influenza nucleoprotein without interfering with viral replication or transcription (Schmidt et al., 2016). This compatibility with native viral processes makes Nb170 an ideal binding partner for structural studies requiring preservation of authentic vRNP architecture.

The rationale for the use of Nb170 in structural studies of the vRNP centers on its potential to reduce local conformational flexibility of nucleoprotein monomers within the ribonucleoprotein complex while still maintaining native vRNP architecture. Influenza nucleoprotein exhibits inherent flexibility, particularly on surface-exposed loop regions, which limits high-resolution structural determination by cryo-EM SPA. Through stable nucleoprotein-nanobody complex formation, Nb170 binding was hypothesised to constrain local nucleoprotein dynamic, thereby reducing structural heterogeneity.

Moreover, the 15 kDa nanobody addition to the vRNP complex was expected to increase the overall signal of the nucleoprotein within the complex and facilitate more accurate computational alignment of nucleoprotein monomers in the vRNP.

Furthermore, it was hypothesised that the binding of Nb170 to the nucleoprotein across the entire vRNP would constrain global vRNP flexibility, reducing the need for extensive conformational sorting and enabling more particles to contribute to a single high-resolution reconstruction. This global rigidification would minimise the conformational heterogeneity that typically requires extensive 3D classification at the cost of particle numbers.

5.2.4.1. Sample Preparation

vRNPs were enriched following the protocol described in section 5.2.1., with modification at the binding partner addition step to incorporate the Nb170 nanobody (figure 5.8). To ensure complete coating of the vRNP, the required nanobody concentration was calculated based on estimated nucleoprotein content within the vRNP preparation.

Nucleoprotein concentration was estimated using the average influenza vRNA segment length of 1613 nucleotides (chapter 1, table 1.3). Given that nucleoprotein binds approximately every 24 nucleotides, each vRNP contains an average 67 nucleoprotein molecules. With a measured vRNP concentration of 0.6 mg/mL and an estimated average molecular weight of 4224 kDa (given 67 nucleoprotein molecules of 56 kDa, one copy of the 270 kDa heterotrimeric influenza polymerase complex, and around 170 kDa of RNA), the molar concentration of vRNPs was calculated as 0.14 µM, corresponding to a nucleoprotein concentration of 9.4 µM.

Purified Nb170 (purification detailed in section 3.2.4.1.) was added to the viral lysate at 1 mg/mL final concentration (67 μM) representing a 7-fold molar excess over estimated nucleoprotein concentration given the assumptions outlined above. The viral lysate and nanobody mixture were allowed to incubate on ice for 30 minutes prior to vRNP fraction separation by ultracentrifugation to ensure complete nanobody-nucleoprotein binding.

Fractions were analysed by silver stain gel to assess nanobody co-enrichment with vRNP complexes (figure 5.11). Input controls containing untreated vRNPs and purified Nb170 were included for band identification.

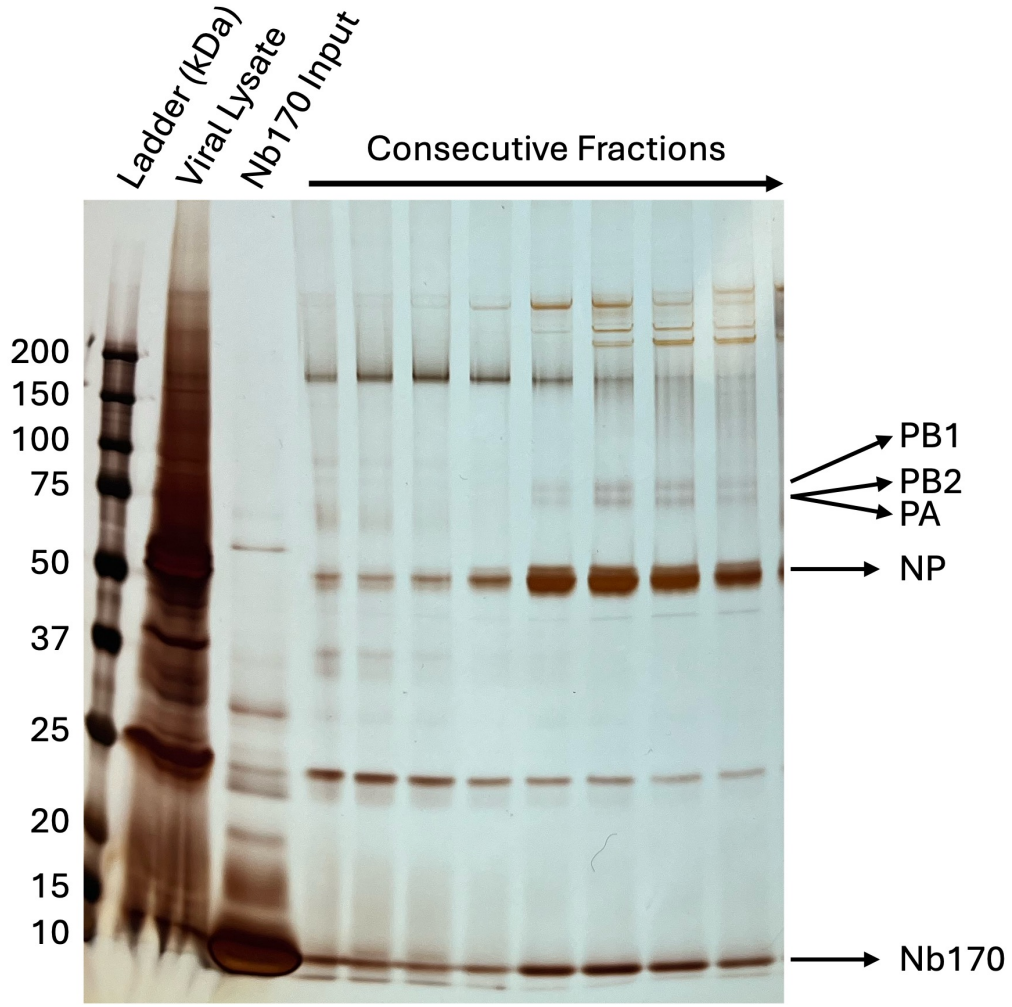


Figure 5.11: Nanobody-Bound vRNP Silver Stain Gel.

Silver stain gel of separated viral lysate fractions with added bound Nb170 nanobody. vRNP (lane 2) and Nb170 (lane 3) inputs are included to demonstrate the contribution of protein from the lysed virions and nanobody stock. Sequential ultracentrifugation fractions show co-enrichment of Nb170 with nucleoprotein and polymerase proteins in vRNP-containing fractions. Boxed fractions were pooled for cryo-EM grid preparation.

Analysis of the silver stain gel above confirmed successful Nb170 binding to the nucleoprotein within the vRNP complex (figure 5.11). vRNP-enriched fractions, characterised by prominent nucleoprotein (56 kDa) and polymerase bands (PA, PB1, PB2 ≈ 85 – 87 kDa), showed concomitant enrichment of the Nb170 band (15 kDa). While some free nanobody was detected across all fractions, the selective co-

enrichment with vRNP proteins in the pooled fractions (figure 5.11, boxed) demonstrated specific complex formation suitable for subsequent cryo-EM analysis.

5.2.4.2. Cryo-EM Screening

Nb170 – vRNP complexes obtained from ultracentrifugation fractionation required additional concentration for cryo-EM analysis. Initial screening revealed that direct application of pooled fractions resulted in insufficient vRNP complex density for worthwhile cryo-EM data collection. Several concentration and sample preparation strategies were systematically evaluated to optimise particle density while preserving complex integrity.

Three initial approaches were tested using standardised grid preparation conditions. In all cases, 2.5 µL of sample was applied to glow-discharged Quantifoil R 2/1 Cu 200 + 2 nm carbon grids, appropriate settings were used on a Vitrobot Mark IV for grid blotting and plunging, and screening was carried out on a ThermoFisher Glacios Transmission Electron Microscope.

First, Nb170-bound vRNPs isolated as described in the previous section were pelleted and resuspended as described in section 5.2.1. (figure 5.8). Ultracentrifugation yielded a visible pellet – notably absent when processing naked vRNPs – suggesting that the Nb170-bound vRNP complex had precipitated. Upon screening, no intact vRNPs were observed within grid holes (figure 5.12, A) indicating that the pelleting process disrupted complex integrity.

Second, 7-fold molar excess Nb170 was added to pelleted and resuspended vRNPs prior to standard grid preparation. This approach resulted in an observed increase in background signal, likely due to free nanobody adsorption to the carbon support film, with no discernible vRNP particles present.

Third, samples underwent spin concentration using a 50 kDa MWCO centrifugal filter (Amicon) at 1,000 x g. Samples were taken every 15 minutes for an hour and grids were made using the standard grid preparation protocol. Whilst material was present on the grids, vRNP degradation was observed which rendered the sample unsuitable for data collection.

To minimise background while obtaining nanobody-bound vRNP complexes, a calculated 1:1 nucleoprotein : nanobody stoichiometry was implemented. Based on the previously determined concentrations (0.14 μM vRNPs containing 9.4 μM nucleoprotein), 0.28 μL of 667 μM Nb170 stock was added to 20 μL aliquots of enriched vRNPs, providing equimolar nanobody without significant excess.

Given a calculated equimolar amount of nucleoprotein and nanobody, the likelihood of complete coating of the vRNP is lower than in the case of an excess of nanobody, but the likelihood of having excess nanobody binding the carbon is greatly reduced.

Example micrographs of excess Nb170 bound to enriched vRNPs before pelleting (figure 5.12, A) and the 1:1 stoichiometric preparation (figure 5.12, B) are shown in the

figure below. Comparably, samples with Nb170 bound to the vRNP which were then pelleted and resuspended demonstrated empty grid holes despite otherwise optimal ice conditions. In contrast, the 1:1 stoichiometric preparation produced well-distributed vRNP particles within holes of appropriate ice thickness suitable for subsequent data collection.

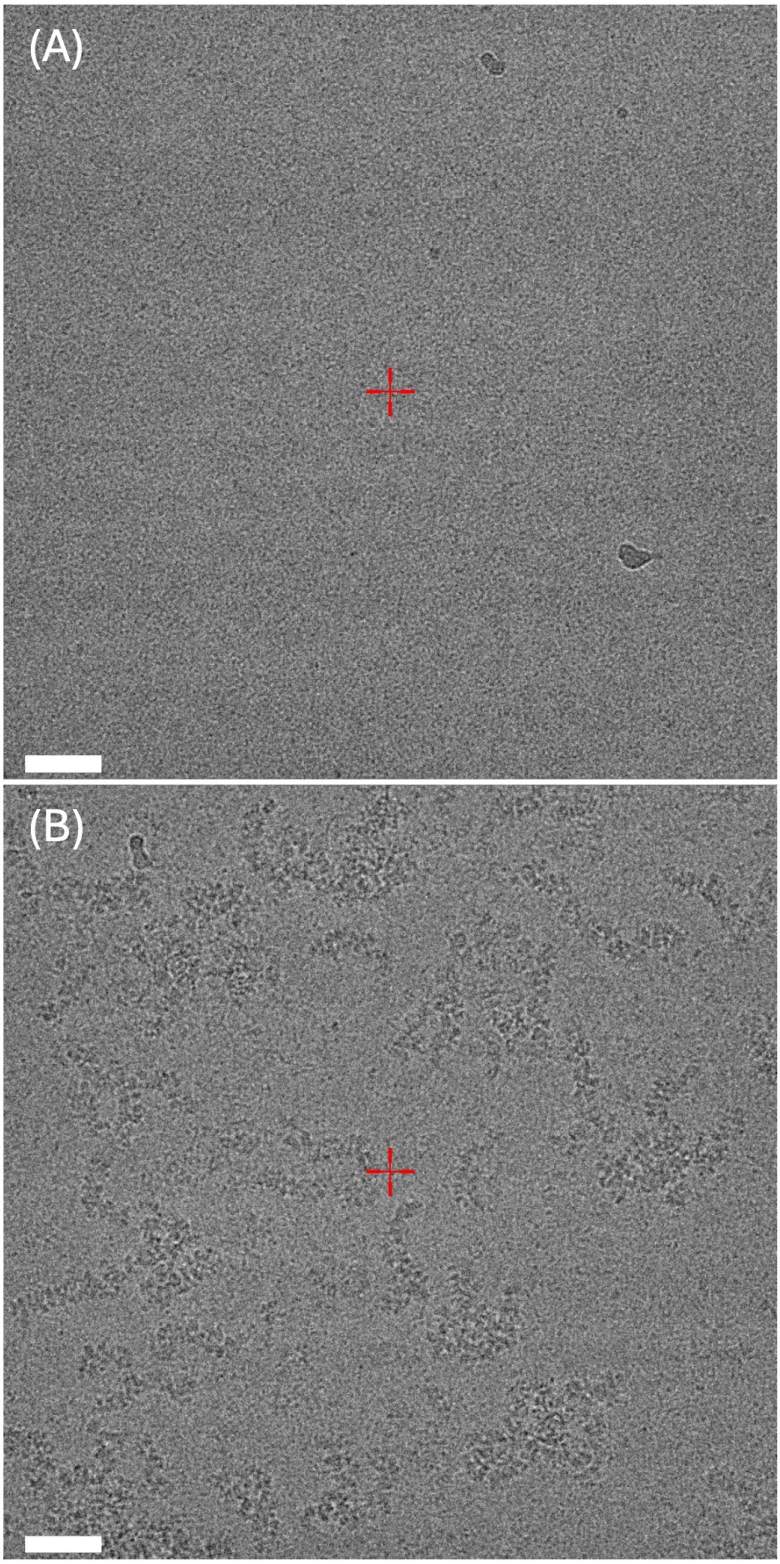


Figure 5.12: Cryo-EM Micrographs of Nanobody-Bound vRNP Samples.

(A) Example micrograph corresponding to the first experiment outlined. Nb170-bound vRNPs were pelleted by ultracentrifugation and applied to cryo-EM grids. (B) Example micrograph corresponding to calculated 1:1 coating of vRNPs with Nb170. Scale bars correspond to 50 nm.

5.2.4.3. Cryo-EM Data Collection

Data collection was performed at eBIC (electron Bio-Imaging Centre, Harwell) using a 300 keV Thermo Fisher Krios transmission electron microscope with the supervision of Dr. Loïc Carrique. Downstream data processing was carried out independently. Data collection and data processing parameters are detailed in table 5.1 below.

The dataset demonstrated a good particle concentration and distribution across the micrographs collected, with individual vRNP complexes being clearly present (figure 5.13). These results validated the optimised sample preparation protocol and provided high-quality data suitable for subsequent structural analysis.

Table 5.1: Data Collection Parameters of Nb170 – vRNP Dataset.

Data Collection – Nb170 – vRNP	
Microscope	Titan Krios
Voltage (keV)	300
Detector	Gatan K3
Magnification	105,000
Pixel Size (Å)	0.831
Total Dose ($e^-/Å^2$)	50
Defocus Range (μm)	-0.5 to -2.5
Frames / Movie	50
Number of Movies	17,520

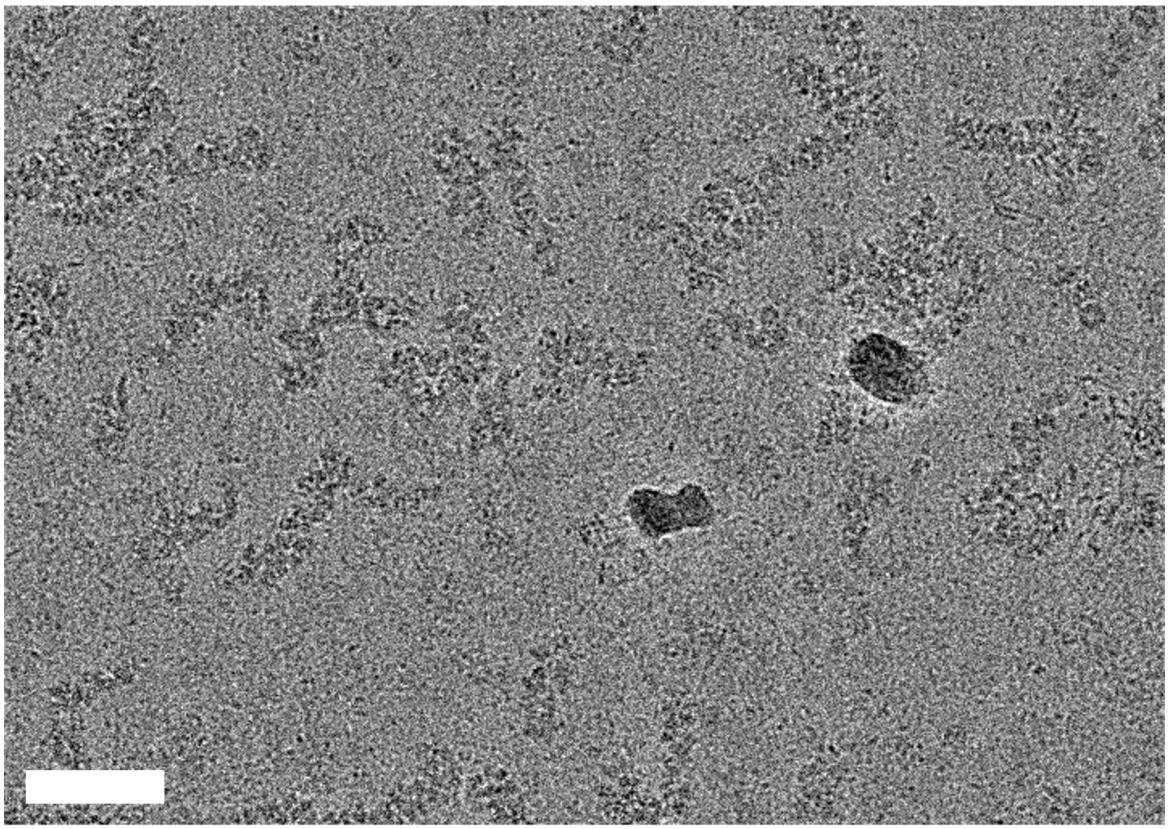


Figure 5.13: Representative Micrograph of Collected Nb170 – vRNP Data.

Representative micrograph of Nb170 – vRNP data collected at eBIC. vRNP complexes are clearly present in the micrograph. Scale bar = 50 nm.

5.2.4.4. Cryo-EM Data Processing

Following motion correction and CTF correction using cryoSPARC Live, initial template-based particle picking was performed using previously generated 2D class averages of vRNP segments as references (Zhu, 2024). From an initial subset of 799 micrographs, 39,993 particles were extracted using a box size of 1024 pixels (corresponding to 1232 Å at a calibrated pixel size of 0.831 Å / pixel). To expedite computational processing during initial screening, particles were binned by a factor of 5.7 to a box size of 180 pixels while maintaining sufficient sampling for initial identification of straight vRNP segments.

Initial 2D classification was performed to remove contaminating particles. Classes exhibiting characteristic vRNP morphology – specifically, clear helical nucleoprotein density with regular periodicity – were selected for further analysis. The high-quality 2D classes resulting from 27,343 particles were subsequently input as templates into a filament tracer job, which used all 17,386 accepted micrographs using a particle separation distance of 60 Å between consecutive picks along identified filaments to identify approximately 3,000,000 particles, corresponding to an average particle density of 173 particles per micrograph.

Quality control filtering based on filament straightness and low sinuosity parameters resulted in the retention of particles suitable for helical reconstruction. The selected particles underwent three rounds of iterative 2D classification to further select for the straightest vRNP segments possible. This selection process yielded a final refined particle stack of 222,628 particles across 20 distinct 2D classes (figure 5.14).

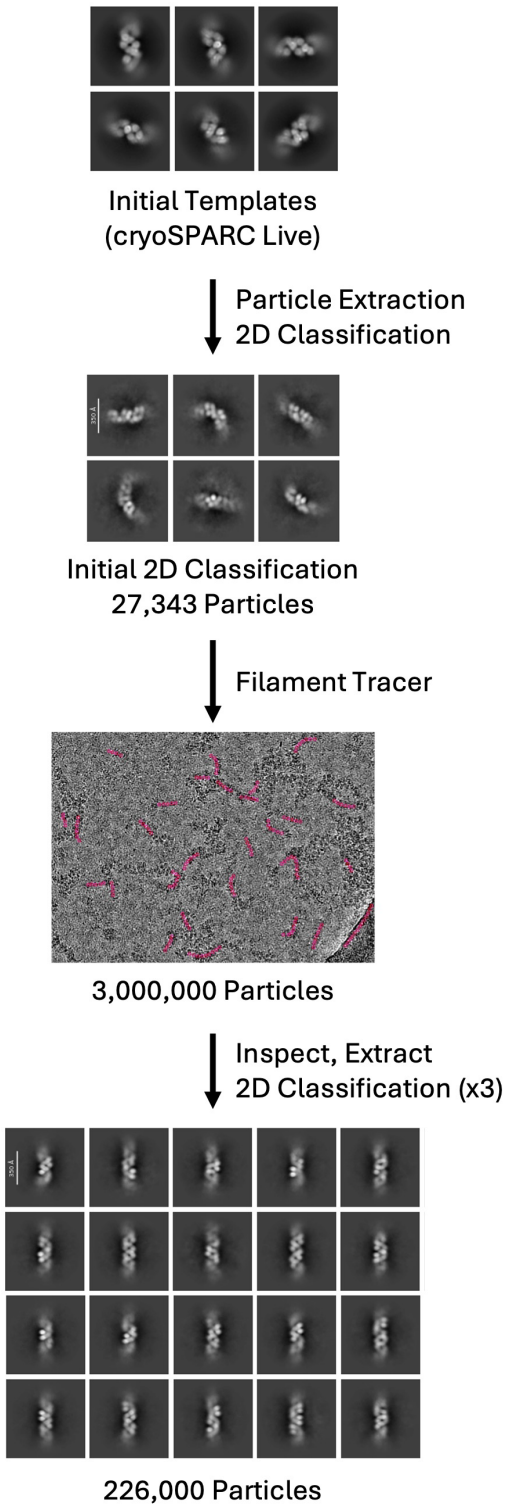


Figure 5.14: Schematic Diagram of Nb170 – vRNP Data Processing.

Schematic diagram of initial processing steps for Nb170-vRNP dataset processing. Example 2D classes or micrographs are shown at each step with relevant numbers of particles noted.

Determination of helical symmetry parameters was attempted using Fourier-Besel analysis. Two-dimensional power spectra were calculated of each 2D class average from the final classification, and the resulting spectra were analysed using Helixplorer (https://rico.ibs.fr/helixplorer/helixplorer/). However, examination of the power spectra revealed limited information on the vRNP sample, precluding robust determination of helical parameters.

Analysis of reflections in the calculated power spectra provided estimates of helical pitch across eight different helical classes, with an average value of 124.69 Å (figure 5.15, B). Independent measurements of the six straightest 2D class averages suggested a pitch of approximately 130 Å (figure 5.15, A). However, the absence of higher-order layer lines and a low signal-to-noise ratio prevented determination of the rise per subunit from the power spectra. Additionally, there was not enough signal present in the 2D class averages obtained to estimate the number of nucleoproteins per turn, and as such, helical parameters such as helical twist could not be estimated.

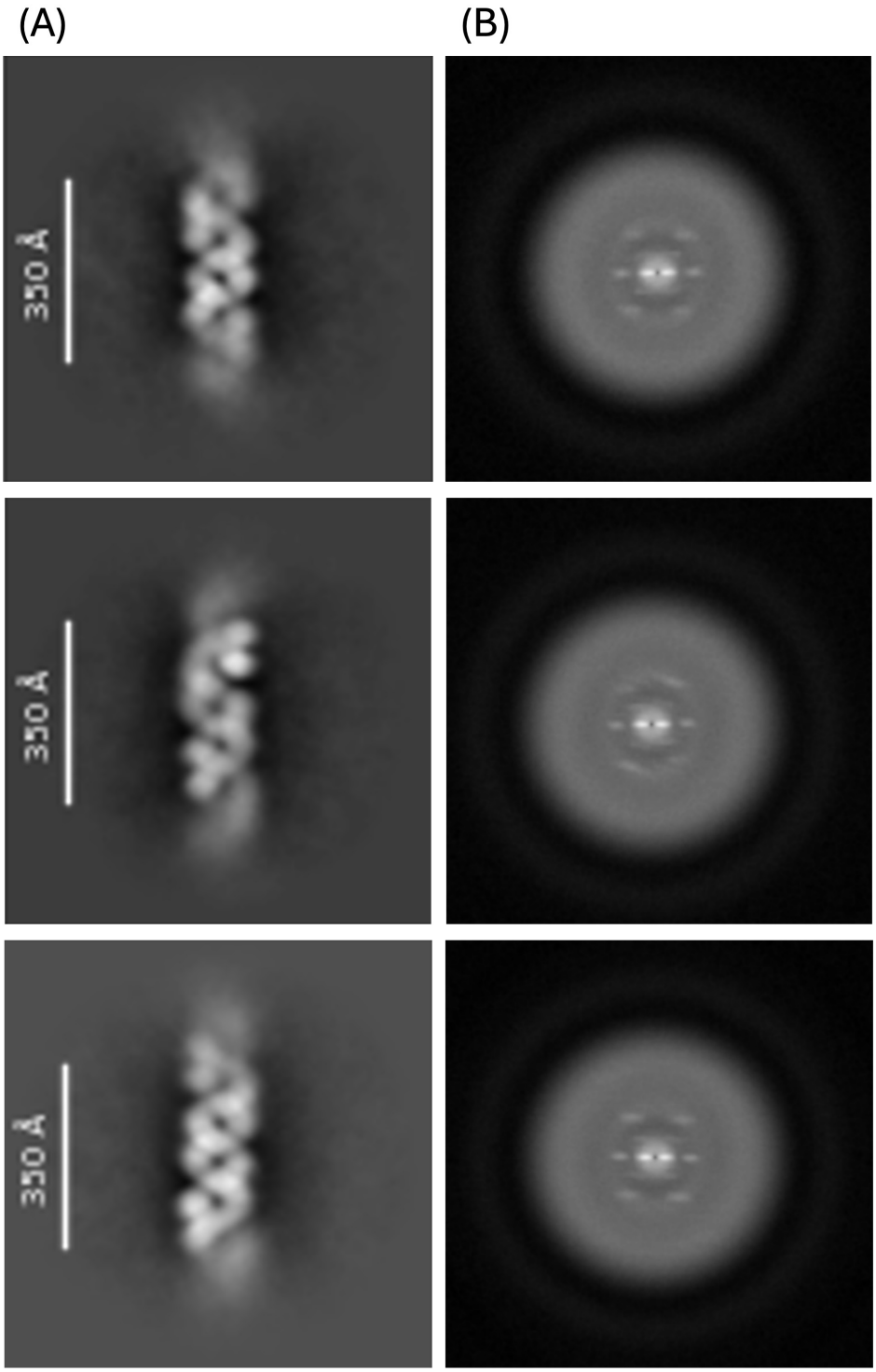


Figure 5.15: Example 2D Class Averages and 2D Power Spectra.

(A) 2D class averages and (B) corresponding 2D power spectra of vRNP segments selected for their straightness. (A) Differences in 2D class averages are observed, corresponding to different specific helical parameters across different vRNP segments. (B) Corresponding central sections of 2D power spectra of the 2D class averages are shown, with clear meridional lines corresponding to pitch, but no other signal useful to the determination of helical symmetry. The number of particles in each 2D class is 3146 (top), 4029 (middle), 4716 (bottom).

To potentially obtain longer-range helical order information and to improve parameter determination, particles were additionally extracted using an enlarged box size of 1500 pixels (corresponding to 1805 Å). Following iterative 2D classification, one class containing 3100 particles was selected based on its optimal straightness. However, analysis of the corresponding 2D power spectrum revealed even more limited helical information than the shorter segments (figure 5.16), with significant degradation in signal attributed to increased structural heterogeneity over the longer filament length.

