TOWARDS UNDERSTANDING THE INFLUENZA VIRUS RIBONUCLEOPROTEIN COMPLEX: NOVEL STRUCTURAL AND FUNCTIONAL APPROACHES

AMANDA ZIHAN ZHU



Sir William Dunn School of Pathology

and

Exeter College

A thesis submitted for the degree of

Doctor of Philosophy

at the University of Oxford, Hilary Term 2024

Abstract

Towards understanding the influenza virus ribonucleoprotein complex:

Novel structural and functional approaches

Amanda Zihan Zhu — Doctor of Philosophy

Exeter College — Hilary Term, 2024

Influenza viruses pose a significant public health concern, given their capacity to cause seasonal epidemics and pandemics like the infamous 1918 influenza pandemic. Central to the viral life cycle is a protein-RNA complex known as viral ribonucleoprotein (vRNP), responsible for viral genome transcription and replication. For decades, studies focused on individual vRNP protein components, namely viral polymerase and nucleoprotein (NP), along with their regulators. Nevertheless, the ultimate goal of visualizing vRNP in various functional states at the molecular level and in real-time remains to be achieved. To progress toward this objective, establishing a minimal system that reconstitutes vRNP functions in a completely cell-free environment would be advantageous. Additionally, there is a need to image vRNPs at a resolution that allows unambiguous localisation of each vRNP component. During my DPhil, my aim was to make advancements toward this goal.

vRNPs isolated from virions transcribe vRNA in a test tube to make viral mRNA when supplied with nucleotides and a cap donor RNA. To recreate the two-step viral genome replication in a simplified and highly controlled *in vitro* system, I initially switched virion-isolated vRNP from mRNA synthesis to cRNA synthesis by adding purified recombinant viral polymerase and host ANP32 protein. Subsequently, I reconstituted a full cycle of viral genome replication in a test tube by further including NP, promoting the second step of replication. This novel system revealed the minimal protein requirements to activate cRNA synthesis in a vRNP and to complete a full viral genome replication cycle. Through introducing mutations to the added proteins, I also managed to dissect their roles and uncover their novel functions. These findings offer original insights into the molecular regulation of vRNP functions. I then utilized a panel of nanobodies with resolved binding sites on the 1918 influenza viral polymerase to identify new inhibitory sites on the viral polymerase and unveil new antiviral mechanisms against influenza virus. Finally, using nanobodies raised against the viral polymerase or NP, I generated megabodies to help structurally resolve the interactions between viral polymerase and NP, and between NP and RNA, in the context of vRNP.

Collectively, this work provides novel insights into the molecular mechanisms of vRNP functions and introduces new approaches to study influenza virus vRNP. Specifically, it lays the groundwork for studying the regulation of vRNP functions systematically and paves the way for imaging viral replication in the context of vRNP. Excitingly, with the system and tools developed in this thesis, we are now closer than ever to visualising a functional vRNP in real-time.

Declaration

I declare that this thesis is entirely my own work and describes my own research, except when stated otherwise.

Amanda Zihan Zhu

Oxford, January 2024

Table of Contents

ABSTRACT I

DECLARATION III

TABLE OF CONTENTS IV

LIST OF FIGURES AND TABLES X

ABBREVIATIONS XIV

CHAPTER 1: INTRODUCTION 1

1.1 A BASIC UNDERSTANDING OF THE INFLUENZA VIRUS 1

1.1.1 Disease and interventions *1*

1.1.2 Classification and phylogenetics *3*

1.1.3 Genome and virion *4*

1.1.4 The viral life cycle *8*

1.2 INFLUENZA VIRUS RIBONUCLEOPROTEIN COMPLEX 12

1.2.1 Influenza virus polymerase *15*

1.2.2 Influenza virus nucleoprotein *22*

1.2.3 Intra-RNP interactions *25*

1.2.3.1 Viral polymerase-NP interactions 25

1.2.3.2 Polymerase-RNA interactions 26

1.2.3.3 NP-RNA interactions 28

1.3 INFLUENZA VIRUS TRANSCRIPTION AND REPLICATION 29

1.3.1 Initiation *30*

1.3.1.1 Primer-dependent initiation in mRNA synthesis 30

1.3.1.2 Primer-independent terminal initiation in cRNA synthesis 31

1.3.1.3 Primer-independent internal initiation in vRNA synthesis 33

1.3.1.4 Polymerase symmetric dimerisation in replication 35

1.3.1.5 Polymerase asymmetric dimerisation in replication 36

1.3.2 Elongation 40

1.3.3 Termination 45

1.3.4 Molecular cues for transcription and replication 49

1.4 NANOBODIES AND THEIR DERIVATIVES 51

1.5 OBJECTIVES 54

CHAPTER 2: MINIMAL COMPONENTS TO SWITCH THE ACTIVITY OF VRNP FROM TRANSCRIPTION TO REPLICATION 55

2.1 INTRODUCTION 55

2.2 RESULTS 58

2.2.1 Purification of active virion-derived vRNP 58

2.2.2 Additional viral polymerase inhibits transcription 62

2.2.3 Viral polymerase and ANP32 are the minimal components to switch the activity of vRNP from transcription to replication 65