The observed deterioration in helical order with increasing segment length, combined with the limited information content in the observed 2D power spectra, indicated that classical Fourier-Bessel analysis was inappropriate for this level of heterogeneity in a helical sample.

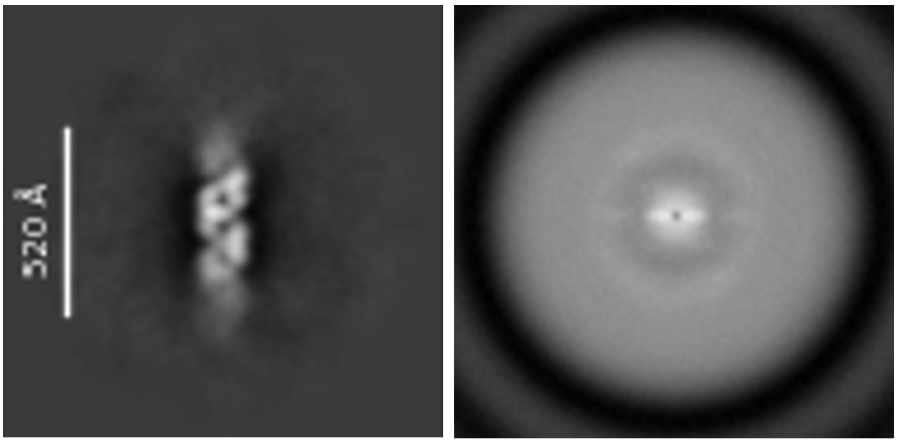


Figure 5.16: 2D Class Average and 2D Power Spectrum with a Larger Box Size.

2D class average and central section of calculated 2D power spectrum of a vRNP segment selected for its straightness in a bigger box (1500 pixels). The blurring of signal at the ends of the vRNP in the 2D class average and the lack of signal in the calculated 2D power spectrum reflect the structural heterogeneity of the vRNP.

An *ab-initio* reconstruction job with three classes and no imposed symmetry was run on the stack of 222,628 particles distributed across 20 distinct 2D classes. One of the resulting volume classes was particularly promising, incorporating 99,401 particles (44.7 % of the total dataset) and exhibiting clear double-helical morphology with well-defined major and minor grooves characteristic of nucleoprotein filaments (figure 5.17). This volume was selected as the reference model for subsequent refinement procedures.

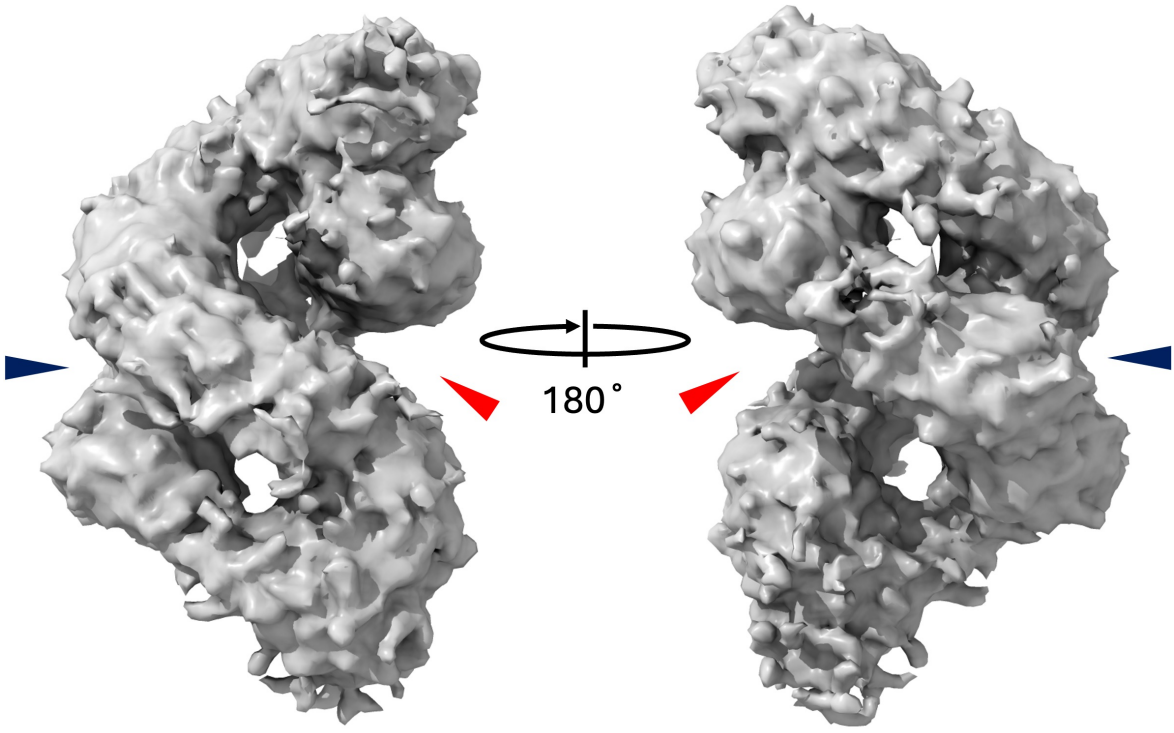


Figure 5.17: vRNP *Ab-initio* Density Map.

Ab-initio density map of influenza vRNP segment. This ab-initio model contains 99,401 particles and displays clear signatures of double helical symmetry indicative of an influenza vRNP, including a major (red) and minor (blue) groove.

Initial helical refinement was performed using starting parameters of helical twist equal to 60˚ and helical rise equal to 20 Å, with broad search ranges (twist: 55 – 65˚,

corresponding to 5.5 – 6.5 nucleoprotein subunits per turn; rise: 10 – 30 Å). The resolution threshold for local symmetry searches was set to 4 Å to permit sufficient helical parameter exploration during the refinement process.

The refinement converged on a helical twist of 57.00˚ and a helical rise of 19.77 Å, corresponding to 6.3 nucleoprotein subunits per turn. The resulting density map demonstrated significant improvement over the *ab-initio* model, with discrete nucleoprotein density resolved across nearly one complete helical turn. However, signal degradation and a loss of helical order were observed at the ends of the density map, consistent with the expected structural heterogeneity of the vRNP sample (figure 5.18).

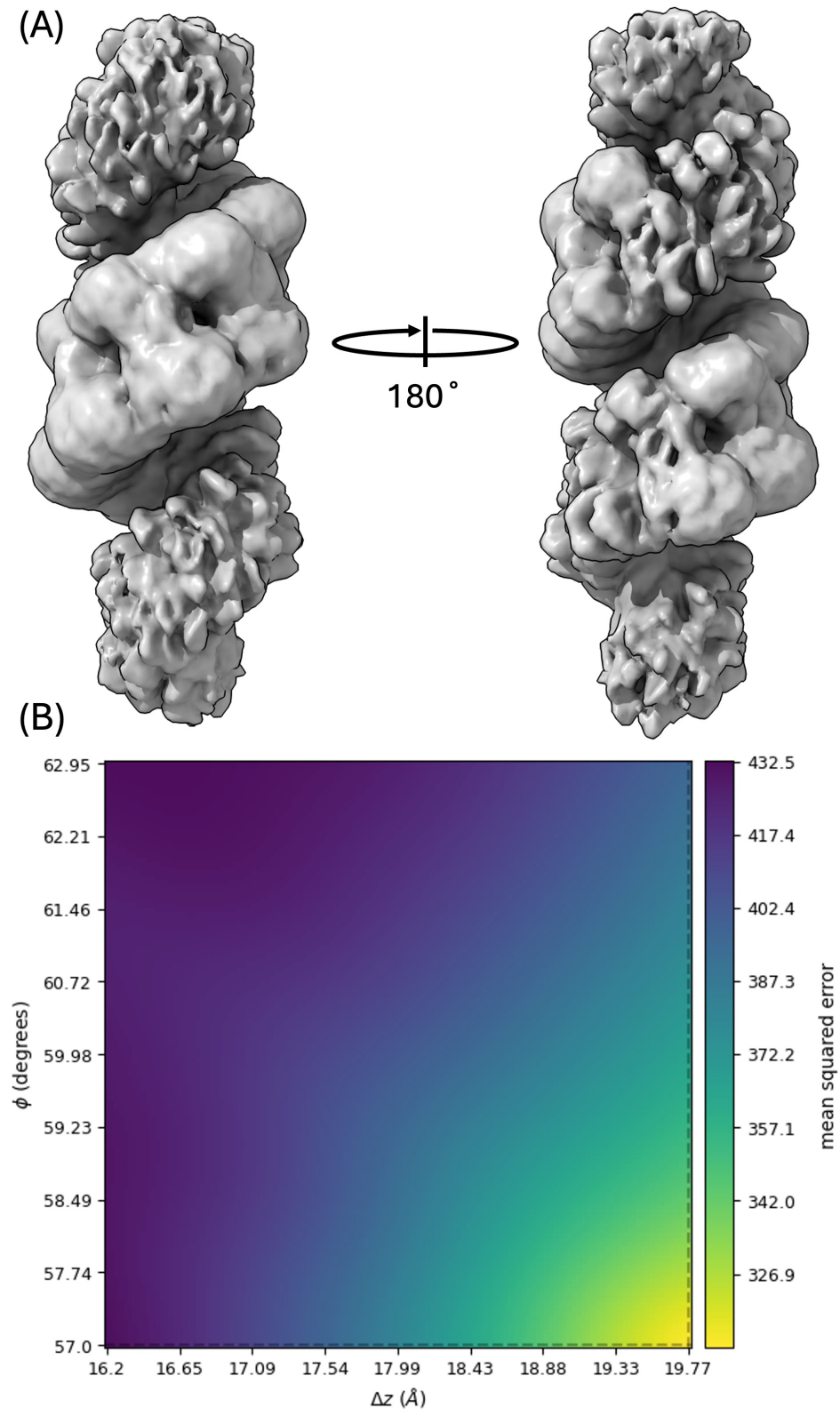


Figure 5.18: Results of Initial Helical Refinement.

(A) Density map of initial helical refinement of influenza vRNP segment. (B) Corresponding helical symmetry search plot of helical rise (Δz) versus helical twist (ϕ).

Within the helical refinement job detailed above, helical symmetry search was carried out (figure 5.18, B). As the initial symmetry search reached the boundary of the parameter space, an additional symmetry search was performed on the resulting volume with narrowed search ranges (helical twist: 55 – 60˚; helical rise: 18 – 22 Å). This analysis yielded optimised parameters of 56.49˚ helical twist and 21.64 Å helical rise, corresponding to 6.7 nucleoprotein subunits per helical turn. These parameters were used in a subsequent helical refinement job. However, the resulting density map showed no significant improvement over the previous reconstruction (figure 5.19), suggesting that the refinement had reached a resolution limit imposed by sample heterogeneity rather than parameter accuracy. Analysis of eight independent helical refinement jobs yielded average parameters of helical twist of 57.68˚ and helical rise of 20.00 Å.

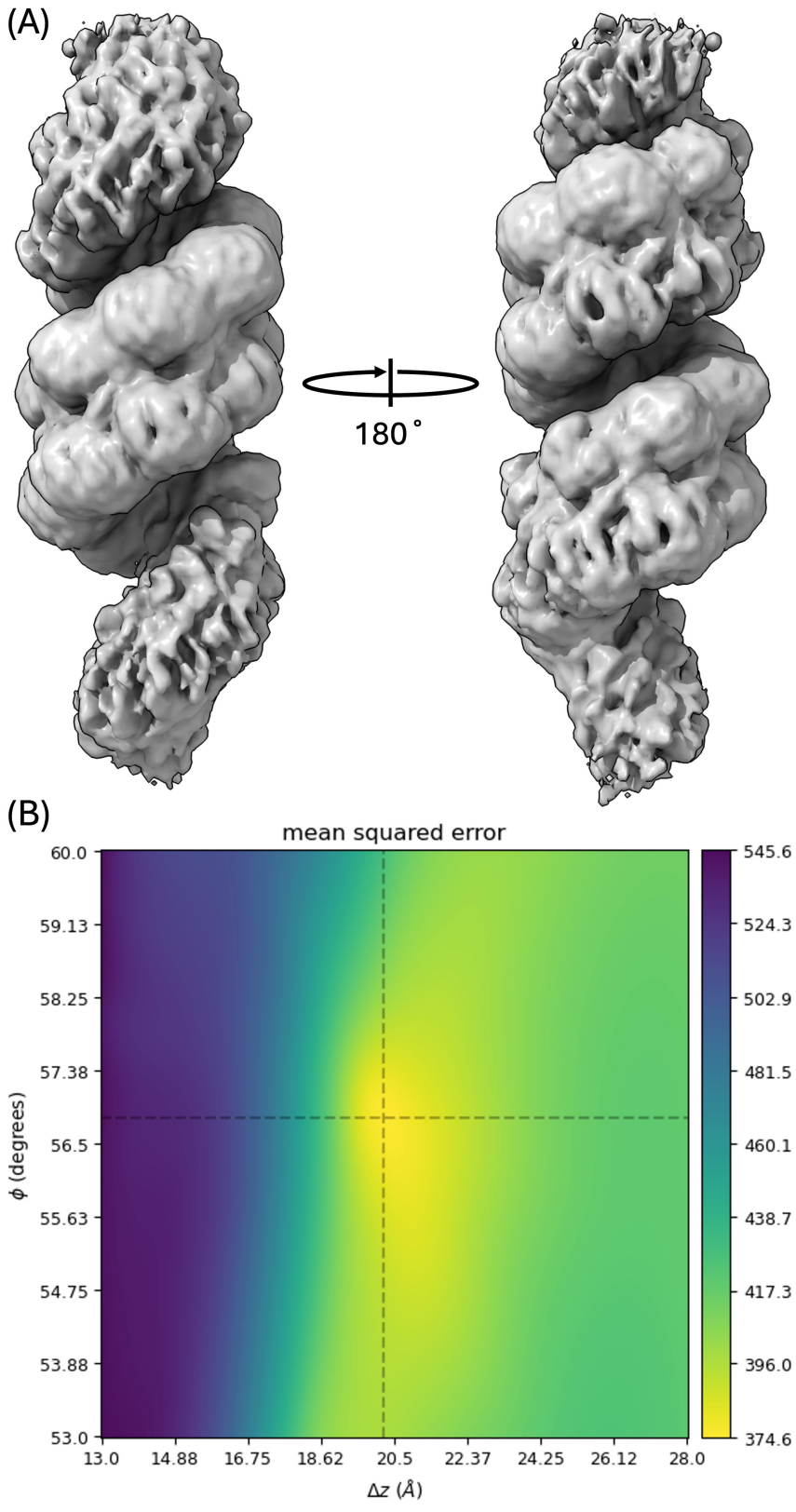


Figure 5.19: Subsequent Helical Refinement.

Subsequent helical refinement job run with improved helical parameters. (A) Density map of initial helical refinement of influenza vRNP segment. (B) Corresponding helical symmetry search plot of helical rise (Δz) versus helical twist (ϕ). Whilst the symmetry search appeared to find a plausible solution, the density map obtained does not represent a significant improvement on the previous helical refinement job.

Given that global helical symmetry imposition was limited by sample heterogeneity, a focused refinement approach was implemented to focus refinement on the most ordered central region of the vRNP filaments. Masks were generated in ChimeraX to isolate specific oligomeric arrangements of nucleoproteins within the vRNP: 6-mer, 4-mer, 3-mer, and 2-mer conformations positioned within the central, most ordered section of the reconstruction.

Particle subtraction was performed using inverted masks created using the cryoSPARC volume tools utility to remove signal contribution from regions outside of the specific area of interest. Local refinement jobs were subsequently executed using the original masks and the subtracted particle stacks to focus reconstruction on the isolated nucleoprotein oligomers (figure 5.20).

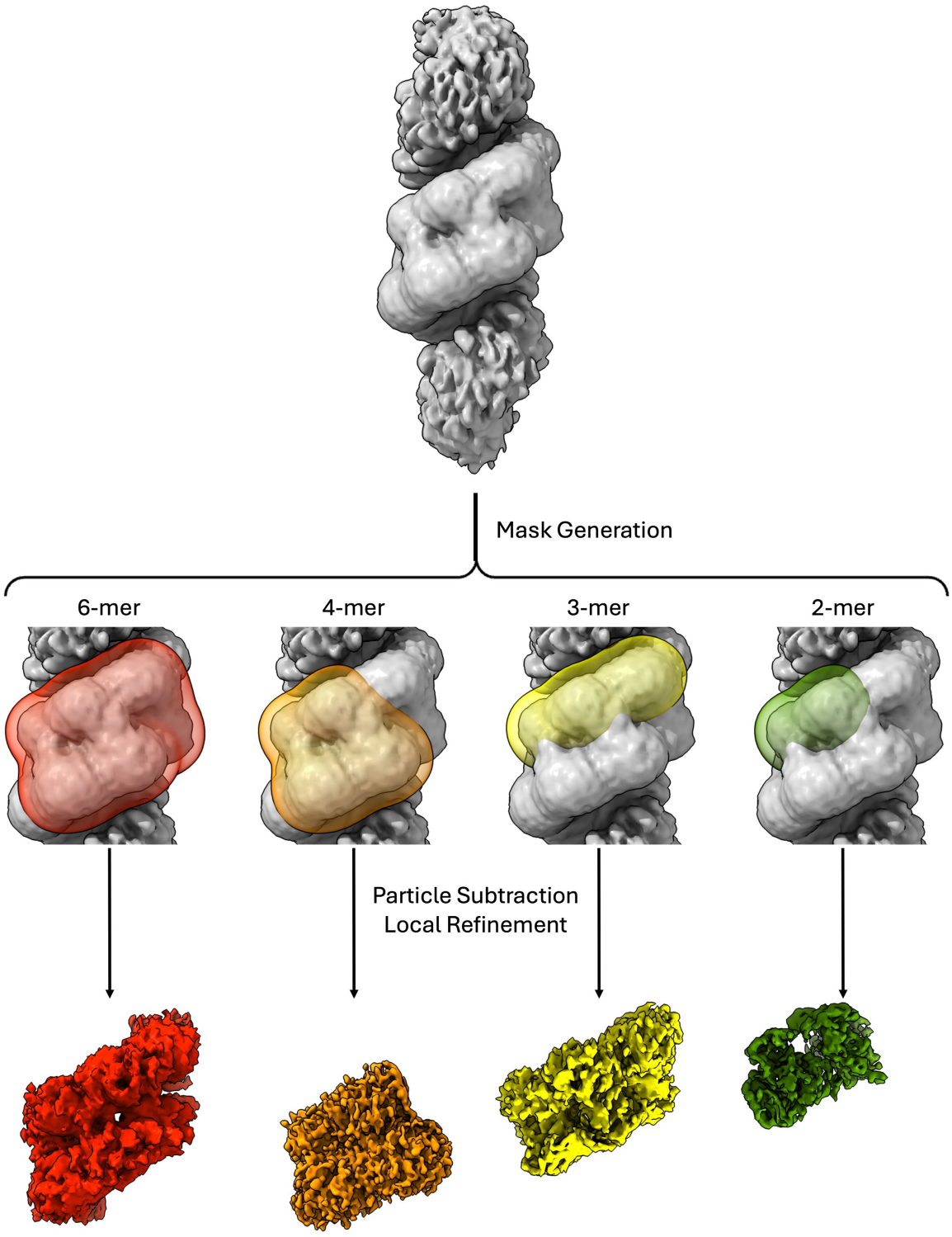


Figure 5.20: Focussed Refinement Schematic.

Schematic diagram of downstream processing. Masks were generated in ChimeraX to isolate a 6-mer (red), 4-mer (orange), 3-mer (yellow), and 2-mer (green) of influenza nucleoprotein for focussed refinement. Particle subtraction was carried out in ChimeraX by generating an inverted mask using the volume tools utility and then running a particle subtraction job. Subsequently, a local refinement job was carried out with the initial mask generated and the particle stack from the particle subtraction job used to focus the refinement on the isolated oligomers of nucleoprotein. The input density map is shown at

the top of the figure. Masks generated are shown coloured with 25 % transparency overlayed with the initial density map, and the density maps resulting from the local refinement jobs are shown below.

While the focused refinement strategy successfully isolated a central section of density regions corresponding to the target oligomeric assembly of nucleoprotein (6/4/3/2-mer, as in figure 5.20), the resulting local refinement outputs, despite showing some discrete nucleoprotein density, did not achieve sufficient resolution for ambiguous rigid body fitting of nucleoprotein structures. The reconstructions reached a nominal resolution of between 5.8 and 8.0 Å, though this is likely an overestimate of resolution given the lack of genuine features in the density maps. In the resulting density maps, significant anisotropy was present with little to no signatures of protein secondary structural elements – likely due to the heterogeneity of the sample.

Critically, no density attributable to the Nb170 nanobody was observed in either 2D class averages or 3D reconstructions. This absence suggests that the 1:1 stoichiometric ratio of Nb170 to calculated nucleoprotein was insufficient to achieve observable nanobody density bound to the nucleoprotein, or that the nanobody binding may be sub-stoichiometric and averaged out during image processing.

These findings highlight the technical challenges associated with structural analysis of the highly heterogeneous influenza vRNP complex and underscore the need for strategies which ensure a high degree of sample homogeneity is enforced in future investigations. The results demonstrate that even sophisticated computational approaches cannot fully compensate for the inherent flexibility and conformational

diversity of native vRNP assemblies, necessitating alternative experimental strategies for high-resolution structural determination.

5.2.5. Functionalised Cryo-EM Grids for 'On-Grid' Sample Concentration

The absence of detectable nanobody signal in the stoichiometric 1:1 Nb170:nucleoprotein approach taken above necessitated the development of an alternative approach to achieve sufficiently high concentrations of vRNPs on cryo-EM grids with a complete coating of binding partner. To address this challenge, a functionalised grid system was implemented as an 'on-grid' concentration method.

Functionalised cryo-EM grids have demonstrated efficacy across diverse structural biology applications to increase sample concentration. Previous successful implementations include ribosome structure determination through immobilisation (Kelly et al., 2008; Kelly et al., 2010), GroEL isolation and structural analysis from crude cell lysate (Benjamin et al., 2016), and high-resolution structure determination of Tulane virus (2.6 Å) from low-yield cell culture preparations (Yu et al., 2016).

Given the challenges encountered with conventional concentration methods for maintaining the structural integrity of influenza vRNPs, the functionalised grid approach offered a promising alternative for achieving the required particle concentration for worthwhile structural analysis.

5.2.5.1. Experimental Design and Implementation

Grid functionalisation and preparation was performed by Dr. Gangshun Yi using Ni-NTA-charged lipid monolayer-coated grids as the immobilisation platform (as in figure 5.7). Two distinct nanobody-based capture approaches were used: a nucleoprotein-targeted approach using the nucleoprotein nanobody, Nb170, as introduced prior, and a polymerase-targeted approach using the influenza polymerase PA linker nanobody Nb8210 (Fan et al., 2019; Keown et al., 2022).

It was hypothesised that nanobody-mediated affinity capture would achieve concentrations of vRNPs on the grid sufficient for structural analysis while maintaining vRNP complex integrity.

5.2.5.2. Nucleoprotein-Targeted Immobilisation Strategy

An initial validation experiment employed degraded vRNP preparation to assess the fundamental capacity for nanobody-mediated sample enrichment. Grids were functionalised as in the materials and methods chapter (section 2.5.4.1.) by establishing a lipid monolayer across a Teflon well and charging with Ni-NTA. 0.06 mg/mL nanobody-containing solution was added and incubated for 2 hours. The grid was laid on the surface of the Teflon for 2 minutes to establish the lipid monolayer on the grid and washed three times with 120 µL of sample buffer (25 mM HEPES-NaOH, pH 7.50, 150 mM NaCl) to remove free nanobody. Up to 10 µL of sample was added and incubated for 15 minutes prior to washing with 120 µL of sample buffer to remove

unbound particles. Grids were then blotted and plunge frozen using a Thermo Fisher Vitrobot IV.

Control grids with no immobilised nanobody exhibited no observable sample enrichment (figure 5.21, top), while functionalised grids demonstrated clear enrichment of degraded, aggregated vRNP material (figure 5.21, bottom). Notably, the filamentous structures corresponding to immobilised vRNA molecules were observed on nanobody-functionalised grids.

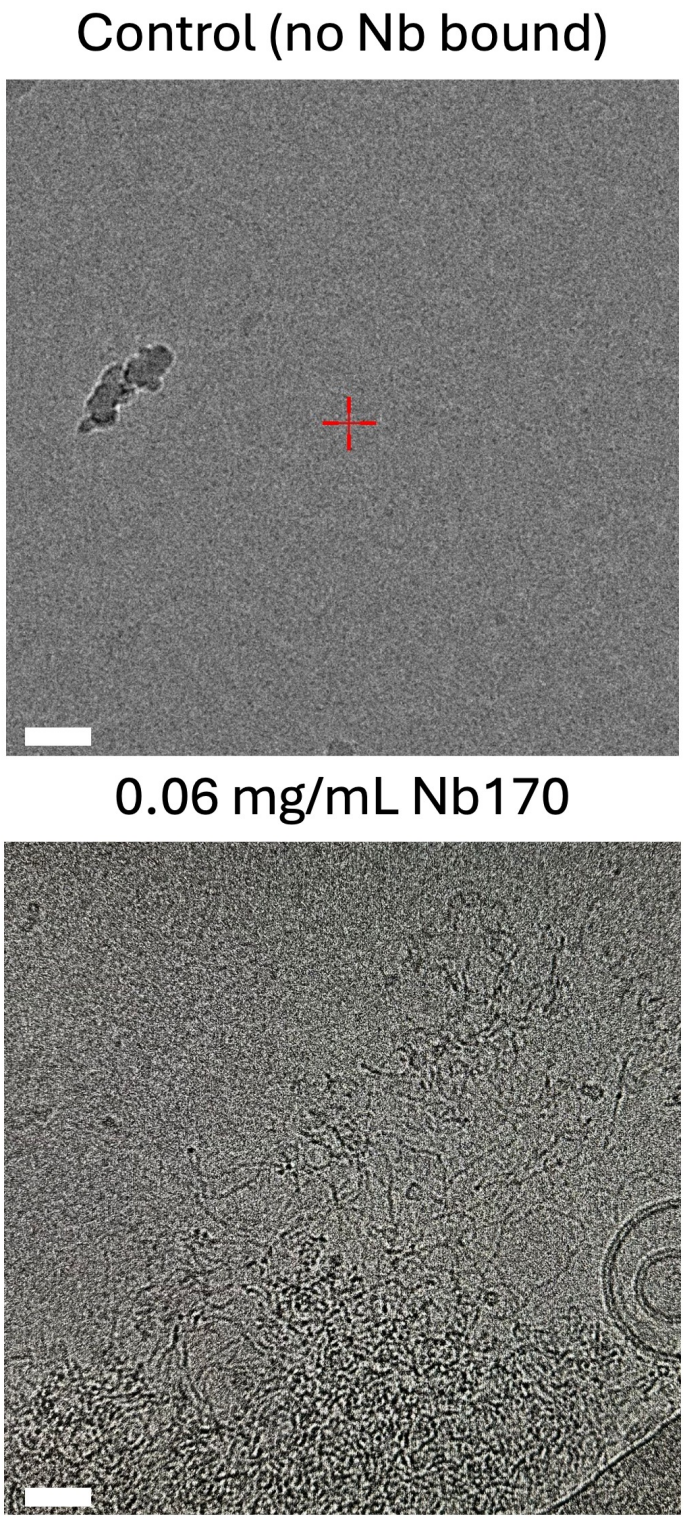


Figure 5.21: Nucleoprotein Nanobody-Mediated Immobilisation Trial.

Representative micrographs of Nb170 immobilisation of degraded vRNP sample on functionalised cryo-EM grids. As an initial test, vRNP sample which had been previously spin-concentrated was applied to the EM grids in order to determine whether sample enrichment was at all possible. (Top) Control affinity grid with no nanobody immobilised. (Bottom) 0.06 mg/mL Nb170 incubated with grid overnight leads to vast enrichment in degraded vRNP material. A significant amount of nucleoprotein is immobilised on the grid alongside vRNA. Scale bar = 50 nm.

Subsequent experiments using freshly prepared vRNP samples evaluated two Nb170 nanobody concentrations (0.02 mg/mL and 0.06 mg/mL) to optimise capture efficiency and sample quality. Both nanobody concentrations achieved effective vRNP enrichment, resulting in good particle concentration and distribution across the grid (figure 5.22). However, when a higher concentration of Nb170 is used, some vRNP aggregation is observed (figure 5.22, lower left micrograph) which is not observed in the lower concentration sample. Additionally, free vRNA is observed in the higher concentration Nb170 grids. This is attributed to indiscriminate nucleoprotein binding to the nanobody-coated surface regardless of assembly state within functional vRNPs.

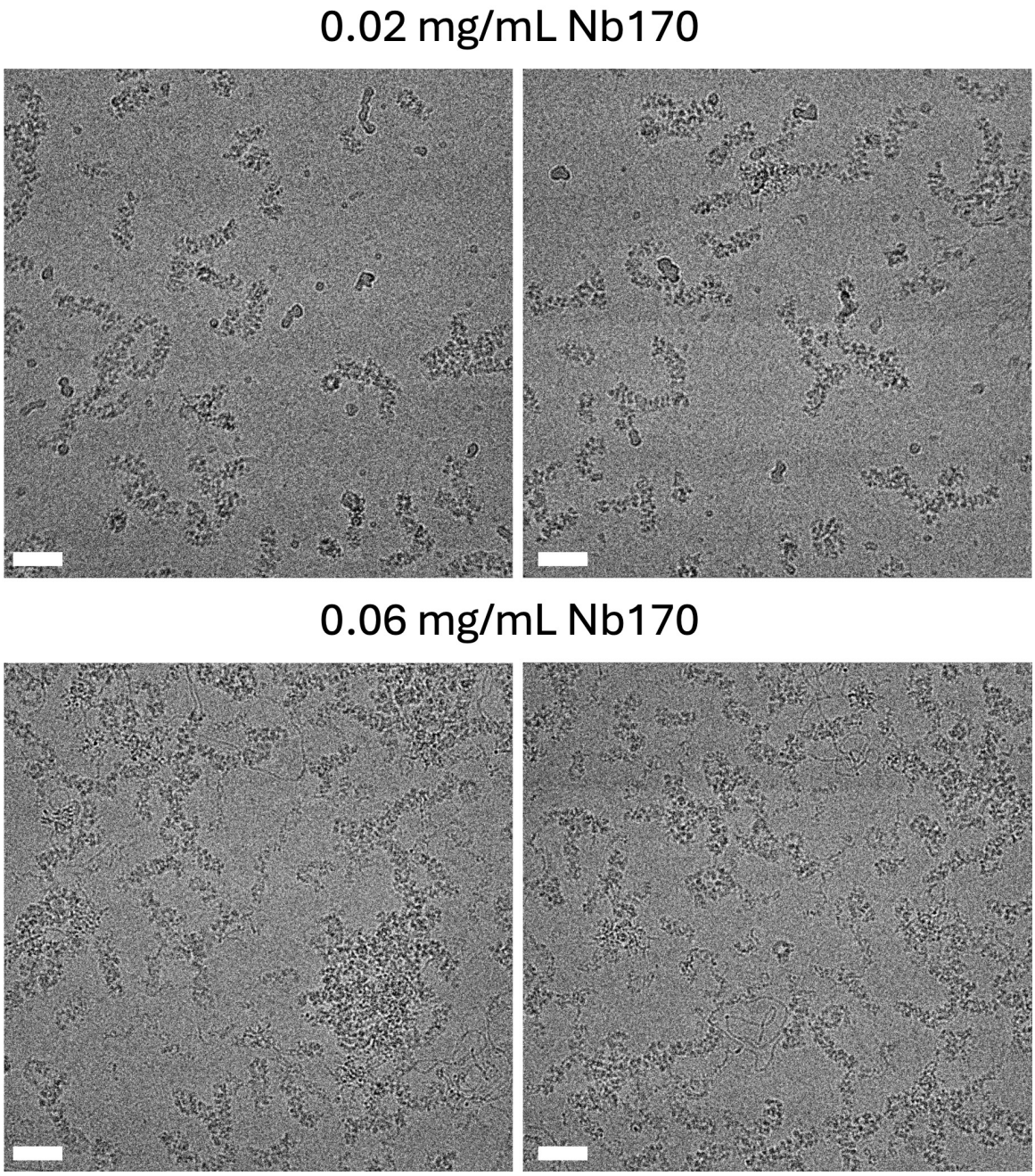


Figure 5.22: Nucleoprotein Nanobody-Mediated Immobilisation of Fresh vRNP Sample.

Representative micrographs of Nb170 immobilisation of fresh vRNP sample on cryo-EM grids. Two concentrations of Nb170 were used, 0.02 mg/mL (top) and 0.06 mg/mL (bottom) whereby both led to good sample concentration and distribution across the grid, but the higher concentration led to aggregation of vRNPs (particularly evident in the bottom left micrograph), and immobilisation of some degraded vRNPs given the presence of vRNA on the grid. Scale bar = 50 nm.

5.2.5.3. Polymerase-Targeted Immobilisation Strategy

The limitations of nucleoprotein-targeted immobilisation prompted the development of an alternative approach using influenza polymerase nanobody Nb8210. This strategy offered multiple advantages.

Firstly, enhanced selectivity for intact vRNP complexes. Since each vRNP contains a single polymerase heterotrimer, polymerase-targeted capture would predominantly immobilise complete vRNPs rather than individual nucleoprotein components or partially disassembled complexes.

Second, by targeting the polymerase rather than the nucleoprotein, surface availability for binding partner interactions of the nucleoprotein is maintained, enabling validation of stable complex purification through biochemical assays prior to cryo-EM preparation.

Finally, targeting the single copy of the polymerase on the vRNP may enhance immobilisation density on the grid compared to multi-site nucleoprotein targeting. Each vRNP contains only one polymerase complex, but many copies of the nucleoprotein. Nucleoprotein-targeted immobilisation may therefore bind multiple sites per vRNP, creating steric hinderance between neighbouring particles and limiting packing density. In contrast, polymerase-targeted immobilisation uses only one binding site per vRNP, allowing closer packing of particles on the grid.

Nb8210 was selected based on its established binding to the PA-linker domain of the polymerase without inhibiting polymerase dimerisation or interfering with viral replication and transcription (Fan et al., 2019; Keown et al., 2022). By choosing another nanobody, it may be possible to lock the polymerase in specific conditions for the study of the replication cycle of influenza virus in subsequent studies.

Recombinant Nb8210 was expressed in BL21 (DE3) *E. coli* and purified as detailed in the materials and methods chapter (section 2.2.5.). As the nanobody construct contains a PelB signal sequence, which is known to direct recombinant proteins to the *E. coli* periplasm after cleavage of the PelB sequence (Power et al., 1992; Su et al., 2015; Shi et al., 2021b), the protein purification was carried out initially with a bacterial osmolysis step, rather than a bacterial lysis step, which leads to the production of very pure protein. After osmolysis, the periplasmically-expressed nanobody was affinity purified via its C-terminal His_6 tag.

Eluted nanobody was then purified further via size-exclusion chromatography. As shown in the SEC trace and accompanying SDS-PAGE gel (figure 5.23, panels B and C), the purified nanobody eluted from the SEC column with a sharp, peak at around 15 mL, and pure protein on the SDS-PAGE gel with a molecular weight of 15 kDa.

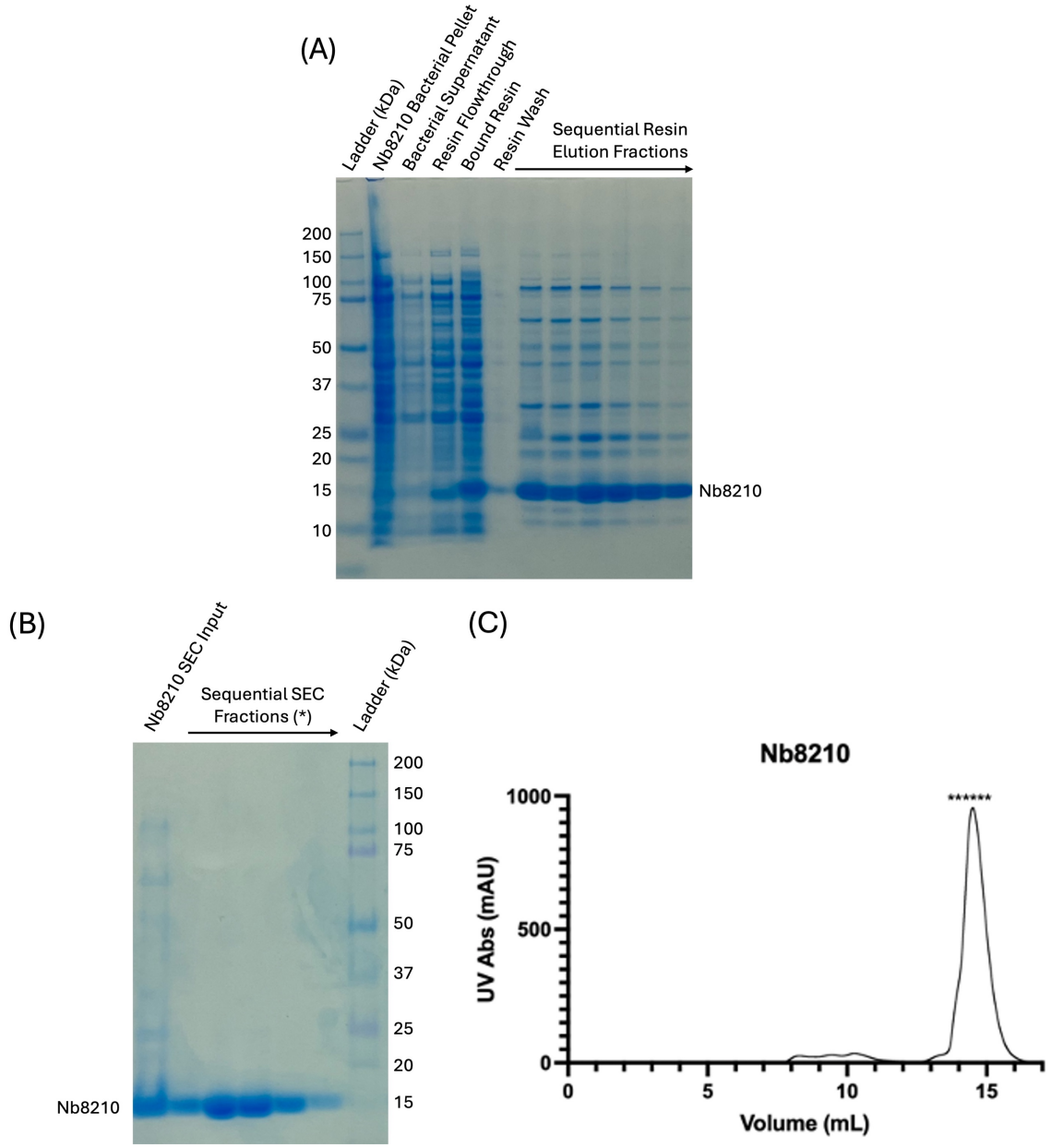


Figure 5.23: Nb8210 Production.

(A) Nb8210 purification SDS-PAGE gel from bacterial pellet through to sequential elution fractions. (B) SDS-PAGE gel of SEC fractions from SEC run of Nb8210. (D) Chromatogram from SEC run of Nb8210 on an S200 10/300 GL Increase column.

Functionalised grids prepared with 0.06 mg/mL Nb8210 demonstrated superior immobilisation quality and concentration compared to the nucleoprotein-targeted approaches outlined earlier. A significant enrichment of vRNP material was observed relative to non-functionalised control grids (figure 5.24, top), and there was an

absence of signatures of aggregation or degradation which were observed with an equivalent concentration of Nb170. Critically, there was no detectable free vRNA, indicating preferential immobilisation of intact vRNP complexes, and good particle distribution and concentration which would facilitate worthwhile data collection.

These results establish nanobody-functionalised cryo-EM grids as a significant advancement over conventional concentration methods for vRNP structural analysis, successfully addressing challenges of achieving adequate sample concentration while maintaining vRNP quality, particularly with bound interaction partners.

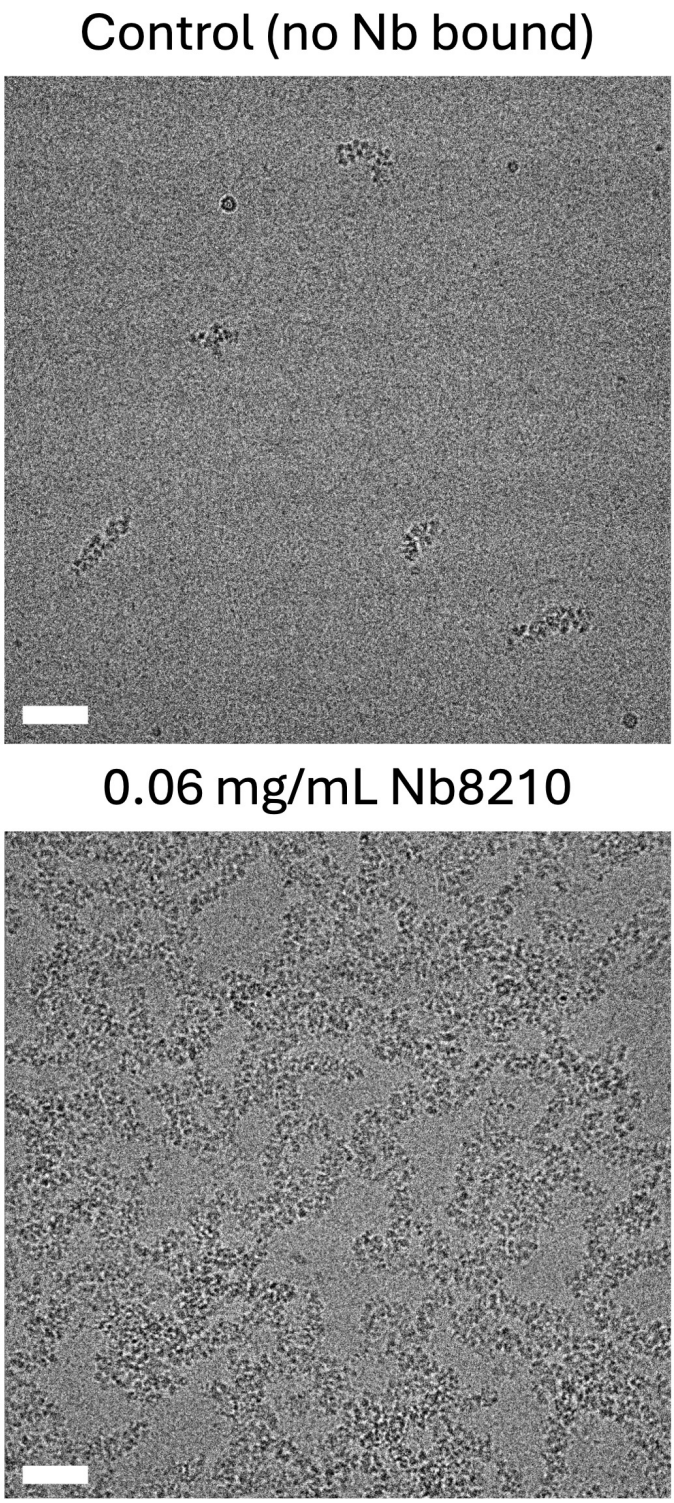


Figure 5.24: Polymerase Nanobody-Mediated Immobilisation of Fresh vRNP Sample.

(Top) Control affinity grid with no nanobody immobilised. (Bottom) 0.06 mg/mL of Nb8210 incubated with a lipid monolayer on a cryo-EM grid overnight and fresh vRNP sample applied and washed. Some vRNPs were visible on the control grid at low concentration. The addition of the polymerase nanobody leads to a significant increase in the amount of vRNP material immobilised, with little to no signatures of vRNP aggregation or degradation present. Scale bar = 50 nm.

5.2.6. MxA Association with Influenza Nucleoprotein in the Context of the Virion-Derived vRNP

In the preceding experiments, all vRNPs were derived from WSN strain influenza A virus (A/WSN/1933 (H1N1)), which contains nucleoprotein residues with canonical MxA resistance mutations. To investigate MxA-vRNP interactions directly, it was therefore necessary to generate vRNPs that could be restricted by MxA. For this purpose, Dr. Kuang-Yu Chen constructed a chimeric influenza A virus using a nine-plasmid reverse genetics system. This chimeric virus comprised all genome segments from WSN except for the nucleoprotein gene, which was replaced with the S09 avian influenza gene (A/Green-Winged Teal/Ohio/175/1966 (H2N1)). This avian nucleoprotein was selected owing to the presence of mutations associated with MxA sensitivity (as discussed in section 3.1.3.) and had been previously validated in recombinant expression studies to be susceptible to MxA binding (as discussed in section 4.2.1.4.).

5.2.6.1. Sample Preparation

Enrichment of vRNPs was performed following the established protocol described in section 5.2.1. Based on the calculated nucleoprotein concentration of 9.4 µM in vRNP preparations, purified MxA protein (section 4.2.1.) was added at a concentration of 2.5 mg/mL (33.3 µM) in a 1:1 volume ratio (50 µL each). This corresponded to an approximate 3.5-fold molar excess of MxA relative to nucleoprotein.

The virus lysate-MxA mixture was incubated on ice for 30 minutes to facilitate binding between MxA and nucleoprotein. Following incubation, vRNPs were isolated from the lysate by ultracentrifugation using the previously established parameters.

Silver staining of fractionated samples revealed enrichment of an additional protein band of approximately 75 kDa, which co-eluted with vRNP-containing fractions (figure 5.25, A), consistent with the approximate molecular weight of MxA. Importantly, this band was not observed in previous silver stain gels where MxA was not added in the viral lysis stage (figure 5.25, B, C).

In order to confirm the identity of the protein, two western blots were performed by Dr. Kuang-Yu Chen using antibodies specific to MxA and influenza nucleoprotein. The MxA-specific western blot demonstrated clear detection of MxA protein in both input samples and co-fractionation with vRNP-containing fractions (figure 5.25, D, top). Parallel nucleoprotein-specific immunoblotting verified the presence of nucleoprotein in the viral lysis input sample (but not the MxA input sample), and in all three boxed fraction samples (figure 5.25, D, bottom), confirming the co-enrichment of MxA with vRNPs. Notably, MxA signal was observed to increase across the three vRNP-containing fractions, while nucleoprotein signal remained relatively constant, suggesting variable MxA-nucleoprotein binding stoichiometries within these samples.

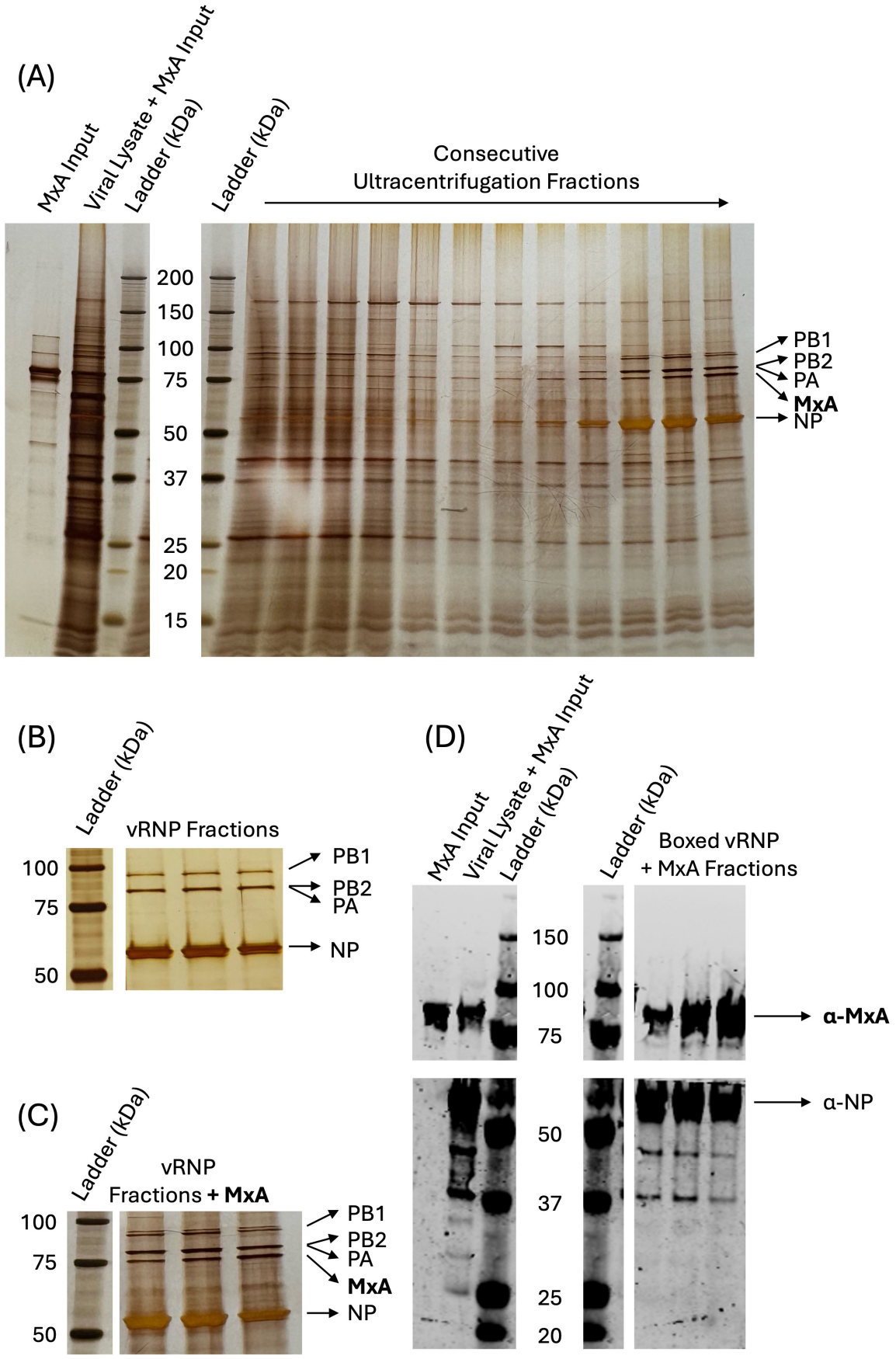

(A)
MxA Input
Viral Lysate + MxA Input
Ladder (kDa)
Ladder (kDa)
Consecutive
Ultracentrifugation Fractions
200
150
100
75
50
37
25
20
15
PB1
PB2
PA
MxA
NP
(B)
Ladder (kDa)
vRNP Fractions
100
75
50
PB1
PB2
PA
NP
(C)
Ladder (kDa)
vRNP
Fractions + MxA
100
75
50
PB1
PB2
PA
MxA
NP
(D)
MxA Input
Viral Lysate + MxA Input
Ladder (kDa)
Ladder (kDa)
Boxed vRNP
+ MxA Fractions
150
100
75
α-MxA
50
37
25
20
α-NP

Figure 5.25: MxA-Bound vRNP Sample Preparation.

(A) Silver stain gel of MxA + vRNP binding experiment. MxA was added in excess to viral lysate and vRNP fractions were separated by ultracentrifugation. Fractions of vRNP were analysed by means of silver stain gel. vRNP-containing fractions are boxed and annotated by protein component. MxA is annotated in bold. (B) A comparison vRNP fraction portion of the silver stain gel in figure 5.8 with the constituent proteins annotated. (C) The vRNP + MxA fraction portion of the silver stain gel in panel (A) for comparison. Again, with MxA highlighted. (D) Two western blots (top) stained with an antibody specific for MxA, (bottom) stained with an antibody for nucleoprotein. MxA is shown to be detected in the input lanes and is observed to be present in the three fractions corresponding to the boxed vRNP + MxA fractions in (A) and (C). Silver stain and western blot courtesy of Dr. Kuang-Yu Chen.

Enriched MxA-vRNP complexes were subsequently prepared for cryo-EM analysis using the functionalised grid methodology described previously. Briefly, cryo-EM grids were functionalised with 0.06 mg/mL Nb8210 nanobody by Dr. Gangshun Yi. This nanobody specifically targets the PA linker of the influenza polymerase complex, enabling selective immobilisation of intact vRNPs while preserving their native architecture.

The MxA-vRNP sample was applied to functionalised grids, blotted, and plunge frozen according to the standard protocol. A representative micrograph (figure 5.26) demonstrated successful immobilisation of vRNPs with bound MxA.

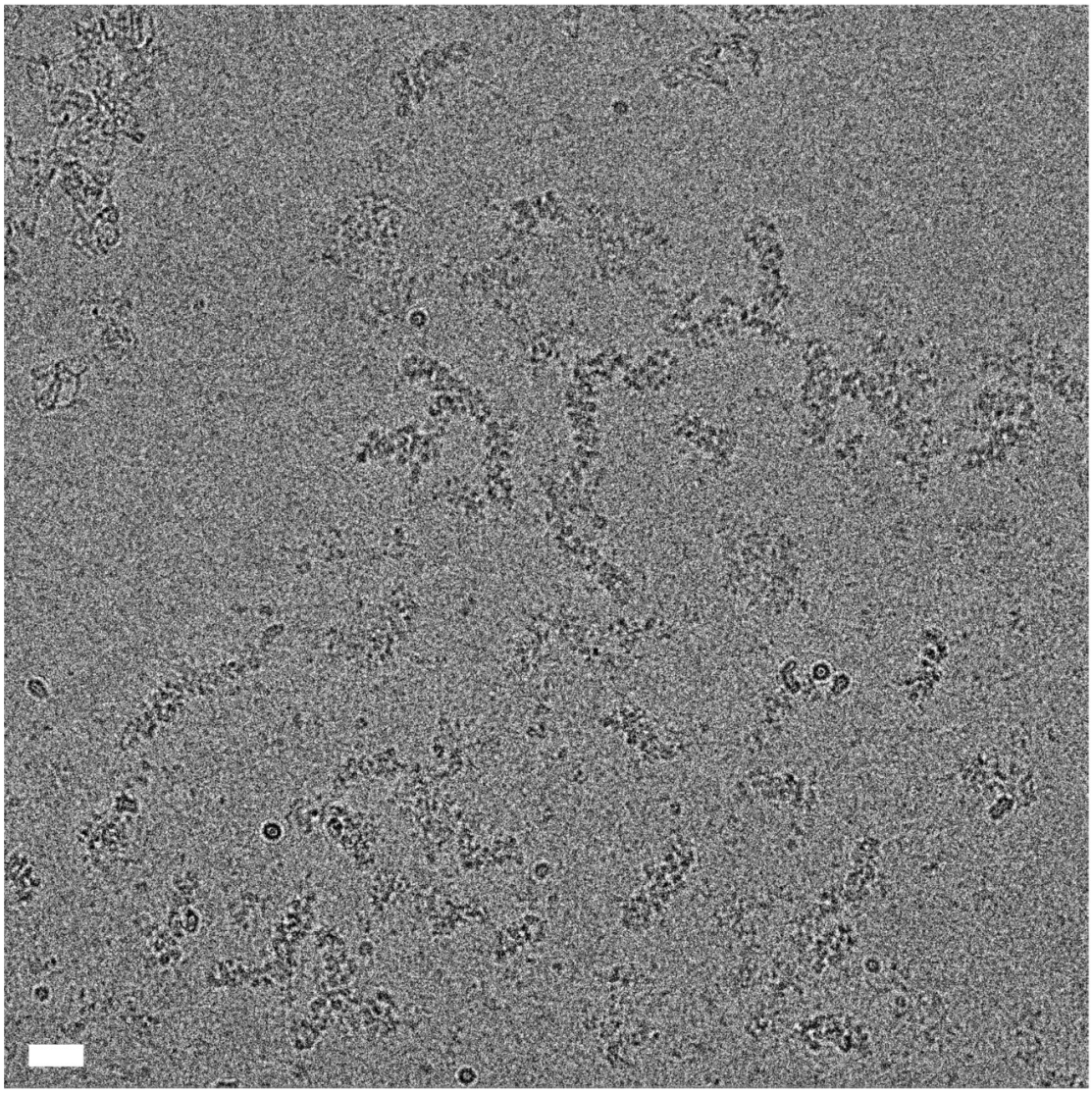


Figure 5.26: Polymerase Nanobody-Mediated Immobilisation of Fresh MxA-Bound vRNP Sample.