2.2.4 ANP32 isoforms each support replication to different levels and ANP32^{LCAR} is required for replication 72

2.2.5 Pol II CTD has different effects on vRNP-bound or RNA-free viral polymerase 77

2.3 DISCUSSION 80

CHAPTER 3: *IN VITRO* RECONSTITUTION OF A COMPLETE CYCLE OF VIRAL GENOME REPLICATION **86**

3.1 INTRODUCTION 86

3.2 RESULTS 88

3.2.1 NP stimulates replicational activity and enables the complete cycle of viral genome replication *88*

3.2.2 Investigating the roles of NP in modulating vRNP activity *96*

3.2.3 Further characterisation of the RNA products from the in vitro assay *103*

3.3 DISCUSSION 111

CHAPTER 4: INHIBITION OF 1918 INFLUENZA VIRUS POLYMERASE USING NANOBODIES **120**

4.1 INTRODUCTION 120

4.2 RESULTS 124

4.2.1 The purified 1918 viral polymerase is functionally active *124*

4.2.2 Nanobodies inhibit the 1918 viral polymerase activity in vitro *126*

4.2.3 Nanobodies inhibit the 1918 viral polymerase activity in cell culture *129*

4.2.4 Nanobodies inhibit viral growth *137*

4.2.5 Inhibitory mechanisms of the nanobodies *143*

4.3 DISCUSSION 147

CHAPTER 5: NEW STRUCTURAL APPROACHES TO STUDY INTRA-RNP INTERACTIONS **153**

5.1 INTRODUCTION 153

5.2 RESULTS 156

5.2.1 Cloning, expression and purification of megabodies *156*

5.2.2 Binding properties of megabodies *161*

5.2.3 Cryo-EM analysis of the megabody-polymerase and the megabody-NP complexes *165*

5.2.4 In vitro assembly of a polymerase-NP complex *171*

5.2.5 Electron microscopy analysis of megabody-bound vRNP *176*

5.3 DISCUSSION 182

CONCLUSIONS AND PERSPECTIVES **186**

CHAPTER 6: MATERIALS AND METHODS **191**

6.1 MATERIALS 191

6.1.1 Plasmids *191*

6.1.2 DNA oligonucleotides *196*

6.1.3 RNA oligonucleotides *199*

6.1.4 Host RNA polymerase II CTD mimic peptides *199*

6.2 MOLECULAR CLONING 201

6.2.1 Preparation of chemically competent cells *201*

6.2.2 Transformation *201*

6.2.3 Preparation of plasmid DNA *202*

6.2.4 Site-directed mutagenesis *202*

6.2.5 Gibson assembly *202*

6.2.6 Megabody cloning *203*

6.2.7 Mutagenesis in pCAGGS plasmids *203*

6.3 CELLS AND VIRUSES 205

6.3.1 Bacterial culture *205*

6.3.2 Insect cell culture *205*

6.3.3 Mammalian cell culture *206*

6.3.4 Influenza virus reverse genetics *206*

6.3.5 Virus propagation *208*

6.3.6 Plaque assay *208*

6.4 CELL-BASED ASSAYS 209

6.4.1 Ribonucleoprotein reconstitution assay *209*

6.4.2 Primary transcription assay *210*

6.4.3 cRNA stabilisation assay *211*

6.4.4 Replication assay *211*

6.4.5 Viral growth assay *212*

6.4.6 Cell viability assay *212*

6.5 PROTEIN EXPRESSION AND PURIFICATION 213

6.5.1 Insect cell expression system *213*

6.5.2 Bacterial cell expression system *214*

6.5.3 Influenza virus polymerase purification *214*

6.5.4 Influenza virus nucleoprotein purification *215*

6.5.5 ANP32 family protein purification *216*

6.5.6 Nanobody purification *217*

6.5.7 Megabody purification *219*

6.6 RNP PURIFICATION 220

6.6.1 Purification of recombinant RNP from mammalian cells *220*

6.6.2 Purification of vRNP from virions *221*

6.7 *IN VITRO* ASSAYS 222

6.7.1 In vitro virion-derived vRNP activity assay *222*

6.7.2 In vitro recombinant RNP activity assay *223*

6.7.3 Polymerase transcription assay *223*

6.7.4 Polymerase replication assay *224*

6.8 PROTEIN ANALYSIS 225

6.8.1 Protein Quantification *225*

6.8.2 SDS polyacrylamide gel electrophoresis (SDS-PAGE), Coomassie or silver staining *225*

6.8.3 Western blots *226*

6.8.4 Size exclusion chromatography (SEC) *227*

6.8.5 CTD-binding assay *227*

6.8.6 Gradient cross-linking assay (Grafix) *228*

6.8.7 Mass spectrometry (MS) *228*

6.9 RNA ANALYSIS 229

6.9.1 RNA purification *229*

6.9.2 Radiolabelling oligonucleotides *230*

6.9.3 Radiolabelling capped oligonucleotides *230*

6.9.5 Denaturing polyacrylamide gel electrophoresis (PAGE) *231*

6.10 ELECTRON MICROSCOPY 232

6.10.1 Negative staining electron microscopy *232*

6.10.2 Cryo-electron microscopy (cryo-EM) *232*

REFERENCES **234**

ACKNOWLEDGEMENTS **264**

List of Figures and Tables

Chapter 1.