Representative micrograph of MxA-bound vRNPs immobilised on a functionalised cryo-EM grid using Nb8210 against the influenza polymerase PA linker. Scale bar = 200 Å.

Although enrichment efficiency was lower when compared with the Nb8210-immobilised vRNP sample lacking MxA (figure 5.24), this functionalised grid approach proved effective for concentrating MxA-bound vRNPs and preserving vRNP integrity. The successful preparation of these complexes established a promising basis for structural investigation and justified subsequent data collection.

5.2.6.2. Cryo-EM/ET Data Collection

Data of the MxA-bound vRNP sample were collected. Given that silver stain analysis (figure 5.25, panel A) revealed weaker MxA signal relative to nucleoprotein, it was unlikely that all vRNPs were saturated with MxA. Consistent with this, comparison of micrographs of vRNPs alone versus MxA-bound vRNP complexes revealed little to no discernible morphological difference, despite expectations that MxA binding might manifest as an additional ring- or helix-like density surrounding the nucleoprotein (Gao et al., 2010; Gao, von der Malsburg, et al., 2011; Haller and Kochs, 2011; Nigg and Pavlovic, 2015).

Owing to this presumed heterogeneity, cryo-ET was prioritised over cryo-EM SPA. Whereas cryo-EM SPA requires a large number of homogeneous particles for high-resolution reconstruction, cryo-ET reconstructs individual particles from tilt series, thereby retaining structural heterogeneity and preserving densities from partially bound states. This makes cryo-ET particularly suitable for sub-stoichiometric binding events, such as the MxA-vRNP binding observed in this sample.

Briefly, cryo-ET data were collected by Dr. Loïc Carrique using a Thermo Fisher Titan Krios 300 keV transmission electron microscope with a Selectris X energy filter and Falcon IVi detector. A dose-symmetrical tilt scheme covering a ± 54˚ range with 2˚ tilt steps using SerialEM, as detailed below (table 5.2). Notably, the zero tilt angle received a dose of 20 $e^-/Å^2$ whereas the non-zero tilt angles had a dose of 3 $e^-/Å^2$ in order to facilitate the initial processing of the zero tilt angles as a small cryo-EM dataset to assess whether MxA could be observed in the dataset.

A constant defocus of – 5.0 μM was used for data collection in order to maximise contrast and improve particle detection and alignment in sub tomogram averaging.

Table 5.2: Cryo-ET Data Collection Parameters for MxA-Bound vRNP Dataset.

Data Collection – vRNPs + MxA	
Microscope	Titan Krios
Voltage (keV)	300
Detector	Falcon IV
Magnification	105,000
Pixel Size (Å)	1.1871
Total Dose ($e^-/Å^2$)	138.6
Dose per Tilt ($e^-/Å^2$)	Zero Tilt = 20; Non-Zero Tilt = 3
Defocus (μm)	-5.0
Number of Tilt Series	104
Tilt Range	-54˚ to +54˚ in 2˚ Increments
Acquisition Scheme	Dose-Symmetric
Data Collection Software	SerialEM-PACE-tomo (Eisenstein et al., 2023)

5.2.6.2.1. Cryo-EM Data Processing

The 104 zero tilt movies were imported to cryoSPARC and underwent motion correction and CTF correction in cryoSPARC Live. Initial template picking was carried out in cryoSPARC Live using 2D class average templates from the previous dataset processed (figure 5.14, bottom panel, figure 5.27).

The exported 5811 particles underwent one round of 2D classification, yielding 2626 particles in 16 selected classes. These 2D class averages were used in a filament tracer job which identified 223,624 particles which were further inspected to remove non-genuine particles, and underwent another round of 2D classification, which yielded 3179 particles in 7 classes selected for their overall straightness (figure 5.27). *Ab-initio* jobs were attempted with these particles, but no density corresponding to

helical vRNP was produced, likely due to the low number of particles and their sub-optimal straightness. Additionally, whilst clear vRNP density was observed in the 2D class averages, at no point in this processing was any extra density attributable to MxA identified. This could be because the binding of MxA to this vRNP sample was sub-stoichiometric, and thus the density was blurred out in the 2D classification process, or because the MxA had dissociated from the vRNP or otherwise was not present in the sample complex.

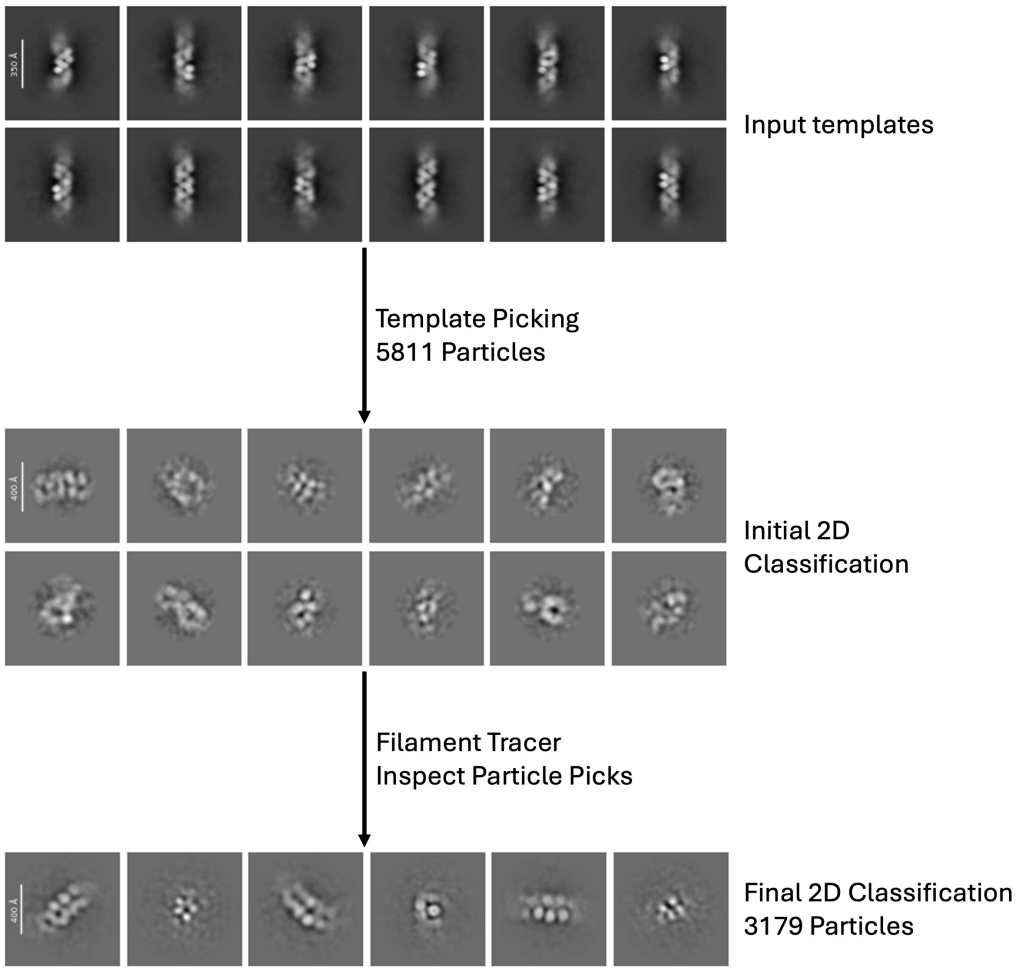


Figure 5.27: Processing Pipeline of Zero-Tilt Exposures of MxA-Bound vRNP Cryo-ET Data in cryoSPARC.

Processing of the cryo-ET data processing was undertaken with assistance from Misha Le Claire.

The 104 tilt series acquired from the MxA – vRNP dataset were pre-processed (motion corrected and CTF corrected) using Warp (Tegunov and Cramer, 2019), as detailed in the materials and methods chapter (section 2.5.5.). Initial tilt series alignment was performed using AreTomo (Zheng et al., 2022) using patch tracking (as no gold fiducials were added to this sample). Aligned tilt series were evaluated for alignment quality and particle distribution in IMOD (Mastronarde and Held, 2017). A subset of tomograms exhibited good sample distribution but poor alignment. These tomograms were manually re-aligned using Etomo (Mastronarde and Held, 2017). At the time of writing, tomogram reconstruction had been carried out, but no downstream processing had been completed.

Representative tomograms (figure 5.28) confirmed the presence of vRNP particles with preserved structural integrity, albeit at a reduced concentration when compared to earlier datasets using the functionalised grid approach (figure 5.24). Some degree of sample degradation was observed, which may be attributable to MxA binding. Given the established role of MxA as an antiviral effector targeting vRNP function, structural perturbation or destabilisation may represent a component of its inhibitory mechanism. However, additional structural analysis will be required to substantiate this hypothesis.

Importantly, the tomographic approach preserved particle heterogeneity and avoided averaging artifacts, thereby enabling future identification of sub-stoichiometric binding events. Compared with *in situ* cryo-ET, the use of purified complexes substantially reduces background noise and simplifies computational processing while maintaining near-native sample conditions.

Future processing of this dataset will focus on tomogram denoising and sub-tomogram averaging to assess the structural basis of MxA-vRNP interaction.

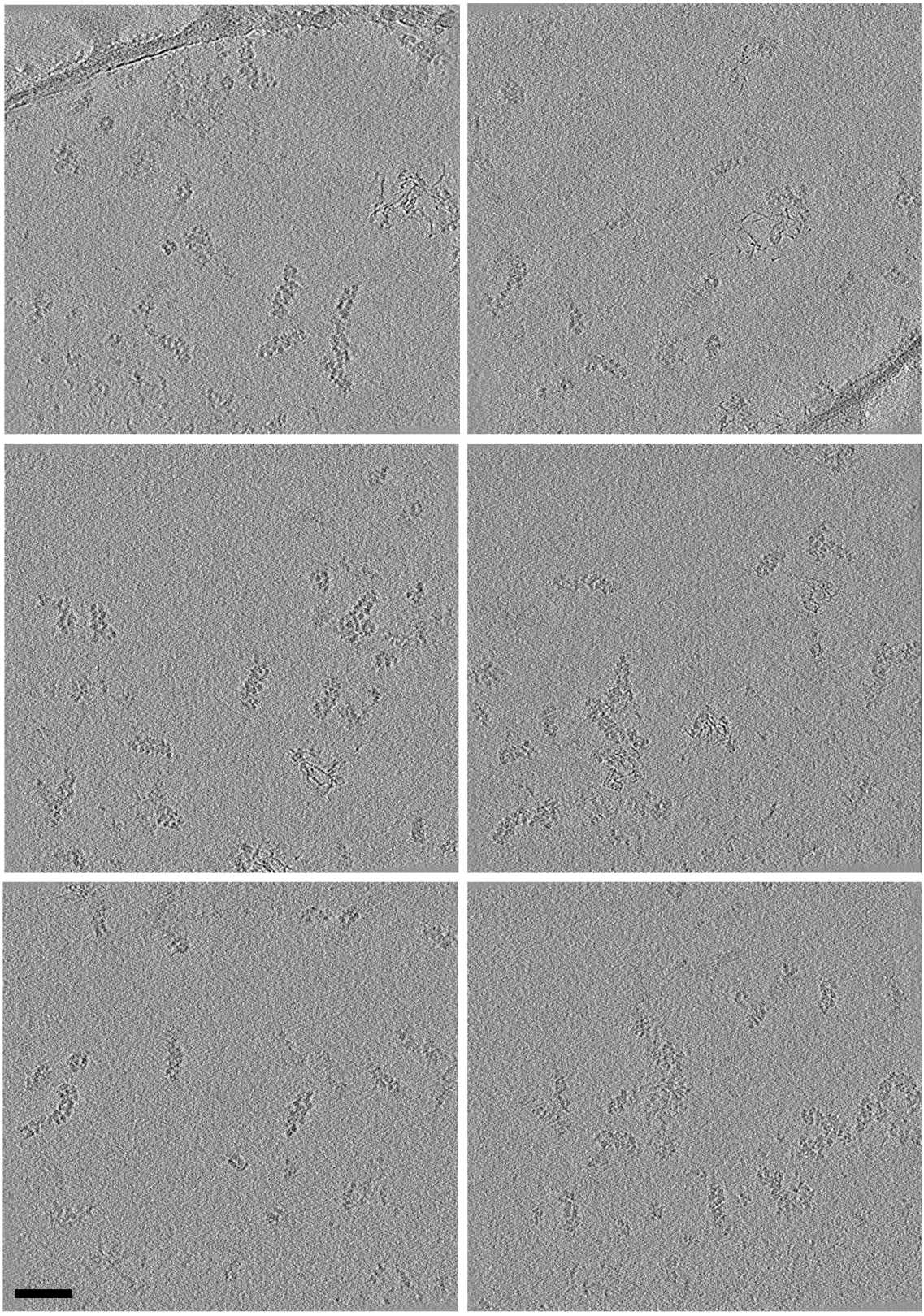


Figure 5.28: Example Reconstructed Tomogram Slices of MxA-Bound vRNP Dataset.

Scale bar conserved across all tomogram slices = 50 nm.

5.3. Discussion

The results presented in this chapter represent methodological progress in the structural characterisation of influenza vRNP complexes using cryo-electron microscopy techniques. This work addresses fundamental challenges in vRNP structural biology, notably the optimisation of sample preparation, concentration, and immobilisation strategies required for high-resolution imaging of these inherently flexible and fragile assemblies.

The enrichment protocol developed here enables the preparation of virus-derived vRNPs from both laboratory-adapted and chimeric influenza viruses, while maintaining the potential for specific binding-partner interactions and preserving native vRNP architecture. Attempts to concentrate enriched fractions demonstrated a clear requirement for high particle concentrations for cryo-EM analysis but also revealed that vRNP complexes are susceptible to degradation when bound by interaction partners such as Nb170 or MxA. These findings underscore the delicate balance between experimental tractability and biological stability in the study of vRNPs.

The implementation of crYOLO represents the first application of this approach to identify vRNP termini. While traditional manual or template-based picking methods performed poorly for filament ends, crYOLO successfully identified vRNP termini across the training dataset. Nevertheless, subsequent 2D classification did not resolve interpretable density corresponding to the influenza polymerase complex,

highlighting the intrinsic difficulty of resolving low-abundance components within heterogeneous samples using cryo-EM.

The use of carbon-support grids offered a strategic means to reduce filament flexibility by restricting motion in the Z-direction, thereby mitigating conformational heterogeneity. This approach proved to be effective in the Nb170-vRNP dataset which displayed sufficient particle concentrations and structural integrity.

However, optimisation of the Nb170-vRNP sample preparation revealed that conventional concentration methods consistently promoted aggregation and degradation, reflecting the vulnerability of vRNP-nanobody complexes to mechanical stress. The use of a 1:1 nanobody-to-nucleoprotein molar ratio provided a compromise between binding efficiency and minimal free nanobody during grid preparation, yielding samples of adequate quality but without clear evidence of nanobody density in reconstructions. This outcome suggests that a vast excess of binding partner is required for complete binding of the vRNP, but that such conditions may further destabilise vRNP complex structural integrity.

The application of the functionalised grid methodology as an ‘on-grid’ concentration approach represents a significant advancement for structural studies of vRNP filaments. This approach demonstrated substantial enrichment of material over control conditions while also avoiding the mechanical stress associated with the other concentration methods attempted earlier.

Comparing the nucleoprotein-mediated and polymerase-mediated immobilisation of vRNPs on functionalised grids, the polymerase-mediated immobilisation method was able to immobilise more sample with fewer signatures of degradation or aggregation. This differential performance is likely due to the specific targeting of the polymerase complex rather than nucleoprotein, as there is only one copy of the polymerase per vRNP, and immobilisation of free polymerase increases the background signal much less than immobilisation of free nucleoprotein.

The functionalised grid method was subsequently adopted for downstream experiments, establishing this approach as a standard component of the influenza vRNP preparation pipeline for analysis by cryo-EM.

The data presented on MxA binding to the vRNP represents the first biochemical evidence for direct interactions between MxA and the vRNP. The integrated data collection strategy, combining single particle cryo-EM analysis of zero-tilt images within a larger cryo-ET data collection represents an efficient approach for maximising information content.

While characteristic vRNP density was observed both on raw micrographs and in 2D class averages, the absence of additional density attributable to MxA indicates sub-stoichiometric binding between MxA and the vRNP. This observation is consistent with the weak band observed in the silver stain gel reported, and the calculated low molar ratio of MxA to nucleoprotein in the sample.

Initial processing of the cryo-ET dataset collected has produced well-aligned tomograms, but no morphological difference between the vRNP – MxA and other vRNP alone datasets have been observed. Data processing of this dataset will continue to produce 3D reconstructions of the vRNPs present and further analysis will be carried out to identify if extra density which could be attributed to MxA is present.

In future experiments, it would be prudent to first isolate highly concentrated vRNPs, determine their concentration, then add a binding partner in excess and repeat the vRNP separation ultracentrifugation step. This may lead to the loss of some sample, but the resulting vRNPs would likely be more enriched, and a saturating amount of MxA would be bound to the vRNP.

Another consideration is the involvement of GTP in MxA target binding. Given the importance of GTP binding and hydrolysis in conformational changes of MxA (but not necessarily binding to influenza nucleoprotein) (Paolo et al., 1999; Kochs et al., 2002; Accola et al., 2002; Haller et al., 2010), it would be well reasoned to add supplemental GTP into the vRNP binding partner binding step. In the experiment presented above, it was assumed a physiological concentration of GTP would be present in the viral lysate (when the vRNP and MxA are first interacting) and that this would be sufficient for GTP binding to MxA, but the effective GTP concentration is lowered significantly by the multiple ultracentrifugation steps.

Since this work was conceived, two relevant studies have been published, providing complementary insights into vRNP and vRNP-like architecture (Chenavier et al., 2025; Peng et al., 2025).

Chenavier and colleagues produced a so-called 'pseudo-vRNP' by recombinant production of an N-terminal truncation mutant influenza nucleoprotein (Δ1-14) and assembling this in a vRNP-like form by the addition of a synthetic fluorescently-tagged RNA molecule ($(UC)_{18}$-fluorescein$^{3'}$(FAM)). This specific assembly was found to produce an antiparallel double helix of influenza nucleoprotein with the $(UC)_{18}$-FAM RNA bound in the basic RNA binding groove. Using a standard reconstruction approach, initial volumes were reconstructed, helical symmetry was determined, and 3D classification in cryoSPARC resulted in the production of a 3.04 Å density map corresponding to two neighbouring copies of the nucleoprotein.

Critically, this work demonstrated that different nucleoprotein truncation variants and synthetic RNA sequences produce drastically different helical assemblies, including single helices, and multiple double helical forms including parallel double helices, and double helices without distinctive major/minor grooves. This variability may reflect the ability of the system to adapt to different RNA sequences and associated fluorophore molecules. However, these varying helical forms of synthetic RNPs observed are not found in the native state of the vRNP, thus emphasising the value of native (or near-native, such as the case of the reconstituted RNP) vRNP preparations, despite the inherent complexity that these systems exhibit.

Peng and colleagues reconstituted the influenza D virus NS-segment RNP using tandem affinity purification, achieving substantial sample enrichment and homogeneity. Their analyses demonstrated an antiparallel double-helical architecture, confirmed by both cryo-EM and cryo-ET sub-tomogram averaging. Despite improved resolution relative to prior studies, the intrinsic flexibility of influenza A vRNPs limited resolution gains. Nevertheless, their work convincingly resolved local nucleoprotein subunit arrangements and provided density suitable for fitting known polymerase structures.

Given such an enriched influenza D RNP sample, from their negative stain data alone, Peng and colleagues were able to identify the loop end of a number of RNP segments, and the polymerase both at the end and in the middle of the RNP.

Carrying out a similar pipeline to the one attempted in this study, the authors were able to carry out a global reconstruction of the RNP complex to reveal a double helical structure limited to nanometre resolution. The authors note that even in the context of such a short RNP, flexibility was still observed. By using both helical reconstruction and particle subtraction, the authors were able to produce a density map focussed on four nucleoprotein subunits resolved globally to 5.1 Å resolution, which facilitated the unambiguous fitting of the crystal structure of influenza D nucleoprotein.

Peng and colleagues additionally carried out cryo-ET and sub-tomogram averaging on virion-derived influenza A vRNPs through a similar density gradient ultracentrifugation protocol, carrying out sub-tomogram averaging on resulting vRNP segments to

compare the architecture of the native influenza A vRNP to the reconstituted influenza D RNP, finding broad similarities in overall double helical architecture, but unable to determine high-resolution structures due to the flexibility of the influenza A vRNP.

Additionally, from tomography experiments performed on the reconstituted influenza D system, Peng and colleagues were able to identify the polymerase associated to the RNP, and convincingly rigid body fit an influenza polymerase structure into the obtained density, further supporting the use of cryo-ET STA for the rigid body fitting of specific structures into resulting density maps from vRNPs.

Together, these studies converge on a consistent architectural model of influenza RNPs: the influenza vRNP adopts an antiparallel double helix in which RNA is encapsidated within the minor groove, with adjacent nucleoprotein subunits connected by flexible tail loops. Importantly, the nucleoprotein – nucleoprotein interfaces revealed in these structures provide a framework for the interpretation of binding mechanisms of both Nb170 and MxA.

Rigid body fitting of the molecular replacement solution of the Nb170 – nucleoprotein complex (and the resulting uncertainty underlying the molecular replacement solution) has already been discussed in section 3.3.4.

As in chapter 3, because a consistent architectural model of nucleoprotein packaging was converged upon across both studies, and because the helical reconstruction of the influenza D RNP offered significantly more information across one whole turn

(compared to much less information in other reconstructions) despite being a different virus, an analysis was made using both the helical packaging of the influenza D RNP and the helical reconstruction of the pseudo-vRNP reported by Chenavier and colleagues. Here, six copies of the S09 NP R416A structure solved in this thesis (section 3.2.5.) were rigid body fitted into each of the atomic models described and the known MxA sensitivity residues (positions 100, 283, and 313) were highlighted (figure 5.29).

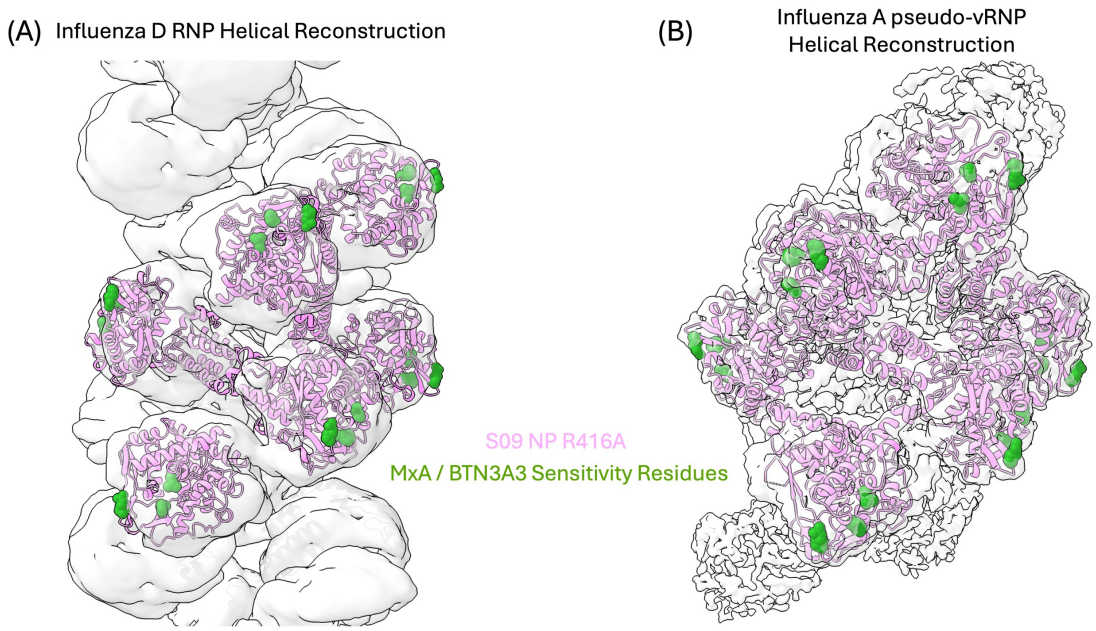


Figure 5.29: MxA Sensitivity Residues Highlighted in the Context of the vRNP.

Six copies of the S09 NP R416A structure solved in this thesis with MxA sensitivity residues (100, 283, 313) displayed as atoms in spheres and coloured in green. (A) Rigid body fitting of the S09 NP R416A as described above into the atomic model reported by Peng and colleagues of the influenza D RNP helical reconstruction with the reported density map coloured in grey and shaded at 30 % opacity (Peng et al., 2025) (PDB: 9C4H; EMDB: 44980). (B) Rigid body fitting of the S09 NP R416A as described above into the atomic model reported by Chenavier and colleagues of the influenza A pseudo-vRNP helical reconstruction (central 6-mer) with the reported density map coloured in grey and shaded at 30 % opacity (Chenavier et al., 2025) (PDB: 9GAT; EMDB: 51188).

In this rigid body fitting analysis, it is notable that multiple copies of the nucleoprotein structure, when fit into the influenza D RNP structure, clash with eachother via their

respective body domains. The packaging of the pseudo-vRNP reported by Chenavier and colleagues was compared, and no clash was shown but the same broad mode of nucleoprotein packaging was observed. Therefore, this observation likely reflects differences in the influenza A and D nucleoproteins and does not negate the conclusions made from this analysis.

Donchet and colleagues compared the structures of influenza virus nucleoproteins, and found that the overall folds of influenza A virus nucleoprotein was very similar to influenza D virus nucleoprotein, but an RMSD of 2.1 Å (356 Cα) was observed between the two nucleoproteins, with the most notable difference being a very short N-terminal unstructured region (thought to be compensated by a very long C-terminal unstructured region) when compared to influenza A virus nucleoprotein (Donchet et al., 2019).

In both cases, the three MxA sensitivity residues are presented on an exposed surface of the vRNP complex. If the mode of MxA oligomerisation and interaction with the vRNP suggested in the literature is correct (Kochs et al., 2002; Accola et al., 2002; Gao et al., 2010; Gao et al., 2011; von der Malsburg et al., 2011; Haller et al., 2015), it is likely that MxA would be able to oligomerise around vRNPs without the necessity for major structural rearrangements – preserving the double helical architecture of the vRNP.

Collectively, the methodological developments presented in this chapter establish a versatile framework for the structural study of the influenza vRNP and its interaction partners. They also highlight the fundamental challenges posed by vRNP heterogeneity,

flexibility, and instability, and emphasise the need for carefully optimised binding conditions to achieve interpretable reconstructions. The integration of native sample preparation, on-grid immobilisation strategies, and tomographic approaches offers a promising pathway toward resolving the structural basis of influenza vRNP regulation and restriction.

6. Concluding Remarks and Future Directions

This thesis presents an investigation of the influenza A virus nucleoprotein structure, and its interactions with innate immune factors, examining the nucleoprotein and its interactions both recombinantly as a monomer, and in the physiologically relevant form of the vRNP. My work encompasses three principal areas of research, outlined in the body of this thesis: the structural characterisation of the monomeric nucleoprotein and the production of a tractable scaffold for interaction studies with binding partners; the biochemical and biophysical investigation of the monomeric nucleoprotein with two interaction partners of the human innate immune system; and the structural investigation of virion-derived vRNPs and their binding to MxA, focussing on methods to address sample heterogeneity.

6.1. Methods to Study Influenza Nucleoprotein Structure

My work studying the structure of influenza nucleoproteins utilised multiple structural techniques. I attempted to crystallise the nucleoprotein both alone and with added RNA, but crystallisation was not possible for the nucleoprotein studied, despite a 98 % sequence identical nucleoprotein being readily able to form protein crystals prior in our laboratory.

I successfully produced and utilised the nucleoprotein nanobody Nb170 as a crystallisation chaperone to facilitate structural determination of the S09 NP R416A nucleoprotein. This nanobody-mediated crystallisation approach reproducibly yielded protein crystals. However, substantial crystal pathologies prevented

successful structure refinement despite a promising initial molecular replacement solution. The increasing R-factors during refinement indicated fundamental crystal quality issues that precluded reliable structure determination.

The di-gluebody platform was employed to attempt to form a stable scaffold of S09 NP R416A nucleoprotein for routine structural determination of novel nucleoproteins, nanobody binding site characterisation, and the study of other binding partner interactions. The di-gluebody platform is a powerful, emerging technology which can be further harnessed to form hetero-di-gluebodies, or higher order gluebody assemblies to act as an orienting fiducial in cryo-EM SPA in order to solve structural questions. However, the nanobodies selected for use in the di-gluebody study in this thesis suffered from significant flexibility, meaning that in this work, the di-gluebody platform was unable to be harnessed to solve the structure of nucleoproteins, or for the formation of reproducible scaffolds to study nucleoprotein binding partner interactions.

The megabody platform successfully enabled structural characterisation through the formation of a stable Mb151-S09 NP R416A complex analysed by cryo-EM SPA. Through focussed refinement, using particle subtraction and local refinement to target the nucleoprotein, I was able to obtain a 3.1 Å composite map of the megabody-nucleoprotein complex. The nanobody exhibited an unconventional mode of binding, predominantly mediated by C'' strand contacts. Key regions of the nucleoprotein including the C-terminus (residues 391-498) were not resolved, precluding analysis of

the flexible C-terminal loop. Globally, the nucleoprotein was highly similar to the previously solved NT60 NP R416A nucleoprotein, as reported.

6.2. Nucleoprotein Interaction Partners of the Human Innate Immune System

My work studying nucleoprotein interactions with human innate immune factors revealed direct binding between the S09 avian influenza nucleoprotein and both the MxA L4 loop and full-length MxA. This confirmed the L4 loop's identity as the minimal binding epitope for nucleoprotein-based restriction of sensitive influenza A virus infection.

While malfunctioning biosensor tips prevented complete kinetic characterisation, comparative Rmax measurements indicated preferential MxA binding to avian-origin nucleoproteins, consistent with literature describing enhanced sensitivity of avian versus human-adapted strains.

My work provides the first evidence for the direct interaction between the BTN3A3 B30.2 domain with sensitive avian influenza A nucleoproteins, confirming my hypothesis that the B30.2 domain mediates this binding interaction. My data additionally provide the first kinetic characterisation of B30.2 binding to susceptible avian-origin influenza A virus nucleoproteins, with notably absent binding to human-adapted variants, providing quantitative support for strain-specific restriction mechanisms.

6.3. Novel Approaches to Study Influenza vRNP Structure and its Interaction with Nucleoprotein Binding Partners

In this thesis, I have presented both historical published literature on influenza A virus vRNP structural biology and efforts previously made by members of this laboratory toward structural determination of the native vRNP structure. The structural heterogeneity of viral-derived vRNPs presents a major challenge for structural biology methods. In this work, I have presented my own efforts to enrich and rigidify virion-derived influenza A virus vRNPs, using a combination of carbon support EM grids and nanobody-mediated rigidification. I outlined a strategy for minimal nanobody usage to coat the vRNP whilst maintaining sufficient vRNP sample density on the EM grids for efficient data collection.

While processed data did not reveal nanobody density in 2D classification – likely due to insufficient nanobody binding – sample heterogeneity limitations prevented achieving resolution adequate for unambiguous nucleoprotein rigid-body fitting.

In sample preparation, the binding partner enrichment method detailed (section 5.2.1.) was validated as a successful method to coat vRNPs with nucleoprotein binding partners. Further, the functionalised grid approach detailed (section 5.2.5.) was validated as a tractable strategy to sufficiently enrich structurally-intact vRNPs for cryo-EM data collection. In using these techniques, the first biochemical evidence of direct MxA binding to enriched vRNPs has been presented. This methodology

represents a significant step forward in vRNP preparation and should be further harnessed with more homogeneous sample preparations to yield structural data of vRNPs and their interaction partners.

6.4. Future Directions

The methodological advances developed in this work provide a foundation for future vRNP structural studies, particularly when combined with more homogeneous sample preparations. The validated binding partner enrichment and functionalised grid approaches should be further developed to yield high-resolution structural data of vRNPs and their interaction partners.

The megabody platform was successfully used to determine the structure of the nucleoprotein, demonstrating its utility as a scaffold for the determination of protein structure at the practical limit of current cryo-EM capabilities. This validates the platform for future investigations of nucleoprotein-binding partner interactions.

The use of nanobodies is now common in structural biology, as demonstrated by the use of nanobodies and their derivatives in this thesis. The specific panel of nucleoprotein-binding nanobodies highlighted is of particular use in the study of influenza A virus biology. However, the use of nanobodies is not limited to structural techniques. For example, Rabouw and colleagues demonstrated their utility in live-cell single molecule microscopy to study the lifecycle of the vRNP in the context of

infection in a technique they name Virus Infection Real-time Imaging Negative-strand (VIRIM-neg) (Rabouw et al., 2025).

In this technique, Rabouw and colleagues employ an influenza A virus nucleoprotein-targeting nanobody generated by Schmidt and colleagues (Schmidt et al., 2016) - which does not interfere with the influenza A virus replication cycle - and fuse it to a green fluorescent protein. Rabouw and colleagues then generate a cell line expressing the chosen nanobody-fluorescent protein construct and were able to map individual vRNPs throughout the infection process. This technique represents a powerful real-time imaging cell biology technique which has the potential to generate impactful data.

For nucleoprotein-binding immune factor studies, where kinetic parameters were not able to be derived for MxA binding to the nucleoprotein, these experiments should be repeated with a wider range of MxA analyte concentrations with enhanced biosensor preparation protocols to minimise handling errors.

A critical limitation requiring attention is the exclusive use of monomeric influenza A virus nucleoprotein constructs. While heterogeneous oligomeric assemblies would preclude reliable kinetic analysis and introduce complex avidity effects – particularly relevant for higher-order oligomers like MxA – future studies should explore more controlled oligomeric systems to better recapitulate physiological binding conditions. Since this is also understood to be one mode of physiological interaction between the nucleoprotein and MxA (in the context of the vRNP), this was studied later with virion-derived vRNPs, but the binding kinetics were not investigated.

Fundamental questions regarding the mode of immune factor binding to the nucleoprotein remain unresolved. Despite demonstrated direct MxA-nucleoprotein interactions and established cellular localisation, no structural data exist for MxA bound to the influenza nucleoprotein. Similarly, while direct BTN3A3 B30.2 domain-nucleoprotein interaction has been demonstrated, structural characterisation remains elusive. Initial strategies should employ vast molar excesses of immune factors to favour stable complex formation, with crosslinking approaches such as gradient fixation (GraFix) additionally available as alternative complex stabilisation methods for cryo-EM analysis.

One critical mechanistic question concerns BTN3A3 nuclear transport. Studies of butyrophilin localisation within the cell have shown that butyrophilin family proteins are not solely limited to plasma membrane localisation (Karunakaran et al., 2023), supporting the idea that the butyrophilin family may in fact be mobile in the cell. Future investigations should employ *in situ* approaches including live cell imaging and cryo-electron tomography to elucidate the proposed transport of BTN3A3 to the nucleus.

As discussed, inherent structural flexibility of virus-derived vRNP complexes is a major limiting factor to high-resolution structure determination. Despite various attempts to minimise this observed heterogeneity, the inclusion of all eight genome segments further contributes to this heterogeneity.

Further processing of the MxA-bound vRNP dataset will be carried out in an attempt to locate MxA bound to the vRNP. If this should not be successful, optimisation of the vRNP coating protocol should be carried out in order to ensure the complete binding of the vRNP with the binding partner of interest.

In the final year of this work, two significant studies were published on the structure of influenza virus genome. As introduced in the body of the thesis, these involved cryo-EM SPA of reconstituted influenza D virus NS segment RNPs, and cryo-ET STA of influenza A virus vRNPs (Peng et al., 2025), and the recombinant formation of influenza A virus pseudo-vRNPs (Chenavier et al., 2025).

These findings present a paradigm shift from earlier head-to-body oligomerisation models (Coloma et al., 2020), revealing that nucleoprotein oligomerisation occurs through flexible loop insertion into the body domain of the neighbouring nucleoprotein. In both studies, a right-handed antiparallel double helical architecture of influenza genome segments is revealed, and similarly that nucleoprotein monomers are positioned with their RNA binding grooves oriented inwards towards the minor groove of the helix.

The novel approach taken by Peng and colleagues to selectively purify the influenza D NS segment led to minimisation of the heterogeneity observed with virion-derived vRNP preparations.

In future work, an influenza A virus NS segment reconstitution combined with the optimised saturation of the resulting RNP with binding partners and finally its enrichment on the grid by means of polymerase nanobody-mediated immobilisation, as developed in this thesis, would be a highly promising method of both determining the structure of the influenza A RNP in a controlled experimental set up, and also determining the mode of binding of immune factors of interest to the RNP. This would ensure minimal sample heterogeneity and maximal RNP occupancy by the binding partner chosen.

Peng and colleagues were additionally able to visualise the polymerase of influenza D virus RNPs using cryo-ET STA, which represents the optimal approach for polymerase visualisation, since signal averaging limitations of cryo-EM SPA are circumvented.

The issue with identifying one specific molecule in a biological complex of many similar molecules on a low-contrast tomogram or micrograph is highly challenging. This could be experimentally aided by the development of a fiducial marker for this specific molecule which suspends the molecule in an obvious way to make particle identification easier.

One potential approach utilises the gluebody technology platform, where higher order nanobody complexes are used to suspend multiple copies of the same molecule with a given orientation, distance, and angular positioning from a defined origin. The penta-gluebody enables the target of a gluebody to be suspended from a pentameric *E. coli* Shiga toxin subunit for the ease of particle identification and amplification of particle

number (Yi et al., 2025) and represents a potential method of accurately obtaining a high concentration of proteins of interest on the cryo-EM grid for structure determination.

Alternatively, atomic force microscopy (AFM) allows determination of ultrastructural information of purified samples including details up to the right-handedness of vRNPs (Nakano et al., 2021; Chenavier et al., 2023) and the ability to study conformational dynamics of recombinant RNPs (Carlero et al., 2024). This methodology could also be employed to determine the low-resolution ultrastructure of interaction partner-bound vRNPs to yield initial low-resolution data on such questions.

In situ cryo-electron tomography continues advancing to determine structural information within the context of live infections (Graham and Zhang, 2023). Unpublished work from the Grimes laboratory achieved sufficient resolution for the accurate determination of the handedness of the vRNP *in situ* (right-handed) by direct comparison to HA structures in the same dataset (Misha Le Claire, personal communication). Future developments in this field will enable improved mechanistic understanding of the influenza lifecycle and vRNP-host factor interactions in physiological contexts.

References

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., Bodenstein, S.W., Evans, D.A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., Beattie, C., Bertolli, O., Bridgland, A., Cherepanov, A., Congreve, M., Cowen-Rivers, A.I., Cowie, A., Figurnov, M., Fuchs, F.B., Gladman, H., Jain, R., Khan, Y.A., Low, C.M.R., Perlin, K., Potapenko, A., Savy, P., Singh, S., Stecula, A., Thillaisundaram, A., Tong, C., Yakneen, S., Zhong, E.D., Zielinski, M., Žídek, A., Bapst, V., Kohli, P., Jaderberg, M., Hassabis, D. and Jumper, J.M. 2024. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. **630**(8016), pp.493–500.

Accola, M.A., Huang, B., Masri, A.A. and McNiven, M.A. 2002. The Antiviral Dynamin Family Member, MxA, Tubulates Lipids and Localizes to the Smooth Endoplasmic Reticulum *. *Journal of Biological Chemistry*. **277**(24), pp.21829–21835.

Ackle, F., Thavarasah, S., Earp, J.C. and Seeger, M.A. 2025. Rigid enlargement of sybodies with antibody fragments for cryo-EM analyses of small membrane proteins. *Scientific Reports*. **15**(1), p.9460.

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C. and Zwart, P.H. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography*. **66**(2), pp.213–221.

Afrache, H., Gouret, P., Ainouche, S., Pontarotti, P. and Olive, D. 2012. The butyrophilin (BTN) gene family: from milk fat to the regulation of the immune response. *Immunogenetics*. **64**(11), pp.781–794.

Ahmad, J., Jiang, J., Boyd, L.F., Zeher, A., Huang, R., Xia, D., Natarajan, K. and Margulies, D.H. 2021. Structures of synthetic nanobody–SARS-CoV-2 receptor-binding domain complexes reveal distinct sites of interaction. *Journal of Biological Chemistry*. **297**(4), p.101202.

Akarsu, H., Burmeister, W.P., Petosa, C., Petit, I., Müller, C.W., Ruigrok, R.W.H. and Baudin, F. 2003. Crystal structure of the M1 protein-binding domain of the influenza A virus nuclear export protein (NEP/NS2). *The EMBO Journal*. **22**(18), pp.4646–4655.

Albertini, A.A.V., Wernimont, A.K., Muziol, T., Ravelli, R.B.G., Clapier, C.R., Schoehn, G., Weissenhorn, W. and Ruigrok, R.W.H. 2006. Crystal Structure of the Rabies Virus Nucleoprotein-RNA Complex. *Science*. **313**(5785), pp.360–363.

Álvarez, L., Haubrich, K., Iselin, L., Gillioz, L., Ruscica, V., Lapouge, K., Augsten, S., Huppertz, I., Choudhury, N.R., Simon, B., Masiewicz, P., Lethier, M., Cusack, S., Rittinger, K., Gabel, F., Leitner, A., Michlewski, G., Hentze, M.W., Allain, F.H.T., Castello, A. and Hennig, J. 2024. The molecular dissection of TRIM25's RNA-binding mechanism provides key insights into its antiviral activity. *Nature Communications*. **15**(1), p.8485.

Amorim, M.J. 2019. A Comprehensive Review on the Interaction Between the Host GTPase Rab11 and Influenza A Virus. *Frontiers in Cell and Developmental Biology*. **6**.

Amorim, M.J., Bruce, E.A., Read, E.K.C., Foeglein, Á., Mahen, R., Stuart, A.D. and Digard, P. 2011. A Rab11- and Microtubule-Dependent Mechanism for Cytoplasmic Transport of Influenza A Virus Viral RNA. *Journal of Virology*. **85**(9), pp.4143–4156.

Aoshi, T., Koyama, S., Kobiyama, K., Akira, S. and Ishii, K.J. 2011. Innate and adaptive immune responses to viral infection and vaccination. *Current Opinion in Virology*. **1**(4), pp.226–232.

Arai, Y., Ibrahim, M.S., Elgendy, E.M., Daidoji, T., Ono, T., Suzuki, Y., Nakaya, T., Matsumoto, K. and Watanabe, Y. 2019. Genetic Compatibility of Reassortants between Avian H5N1 and H9N2 Influenza Viruses with Higher Pathogenicity in Mammals. *Journal of Virology*. **93**(4), pp.e01969-18.

Area, E., Martín-Benito, J., Gastaminza, P., Torreira, E., Valpuesta, J.M., Carrascosa, J.L. and Ortín, J. 2004. 3D structure of the influenza virus polymerase complex: Localization of subunit domains. *Proceedings of the National Academy of Sciences of the United States of America*. **101**(1), pp.308–313.

Arnett, H.A. and Viney, J.L. 2014. Immune modulation by butyrophilins. *Nature Reviews Immunology*. **14**(8), pp.559–569.

Arragain, B., Krischuns, T., Pelosse, M., Drncova, P., Blackledge, M., Naffakh, N. and Cusack, S. 2024. Structures of influenza A and B replication complexes give insight into avian to human host adaptation and reveal a role of ANP32 as an electrostatic chaperone for the apo-polymerase. *Nature Communications*. **15**(1), p.6910.

Arragain, B., Reguera, J., Desfosses, A., Gutsche, I., Schoehn, G. and Malet, H. 2019. High resolution cryo-EM structure of the helical RNA-bound Hantaan virus nucleocapsid reveals its assembly mechanisms E. H. Egelman & J. Kuriyan, eds. *eLife*. **8**, p.e43075.

Arranz, R., Coloma, R., Chichón, F.J., Conesa, J.J., Carrascosa, J.L., Valpuesta, J.M., Ortín, J. and Martín-Benito, J. 2012. The Structure of Native Influenza Virion Ribonucleoproteins. *Science*. **338**(6114), pp.1634–1637.

Ashenberg, O., Padmakumar, J., Doud, M.B. and Bloom, J.D. 2017. Deep mutational scanning identifies sites in influenza nucleoprotein that affect viral inhibition by MxA. *PLOS Pathogens*. **13**(3), p.e1006288.

Bailey, E.S., Choi, J.Y., Fieldhouse, J.K., Borkenhagen, L.K., Zemke, J., Zhang, D. and Gray, G.C. 2018. The continual threat of influenza virus infections at the human–animal interface: What is new from a one health perspective? *Evolution, Medicine, and Public Health*. **2018**(1), pp.192–198.

Bannas, P., Hambach, J. and Koch-Nolte, F. 2017. Nanobodies and Nanobody-Based Human Heavy Chain Antibodies As Antitumor Therapeutics. *Frontiers in Immunology*. **8**.

Baudin, F., Bach, C., Cusack, S. and Ruigrok, R.W. 1994. Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *The EMBO journal*. **13**(13), pp.3158–3165.

Beghein, E. and Gettemans, J. 2017. Nanobody Technology: A Versatile Toolkit for Microscopic Imaging, Protein–Protein Interaction Analysis, and Protein Function Exploration. *Frontiers in Immunology*. **8**, p.771.

Benfield, C.T.O., Lyall, J.W., Kochs, G. and Tiley, L.S. 2008. Asparagine 631 Variants of the Chicken Mx Protein Do Not Inhibit Influenza Virus Replication in Primary Chicken Embryo Fibroblasts or In Vitro Surrogate Assays. *Journal of Virology*. **82**(15), pp.7533–7539.

Benjamin, C.J., Wright, K.J., Bolton, S.C., Hyun, S.-H., Krynski, K., Grover, M., Yu, G., Guo, F., Kinzer-Ursem, T.L., Jiang, W. and Thompson, D.H. 2016. Selective Capture of Histidine-tagged Proteins from Cell Lysates Using TEM grids Modified with NTA-Graphene Oxide. *Scientific Reports*. **6**(1), p.32500.

Bepler, T., Kelley, K., Noble, A.J. and Berger, B. 2020. Topaz-Denoise: general deep denoising models for cryoEM and cryoET. *Nature Communications*. **11**(1), p.5208.

Bhattacharya, M., Chatterjee, S., Lee, S.-S. and Chakraborty, C. 2023. Therapeutic applications of nanobodies against SARS-CoV-2 and other viral infections: Current update. *International Journal of Biological Macromolecules*. **229**, pp.70–80.

Bishop, D.H.L., Obijeski, J.F. and Simpson, R.W. 1971. Transcription of the Influenza Ribonucleic Acid Genome by a Virion Polymerase I. Optimal Conditions for In Vitro Activity of the Ribonucleic Acid-Dependent Ribonucleic Acid Polymerase. *Journal of Virology*. **8**(1), pp.66–73.

Blazquez, J.-L., Benyamine, A., Pasero, C. and Olive, D. 2018. New Insights Into the Regulation of γδ T Cells by BTN3A and Other BTN/BTNL in Tumor Immunity. *Frontiers in Immunology*. **9**.

Boivin, S., Kozak, S. and Meijers, R. 2013. Optimization of protein purification and characterization using Thermofluor screens. *Protein Expression and Purification*. **91**(2), pp.192–206.

Boudinot, P., Langevin, C., Secombes, C.J. and Levraud, J.-P. 2016. The Peculiar Characteristics of Fish Type I Interferons. *Viruses*. **8**(11), p.298.

Bouvier, N.M. and Palese, P. 2008. The biology of influenza viruses. *Vaccine*. **26**, pp.D49–D53.

Brocchieri, L. and Karlin, S. 2005. Protein length in eukaryotic and prokaryotic proteomes. *Nucleic Acids Research*. **33**(10), pp.3390–3400.

Brunotte, L., Flies, J., Bolte, H., Reuther, P., Vreede, F. and Schwemmle, M. 2014. The Nuclear Export Protein of H5N1 Influenza A Viruses Recruits Matrix 1 (M1) Protein to the Viral Ribonucleoprotein to Mediate Nuclear Export*. *Journal of Biological Chemistry*. **289**(29), pp.20067–20077.

Buchholz, T.-O., Jordan, M., Pigino, G. and Jug, F. 2019. Cryo-CARE: Content-Aware Image Restoration for Cryo-Transmission Electron Microscopy Data *In*: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)* [Online]., pp.502–506. [Accessed 31 July 2025]. Available from: https://ieeexplore.ieee.org/abstract/document/8759519.

Budziszewski, G.R., Snell, M.E., Wright, T.R., Lynch, M.L. and Bowman, S.E.J. 2023. High-Throughput Screening to Obtain Crystal Hits for Protein Crystallography. *Journal of visualized experiments : JoVE*. (193), 10.3791/65211.

Busiek, K.K. and Margolin, W. 2015. Bacterial Actin and Tubulin Homologs in Cell Growth and Division. *Current Biology*. **25**(6), pp.R243–R254.

Carlero, D., Fukuda, S., Bocanegra, R., Ando, T., Martin-Benito, J. and Ibarra, B. 2024. Conformational Dynamics of Influenza A Virus Ribonucleoprotein Complexes during RNA Synthesis. *ACS Nano*. **18**(30), pp.19518–19527.

Carrique, L., Fan, H., Walker, A.P., Keown, J.R., Sharps, J., Staller, E., Barclay, W.S., Fodor, E. and Grimes, J.M. 2020. Host ANP32A mediates the assembly of the influenza virus replicase. *Nature*. **587**(7835), pp.638–643.

Castells-Graells, R., Meador, K., Arbing, M.A., Sawaya, M.R., Gee, M., Cascio, D., Gleave, E., Debreczeni, J.É., Breed, J., Leopold, K., Patel, A., Jahagirdar, D., Lyons, B., Subramaniam, S., Phillips, C. and Yeates, T.O. 2023. Cryo-EM structure determination of small therapeutic protein targets at 3 Å-resolution using a rigid imaging scaffold. *Proceedings of the National Academy of Sciences*. **120**(37), p.e2305494120.

Chan, W.-H., Ng, A.K.-L., Robb, N.C., Lam, M.K.-H., Chan, P.K.-S., Au, S.W.-N., Wang, J.-H., Fodor, E. and Shaw, P.-C. 2010. Functional Analysis of the Influenza Virus H5N1 Nucleoprotein Tail Loop Reveals Amino Acids That Are Crucial for

Oligomerization and Ribonucleoprotein Activities. *Journal of Virology*. **84**(14), pp.7337–7345.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. and Richardson, D.C. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*. **66**(1), pp.12–21.

Chen, Y., Zhang, L., Graf, L., Yu, B., Liu, Y., Kochs, G., Zhao, Y. and Gao, S. 2017. Conformational dynamics of dynamin-like MxA revealed by single-molecule FRET. *Nature Communications*. **8**(1), p.15744.

Chenavas, S., Estrozi, L.F., Slama-Schwok, A., Delmas, B., Primo, C.D., Baudin, F., Li, X., Crépin, T. and Ruigrok, R.W.H. 2013. Monomeric Nucleoprotein of Influenza A Virus. *PLOS Pathogens*. **9**(3), p.e1003275.

Chenavier, F., Estrozi, L.F., Teulon, J.-M., Zarkadas, E., Freslon, L.-L., Pellequer, J.-L., Ruigrok, R.W.H., Schoehn, G., Ballandras-Colas, A. and Crépin, T. 2023. Cryo-EM structure of influenza helical nucleocapsid reveals NP-NP and NP-RNA interactions as a model for the genome encapsidation. *Science Advances*. **9**(50), p.eadj9974.

Chenavier, F., Zarkadas, E., Freslon, L.-L., Stelfox, A.J., Schoehn, G., Ruigrok, R.W.H., Ballandras-Colas, A. and Crépin, T. 2025. Influenza a virus antiparallel helical nucleocapsid-like pseudo-atomic structure. *Nucleic Acids Research*. **53**(3), p.gkae1211.

Chiu, M.L., Goulet, D.R., Teplyakov, A. and Gilliland, G.L. 2019. Antibody Structure and Function: The Basis for Engineering Therapeutics. *Antibodies*. **8**(4), p.55.

Chlanda, P., Schraidt, O., Kummer, S., Riches, J., Oberwinkler, H., Prinz, S., Kräusslich, H.-G. and Briggs, J.A.G. 2015. Structural Analysis of the Roles of Influenza A Virus Membrane-Associated Proteins in Assembly and Morphology. *Journal of Virology*. **89**(17), pp.8957–8966.

Chou, Y., Heaton, N.S., Gao, Q., Palese, P., Singer, R. and Lionnet, T. 2013. Colocalization of Different Influenza Viral RNA Segments in the Cytoplasm before Viral Budding as Shown by Single-molecule Sensitivity FISH Analysis. *PLOS Pathogens*. **9**(5), p.e1003358.

Ciminski, K., Pulvermüller, J., Adam, J. and Schwemmle, M. 2019. Human MxA is a potent interspecies barrier for the novel bat-derived influenza A-like virus H18N11. *Emerging Microbes & Infections*. **8**(1), pp.556–563.

Coloma, R., Arranz, R., de la Rosa-Trevín, J.M., Sorzano, C.O.S., Munier, S., Carlero, D., Naffakh, N., Ortín, J. and Martín-Benito, J. 2020. Structural insights into influenza A virus ribonucleoproteins reveal a processive helical track as transcription mechanism. *Nature Microbiology*. **5**(5), pp.727–734.

Coloma, R., Valpuesta, J.M., Arranz, R., Carrascosa, J.L., Ortín, J. and Martín-Benito, J. 2009. The Structure of a Biologically Active Influenza Virus Ribonucleoprotein Complex. *PLoS Pathogens*. **5**(6), p.e1000491.

Compans, R.W., Content, J. and Duesberg, P.H. 1972. Structure of the Ribonucleoprotein of Influenza Virus. *Journal of Virology*. **10**(4), pp.795–800.

Dadonaite, B., Gilbertson, B., Knight, M.L., Trifkovic, S., Rockman, S., Laederach, A., Brown, L.E., Fodor, E. and Bauer, D.L.V. 2019. The structure of the influenza A virus genome. *Nature Microbiology*. **4**(11), pp.1781–1789.