Figure 1.1 Viral genome and virion structures.

Figure 1.2 Influenza virus life cycle.

Figure 1.3 Structure of the influenza virus ribonucleoprotein complex.

Figure 1.4 Structure and conformations of influenza viral polymerase.

Figure 1.5 Structure of influenza virus nucleoprotein.

Figure 1.6 Structures of vRNA and cRNA promoters.

Figure 1.7 Cartoon models of transcription and replication initiation.

Figure 1.8 Structural model of an influenza virus genome replication platform.

Figure 1.9 NP dynamics during replication.

Figure 1.10 Cartoon model of transcription termination.

Figure 1.11 Cartoon model of replication termination.

Figure 1.12 Molecular factors regulating the activity of the influenza virus polymerase.

Figure 1.13 Nanobody and megabody structures.

Chapter 2.

Figure 2.1 Purification, semi-quantification and activity test of virion-derived vRNP.

Figure 2.2 Free viral polymerase inhibits transcriptional activity of vRNP-associated polymerase.

Figure 2.3 Viral polymerase and human ANP32B together promote vRNP replication.

Figure 2.4 Further investigation into the requirements for the polymerase to promote replication.

Figure 2.5 Effect of ANP32 isoforms on vRNP replication.

Figure 2.6 Effect of chicken ANP32 ^{LCAR} serial truncation on vRNP activity.

Figure 2.7 Effect of host Pol II CTD on vRNP activity.

Chapter 3.

Figure 3.1 Viral NP promotes vRNP replication and enables a complete cycle of viral genome replication.

Figure 3.2 Viral polymerase, ANP32 protein and viral NP enable a full replication cycle in different influenza virus genome segments.

Figure 3.3 Effect of viral polymerase from different types and subtypes of influenza viruses on vRNP activity.

Figure 3.4 Effect of the oligomerisation ability and the RNA-binding ability of NP on its functions.

Figure 3.5 Relative amount of viral polymerase and ANP32 protein impact NP in promoting replication.

Figure 3.6 Input vRNA and cRNA produced in the presence of additional viral polymerase contains truncated 5' termini.

Figure 3.7 The addition of viral and host factors results in the synthesis of aberrant replication products.

Figure 3.8 Viral RNA in the *in vitro* vRNP replication reconstitution assay contains heterogeneous 5' termini.

Figure 3.9 Model summarising the modulation of vRNP activity by viral and host factors in the *in vitro* vRNP replication reconstitution assay.

Chapter 4.

Table 4.1 Summary of the interaction interfaces on influenza A viral polymerase important for transcription and replication.

Table 4.2 Summary of the nanobody binding sites on the 1918 viral polymerase and the expected effects on viral polymerase based on binding information.

Figure 4.1 Nanobody phylogenetics and binding site information.

Figure 4.2 Activity analysis of the purified 1918 viral polymerase.

Figure 4.3 Nanobodies inhibit the 1918 viral polymerase activity *in vitro*.

Figure 4.4 Viral subtypes and plasmid backbones affect nanobody inhibition of polymerase activity in cell culture.

Figure 4.5 Nanobodies inhibit the 1918 viral polymerase activity in cell culture.

Figure 4.6 Nanobodies reduce viral growth in infection.

Figure 4.7 Nanobodies inhibit viral polymerase activity in infection.

Figure 4.8 Further investigation on the selected inhibitory nanobodies.

Figure 4.9 Investigating the inhibitory mechanisms of selected nanobodies.

Figure 4.10 Inhibitory sites and potential inhibitory mechanisms of selected nanobodies on the 1918 viral polymerase.

Chapter 5.

Table 5.1 Summary of megabody constructs generated in this study.

Figure 5.1 Cloning, expression and purification of megabodies.

Figure 5.2 Binding of megabodies to NP.

Figure 5.3 Binding of megabodies to viral polymerase.

Figure 5.4 Cryo-EM micrographs and initial structural model of the megabody-NP samples.

Figure 5.5 Cryo-EM micrographs and structural models of the megabody-viral polymerase samples.

Figure 5.6 *In vitro* assembly of a viral polymerase-NP complex for structural analysis.

Figure 5.7 Binding of megabodies to virion-purified vRNP.

Figure 5.8 Electron micrographs and analysis of the megabody-bound vRNP samples.

Chapter 6.

Table 6.1. List and description of plasmids used in this study.

Table 6.2 List and sequences of DNA oligonucleotides used in this study.

Table 6.3 List and sequences of RNA oligonucleotides used in this study.

Table 6.4 List and sequences of the synthetic DNA-dependent RNA polymerase II CTD peptides used in this study.

Table 6.5 List and origins of the bacterial strains used in this study.

Table 6.6. List and descriptions of viruses produced by reverse genetics in this study.

Table 6.7. List and descriptions of baculoviruses used in this study for protein expression in Sf9 cells.

Abbreviations

UNITS

In addition to the SI units of measurement, the following additional units are used in this thesis:

Å	Ångström
bp	base pair
Ci	Curie (1 Ci = 3.7 ×10^{10} becquerels)
Da	Dalton
eV	Electronvolt
g	standard gravity (9.80665 m/s^2)
h	Hour
hpi	Hours post infection
kb	Kilobases
kDa	kiloDalton
min	minute
MOI	multiplicity of infection
nt	nucleotide
PFU	plaque-forming unit
pH	-log10 c, where c is [H+] in mol/L
rpm	revolutions per minute
s	second
U	enzyme unit

OTHER ABBREVIATIONS

In addition to the amino acid one-letter code, the following additional abbreviations are used in this thesis:

1918	influenza A/Brevig Mission/1/1918 (H1N1) virus
3D	three-dimentional
A260/280	absorbance at 260 nm and 280 nm
ADP	adenosine diphosphate
ANOVA	analysis of variance
ANP32A	acidic nuclear phosphoprotein 32 family member A
ANP32B	acidic nuclear phosphoprotein 32 family member B
ANP32E	acidic nuclear phosphoprotein 32 family member E
AP	alkaline phosphatase
APS	ammonium persulfate
ATP	adenosine triphosphate
BSA	bovine serum albumin
$CaCl_2$	calcium chloride
cDNA	complementary DNA
chANP32A	chicken acidic nuclear phosphoprotein 32 family member A
chANP32B	chicken acidic nuclear phosphoprotein 32 family member B
CO_2	carbon dioxide
CRM1	chromosomal region maintenance 1
cRNA	complementary RNA
cRNP	complementary ribonucleoprotein
cryoEM	cryogenic electron microscopy

CTD	(host polymerase II) C-terminal domain
CTF	contrast transfer function
CTP	cytidine triphosphate
ddH2O	double distilled water
dH2O	distilled water
DMEM	Dulbecco′s modified Eagle′s medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
DTT	dithiothreitol
ECL	enhanced chemiluminescence
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EM	electron microscopy
ER	endoplasmic reticulum
FCS	foetal calf serum
Fj	Influenza A/duck/Fujian/01/2002 (H5N1) virus
FluPol	influenza virus RNA polymerase
FluPolA	influenza A virus RNA polymerase
FluPolB	influenza B virus RNA polymerase
FluPolC	influenza C virus RNA polymerase
FluPolD	influenza D virus RNA polymerase
GST	glutathione S-transferase
GTP	guanosine triphosphate
HA	haemagglutinin

HCl	hydrochloric acid
HEK	human embryonic kidney (cells)
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine
HIV-1	human immunodeficiency virus 1
hpi	hour(s) post infection
HRP	horseradish peroxidase
huANP32A	human acidic nuclear phosphoprotein 32 family member A
huANP32B	human acidic nuclear phosphoprotein 32 family member B
huANP32E	human acidic nuclear phosphoprotein 32 family member E
IgG	immunoglobulin G
IP	immunoprecipitation
IPTG	isopropyl β-d-1-thiogalactopyranoside
KCl	potassium chloride
L	litre
LB	Luria broth
LCAR	low-complexity acidic region
LRR	leucine-rich repeat
M1	matrix protein 1
M2	matrix protein 2
m7G	7-methylguanosine
Mb	megabody
MDBK	Madin-Darby bovine kidney (cells)
MDCK	Madin-Darby canine kidney (cells)
MEM	minimum essential medium

mFluPol	monomeric viral polymerase
$MgCl_2$	magnesium chloride
mNP	monomeric nucleoprotein
MOI	multiplicity of infection
mRNA	messenger RNA
MS	mass spectrometry
MTOC	microtubule organising centre
MVP	most valuable player
NA	neuraminidase
NaCl	sodium chloride
NaOH	sodium hydroxide
Nb	nanobody
NEP	nuclear export protein
NLS	nuclear localisation signal
NP	nucleoprotein
NS1	non-structural protein 1 (influenza virus)
NT60	Influenza A/Northern Territory/60/1968 (H3N2) virus
NTP	nucleoside triphosphate
N/A	not applicable
ORF	open reading frame
OTG	octylthioglucoside
P3	polymerase 3 (PA equivalent in influenza C virus)
PA	polymerase acidic protein
PAGE	polyacrylamide gel electrophoresis
PB1	polymerase basic protein 1

PB2	polymerase basic protein 2
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
Pi	inorganic phosphate
PMSF	phenylmethylsulfonyl fluoride
PNK	T4 polynucleotide kinase
Pol II	host RNA polymerase II
poly(A)	polyadenylic acid
poly(U)	polyuridylic acid
pppApG	5' triphosphorylated ApG
PR8	influenza A/Puerto Rico/8/1934 (H1N1) virus
pS5	serine-5-phosphorylated
pS2	serine-2-phosphorylated
PTM	Post-translational modification
Ran	Ras-related nuclear protein
RdRp	RNA-dependent RNA polymerase
RIG-I	retinoic acid-inducible gene I
RNA	ribonucleic acid
RNaseA	ribonuclease A
RNP	ribonucleoprotein
rNTP	ribonucleotide triphosphate
rRNA	ribosomal RNA
RT	reverse transcription
s.e.m.	standard error of the mean