D'Cruz, A.A., Babon, J.J., Norton, R.S., Nicola, N.A. and Nicholson, S.E. 2013. Structure and function of the SPRY/B30.2 domain proteins involved in innate immunity. *Protein Science*. **22**(1), pp.1–10.

De Felice, S., Romanyuk, Z., Chinellato, M., Zoia, G., Linciano, S., Kumada, Y., Pardon, E., Steyaert, J., Angelini, A. and Cendron, L. 2024. Crystal structure of human serum albumin in complex with megabody reveals unique human and murine cross-reactive binding site. *Protein Science*. **33**(2), p.e4887.

De Marco, M.A., Delogu, M. and Cotti, C. 2023. Special Issue "Ecology of Influenza A Viruses": Editorial. *Microorganisms*. **11**(5), p.1287.

De Rosier, D.J. and Klug, A. 1968. Reconstruction of Three Dimensional Structures from Electron Micrographs. *Nature*. **217**(5124), pp.130–134.

Deeg, C.M., Hassan, E., Mutz, P., Rheinemann, L., Götz, V., Magar, L., Schilling, M., Kallfass, C., Nürnberger, C., Soubies, S., Kochs, G., Haller, O., Schwemmle, M. and Staeheli, P. 2017. In vivo evasion of MxA by avian influenza viruses requires human signature in the viral nucleoprotein. *Journal of Experimental Medicine*. **214**(5), pp.1239–1248.

Denz, P.J., Speaks, S., Kenney, A.D., Eddy, A.C., Papa, J.L., Roettger, J., Scace, S.C., Rubrum, A., Hemann, E.A., Forero, A., Webby, R.J., Bowman, A.S. and Yount, J.S. 2024. Innate immune control of influenza virus interspecies adaptation via IFITM3. *Nature Communications*. **15**(1), p.9375.

Desfosses, A., Milles, S., Jensen, M.R., Guseva, S., Colletier, J.-P., Maurin, D., Schoehn, G., Gutsche, I., Ruigrok, R.W.H. and Blackledge, M. 2019. Assembly and cryo-EM structures of RNA-specific measles virus nucleocapsids provide mechanistic insight into paramyxoviral replication. *Proceedings of the National Academy of Sciences*. **116**(10), pp.4256–4264.

Dick, A., Graf, L., Olal, D., Malsburg, A. von der, Gao, S., Kochs, G. and Daumke, O. 2015. Role of Nucleotide Binding and GTPase Domain Dimerization in Dynamin-like Myxovirus Resistance Protein A for GTPase Activation and Antiviral Activity *. *Journal of Biological Chemistry*. **290**(20), pp.12779–12792.

Dick, A., Mikirtumov, V., Fuchs, J., Krupp, F., Olal, D., Bendl, E., Sprink, T., Diebolder, C., Kudryashev, M., Kochs, G., Roske, Y. and Daumke, O. 2024. Structural characterization of Thogoto Virus nucleoprotein provides insights into viral RNA encapsidation and RNP assembly. *Structure*. **32**(8), pp.1068-1078.e5.

Donchet, A., Oliva, J., Labaronne, A., Tengo, L., Miloudi, M., C. A. Gerard, F., Mas, C., Schoehn, G., W. H. Ruigrok, R., Ducatez, M. and Crépin, T. 2019. The structure of the nucleoprotein of Influenza D shows that all Orthomyxoviridae nucleoproteins have a similar NPCORE, with or without a NPTAIL for nuclear transport. *Scientific Reports*. **9**(1), p.600.

Dornfeld, D., Petric, P.P., Hassan, E., Zell, R. and Schwemmle, M. 2019. Eurasian Avian-Like Swine Influenza A Viruses Escape Human MxA Restriction through Distinct Mutations in Their Nucleoprotein. *Journal of Virology*. **93**(2), pp.e00997-18.

Drulyte, I., Johnson, R.M., Hesketh, E.L., Hurdiss, D.L., Scarff, C.A., Porav, S.A., Ranson, N.A., Muench, S.P. and Thompson, R.F. 2018. Approaches to altering particle distributions in cryo-electron microscopy sample preparation. *Acta Crystallographica Section D: Structural Biology*. **74**(6), pp.560–571.

Dunbar, J., Krawczyk, K., Leem, J., Marks, C., Nowak, J., Regep, C., Georges, G., Kelm, S., Popovic, B. and Deane, C.M. 2016. SAbPred: a structure-based antibody prediction server. *Nucleic Acids Research*. **44**(W1), pp.W474–W478.

Egelman, E.H. 2015. Three-dimensional reconstruction of helical polymers. *Archives of Biochemistry and Biophysics*. **581**, pp.54–58.

Eisenstein, F., Yanagisawa, H., Kashihara, H., Kikkawa, M., Tsukita, S. and Danev, R. 2023. Parallel cryo electron tomography on in situ lamellae. *Nature Methods*. **20**(1), pp.131–138.

Elton, D., Medcalf, L., Bishop, K., Harrison, D. and Digard, P. 1999. Identification of amino acid residues of influenza virus nucleoprotein essential for RNA binding. *Journal of Virology*. **73**(9), pp.7357–7367.

Emsley, P. and Cowtan, K. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D: Biological Crystallography*. **60**(12), pp.2126–2132.

Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. 2010. Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*. **66**(4), pp.486–501.

Eshak, F. and Goupil-Lamy, A. 2025. Advancements in Nanobody Epitope Prediction: A Comparative Study of AlphaFold2Multimer vs AlphaFold3. *Journal of Chemical Information and Modeling*. **65**(4), pp.1782–1797.

Fan, H., Walker, A.P., Carrique, L., Keown, J.R., Serna Martin, I., Karia, D., Sharps, J., Hengrung, N., Pardon, E., Steyaert, J., Grimes, J.M. and Fodor, E. 2019.

Structures of influenza A virus RNA polymerase offer insight into viral genome replication. *Nature*. **573**(7773), pp.287–290.

Fodor, E., Devenish, L., Engelhardt, O.G., Palese, P., Brownlee, G.G. and García-Sastre, A. 1999. Rescue of Influenza A Virus from Recombinant DNA. *Journal of Virology*. **73**(11), pp.9679–9682.

Fontana, J. and Steven, A.C. 2013. At Low pH, Influenza Virus Matrix Protein M1 Undergoes a Conformational Change Prior to Dissociating from the Membrane. *Journal of Virology*. **87**(10), pp.5621–5628.

Fournier, E., Moules, V., Essere, B., Paillart, J.-C., Sirbat, J.-D., Isel, C., Cavalier, A., Rolland, J.-P., Thomas, D., Lina, B. and Marquet, R. 2012. A supramolecular assembly formed by influenza A virus genomic RNA segments. *Nucleic Acids Research*. **40**(5), pp.2197–2209.

Franke, W.W., Heid, H.W., Grund, C., Winter, S., Freudenstein, C., Schmid, E., Jarasch, E.D. and Keenan, T.W. 1981. Antibodies to the major insoluble milk fat globule membrane-associated protein: specific location in apical regions of lactating epithelial cells. *The Journal of Cell Biology*. **89**(3), pp.485–494.

Franklin, R.E. 1955. Structure of Tobacco Mosaic Virus. *Nature*. **175**(4452), pp.379–381.

Franklin, R.E. 1956. Structure of Tobacco Mosaic Virus: Location of the Ribonucleic Acid in the Tobacco Mosaic Virus Particle. *Nature*. **177**(4516), pp.928–930.

Fribourgh, J.L., Nguyen, H.C., Matreyek, K.A., Alvarez, F.J.D., Summers, B.J., Dewdney, T.G., Aiken, C., Zhang, P., Engelman, A. and Xiong, Y. 2014. Structural Insight into HIV-1 Restriction by MxB. *Cell Host & Microbe*. **16**(5), pp.627–638.

Fridy, P.C., Li, Y., Keegan, S., Thompson, M.K., Nudelman, I., Scheid, J.F., Oeffinger, M., Nussenzweig, M.C., Fenyö, D., Chait, B.T. and Rout, M.P. 2014. A robust pipeline for rapid production of versatile nanobody repertoires. *Nature Methods*. **11**(12), pp.1253–1260.

Gabriel, G., Klingel, K., Otte, A., Thiele, S., Hudjetz, B., Arman-Kalcek, G., Sauter, M., Shmidt, T., Rother, F., Baumgarte, S., Keiner, B., Hartmann, E., Bader, M., Brownlee, G.G., Fodor, E. and Klenk, H.-D. 2011. Differential use of importin-α isoforms governs cell tropism and host adaptation of influenza virus. *Nature Communications*. **2**(1), p.156.

Gack, M.U., Albrecht, R.A., Urano, T., Inn, K.-S., Huang, I.-C., Carnero, E., Farzan, M., Inoue, S., Jung, J.U. and García-Sastre, A. 2009. Influenza A Virus NS1 Targets the Ubiquitin Ligase TRIM25 to Evade Recognition by the Host Viral RNA Sensor RIG-I. *Cell Host & Microbe*. **5**(5), pp.439–449.

Gallagher, J.R., Torian, U., McCraw, D.M. and Harris, A.K. 2017. Structural studies of influenza virus RNPs by electron microscopy indicate molecular contortions within NP supra-structures. *Journal of Structural Biology*. **197**(3), pp.294–307.

Gamblin, S.J., Vachieri, S.G., Xiong, X., Zhang, J., Martin, S.R. and Skehel, J.J. 2021. Hemagglutinin Structure and Activities. *Cold Spring Harbor Perspectives in Medicine*. **11**(10), p.a038638.

Gao, S., von der Malsburg, A., Haller, O., Kochs, G. and Daumke, O. 2011. Structural basis of oligomerization and the mechano-chemical function in dynamin-like MxA. *The FASEB Journal*. **25**(S1), 752.4-752.4.

Gao, S., von der Malsburg, A., Paeschke, S., Behlke, J., Haller, O., Kochs, G. and Daumke, O. 2010. Structural basis of oligomerization in the stalk region of dynamin-like MxA. *Nature*. **465**(7297), pp.502–506.

Gao, S., von der Malsburg, A., Dick, A., Faelber, K., Schröder, G.F., Haller, O., Kochs, G. and Daumke, O. 2011. Structure of Myxovirus Resistance Protein A Reveals Intra- and Intermolecular Domain Interactions Required for the Antiviral Function. *Immunity*. **35**(4), pp.514–525.

Garg, S., Reed, C., Davis, C.T., Uyeki, T.M., Behravesh, C.B., Kniss, K., Budd, A., Biggerstaff, M., Adjemian, J., Barnes, J.R., Kirby, M.K., Basler, C., Szablewski, C.M., Richmond-Crum, M., Burns, E., Limbago, B., Daskalakis, D.C., Armstrong, K., Boucher, D., Shimabukuro, T.T., Jhung, M.A., Olsen, S.J. and Dugan, V. 2024. Outbreak of Highly Pathogenic Avian Influenza A(H5N1) Viruses in U.S. Dairy Cattle and Detection of Two Human Cases - United States, 2024. *MMWR. Morbidity and mortality weekly report*. **73**(21), pp.501–505.

Ge, P., Tsao, J., Schein, S., Green, T.J., Luo, M. and Zhou, Z.H. 2010. Cryo-EM Model of the Bullet-Shaped Vesicular Stomatitis Virus. *Science*. **327**(5966), pp.689–693.

Gildea, R.J., Beilsten-Edmands, J., Axford, D., Horrell, S., Aller, P., Sandy, J., Sanchez-Weatherby, J., Owen, C.D., Lukacik, P., Strain-Damerell, C., Owen, R.L., Walsh, M.A. and Winter, G. 2022. xia2.multiplex: a multi-crystal data-analysis pipeline. *Acta Crystallographica Section D: Structural Biology*. **78**(6), pp.752–769.

Glaeser, R.M. 2018. Proteins, interfaces, and cryo-EM grids. *Current Opinion in Colloid & Interface Science*. **34**, pp.1–8.

Goot, J. a. V.D., Jong, M.C.M.D., Koch, G. and Boven, M.V. 2003. Comparison of the transmission characteristics of low and high pathogenicity avian influenza A virus (H5N2). *Epidemiology & Infection*. **131**(2), pp.1003–1013.

Götz, V., Magar, L., Dornfeld, D., Giese, S., Pohlmann, A., Höper, D., Kong, B.-W., Jans, D.A., Beer, M., Haller, O. and Schwemmle, M. 2016. Influenza A viruses escape from MxA restriction at the expense of efficient nuclear vRNP import. *Scientific Reports*. **6**(1), p.23138.

Goujon, C., Moncorgé, O., Bauby, H., Doyle, T., Ward, C.C., Schaller, T., Hué, S., Barclay, W.S., Schulz, R. and Malim, M.H. 2013. Human MX2 is an interferon-induced post-entry inhibitor of HIV-1 infection. *Nature*. **502**(7472), pp.559–562.

Graham, M. and Zhang, P. 2023. Cryo-electron tomography to study viral infection. *Biochemical Society Transactions*. **51**(4), pp.1701–1711.

Green, T.J., Zhang, X., Wertz, G.W. and Luo, M. 2006. Structure of the Vesicular Stomatitis Virus Nucleoprotein-RNA Complex. *Science*. **313**(5785), pp.357–360.

Gu, S., Borowska, M.T., Boughter, C.T. and Adams, E.J. 2018. Butyrophilin3A proteins and Vγ9Vδ2 T cell activation. *Seminars in Cell & Developmental Biology*. **84**, pp.65–74.

Gutsche, I., Desfosses, A., Effantin, G., Ling, W.L., Haupt, M., Ruigrok, R.W.H., Sachse, C. and Schoehn, G. 2015. Near-atomic cryo-EM structure of the helical measles virus nucleocapsid. *Science*. **348**(6235), pp.704–707.

Haas, M., Mills, J.T., Kelley, C., Das, A., Hóf, H., Niemel, C., Donselaar, T., Drulyte, I., Kuppeveld, F.J.M. van, Dulin, D., Snijder, J., Herod, M.R. and Hurdiss, D.L. 2025. Integrative Structure of Norovirus NS3 Suggests a Role in RNA Transport. , 2025.06.17.659504.

Haller, O., Gao, S., Malsburg, A. von der, Daumke, O. and Kochs, G. 2010. Dynamin-like MxA GTPase: Structural Insights into Oligomerization and Implications for Antiviral Activity *. *Journal of Biological Chemistry*. **285**(37), pp.28419–28424.

Haller, O. and Kochs, G. 2011. Human MxA protein: an interferon-induced dynamin-like GTPase with broad antiviral activity. *Journal of Interferon & Cytokine Research: The Official Journal of the International Society for Interferon and Cytokine Research*. **31**(1), pp.79–87.

Haller, O. and Kochs, G. 2020. Mx genes: host determinants controlling influenza virus infection and trans-species transmission. *Human Genetics*. **139**(6), pp.695–705.

Haller, O., Staeheli, P., Schwemmle, M. and Kochs, G. 2015. Mx GTPases: dynamin-like antiviral machines of innate immunity. *Trends in Microbiology*. **23**(3), pp.154–163.

Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hammers, C., Songa, E.B., Bendahman, N. and Hammers, R. 1993. Naturally occurring antibodies devoid of light chains. *Nature*. **363**(6428), pp.446–448.

Hanke, L., Knockenhauer, K.E., Brewer, R.C., van Diest, E., Schmidt, F.I., Schwartz, T.U. and Ploegh, H.L. 2016. The Antiviral Mechanism of an Influenza A Virus Nucleoprotein-Specific Single-Domain Antibody Fragment. *mBio*. **7**(6), 10.1128/mbio.01569-16.

Harly, C., Guillaume, Y., Nedellec, S., Peigné, C.-M., Mönkkönen, H., Mönkkönen, J., Li, J., Kuball, J., Adams, E.J., Netzer, S., Déchanet-Merville, J., Léger, A., Herrmann, T., Breathnach, R., Olive, D., Bonneville, M. and Scotet, E. 2012. Key implication

of CD277/butyrophilin-3 (BTN3A) in cellular stress sensing by a major human γδ T-cell subset. *Blood*. **120**(11), pp.2269–2279.

Hefti, H.P., Frese, M., Landis, H., Di Paolo, C., Aguzzi, A., Haller, O. and Pavlovic, J. 1999. Human MxA Protein Protects Mice Lacking a Functional Alpha/Beta Interferon System against La Crosse Virus and Other Lethal Viral Infections. *Journal of Virology*. **73**(8), pp.6984–6991.

Henderson, R. 1995. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Quarterly Reviews of Biophysics*. **28**(2), pp.171–193.

Hengrung, N., El Omari, K., Serna Martin, I., Vreede, F.T., Cusack, S., Rambo, R.P., Vonrhein, C., Bricogne, G., Stuart, D.I., Grimes, J.M. and Fodor, E. 2015. Crystal structure of the RNA-dependent RNA polymerase from influenza C virus. *Nature*. **527**(7576), pp.114–117.

Herzik, M.A., Wu, M. and Lander, G.C. 2019. High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM. *Nature Communications*. **10**(1), p.1032.

Hoffmann, E., Neumann, G., Hobom, G., Webster, R.G. and Kawaoka, Y. 2000. “Ambisense” Approach for the Generation of Influenza A Virus: vRNA and mRNA Synthesis from One Template. *Virology*. **267**(2), pp.310–317.

Holzinger, D., Jorns, C., Stertz, S., Boisson-Dupuis, S., Thimme, R., Weidmann, M., Casanova, J.-L., Haller, O. and Kochs, G. 2007. Induction of MxA Gene Expression by Influenza A Virus Requires Type I or Type III Interferon Signaling. *Journal of Virology*. **81**(14), pp.7776–7785.

Honda, A., Uéda, K., Nagata, K. and Ishihama, A. 1988. RNA polymerase of influenza virus: role of NP in RNA chain elongation. *Journal of Biochemistry*. **104**(6), pp.1021–1026.

Hopkins, F.R., Álvarez-Rodríguez, B., Heath, G.R., Panayi, K., Hover, S., Edwards, T.A., Barr, J.N. and Fontana, J. 2022. The Native Orthobunyavirus Ribonucleoprotein Possesses a Helical Architecture. *mBio*. **13**(4), pp.e01405-22.

Hu, Y.-B., Dammer, E.B., Ren, R.-J. and Wang, G. 2015. The endosomal-lysosomal system: from acidification and cargo sorting to neurodegeneration. *Translational Neurodegeneration*. **4**, p.18.

Huang, S., Chen, Jingjing, Chen, Q., Wang, H., Yao, Y., Chen, Jianjun and Chen, Z. 2013. A Second CRM1-Dependent Nuclear Export Signal in the Influenza A Virus NS2 Protein Contributes to the Nuclear Export of Viral Ribonucleoproteins. *Journal of Virology*. **87**(2), pp.767–778.

Huotari, J. and Helenius, A. 2011. Endosome maturation. *The EMBO Journal*. **30**(17), pp.3481–3500.

Hutchinson, E.C. and Fodor, E. 2013. Transport of the Influenza Virus Genome from Nucleus to Nucleus. *Viruses*. **5**(10), pp.2424–2446.

Iuliano, A.D., Roguski, K.M., Chang, H.H., Muscatello, D.J., Palekar, R., Tempia, S., Cohen, C., Gran, J.M., Schanzer, D., Cowling, B.J., Wu, P., Kyncl, J., Ang, L.W., Park, M., Redlberger-Fritz, M., Yu, H., Espenhain, L., Krishnan, A., Emukule, G., van Asten, L., Pereira da Silva, S., Aungkulanon, S., Buchholz, U., Widdowson, M.-A., Bresee, J.S., and Global Seasonal Influenza-associated Mortality Collaborator Network 2018. Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet (London, England)*. **391**(10127), pp.1285–1300.

Jamali, K., Käll, L., Zhang, R., Brown, A., Kimanius, D. and Scheres, S.H.W. 2024. Automated model building and protein identification in cryo-EM maps. *Nature*. **628**(8007), pp.450–457.

Janzen, C., Kochs, G. and Haller, O. 2000. A Monomeric GTPase-Negative MxA Mutant with Antiviral Activity. *Journal of Virology*. **74**(17), pp.8202–8206.

Javanian, M., Barary, M., Ghebrehewet, S., Koppolu, V., Vasigala, V. and Ebrahimpour, S. 2021. A brief review of influenza virus infection. *Journal of Medical Virology*. **93**(8), pp.4638–4646.

Jenni, S., Horwitz, J.A., Bloyet, L.-M., Whelan, S.P.J. and Harrison, S.C. 2022. Visualizing molecular interactions that determine assembly of a bullet-shaped vesicular stomatitis virus particle. *Nature Communications*. **13**, p.4802.

Jimah, J.R. and Hinshaw, J.E. 2019. Structural Insights into the Mechanism of Dynamin Superfamily Proteins. *Trends in Cell Biology*. **29**(3), pp.257–273.

Jin, B., Odongo, S., Radwanska, M. and Magez, S. 2023. NANOBODIES®: A Review of Generation, Diagnostics and Therapeutics. *International Journal of Molecular Sciences*. **24**(6), p.5994.

Jin, H., Leser, G.P., Zhang, J. and Lamb, R.A. 1997. Influenza virus hemagglutinin and neuraminidase cytoplasmic tails control particle shape. *The EMBO Journal*. **16**(6), pp.1236–1247.

Johnson, N.P.A.S. and Mueller, J. 2002. Updating the Accounts: Global Mortality of the 1918-1920 'Spanish' Influenza Pandemic. *Bulletin of the History of Medicine*. **76**(1), pp.105–115.

Joseph, U., Su, Y.C.F., Vijaykrishna, D. and Smith, G.J.D. 2017. The ecology and adaptive evolution of influenza A interspecies transmission. *Influenza and Other Respiratory Viruses*. **11**(1), pp.74–84.

Jovčevska, I. and Muyldermans, S. 2020. The Therapeutic Potential of Nanobodies. *BioDrugs*. **34**(1), pp.11–26.

Kane, M., Yadav, S.S., Bitzegeio, J., Kutluay, S.B., Zang, T., Wilson, S.J., Schoggins, J.W., Rice, C.M., Yamashita, M., Hatziioannou, T. and Bieniasz, P.D. 2013. MX2 is an interferon-induced inhibitor of HIV-1 infection. *Nature*. **502**(7472), pp.563–566.

Kao, R.Y., Yang, D., Lau, L.-S., Tsui, W.H.W., Hu, L., Dai, J., Chan, M.-P., Chan, C.-M., Wang, P., Zheng, B.-J., Sun, J., Huang, J.-D., Madar, J., Chen, G., Chen, H., Guan, Y. and Yuen, K.-Y. 2010. Identification of influenza A nucleoprotein as an antiviral target. *Nature Biotechnology*. **28**(6), pp.600–605.

Karasik, A. and Guydosh, N.R. 2024. The Unusual Role of Ribonuclease L in Innate Immunity. *Wiley Interdisciplinary Reviews. RNA*. **15**(6), p.e1878.

Karunakaran, M.M., Subramanian, H., Jin, Y., Mohammed, F., Kimmel, B., Juraske, C., Starick, L., Nöhren, A., Länder, N., Willcox, C.R., Singh, R., Schamel, W.W., Nikolaev, V.O., Kunzmann, V., Wiemer, A.J., Willcox, B.E. and Herrmann, T. 2023. A distinct topology of BTN3A IgV and B30.2 domains controlled by juxtamembrane regions favors optimal human γδ T cell phosphoantigen sensing. *Nature Communications*. **14**(1), p.7617.

Kawaguchi, A., Matsumoto, K. and Nagata, K. 2012. YB-1 Functions as a Porter To Lead Influenza Virus Ribonucleoprotein Complexes to Microtubules. *Journal of Virology*. **86**(20), pp.11086–11095.

Kelly, D.F., Abeyrathne, P.D., Dukovski, D. and Walz, T. 2008. The Affinity Grid: A Pre-fabricated EM Grid for Monolayer Purification. *Journal of Molecular Biology*. **382**(2), pp.423–433.

Kelly, D.F., Dukovski, D. and Walz, T. 2010. Strategy for the Use of Affinity Grids to Prepare Non-His-Tagged Macromolecular Complexes for Single-Particle Electron Microscopy. *Journal of Molecular Biology*. **400**(4), pp.675–681.

Keown, J.R., Zhu, Z., Carrique, L., Fan, H., Walker, A.P., Serna Martin, I., Pardon, E., Steyaert, J., Fodor, E. and Grimes, J.M. 2022. Mapping inhibitory sites on the RNA polymerase of the 1918 pandemic influenza virus using nanobodies. *Nature Communications*. **13**(1), p.251.

Kim, H., Webster, R.G. and Webby, R.J. 2018. Influenza Virus: Dealing with a Drifting and Shifting Pathogen. *Viral Immunology*. **31**(2), pp.174–183.

Klein, S., Golani, G., Lolicato, F., Lahr, C., Beyer, D., Herrmann, A., Wachsmuth-Melm, M., Reddmann, N., Brecht, R., Hosseinzadeh, M., Kolovou, A., Makroczyova, J., Peterl, S., Schorb, M., Schwab, Y., Brügger, B., Nickel, W., Schwarz, U.S. and Chlanda, P. 2023. IFITM3 blocks influenza virus entry by sorting lipids and stabilizing hemifusion. *Cell Host & Microbe*. **31**(4), pp.616-633.e20.

Knight, M.L., Fan, H., Bauer, D.L.V., Grimes, J.M., Fodor, E. and Keown, J.R. 2021. Structure of an H3N2 influenza virus nucleoprotein. *Acta Crystallographica Section F: Structural Biology Communications*. **77**(7), pp.208–214.

Kochs, G., Janzen, C., Hohenberg, H. and Haller, O. 2002. Antivirally active MxA protein sequesters La Crosse virus nucleocapsid protein into perinuclear complexes. *Proceedings of the National Academy of Sciences*. **99**(5), pp.3153–3158.

Kolb, E., Laine, E., Strehler, D. and Staeheli, P. 1992. Resistance to influenza virus infection of Mx transgenic mice expressing Mx protein under the control of two constitutive promoters. *Journal of Virology*. **66**(3), pp.1709–1716.

Koliopoulos, M.G., Lethier, M., van der Veen, A.G., Haubrich, K., Hennig, J., Kowalinski, E., Stevens, R.V., Martin, S.R., Reis e Sousa, C., Cusack, S. and Rittinger, K. 2018. Molecular mechanism of influenza A NS1-mediated TRIM25 recognition and inhibition. *Nature Communications*. **9**(1), p.1820.

Kouba, T., Drncová, P. and Cusack, S. 2019. Structural snapshots of actively transcribing influenza polymerase. *Nature Structural & Molecular Biology*. **26**(6), pp.460–470.

Koyama, S., Ishii, K.J., Coban, C. and Akira, S. 2008. Innate immune response to viral infection. *Cytokine*. **43**(3), pp.336–341.

Kreutzberger, M.A.B., Sonani, R.R. and Egelman, E.H. 2024. Cryo-EM reconstruction of helical polymers: Beyond the simple cases. *Quarterly Reviews of Biophysics*. **57**, p.e16.

Krissinel, E. and Henrick, K. 2007. Inference of Macromolecular Assemblies from Crystalline State. *Journal of Molecular Biology*. **372**(3), pp.774–797.

Kueh, H.Y. and Mitchison, T.J. 2009. Structural Plasticity in Actin and Tubulin Polymer Dynamics. *Science*. **325**(5943), pp.960–963.

Kumlin, U., Olofsson, S., Dimock, K. and Arnberg, N. 2008. Sialic acid tissue distribution and influenza virus tropism. *Influenza and Other Respiratory Viruses*. **2**(5), pp.147–154.

Labaronne, A., Swale, C., Monod, A., Schoehn, G., Crépin, T. and Ruigrok, R.W.H. 2016. Binding of RNA by the Nucleoproteins of Influenza Viruses A and B. *Viruses*. **8**(9), p.247.