SEC	size exclusion chromatography
SD	standard deviation
SDS	sodium dodecyl sulfate
Sf9	*Spodoptera frugiperda* clone 9
siRNA	small interfering ribonucleic acid
ssRNA	single-stranded RNA
svRNA	small viral RNA
TAP	tandem affinity purification
TBE	tris-borate-ethylenediaminetetraacetic acid
TEM	transmission electron microscopy
TEMED	N,N,N′,N′-tetramethylethylenediamine
TEV	tobacco etch virus protease
Tris	tris(hydroxymethyl)aminomethane
UTP	uridine triphosphate
UTR	untranslated region
UV	ultraviolet
vRNA	viral RNA
vRNP	viral ribonucleoprotein
VSV	vesicular stomatitis virus
WCL	whole cell lysate
WSN	Wilson Smith neurotropic

Chapter 1: Introduction

1.1 A BASIC UNDERSTANDING OF THE INFLUENZA VIRUS

1.1.1 Disease and interventions

Influenza virus presents a significant public health concern due to its ability to cause seasonal influenza epidemics and occasional pandemics (Carrat et al., 2008). Also known as flu, influenza is a highly contagious respiratory illness characterized by symptoms such as fever, cough, muscle fatigue, and nasal congestion. In a small portion of cases, influenza can lead to severe illness, for instance through a direct effect of pneumonia, but more often indirect consequences, such as bacterial infection and acceleration of pre-existing respiratory or cardiovascular illness (Taubenberger & Morens, 2008). In a more adverse scenario, when an unexpected variant of influenza virus adapts to transmit efficiently between human hosts, an influenza outbreak can lead to pandemic. An infamous example of this is the 1918 influenza pandemic, commonly but unfairly known as the Spanish flu. As one of the deadliest pandemics in history, the 1918 influenza pandemic infected an estimated one-third of the global population at the time, resulted in millions of deaths, and mothered almost all influenza A viruses that later caused worldwide infections in humans (Taubenberger & Morens, 2006).

The ability of the influenza virus to cause epidemics and pandemics is attributed to several factors. One important factor is the zoonotic potential of the virus: the ability of the virus to

infect a wide range of animal species in addition to humans, such as birds, pigs, and bats. This property of influenza viruses increases the likelihood of interspecies transmission and the emergence of new strains, against which humans have no natural immunity (Short et al., 2015). Another key factor is their high mutation rate. Frequent mutations leads to faster evolution, which helps the virus to escape from host immune surveillance and alters pathogenicity via antigenic drift (Shao et al., 2017). In addition, the influenza virus carries a segmented genome, which is known to undergo reassortment during co-infection. Reassortment in segmented viruses allows the exchange of genetic material between different strains of the same virus – a primary cause of antigenic shift (Shao et al., 2017). As a result, during an outbreak, a newly-emerged highly pathogenic influenza virus can spread rapidly with its high transmissibility, leading to a significant number of infections and deaths within a short period of time. Due to the significant threat posed to human health by influenza viruses, characterizing viral infections at various levels is crucial for understanding their virological properties and host dependency, thereby developing effective protective measures.

Studies of influenza virus at epidemiological, immunological, and biochemical levels have led to the development of strategies against the virus at multiple levels. Influenza surveillance systems, such as Google Flu Trends, have been launched to monitor the spread and impact of the virus; however, currently they still lack accuracy (Jabour et al., 2021). On the other hand, vaccination has been a key preventive method against influenza viruses by inducing immune responses against predicted circulating strains. However, due to antigenic drift and antigenic shift, influenza vaccines need to be updated annually based on predictions. The effectiveness of vaccination can vary each year, influenced by factors such as the match between predicted strains and the circulating strains, as well as individual immune response (Osterholm et al., 2012). When influenza surveillance and vaccination fail, antiviral drugs are the next in line to

manage the disease. While antivirals cannot prevent infection, they restrict viral growth and viral spread within and between individuals, hence containing an outbreak and reducing the cases of hospitalization. Antivirals used against influenza virus target viral entry, release and replication by blocking viral proteins such as the matrix protein 2 (M2) ion channel, the neuraminidase (NA) surface protein or the viral RNA-dependent RNA polymerase (RdRp). However, the effectiveness of antiviral drugs may vary depending on factors such as the timing of treatment and the development of drug resistance in the circulating virus (Liu et al., 2021). Therefore, it is important to acquire a deeper understanding of the influenza virus, with the aim to improve and combine current strategies and to develop new preventive methods against our ever-evolving enemy.

1.1.2 Classification and phylogenetics

Rooted in "myxa," the Greek word for 'mucus,' negative-sense RNA viruses that historically belong to the Myxoviridae family are commonly known for their ability to bind to and degrade mucins (Bang & Andrewes, 1955). This viral family was later found to consist two evolutionarily unrelated families: Orthomyxoviridae family, which has a segmented genome that replicates inside the host nucleus, and Paramyxoviridae family, which contains a single molecule of RNA that is copied in the cytoplasm (Waterson, 1962).

Influenza viruses are a collection of four viral genera within the Orthomyxoviridae family, namely, influenza A, B, C, and D viruses. The four types of influenza viruses are genetically distinct from each other due to the accumulation of substantial differences in each genome segment over the course of evolution, including a change in genome segment number (eight

segments for influenza A and B viruses and seven segments for influenza C and D viruses). They also exhibit different infection profiles. Influenza A viruses have also been the primary drivers of pandemics in the past. Both influenza A and B viruses infect a range of hosts, including birds, pigs, seals, bats, and humans, and are responsible for causing seasonal epidemics in humans. Influenza C virus mostly infects pigs and humans but is not known to cause severe disease in humans, while influenza D virus only causes mild disease in cattle and pigs (Blümel et al., 2009).

Influenza A viruses are further classified into subtypes based on their antigenic properties. Influenza A viruses are classified based on different combinations of two viral surface glycoproteins: haemagglutinin (HA) and NA (Blümel et al., 2009). So far, there are 19 known HA and 11 known NA subtypes in influenza A viruses, resulting in various combinations of surface proteins, such as H1N1, H3N2, and H5N1 (Fereidouni et al., 2023). On the other hand, influenza B virus is divided into two main lineages, Victoria and Yamagata, based on genetic and antigenic differences in the HA gene. Both lineages had been circulating in our population until the extinction of the B/Yamagata lineage was reported during the COVID-19 pandemic (Koutsakos et al., 2021; Paget et al., 2022; Xu et al., 2014).

1.1.3 Genome and virion

As mentioned earlier, the genomes of influenza A and B viruses consist of eight single-stranded, negative-sense RNA segments, while influenza C and D viruses each have seven segments. A segmented genome allows the members of the influenza virus family to exchange genetic materials upon co-infection via genetic reassortment, which plays an important role in

generating viral diversity over the course of evolution (Compans & Oldstone, 2014; Shao et al., 2017). Each genomic segment of an influenza virus encodes one or two major viral proteins that orchestrate various stages of the viral life cycle (Vasin et al., 2014). Although the length of each individual segment varies from around 900 nt to 2400 nt, their general structures are similar. Each vRNA encodes one or two open reading frames for viral protein translation, flanked by untranslated regions and further flanked by 3' non-polyadenylated terminus and 5' triphosphorylated terminus (Figure 1.1A). The two termini of a vRNA segment are collectively known as 'the promoter' and are bound by a viral polymerase (Pflug et al., 2014; Robb et al., 2019; Thierry et al., 2016). The rest of the vRNA is packaged and protected by nucleoprotein (NP), with an estimated ratio of one NP per 24 nt (Ortega et al., 2000). Together with a viral polymerase and numerous molecules of NP, each vRNA segment forms a viral ribonucleoprotein (vRNP) complex. One of each vRNA segment, in the form of vRNP, is selectively incorporated into a virion through RNA-RNA interactions, and these vRNPs are arranged into a '1+7' pattern through potential interactions with each other and with the matrix protein 1 (M1) (Figure 1.1D). Peculiarly, despite the difference in genome segment numbers, all types of influenza viruses are found to prefer this '1+7' pattern. Host RNA has been reported to be incorporated as the eighth vRNP when only seven segments were present, suggesting the potential importance of this pattern in the viral genome packaging mechanism (Noda et al., 2018).

Influenza virions are pleomorphic. Specifically, they can adopt spherical, filamentous, and bacilliform morphologies (Calder et al., 2010) (Figure 1.1C). While the average diameter of an H3N2 strain of influenza A virus virion was estimated to be around 100 nanometres, the exact dimensions of an influenza virus virion can vary due to their pleiomorphic nature, which has hindered the comprehensive understanding of virion structure (Harris et al., 2006).

Influenza viruses belong to the group of enveloped viruses. Each viral particle consists of an outer most host-derived lipid bilayer membrane embedded with two types of viral glycoproteins, HA and NA, and the ion channel protein, M2. HA is responsible for viral attachment to host cells and facilitates viral entry, while NA aids in the release of newly formed virions from infected cells. Additionally, M1 forms a layer that underlies the lipid envelope, defining the virion cavity and providing structural support to the virion. Various non-structural proteins and protein complexes are packaged within the virion, including vRNPs, nuclear export protein (NEP), and non-structural protein 1 (NS1) (Hutchinson et al., 2014). NEP is involved in the export of vRNP during virion assembly, while NS1 is a multifunctional protein that plays a role in inhibiting the host immune response to viral infection (Bouvier & Palese, 2008) (Figure 1.1B). In addition to the viral proteins, proteomic analysis of purified influenza virus particles has also revealed the presence of more than 30 host-encoded proteins, including cytoskeletal proteins such as tubulin and actin, and glycolytic enzymes such as enolase and GAPDH. The active roles of these host factors packaged inside a virion remain under debate, but some are likely specifically recruited into the virion and are actively involved in the viral life cycle (Hutchinson et al., 2014).