Lakdawala, S.S., Wu, Y., Wawrzusin, P., Kabat, J., Broadbent, A.J., Lamirande, E.W., Fodor, E., Altan-Bonnet, N., Shroff, H. and Subbarao, K. 2014. Influenza A Virus Assembly Intermediates Fuse in the Cytoplasm. *PLOS Pathogens*. **10**(3), p.e1003971.

Langley, C.A., Dietzen, P.A., Emerman, M., Tenthorey, J.L. and Malik, H.S. 2025. Antiviral Mx proteins have an ancient origin and widespread distribution among eukaryotes. *Proceedings of the National Academy of Sciences*. **122**(4), p.e2416811122.

Lara, B., Rico, A.I., Petruzzelli, S., Santona, A., Dumas, J., Biton, J., Vicente, M., Mingorance, J. and Massidda, O. 2005. Cell division in cocci: localization and properties of the Streptococcus pneumoniae FtsA protein. *Molecular Microbiology*. **55**(3), pp.699–711.

Lebrero Fernández, C. 2016. *Butyrophilin- and Butyrophilin-like genes and their role in epithelial cell-intraepithelial T lymphocyte cross-talk* [Online]. [Accessed 19 September 2025]. Available from: https://gupea.ub.gu.se/handle/2077/41848.

Li, N., Zhang, L., Chen, L., Feng, W., Xu, Y., Chen, F., Liu, X., Chen, Z. and Liu, W. 2012. MxA inhibits hepatitis B virus replication by interaction with hepatitis B core antigen. *Hepatology*. **56**(3), pp.803–811.

Li, X., Gu, M., Zheng, Q., Gao, R. and Liu, X. 2021. Packaging signal of influenza A virus. *Virology Journal*. **18**(1), p.36.

Li, Y., Wang, X., Blau, D.M., Caballero, M.T., Feikin, D.R., Gill, C.J., Madhi, S.A., Omer, S.B., Simões, E.A.F., Campbell, H., Pariente, A.B., Bardach, D., Bassat, Q., Casalegno, J.-S., Chakhunashvili, G., Crawford, N., Danilenko, D., Do, L.A.H., Echavarria, M., Gentile, A., Gordon, A., Heikkinen, T., Huang, Q.S., Jullien, S., Krishnan, A., Lopez, E.L., Markić, J., Mira-Iglesias, A., Moore, H.C., Moyes, J., Mwananyanda, L., Nokes, D.J., Noordeen, F., Obodai, E., Palani, N., Romero, C., Salimi, V., Satav, A., Seo, E., Shchomak, Z., Singleton, R., Stolyarov, K., KStoszek, S., von Gottberg, A., Wurzel, D., Yoshida, L.-M., Yung, C.F., Zar, H.J. and Nair, H. 2022. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in children younger than 5 years in 2019: a systematic analysis. *Lancet (London, England)*. **399**(10340), pp.2047–2064.

Li, Z., Watanabe, T., Hatta, M., Watanabe, S., Nanbo, A., Ozawa, M., Kakugawa, S., Shimojima, M., Yamada, S., Neumann, G. and Kawaoka, Y. 2009. Mutational analysis of conserved amino acids in the influenza A virus nucleoprotein. *Journal of Virology*. **83**(9), pp.4153–4162.

Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A.J., Moriarty, N.W., Oeffner, R.D., Poon, B.K., Prisant, M.G., Read, R.J., Richardson, J.S., Richardson, D.C., Sammito, M.D., Sobolev, O.V., Stockwell, D.H., Terwilliger, T.C., Urzhumtsev, A.G., Videau, L.L., Williams, C.J. and Adams, P.D. 2019. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica Section D: Structural Biology*. **75**(10), pp.861–877.

Lindenmann, J. 2005. Of Mice and Men – The Mx Connection *In*: *Comprehensive Biochemistry* [Online]. Elsevier, pp.267–295. [Accessed 18 August 2025]. Available from: https://www.sciencedirect.com/science/article/pii/S0069803205440048.

Lindenmann, J. 1962. Resistance of mice to mouse-adapted influenza A virus. *Virology*. **16**(2), pp.203–204.

Liu, G., Park, H.-S., Pyo, H.-M., Liu, Q. and Zhou, Y. 2015. Influenza A Virus Panhandle Structure Is Directly Involved in RIG-I Activation and Interferon Induction. *Journal of Virology*. **89**(11), pp.6067–6079.

Long, J.S., Mistry, B., Haslam, S.M. and Barclay, W.S. 2019. Host and viral determinants of influenza A virus species specificity. *Nature Reviews Microbiology*. **17**(2), pp.67–81.

Lorieau, J.L., Louis, J.M. and Bax, A. 2010. The complete influenza hemagglutinin fusion domain adopts a tight helical hairpin arrangement at the lipid:water interface. *Proceedings of the National Academy of Sciences of the United States of America*. **107**(25), pp.11341–11346.

Lousa, D. and Soares, C.M. 2021. Molecular mechanisms of the influenza fusion peptide: insights from experimental and simulation studies. *FEBS Open Bio*. **11**(12), pp.3253–3261.

Lövestam, S., Li, D., Wagstaff, J.L., Kotecha, A., Kimanius, D., McLaughlin, S.H., Murzin, A.G., Freund, S.M.V., Goedert, M. and Scheres, S.H.W. 2024. Disease-specific tau filaments assemble via polymorphic intermediates. *Nature*. **625**(7993), pp.119–125.

Lu, A., Magupalli, V., Ruan, J., Yin, Q., Atianand, M.K., Vos, M., Schröder, G.F., Fitzgerald, K.A., Wu, H. and Egelman, E.H. 2014. Unified Polymerization Mechanism for the Assembly of ASC-dependent Inflammasomes. *Cell*. **156**(6), pp.1193–1206.

Luo, M., Terrell, J.R. and Mcmanus, S.A. 2020. Nucleocapsid Structure of Negative Strand RNA Virus. *Viruses*. **12**(8), p.835.

MacCosham, A., Vasiliu, A.G. and Atchessi, N. 2025. A rapid review of the avian influenza PB2 E627K mutation in human infection studies. *Canada Communicable Disease Report*. **51**(4), pp.137–144.

Malawski, G.A., Hillig, R.C., Monteclaro, F., Eberspaecher, U., Schmitz, A.A.P., Crusius, K., Huber, M., Egner, U., Donner, P. and Müller-Tiemann, B. 2006. Identifying protein construct variants with increased crystallization propensity—A case study. *Protein Science : A Publication of the Protein Society*. **15**(12), pp.2718–2728.

von der Malsburg, A., Abutbul-Ionita, I., Haller, O., Kochs, G. and Danino, D. 2011. Stalk domain of the dynamin-like MxA GTPase protein mediates membrane binding and liposome tubulation via the unstructured L4 loop. *The Journal of Biological Chemistry*. **286**(43), pp.37858–37865.

Manso, T., Folch, G., Giudicelli, V., Jabado-Michaloud, J., Kushwaha, A., Nguefack Ngoune, V., Georga, M., Papadaki, A., Debbagh, C., Pégorier, P., Bertignac, M., Hadi-Saljoqi, S., Chentli, I., Cherouali, K., Aouinti, S., El Hamwi, A., Albani, A., Elazami Elhassani, M., Viart, B., Goret, A., Tran, A., Sanou, G., Rollin, M., Duroux, P. and Kossida, S. 2022. IMGT® databases, related tools and web resources

through three main axes of research and development. *Nucleic Acids Research*. **50**(D1), pp.D1262–D1272.

Mänz, B., Dornfeld, D., Götz, V., Zell, R., Zimmermann, P., Haller, O., Kochs, G. and Schwemmle, M. 2013. Pandemic Influenza A Viruses Escape from Restriction by Human MxA through Adaptive Mutations in the Nucleoprotein. *PLOS Pathogens*. **9**(3), p.e1003279.

Marques, M.A., Purdy, M.D. and Yeager, M. 2019. CryoEM maps are full of potential. *Current Opinion in Structural Biology*. **58**, pp.214–223.

Martin, I.S., Hengrung, N., Renner, M., Sharps, J., Martínez-Alonso, M., Masiulis, S., Grimes, J.M. and Fodor, E. 2018. A Mechanism for the Activation of the Influenza Virus Transcriptase. *Molecular Cell*. **70**(6), pp.1101-1110.e4.

Martín-Benito, J., Area, E., Ortega, J., Llorca, O., Valpuesta, J.M., Carrascosa, J.L. and Ortín, J. 2001. Three-dimensional reconstruction of a recombinant influenza virus ribonucleoprotein particle. *EMBO Reports*. **2**(4), pp.313–317.

Mastronarde, D.N. and Held, S.R. 2017. Automated tilt series alignment and tomographic reconstruction in IMOD. *Journal of Structural Biology*. **197**(2), pp.102–113.

Matreyek, K.A., Wang, W., Serrao, E., Singh, P.K., Levin, H.L. and Engelman, A. 2014. Host and viral determinants for MxB restriction of HIV-1 infection. *Retrovirology*. **11**(1), p.90.

Matzinger, S.R., Carroll, T.D., Dutra, J.C., Ma, Z.-M. and Miller, C.J. 2013. Myxovirus Resistance Gene A (MxA) Expression Suppresses Influenza A Virus Replication in Alpha Interferon-Treated Primate Cells. *Journal of Virology*. **87**(2), pp.1150–1158.

Maveyraud, L. and Mourey, L. 2020. Protein X-ray Crystallography and Drug Discovery. *Molecules*. **25**(5), p.1030.

McAuley, J.L., Gilbertson, B.P., Trifkovic, S., Brown, L.E. and McKimm-Breschkin, J.L. 2019. Influenza Virus Neuraminidase Structure and Functions. *Frontiers in Microbiology*. **10**.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. 2007. Phaser crystallographic software. *Journal of Applied Crystallography*. **40**(4), pp.658–674.

McKellar, J., Arnaud-Arnould, M., Chaloin, L., Tauziet, M., Arpin-André, C., Pourcelot, O., Blaise, M., Moncorgé, O. and Goujon, C. 2023. An evolutionarily conserved N-terminal leucine is essential for MX1 GTPase antiviral activity against different families of RNA viruses. *Journal of Biological Chemistry*. **299**(1).

McKellar, J., Cadènes, J., García de Gracia, F., Aubé, C., Chaves Valadão, A.L., Tauziet, M., Arnaud-Arnould, M., Rebendenne, A., Kociánová, L., Labaronne, E., Ricci, E.P., Delaval, B., Gaudin, R., Naffakh, N., Gallois-Montbrun, S., Moncorgé, O. and Goujon, C. 2025. Human MX1 orchestrates the cytoplasmic sequestration of neosynthesized influenza A virus vRNPs. *Proceedings of the National Academy of Sciences*. **122**(41), p.e2418935122.

McKellar, J., Gracia, F.G. de, Aubé, C., Valadão, A.L.C., Tauziet, M., Arnaud-Arnould, M., Rebendenne, A., Neyret, A., Labaronne, E., Ricci, E., Delaval, B., Gaudin, R., Naffakh, N., Gallois-Montbrun, S., Moncorgé, O. and Goujon, C. 2024. Human MX1 orchestrates the cytoplasmic sequestration of neo-synthesized influenza A virus vRNPs. , 2024.02.22.581565.

Meyerson, N.R., Zhou, L., Guo, Y.R., Zhao, C., Tao, Y.J., Krug, R.M. and Sawyer, S.L. 2017. Nuclear TRIM25 Specifically Targets Influenza Virus Ribonucleoproteins to Block the Onset of RNA Chain Elongation. *Cell Host & Microbe*. **22**(5), pp.627-638.e7.

Mibayashi, M., Nakad, K. and Nagata, K. 2002. Promoted cell death of cells expressing human MxA by influenza virus infection. *Microbiology and Immunology*. **46**(1), pp.29–36.

Minta, A.A. 2024. Progress Toward Measles Elimination — Worldwide, 2000–2023. *MMWR. Morbidity and Mortality Weekly Report*. **73**.

Mitchell, L.S. and Colwell, L.J. 2018a. Analysis of nanobody paratopes reveals greater diversity than classical antibodies. *Protein engineering, design & selection: PEDS*. **31**(7–8), pp.267–275.

Mitchell, L.S. and Colwell, L.J. 2018b. Comparative analysis of nanobody sequence and structure data. *Proteins*. **86**(7), pp.697–706.

Mitchell, P.S., Patzina, C., Emerman, M., Haller, O., Malik, H.S. and Kochs, G. 2012. Evolution-Guided Identification of Antiviral Specificity Determinants in the Broadly Acting Interferon-Induced Innate Immunity Factor MxA. *Cell Host & Microbe*. **12**(4), pp.598–604.

Modrego, A., Carlero, D., Arranz, R. and Martín-Benito, J. 2023. CryoEM of Viral Ribonucleoproteins and Nucleocapsids of Single-Stranded RNA Viruses. *Viruses*. **15**(3), p.653.

Moeller, A., Kirchdoerfer, R.N., Potter, C.S., Carragher, B. and Wilson, I.A. 2012. Organization of the influenza virus replication machinery. *Science (New York, N.Y.)*. **338**(6114), pp.1631–1634.

Monto, A.S. and Fukuda, K. 2020. Lessons From Influenza Pandemics of the Last 100 Years. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America*. **70**(5), pp.951–957.

Moritz, B. and Stracke, J.O. 2017. Assessment of disulfide and hinge modifications in monoclonal antibodies. *ELECTROPHORESIS*. **38**(6), pp.769–785.

Morris, S.J., Price, G.E., Barnett, J.M., Hiscox, S.A., Smith, H. and Sweet, C. 1999. Role of neuraminidase in influenza virus-induced apoptosis. *Journal of General Virology*. **80**(1), pp.137–146.

Mostafa, A., Naguib, M.M., Nogales, A., Barre, R.S., Stewart, J.P., García-Sastre, A. and Martinez-Sobrido, L. 2024. Avian influenza A (H5N1) virus in dairy cattle: origin, evolution, and cross-species transmission. *mBio*. **15**(12), pp.e02542-24.

Muyldermans, S. 2013. Nanobodies: Natural Single-Domain Antibodies. *Annual Review of Biochemistry*. **82**(Volume 82, 2013), pp.775–797.

Nakajima, N., Hata, S., Sato, Y., Tobiume, M., Katano, H., Kaneko, K., Nagata, N., Kataoka, M., Ainai, A., Hasegawa, H., Tashiro, M., Kuroda, M., Odai, T., Urasawa, N., Ogino, T., Hanaoka, H., Watanabe, M. and Sata, T. 2010. The First Autopsy Case of Pandemic Influenza (A/H1N1pdm) Virus Infection in Japan: Detection of a High Copy Number of the Virus in Type II Alveolar Epithelial Cells by Pathological and Virological Examination. *Japanese Journal of Infectious Diseases*. **63**(1), pp.67–71.

Nakano, M., Sugita, Y., Kodera, N., Miyamoto, S., Muramoto, Y., Wolf, M. and Noda, T. 2021. Ultrastructure of influenza virus ribonucleoprotein complexes during viral RNA synthesis. *Communications Biology*. **4**(1), pp.1–10.

Neumann, G., Eisfeld, A.J. and Kawaoka, Y. 2025. Viral factors underlying the pandemic potential of influenza viruses. *Microbiology and molecular biology reviews: MMBR*. **89**(2), p.e0006624.

Neumann, G. and Kawaoka, Y. 2024. Highly pathogenic H5N1 avian influenza virus outbreak in cattle: the knowns and unknowns. *Nature Reviews Microbiology*. **22**(9), pp.525–526.

Neumann, G., Noda, T. and Kawaoka, Y. 2009. Emergence and pandemic potential of swine-origin H1N1 influenza virus. *Nature*. **459**(7249), pp.931–939.

Ng, A.K.-L., Chan, W.-H., Choi, S.-T., Lam, M.K.-H., Lau, K.-F., Chan, P.K.-S., Au, S.W.-N., Fodor, E. and Shaw, P.-C. 2012. Influenza Polymerase Activity Correlates with the Strength of Interaction between Nucleoprotein and PB2 through the Host-Specific Residue K/E627. *PLoS ONE*. **7**(5), p.e36415.

Ng, A.K.-L., Lam, M.K.-H., Zhang, H., Liu, J., Au, S.W.-N., Chan, P.K.-S., Wang, J. and Shaw, P.-C. 2012. Structural Basis for RNA Binding and Homo-Oligomer Formation by Influenza B Virus Nucleoprotein. *Journal of Virology*. **86**(12), pp.6758–6767.

Ng, A.K.-L., Zhang, H., Tan, K., Li, Z., Liu, J., Chan, P.K.-S., Li, S.-M., Chan, W.-Y., Au, S.W.-N., Joachimiak, A., Walz, T., Wang, J.-H. and Shaw, P.-C. 2008. Structure of

the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *The FASEB Journal*. **22**(10), pp.3638–3647.

Nigg, P.E. and Pavlovic, J. 2015. Oligomerization and GTP-binding Requirements of MxA for Viral Target Recognition and Antiviral Activity against Influenza A Virus *. *Journal of Biological Chemistry*. **290**(50), pp.29893–29906.

Niu, F., Shaw, N., Wang, Y.E., Jiao, L., Ding, W., Li, X., Zhu, P., Upur, H., Ouyang, S., Cheng, G. and Liu, Z.-J. 2013. Structure of the Leanyer orthobunyavirus nucleoprotein–RNA complex reveals unique architecture for RNA encapsidation. *Proceedings of the National Academy of Sciences of the United States of America*. **110**(22), pp.9054–9059.

Noda, T., Sagara, H., Yen, A., Takada, A., Kida, H., Cheng, R.H. and Kawaoka, Y. 2006. Architecture of ribonucleoprotein complexes in influenza A virus particles. *Nature*. **439**(7075), pp.490–492.

Ogg, S.L., Weldon, A.K., Dobbie, L., Smith, A.J.H. and Mather, I.H. 2004. Expression of butyrophilin (Btn1a1) in lactating mammary gland is essential for the regulated secretion of milk–lipid droplets. *Proceedings of the National Academy of Sciences*. **101**(27), pp.10084–10089.

Olal, D., Dick, A., Woods, V.L., Liu, T., Li, S., Devignot, S., Weber, F., Saphire, E.O. and Daumke, O. 2014. Structural insights into RNA encapsidation and helical assembly of the Toscana virus nucleoprotein. *Nucleic Acids Research*. **42**(9), pp.6025–6037.

Ortega, J., Martín-Benito, J., Zürcher, T., Valpuesta, J.M., Carrascosa, J.L. and Ortín, J. 2000. Ultrastructural and functional analyses of recombinant influenza virus ribonucleoproteins suggest dimerization of nucleoprotein during virus amplification. *Journal of Virology*. **74**(1), pp.156–163.

Oxford, J.S., Sefton, A., Jackson, R., Innes, W., Daniels, R.S. and Johnson, N. 2002. World War I may have allowed the emergence of “Spanish” influenza. *The Lancet Infectious Diseases*. **2**(2), pp.111–114.

Ozato, K., Shin, D.-M., Chang, T.-H. and Morse, H.C. 2008. TRIM family proteins and their emerging roles in innate immunity. *Nature Reviews. Immunology*. **8**(11), pp.849–860.

Palese, P., Tobita, K., Ueda, M. and Compans, R.W. 1974. Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. *Virology*. **61**(2), pp.397–410.

Palovcak, E., Wang, F., Zheng, S.Q., Yu, Z., Li, S., Betegon, M., Bulkley, D., Agard, D.A. and Cheng, Y. 2018. A simple and robust procedure for preparing graphene-oxide cryo-EM grids. *Journal of Structural Biology*. **204**(1), pp.80–84.

Pantelic, R.S., Meyer, J.C., Kaiser, U., Baumeister, W. and Plitzko, J.M. 2010. Graphene oxide: A substrate for optimizing preparations of frozen-hydrated samples. *Journal of Structural Biology*. **170**(1), pp.152–156.

Paolo, C.D., Hefti, H.P., Meli, M., Landis, H. and Pavlovic, J. 1999. Intramolecular Backfolding of the Carboxyl-terminal End of MxA Protein Is a Prerequisite for Its Oligomerization *. *Journal of Biological Chemistry*. **274**(45), pp.32071–32078.

Passmore, L.A. and Russo, C.J. 2016. Chapter Three - Specimen Preparation for High-Resolution Cryo-EM *In*: R. A. Crowther, ed. *Methods in Enzymology* [Online]. The Resolution Revolution: Recent Advances In cryoEM. Academic Press, pp.51–86. [Accessed 28 September 2025]. Available from: https://www.sciencedirect.com/science/article/pii/S0076687916300295.

Patrono, L.V., Vrancken, B., Budt, M., Düx, A., Lequime, S., Boral, S., Gilbert, M.T.P., Gogarten, J.F., Hoffmann, L., Horst, D., Merkel, K., Morens, D., Prepoint, B., Schlotterbeck, J., Schuenemann, V.J., Suchard, M.A., Taubenberger, J.K., Tenkhoff, L., Urban, C., Widulin, N., Winter, E., Worobey, M., Schnalke, T., Wolff, T., Lemey, P. and Calvignac-Spencer, S. 2022. Archival influenza virus genomes from Europe reveal genomic variability during the 1918 pandemic. *Nature Communications*. **13**(1), p.2314.

Patzina, C., Haller, O. and Kochs, G. 2014. Structural Requirements for the Antiviral Activity of the Human MxA Protein against Thogoto and Influenza A Virus *. *Journal of Biological Chemistry*. **289**(9), pp.6020–6027.

Pavlovic, J., Haller, O. and Staeheli, P. 1992. Human and mouse Mx proteins inhibit different steps of the influenza virus multiplication cycle. *Journal of Virology*. **66**(4), pp.2564–2569.

Peng, Q., Liu, Y., Peng, R., Wang, M., Yang, W., Song, H., Chen, Y., Liu, S., Han, M., Zhang, X., Wang, P., Yan, J., Zhang, B., Qi, J., Deng, T., Gao, G.F. and Shi, Y. 2019. Structural insight into RNA synthesis by influenza D polymerase. *Nature Microbiology*. **4**(10), pp.1750–1759.

Peng, R., Xu, X., Nepal, B., Gong, Y., Li, F., Ferretti, M.B., Zhou, M., Lynch, K.W., Burslem, G.M., Kortagere, S., Marmorstein, R. and Chang, Y.-W. 2025. Molecular basis of influenza ribonucleoprotein complex assembly and processive RNA synthesis. *Science (New York, N.Y.)*. **388**(6748), p.eadq7597.

Perfetto, L., Gherardini, P.F., Davey, N.E., Diella, F., Helmer-Citterich, M. and Cesareni, G. 2013. Exploring the diversity of SPRY/B30.2-mediated interactions. *Trends in Biochemical Sciences*. **38**(1), pp.38–46.

Pertel, T., Hausmann, S., Morger, D., Züger, S., Guerra, J., Lascano, J., Reinhard, C., Santoni, F.A., Uchil, P.D., Chatel, L., Bisiaux, A., Albert, M.L., Strambio-De-Castillia, C., Mothes, W., Pizzato, M., Grütter, M.G. and Luban, J. 2011. TRIM5 is an innate immune sensor for the retrovirus capsid lattice. *Nature*. **472**(7343), pp.361–365.

Peterl, S., Wachsmuth-Melm, M. and Chlanda, P. 2025. Sample Preparation for Cryo-Electron Tomography of Influenza A Virus and Infected Cells *In*: Y. Yamauchi and M. J. Amorim, eds. *Influenza Virus: Methods and Protocols* [Online]. New York, NY: Springer US, pp.169–184. [Accessed 4 March 2025]. Available from: https://doi.org/10.1007/978-1-0716-4326-6_9.

Petersen, R.L. 2017. Strategies Using Bio-Layer Interferometry Biosensor Technology for Vaccine Research and Development. *Biosensors*. **7**(4), p.49.

Petric, P.P., Schwemmle, M. and Graf, L. 2023. Anti-influenza A virus restriction factors that shape the human species barrier and virus evolution. *PLOS Pathogens*. **19**(7), p.e1011450.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H. and Ferrin, T.E. 2021. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*. **30**(1), pp.70–82.

Peukes, J., Xiong, X., Erlendsson, S., Qu, K., Wan, W., Calder, L.J., Schraidt, O., Kummer, S., Freund, S.M.V., Kräusslich, H.-G. and Briggs, J.A.G. 2020. The native structure of the assembled matrix protein 1 of influenza A virus. *Nature*. **587**(7834), pp.495–498.

Pflug, A., Guilligay, D., Reich, S. and Cusack, S. 2014. Structure of influenza A polymerase bound to the viral RNA promoter. *Nature*. **516**(7531), pp.355–360.

Pinto, R.M., Bakshi, S., Lytras, S., Zakaria, M.K., Swingler, S., Worrell, J.C., Herder, V., Hargrave, K.E., Varjak, M., Cameron-Ruiz, N., Collados Rodriguez, M., Varela, M., Wickenhagen, A., Loney, C., Pei, Y., Hughes, J., Valette, E., Turnbull, M.L., Furnon, W., Gu, Q., Orr, L., Taggart, A., Diebold, O., Davis, C., Boutell, C., Grey, F., Hutchinson, E., Digard, P., Monne, I., Wootton, S.K., MacLeod, M.K.L., Wilson, S.J. and Palmarini, M. 2023. BTN3A3 evasion promotes the zoonotic potential of influenza A viruses. *Nature*. **619**(7969), pp.338–347.

Plotch, S.J., Bouloy, M., Ulmanen, I. and Krug, R.M. 1981. A unique cap(m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell*. **23**(3), pp.847–858.

Poole, E., Elton, D., Medcalf, L. and Digard, P. 2004. Functional domains of the influenza A virus PB2 protein: identification of NP- and PB1-binding sites. *Virology*. **321**(1), pp.120–133.

Poon, L.L.M., Pritlove, D.C., Fodor, E. and Brownlee, G.G. 1999. Direct Evidence that the Poly(A) Tail of Influenza A Virus mRNA Is Synthesized by Reiterative Copying of a U Track in the Virion RNA Template. *Journal of Virology*. **73**(4), pp.3473–3476.

Power, B.E., Ivancic, N., Harley, V.R., Webster, R.G., Kortt, A.A., Irving, R.A. and Hudson, P.J. 1992. High-level temperature-induced synthesis of an antibody VH-domain in *Escherichia coli* using the PelB secretion signal. *Gene*. **113**(1), pp.95–99.

Punjani, A., Rubinstein, J.L., Fleet, D.J. and Brubaker, M.A. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods*. **14**(3), pp.290–296.

Rabouw, H.H., Schokolowski, J., Müller, M., Baars, M.J.D., Dost, A.F.M., Bestebroer, T.M., Püschel, J., Clevers, H., Fouchier, R.A.M. and Tanenbaum, M.E. 2025. Mapping the complete influenza A virus infection cycle through single vRNP imaging. , 2025.01.20.633851.

Ramachandran, R. and Schmid, S.L. 2018. The dynamin superfamily. *Current Biology*. **28**(8), pp.R411–R416.

Ramlaul, K., Feng, Z., Canavan, C., Garrido, M.N. de, Carreño, D., Crone, M., Jensen, K.E., Li, B., Barnet, H., Riglar, D.T., Freemont, P.S., Miller, D. and Aylett, C.H.S. 2023. A 3D-printed flow-cell for on-grid purification of electron microscopy samples directly from lysate. , 2023.03.17.533159.

Rasmussen, S.G.F., DeVree, B.T., Zou, Y., Kruse, A.C., Chung, K.Y., Kobilka, T.S., Thian, F.S., Chae, P.S., Pardon, E., Calinski, D., Mathiesen, J.M., Shah, S.T.A., Lyons, J.A., Caffrey, M., Gellman, S.H., Steyaert, J., Skiniotis, G., Weis, W.I., Sunahara, R.K. and Kobilka, B.K. 2011. Crystal structure of the β2 adrenergic receptor–Gs protein complex. *Nature*. **477**(7366), pp.549–555.