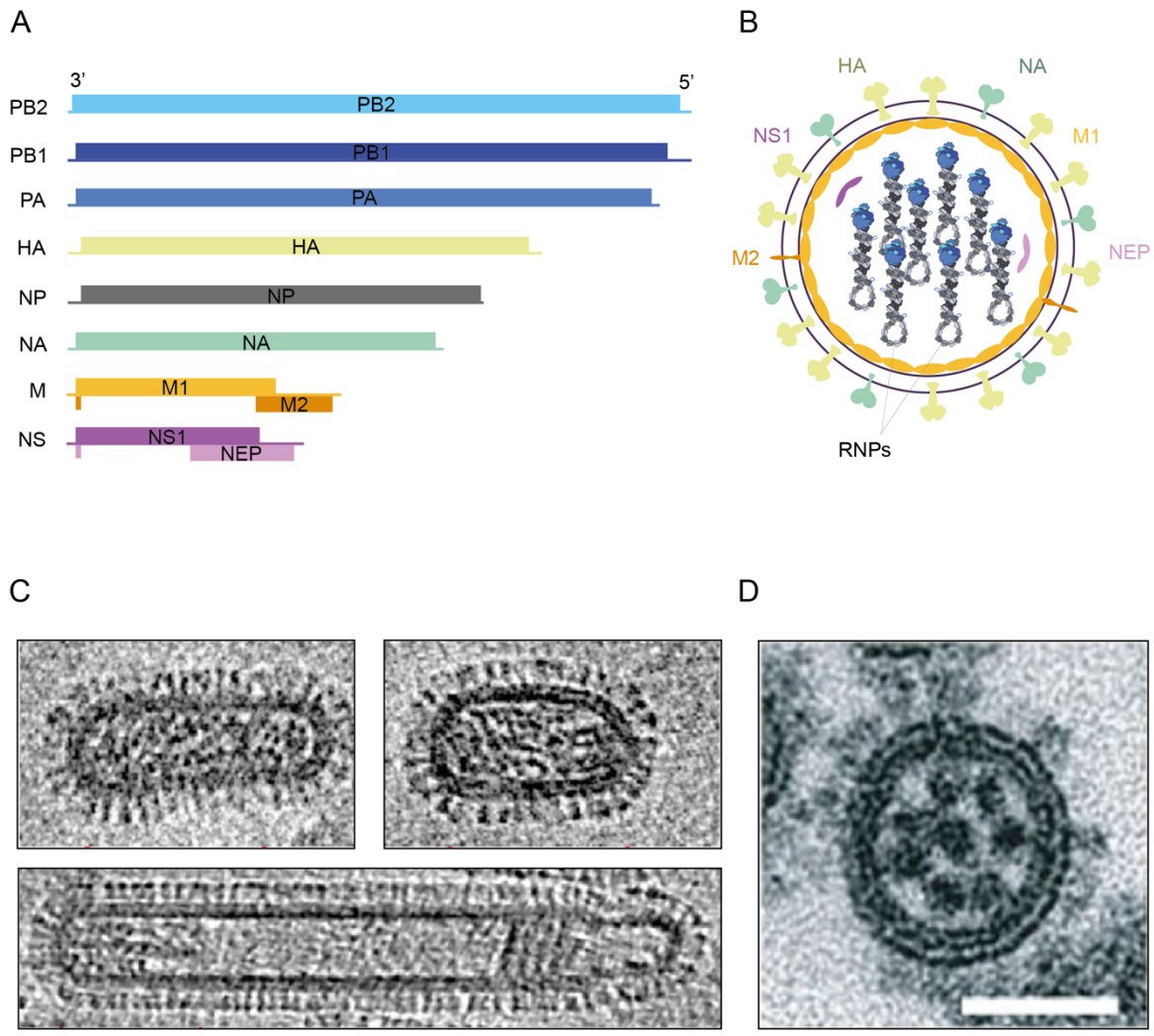


Figure 1.1 Viral genome and virion structures. *A. Schematics of influenza A virus genome segments arranged top to bottom according to lengths in descending order. Open reading frames for major viral proteins are highlighted as rectangles above and beneath the gene. B. Cartoon diagram of an influenza virus particle showing all viral proteins and vRNP complexes. C. Electron micrograph of influenza viral particles of spherical (top elft), filamentous (bottom), and bacilliform (top right) morphologies. Adapted from Calder et al., 2010. D. Electron micrograph of a spherical influenza A virus virion, showing the '7+1' arrangement of RNPs within a viral particle. Adapted from Noda et al., 2006. Scale bars: 50 nm.*

1.1.4 The viral life cycle

Like other viruses, influenza viruses rely on host cells to amplify and display cell-type tropism during infection. The ability of influenza viruses to infect certain cell types in specific hosts grants them access to essential host factors that support the viral life cycle. In mammalian hosts, influenza viruses primarily target respiratory epithelial cells, such as those in the trachea and lungs. Avian influenza viruses can infect the respiratory tract and reproductive organs, but primarily replicate in the gut, targeting epithelial cells in the intestine (Van Riel et al., 2007). While the virus cannot determine which host organism to enter, compatibility between viral surface proteins and host cell receptors directs the process of cell-type tropism. In particular, for the virus to gain cellular entry, the viral membrane protein HA has to bind to specific host cell membrane receptors called sialic acid (SA). Avian influenza viruses typically have a preference for avian-specific receptors, such as α2,3-linked sialic acids, whereas mammalian influenza viruses have a preference for α2,6-linked sialic acid (Kumlin et al., 2008).