Reguera, J., Malet, H., Weber, F. and Cusack, S. 2013. Structural basis for encapsidation of genomic RNA by La Crosse Orthobunyavirus nucleoprotein. *Proceedings of the National Academy of Sciences*. **110**(18), pp.7246–7251.

Rehwinkel, J. and Gack, M.U. 2020. RIG-I-like receptors: their regulation and roles in RNA sensing. *Nature Reviews Immunology*. **20**(9), pp.537–551.

Rehwinkel, J., Tan, C.P., Goubau, D., Schulz, O., Pichlmair, A., Bier, K., Robb, N., Vreede, F., Barclay, W., Fodor, E. and Sousa, C.R. e 2010. RIG-I Detects Viral Genomic RNA during Negative-Strand RNA Virus Infection. *Cell*. **140**(3), pp.397–408.

Reich, S., Guilligay, D., Pflug, A., Malet, H., Berger, I., Crépin, T., Hart, D., Lunardi, T., Nanao, M., Ruigrok, R.W.H. and Cusack, S. 2014. Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature*. **516**(7531), pp.361–366.

Reid, A.H., Fanning, T.G., Janczewski, T.A., Lourens, R.M. and Taubenberger, J.K. 2004. Novel Origin of the 1918 Pandemic Influenza Virus Nucleoprotein Gene. *Journal of Virology*. **78**(22), pp.12462–12470.

Rennie, M.L., McKelvie, S.A., Bulloch, E.M.M. and Kingston, R.L. 2014. Transient Dimerization of Human MxA Promotes GTP Hydrolysis, Resulting in a Mechanical Power Stroke. *Structure*. **22**(10), pp.1433–1445.

Rhodes, D.A., Reith, W. and Trowsdale, J. 2016. Regulation of Immunity by Butyrophilins. *Annual Review of Immunology*. **34**(Volume 34, 2016), pp.151–172.

Rhodes, D.A. and Trowsdale, J. 2007. TRIM21 is a trimeric protein that binds IgG Fc via the B30.2 domain. *Molecular Immunology*. **44**(9), pp.2406–2414.

Robenek, H., Hofnagel, O., Buers, I., Lorkowski, S., Schnoor, M., Robenek, M.J., Heid, H., Troyer, D. and Severs, N.J. 2006. Butyrophilin controls milk fat globule secretion. *Proceedings of the National Academy of Sciences*. **103**(27), pp.10385–10390.

Robert, X. and Gouet, P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*. **42**(W1), pp.W320–W324.

de Rozières, C.M., Pequeno, A., Shahabi, S., Lucas, T.M., Godula, K., Ghosh, G. and Joseph, S. 2022. PABP1 Drives the Selective Translation of Influenza A Virus mRNA. *Journal of Molecular Biology*. **434**(5), p.167460.

Ruigrok, R.W., Crépin, T. and Kolakofsky, D. 2011. Nucleoproteins and nucleocapsids of negative-strand RNA viruses. *Current Opinion in Microbiology*. **14**(4), pp.504–510.

Ruigrok, R.W.H. and Baudin, F. 1995. Structure of influenza virus ribonucleoprotein particles. II. Purified RNA-free influenza virus ribonucleoprotein forms structures that are indistinguishable from the intact influenza virus ribonucleoprotein particles. *Journal of General Virology*. **76**(4), pp.1009–1014.

Sabsay, K.R. and te Velthuis, A.J.W. 2024. Using structure prediction of negative sense RNA virus nucleoproteins to assess evolutionary relationships. *Virus Evolution*. **10**(1), p.veae058.

Sachdev, S., Roy, S., Saha, S.J., Zhao, G., Kumariya, R., Creemer, B.A., Yin, R., Pierce, B.G., Bewley, C.A. and Cheloha, R.W. 2025. Evaluation of AlphaFold modeling for elucidation of nanobody–peptide epitope interactions. *Journal of Biological Chemistry*. **301**(7), p.110268.

Sacristán, C., Ewbank, A.C., Ibáñez Porras, P., Pérez-Ramírez, E., de la Torre, A., Briones, V. and Iglesias, I. 2024. Novel Epidemiologic Features of High Pathogenicity Avian Influenza Virus A H5N1 2.3.3.4b Panzootic: A Review. *Transboundary and Emerging Diseases*. **2024**(1), p.5322378.

Schedler, B., Yukhnovets, O., Lindner, L., Meyer, A. and Fitter, J. 2023. The Thermodynamic Fingerprints of Ultra-Tight Nanobody–Antigen Binding Probed via Two-Color Single-Molecule Coincidence Detection. *International Journal of Molecular Sciences*. **24**(22), p.16379.

Schmidt, F.I., Hanke, L., Morin, B., Brewer, R., Brusic, V., Whelan, S.P.J. and Ploegh, H.L. 2016. Phenotypic lentivirus screens to identify functional single domain antibodies. *Nature Microbiology*. **1**(8), pp.1–10.

Schneider-Schaulies, S., Schneider-Schaulies, J., Schuster, A., Bayer, M., Pavlovic, J. and ter Meulen, V. 1994. Cell type-specific MxA-mediated inhibition of measles

virus transcription in human brain cells. *Journal of Virology*. **68**(11), pp.6910–6917.

Schoggins, J.W. 2019. Interferon-Stimulated Genes: What Do They All Do? *Annual Review of Virology*. **6**(Volume 6, 2019), pp.567–584.

Schusser, B., Reuter, A., von der Malsburg, A., Penski, N., Weigend, S., Kaspers, B., Staeheli, P. and Härtle, S. 2011. Mx Is Dispensable for Interferon-Mediated Resistance of Chicken Cells against Influenza A Virus. *Journal of Virology*. **85**(16), pp.8307–8315.

Schwemmle, M., Weining, K.C., Richter, M.F., Schumacher, B. and Staeheli, P. 1995. Vesicular stomatitis virus transcription inhibited by purified MxA protein. *Virology*. **206**(1), pp.545–554.

Scott, C.C., Vacca, F. and Gruenberg, J. 2014. Endosome maturation, transport and functions. *Seminars in Cell & Developmental Biology*. **31**, pp.2–10.

Shah, N.B. and Duncan, T.M. 2014. Bio-layer Interferometry for Measuring Kinetics of Protein-protein Interactions and Allosteric Ligand Effects. *Journal of Visualized Experiments (JoVE)*. (84), p.e51383.

Shi, L., Liu, H., Gao, S., Weng, Y. and Zhu, L. 2021a. Enhanced Extracellular Production of IsPETase in Escherichia coli via Engineering of the pelB Signal Peptide. *Journal of Agricultural and Food Chemistry*. **69**(7), pp.2245–2252.

Shi, L., Liu, H., Gao, S., Weng, Y. and Zhu, L. 2021b. Enhanced Extracellular Production of IsPETase in Escherichia coli via Engineering of the pelB Signal Peptide. *Journal of Agricultural and Food Chemistry*. **69**(7), pp.2245–2252.

Shimizu, T., Takizawa, N., Watanabe, K., Nagata, K. and Kobayashi, N. 2011. Crucial role of the influenza virus NS2 (NEP) C-terminal domain in M1 binding and nuclear export of vRNP. *FEBS Letters*. **585**(1), pp.41–46.

Spreeuwenberg, P., Kroneman, M. and Paget, J. 2018. Reassessing the Global Mortality Burden of the 1918 Influenza Pandemic. *American Journal of Epidemiology*. **187**(12), pp.2561–2567.

Stevens, T.A., Tomaleri, G.P., Hazu, M., Wei, S., Nguyen, V.N., DeKalb, C., Voorhees, R.M. and Pleiner, T. 2024. A nanobody-based strategy for rapid and scalable purification of human protein complexes. *Nature Protocols*. **19**(1), pp.127–158.

Su, X.-D., Zhang, H., Terwilliger, T.C., Liljas, A., Xiao, J. and Dong, Y. 2015. Protein Crystallography from the Perspective of Technology Developments. *Crystallography reviews*. **21**(1–2), pp.122–153.

Su, Z., Wu, C., Shi, L., Luthra, P., Pintilie, G.D., Johnson, B., Porter, J.R., Ge, P., Chen, M., Liu, G., Frederick, T.E., Binning, J.M., Bowman, G.R., Zhou, Z.H., Basler, C.F., Gross, M.L., Leung, D.W., Chiu, W. and Amarasinghe, G.K. 2018. Electron cryo-

microscopy structure of Ebola nucleoprotein reveals a mechanism for nucleocapsid-like assembly. *Cell*. **172**(5), pp.966-978.e12.

Sugita, Y., Matsunami, H., Kawaoka, Y., Noda, T. and Wolf, M. 2018. Cryo-EM structure of the Ebola virus nucleoprotein-RNA complex at 3.6 Å resolution. *Nature*. **563**(7729), pp.137–140.

Sultana, A. and Lee, J.E. 2015. Measuring protein-protein and protein-nucleic Acid interactions by biolayer interferometry. *Current Protocols in Protein Science*. **79**, 19.25.1-19.25.26.

Sun, Y. 2017. The antiviral mechanism of MxA revealed by a molecular docking-based structural model of the complex of MxA and influenza A virus NP/RNPs. , p.219915.

Szeto, W.-C., Hsia, H.-P., Tang, Y.-S. and Shaw, P.-C. 2020. Interaction between influenza A virus nucleoprotein and PB2 cap-binding domain is mediated by RNA. *PLoS ONE*. **15**(9), p.e0239899.

Taft, B.R., Hesse, M.J., Mamo, M., Bussiere, D.E., Huang, R., Lee, P.S., Wedel, L., Growcott, E., Wolff, K.C., Kuhen, K., Abend, J., Wong, K.A., Ganem, D., Leonard, V.H.J. and Tully, D.C. 2025. Discovery and Optimization of a Novel Series of Influenza A Virus Replication Inhibitors Targeting the Nucleoprotein Protein–Protein Interaction. *Journal of Medicinal Chemistry*. **68**(15), pp.16349–16370.

Takeuchi, O. and Akira, S. 2009. Innate immunity to virus infection. *Immunological Reviews*. **227**(1), pp.75–86.

Tang, Y.-S., Xu, S., Chen, Y.-W., Wang, J.-H. and Shaw, P.-C. 2021. Crystal structures of influenza nucleoprotein complexed with nucleic acid provide insights into the mechanism of RNA interaction. *Nucleic Acids Research*. **49**(7), pp.4144–4154.

Tarus, B., Bakowiez, O., Chenavas, S., Duchemin, L., Estrozi, L.F., Bourdieu, C., Lejal, N., Bernard, J., Moudjou, M., Chevalier, C., Delmas, B., Ruigrok, R.W.H., Di Primo, C. and Slama-Schwok, A. 2012. Oligomerization paths of the nucleoprotein of influenza A virus. *Biochimie*. **94**(3), pp.776–785.

Taubenberger, J.K. and Kash, J.C. 2010. Influenza Virus Evolution, Host Adaptation and Pandemic Formation. *Cell host & microbe*. **7**(6), pp.440–451.

Taubenberger, J.K. and Morens, D.M. 2006. 1918 Influenza: The mother of all pandemics. *Revista Biomédica*. **17**(1), pp.69–79.

Taubenberger, J.K. and Morens, D.M. 2008. The Pathology of Influenza Virus Infections. *Annual Review of Pathology: Mechanisms of Disease*. **3**(Volume 3, 2008), pp.499–522.

Tawar, R.G., Duquerroy, S., Vonrhein, C., Varela, P.F., Damier-Piolle, L., Castagné, N., MacLellan, K., Bedouelle, H., Bricogne, G., Bhella, D., Eléouët, J.-F. and Rey, F.A.

2009. Crystal Structure of a Nucleocapsid-Like Nucleoprotein-RNA Complex of Respiratory Syncytial Virus. *Science*. **326**(5957), pp.1279–1283.

Tegunov, D. and Cramer, P. 2019. Real-time cryo-electron microscopy data preprocessing with Warp. *Nature Methods*. **16**(11), pp.1146–1152.

Thompson, M.R., Kaminski, J.J., Kurt-Jones, E.A. and Fitzgerald, K.A. 2011. Pattern Recognition Receptors and the Innate Immune Response to Viral Infection. *Viruses*. **3**(6), pp.920–940.

Thompson, R.F., Walker, M., Siebert, C.A., Muench, S.P. and Ranson, N.A. 2016. An introduction to sample preparation and imaging by cryo-electron microscopy for structural biology. *Methods*. **100**, pp.3–15.

Tohidi, E., Ghaemi, M. and Golvajouei, M.S. 2024. A review on camelid nanobodies with potential application in veterinary medicine. *Veterinary Research Communications*. **48**(4), pp.2051–2068.

Tomar, D. and Singh, R. 2015. TRIM family proteins: emerging class of RING E3 ligases as regulator of NF-κB pathway. *Biology of the Cell*. **107**(1), pp.22–40.

Tumpey, T.M., Basler, C.F., Aguilar, P.V., Zeng, H., Solórzano, A., Swayne, D.E., Cox, N.J., Katz, J.M., Taubenberger, J.K., Palese, P. and García-Sastre, A. 2005. Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science (New York, N.Y.)*. **310**(5745), pp.77–80.

Turrell, L., Lyall, J.W., Tiley, L.S., Fodor, E. and Vreede, F.T. 2013. The role and assembly mechanism of nucleoprotein in influenza A virus ribonucleoprotein complexes. *Nature Communications*. **4**(1), p.1591.

Uchański, T., Masiulis, S., Fischer, B., Kalichuk, V., López-Sánchez, U., Zarkadas, E., Weckener, M., Sente, A., Ward, P., Wohlkönig, A., Zögg, T., Remaut, H., Naismith, J.H., Nury, H., Vranken, W., Aricescu, A.R., Pardon, E. and Steyaert, J. 2021. Megabodies expand the nanobody toolkit for protein structure determination by single-particle cryo-EM. *Nature Methods*. **18**(1), pp.60–68.

Uyeki, T.M., Milton, S., Abdul Hamid, C., Reinoso Webb, C., Presley, S.M., Shetty, V., Rollo, S.N., Martinez, D.L., Rai, S., Gonzales, E.R., Kniss, K.L., Jang, Y., Frederick, J.C., De La Cruz, J.A., Liddell, J., Di, H., Kirby, M.K., Barnes, J.R. and Davis, C.T. 2024. Highly Pathogenic Avian Influenza A(H5N1) Virus Infection in a Dairy Farm Worker. *The New England Journal of Medicine*. **390**(21), pp.2028–2029.

Vasylieva, N., Kitamura, S., Dong, J., Barnych, B., Hvorecny, K.L., Madden, D.R., Gee, S.J., Wolan, D.W., Morisseau, C. and Hammock, B.D. 2019. Nanobody-based binding assay for the discovery of potent inhibitors of CFTR inhibitory factor (Cif). *Analytica Chimica Acta*. **1057**, pp.106–113.

Veler, H., Fan, H., Keown, J.R., Sharps, J., Fournier, M., Grimes, J.M. and Fodor, E. 2022. The C-Terminal Domains of the PB2 Subunit of the Influenza A Virus RNA

Polymerase Directly Interact with Cellular GTPase Rab11a. *Journal of Virology*. **96**(5), pp.e01979-21.

Verhelst, J., Spitaels, J., Nürnberger, C., De Vlieger, D., Ysenbaert, T., Staeheli, P., Fiers, W. and Saelens, X. 2015. Functional Comparison of Mx1 from Two Different Mouse Species Reveals the Involvement of Loop L4 in the Antiviral Activity against Influenza A Viruses. *Journal of Virology*. **89**(21), pp.10879–10890.

Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T. and Bricogne, G. 2011. Data processing and analysis with the autoPROC toolbox. *Acta Crystallographica Section D: Biological Crystallography*. **67**(4), pp.293–302.

Vonrhein, C., Tickle, I.J., Flensburg, C., Keller, P., Paciorek, W., Sharff, A. and Bricogne, G. 2018. Advances in automated data analysis and processing within autoPROC, combined with improved characterisation, migration and visualisation of the anisotropy if diffraction limits using STARANISO. *Acta Crystallographica Section A*. **A74**(a360).

Vreede, F.T. and Brownlee, G.G. 2007. Influenza Virion-Derived Viral Ribonucleoproteins Synthesize both mRNA and cRNA In Vitro. *Journal of Virology*. **81**(5), pp.2196–2204.

Wachsmuth-Melm, M., Peterl, S., O'Riain, A., Makroczyová, J., Fischer, K., Krischuns, T., Vale-Costa, S., Amorim, M.J. and Chlanda, P. 2025. Visualizing influenza A virus assembly by in situ cryo-electron tomography. *Nature Communications*. **16**(1), p.9394.

Wagner, T., Merino, F., Stabrin, M., Moriya, T., Antoni, C., Apelbaum, A., Hagel, P., Sitsel, O., Raisch, T., Prumbaum, D., Quentin, D., Roderer, D., Tacke, S., Siebolds, B., Schubert, E., Shaikh, T.R., Lill, P., Gatsogiannis, C. and Raunser, S. 2019. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Communications Biology*. **2**(1), p.218.

Wagner, T. and Raunser, S. 2020. The evolution of SPHIRE-crYOLO particle picking and its application in automated cryo-EM processing workflows. *Communications Biology*. **3**(1), p.61.

Walker, A.P. and Fodor, E. 2019. Interplay between Influenza Virus and the Host RNA Polymerase II Transcriptional Machinery. *Trends in Microbiology*. **27**(5), pp.398–407.

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.E., Grimes, J.M., Stammers, D.K., Esnouf, R.M., Jones, E.Y., Owens, R.J., Stuart, D.I. and Harlos, K. 2005. A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization. *Acta Crystallographica Section D: Biological Crystallography*. **61**(6), pp.651–657.

Wandzik, J.M., Kouba, T. and Cusack, S. 2021. Structure and Function of Influenza Polymerase. *Cold Spring Harbor Perspectives in Medicine*. **11**(9), p.a038372.

Wang, Fengbin, Gnewou, O., Solemanifar, A., Conticello, V.P. and Egelman, E.H. 2022. Cryo-EM of Helical Polymers. *Chemical Reviews*. **122**(17), pp.14055–14065.

Wang, Fangzheng, Sheppard, C.M., Mistry, B., Staller, E., Barclay, W.S., Grimes, J.M., Fodor, E. and Fan, H. 2022. The C-terminal LCAR of host ANP32 proteins interacts with the influenza A virus nucleoprotein to promote the replication of the viral RNA genome. *Nucleic Acids Research*. **50**(10), pp.5713–5725.

Watson, J.D. 1954. The structure of tobacco mosaic virus: I. X-ray evidence of a helical arrangement of sub-units around the longitudinal axis. *Biochimica et Biophysica Acta*. **13**, pp.10–19.

Wentinck, K., Gogou, C. and Meijer, D.H. 2022. Putting on molecular weight: Enabling cryo-EM structure determination of sub-100-kDa proteins. *Current Research in Structural Biology*. **4**, p.332.

Whitehead, J.D., Decool, H., Leyrat, C., Carrique, L., Fix, J., Eléouët, J.-F., Galloux, M. and Renner, M. 2023. Structure of the N-RNA/P interface indicates mode of L/P recruitment to the nucleocapsid of human metapneumovirus. *Nature Communications*. **14**(1), p.7627.

Williams, G.D., Townsend, D., Wylie, K.M., Kim, P.J., Amarasinghe, G.K., Kutluay, S.B. and Boon, A.C.M. 2018. Nucleotide resolution mapping of influenza A virus nucleoprotein-RNA interactions reveals RNA features required for replication. *Nature Communications*. **9**, p.465.

Woof, J.M. and Burton, D.R. 2004. Human antibody–Fc receptor interactions illuminated by crystal structures. *Nature Reviews Immunology*. **4**(2), pp.89–99.

Wu, X. and Rapoport, T.A. 2021. Cryo-EM structure determination of small proteins by nanobody-binding scaffolds (Legobodies). *Proceedings of the National Academy of Sciences of the United States of America*. **118**(41), p.e2115001118.

Xiao, H., Killip, M.J., Staeheli, P., Randall, R.E. and Jackson, D. 2013. The Human Interferon-Induced MxA Protein Inhibits Early Stages of Influenza A Virus Infection by Retaining the Incoming Viral Genome in the Cytoplasm. *Journal of Virology*. **87**(23), pp.13053–13058.

Yang, F., Pang, B., Lai, K.K., Cheung, N.N., Dai, J., Zhang, W., Zhang, J., Chan, K.-H., Chen, H., Sze, K.-H., Zhang, H., Hao, Q., Yang, D., Yuen, K.-Y. and Kao, R.Y. 2021. Discovery of a Novel Specific Inhibitor Targeting Influenza A Virus Nucleoprotein with Pleiotropic Inhibitory Effects on Various Steps of the Viral Life Cycle. *Journal of Virology*. **95**(9), 10.1128/jvi.01432-20.

Yang, X., Steukers, L., Forier, K., Xiong, R., Braeckmans, K., Reeth, K.V. and Nauwynck, H. 2014. A Beneficiary Role for Neuraminidase in Influenza Virus Penetration through the Respiratory Mucus. *PLOS ONE*. **9**(10), p.e110026.

Yang, Y., Li, F. and Du, L. 2024. Therapeutic nanobodies against SARS-CoV-2 and other pathogenic human coronaviruses. *Journal of Nanobiotechnology*. **22**(1), p.304.

Ye, M., Makola, M., Newman, J.A., Fairhead, M., Maclean, E., Wright, N.D., Koekemoer, L., Thompson, A., Bezerra, G.A., Yi, G., Li, H., Rangel, V.L., Mamalis, D., Aitkenhead, H., Davis, B.G., Gilbert, R.J.C., Duerr, K., Gileadi, O. and Delft, F. von 2024. Gluebodies improve crystal reliability and diversity through transferable nanobody mutations that introduce constitutive close contacts. , 2022.07.26.501559.

Ye, M., Makola, M., Newman, J.A., Fairhead, M., Maclean, E., Wright, N.D., Koekemoer, L., Thompson, A., Bezerra, G.A., Yi, G., Li, H., Rangel, V.L., Mamalis, D., Aitkenhead, H., Davis, B.G., Gilbert, R.J.C., Duerr, K., Gileadi, O. and Delft, F. von 2022. Gluebodies improve crystal reliability and diversity through transferable nanobody mutations that introduce constitutive crystal contacts. , 2022.07.26.501559.

Ye, Q., Guu, T.S.Y., Mata, D.A., Kuo, R.-L., Smith, B., Krug, R.M. and Tao, Y.J. 2012. Biochemical and Structural Evidence in Support of a Coherent Model for the Formation of the Double-Helical Influenza A Virus Ribonucleoprotein. *mBio*. **4**(1), 10.1128/mbio.00467-12.

Ye, Q., Krug, R.M. and Tao, Y.J. 2006. The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature*. **444**(7122), pp.1078–1082.

Yi, G., Mamalis, D., Ye, M., Carrique, L., Fairhead, M., Li, H., Duerr, K.L., Zhang, P., Sauer, D.B., von Delft, F., Davis, B.G. and Gilbert, R.J.C. 2025. Covalently constrained 'Di-Gembodies' enable parallel structure solutions by cryo-EM. *Nature Chemical Biology*., pp.1–8.

Yoon, J., Zhang, Y.M., Her, C., Grant, R.A., Ponomarenko, A.I., Ackermann, B.E., Hui, T., Lin, Y.-S., Debelouchina, G.T. and Shoulders, M.D. 2024. The immune-evasive proline-283 substitution in influenza nucleoprotein increases aggregation propensity without altering the native structure. *Science Advances*. **10**(16), p.eadl6144.

York, A., Hengrung, N., Vreede, F.T., Huiskonen, J.T. and Fodor, E. 2013. Isolation and characterization of the positive-sense replicative intermediate of a negative-strand RNA virus. *Proceedings of the National Academy of Sciences*. **110**(45), pp.E4238–E4245.

Yu, G., Li, K. and Jiang, W. 2016. Antibody-based affinity cryo-EM grid. *Methods*. **100**, pp.16–24.

Zavrtanik, U., Lukan, J., Loris, R., Lah, J. and Hadži, S. 2018. Structural Basis of Epitope Recognition by Heavy-Chain Camelid Antibodies. *Journal of Molecular Biology*. **430**(21), pp.4369–4386.

Zhang, H., Li, X., Guo, J., Li, L., Chang, C., Li, Y., Bian, C., Xu, K., Chen, H. and Sun, B. 2014. The PB2 E627K mutation contributes to the high polymerase activity and enhanced replication of H7N9 influenza virus. *Journal of General Virology*. **95**(4), pp.779–786.

Zhang, M., Wang, Y., Cai, N., Qu, Y., Ma, X., Xue, J., Chen, X., Zhang, X., Xiao, J. and Zhang, Y. 2025. Structures of butyrophilin multimers reveal a plier-like mechanism for Vγ9Vδ2 T cell receptor activation. *Immunity*. **58**(7), pp.1660-1669.e7.

Zhao, C. and Pu, J. 2022. Influence of Host Sialic Acid Receptors Structure on the Host Specificity of Influenza Viruses. *Viruses*. **14**(10), p.2141.

Zheng, S., Wolff, G., Greenan, G., Chen, Z., Faas, F.G.A., Bárcena, M., Koster, A.J., Cheng, Y. and Agard, D.A. 2022. AreTomo: An integrated software package for automated marker-free, motion-corrected cryo-electron tomographic alignment and reconstruction. *Journal of Structural Biology: X*. **6**, p.100068.

Zhu, A.Z. 2024. *Towards understanding the influenza virus ribonucleoprotein complex: novel structural and functional approaches*.[Online] http://purl.org/dc/dcmitype/Text, University of Oxford. [Accessed 17 July 2025]. Available from: https://ora.ox.ac.uk/objects/uuid:63834db4-efdc-4b7a-b4eb-f4df5e84742d.

Zhu, Z., Fodor, E. and Keown, J.R. 2023. A structural understanding of influenza virus genome replication. *Trends in Microbiology*. **31**(3), pp.308–319.

Zimmermann, P., Mänz, B., Haller, O., Schwemmle, M. and Kochs, G. 2011. The viral nucleoprotein determines Mx sensitivity of influenza A viruses. *Journal of Virology*. **85**(16), pp.8133–8140.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E. and Scheres, S.H. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3 E. H. Egelman & J. Kuriyan, eds. *eLife*. **7**, p.e42166.

Zürcher, T., Pavlovic, J. and Staeheli, P. 1992. Mechanism of human MxA protein action: variants with changed antiviral properties. *The EMBO Journal*. **11**(4), pp.1657–1661.

7. Appendix

7.1. Nanobody and Gluebody Numbering Scheme

All nanobodies were donated by Dr. Florian Schmidt, and published in (Schmidt et al., 2016) with the exception of the final two nanobodies in the table below, VHH1_Hanke and VHH2_Hanke, which were presented in (Hanke et al., 2016). In these two cases, Hanke and colleagues additionally confirmed that these nanobodies, VHH1 and VHH2, block influenza infection. One megabody produced by Dr. Amanda Zhu was used in this work, based on the nanobody VHH151, Mb151.

Table 7.1: Nanobody and Gluebody Numbering Scheme

Nanobody number (VHH/Nb)	Gluebody number	vRNP Nuclear Import Blocker	NP Nuclear Import Blocker	vRNP Activity Blocker
22	44M	Y	N	Y
28	45M	Y	N	N
52	47M	Y	N	N
77	46M	Y	N	N
103	50M	Y	N	N
108	51M	Y	N	N
135	52M	Y	N	N
151	53S	Y	N	Y
170	54S	N	N	N
191	55S	Y	N	N
296	56S	Y	N	N
341	57S	Y	N	N
355	58S	Y	N	N
495	59S	Y	N	Y
508	60S	Y	N	N
1_Hanke	48S	Y	N	Y
2_Hanke	49S	Y	N	N