Upon attachment to the host cell membrane via the HA and SA interaction, the virion enters the host cell by triggering receptor-mediated endocytosis (Figure 1.2A-B). Thereafter, the early endosome containing viral particles becomes acidified by the endosomal membrane proton ATPase, triggering a conformational change in HA. This conformational rearrangement of HA enables the insertion of a hydrophobic fusion peptide into the endosomal membrane, followed by the fusion of viral and endosomal membranes. It is worth noting that this conformational change in HA, in response to a drop in pH, happens only in matured HA, formed by the two peptides resulted from the cleavage of immature HA by a trypsin-like serine endoprotease. Some laboratory-adapted strains of influenza virus, such as influenza A/WSN/33, have acquired mutations that allow auto-cleavage of HA, hence increasing the ease of passaging

(Goto & Kawaoka, 1998). Endosomal acidification also enables the viral M2 ion channel to subsequently acidify the interior of the virion, causing a disruption in the interaction between vRNP and M1, as well as within the vRNP bundle and the M1 layer themselves. Together, membrane fusion and destabilization of the virion interior organisation allow the vRNPs to be released into the cytosol (Figure 1.2C), where host importin-α1 or -α5 recognises the nuclear localisation signal (NLS) of NP on the vRNP and transports the vRNPs into the cell nucleus (Bouvier & Palese, 2008) (Figure 1.2D). This process of vRNP release and nuclear import has been visualised in real-time in a recent single-molecule microscopic study at single-RNP resolution (Tanenbaum Lab unpublished data).

Once inside the nucleus, the vRNPs begin the crucial process of viral transcription and replication. Viral transcription and replication are temporally segregated. Early in infection, the RNP-associated viral polymerase transcribes the negative-sense genome into positive-sense viral messenger RNA (mRNA) to produce viral proteins (Kawakami et al., 2011; Phan et al., 2021). To avoid triggering host innate immune response and for viral mRNA to be translated by host ribosomes, vRNPs produce viral mRNA that resembles host mRNA, with a 5' 7-methyl-guanosine (m7G) cap and a 3' polyadenylated (polyA) tail of varying length (Etkind & Krug, 1974; Krug et al., 1976). Influenza virus obtains its mRNA cap from host capped RNA through a famously known process called cap-snatching and produces a polyA tail as the polymerase stutters on an oligoU stretch close to the template 5' terminus (Plotch et al., 1981; Poon et al., 1999; Walker & Fodor, 2019) (Figure 1.2E). These viral mRNAs are then exported from the nucleus, likely through binding to host mRNA binding proteins, into the cytosol, where they hijack host ribosomal pathways (Figure 1.2F). As viral proteins accumulate, and are selectively imported into the nucleus, replication of the viral genome initiates. Viral genome is faithfully copied over a two-step process. The first step of replication produces a

positive-sense 5' triphosphorylated and 3' non-polyadenylated reverse-complementary copy of vRNA templates, namely the complementary RNA (cRNA). Together with newly synthesised viral polymerase and NP, these intermediate cRNAs are packaged into complementary RNPs (cRNPs), which are structurally similar to vRNPs and act as templates for *de novo* vRNA synthesis during the second step of replication. Progeny vRNAs, faithful copies of the parental vRNA, are subsequently packaged into vRNPs, which can carry out secondary mRNA or cRNA synthesis (Klumpp et al., 1997; Zhu et al., 2023) (Figure 1.2G).

Alternatively, progeny vRNPs can be trafficked to the cell membrane for virion assembly. The molecular mechanism of vRNP nuclear export remains elusive; however, current literature suggests that the export of vRNP is a signalling-induced event and begins with the formation of a vRNP export complex, consisting of vRNP, NEP, and M1 proteins, which assembles at the dense chromatin (Chase et al., 2011; Takizawa et al., 2006). CRM1, part of the cellular export machinery, subsequently recognises the N-terminal nuclear export signal on NEP and directs the nuclear export of vRNP through the nuclear pore complex (Eisfeld et al., 2015). Thereafter, the vRNPs accumulate at the microtubule organising centre (MTOC) close to the nucleus and are trafficked to the cell plasma membrane by attaching to recycling endosomes in a Rab11-dependent manner (De Castro Martin et al., 2017) (Figure 1.2H).

At the plasma membrane, bundles of vRNP assemble at budding loci, where newly synthesised HA, NA, and M2 proteins accumulate on the host cell membrane and associate with M1 proteins in the cytosol. These vRNP bundles form selectively via RNA-RNA interactions between segments to ensure at least one of each of the influenza virus genome segments is included per bundle (Dadonaite et al., 2019; Fournier et al., 2012; Gavazzi et al., 2013; Hutchinson et al., 2010) (Figure 1.2I). vRNPs within a bundle are organised into the '7+1'

formation and are packaged into a budding virion as M1 polymerises into a helical structure (Peukes et al., 2020). As the membrane scission and virion budding occur, the virion surface NA cleaves the SA linkers on the surrounding plasma membrane to prevent reattachment or aggregation of viruses (Bouvier & Palese, 2008) (Figure 1.2J).

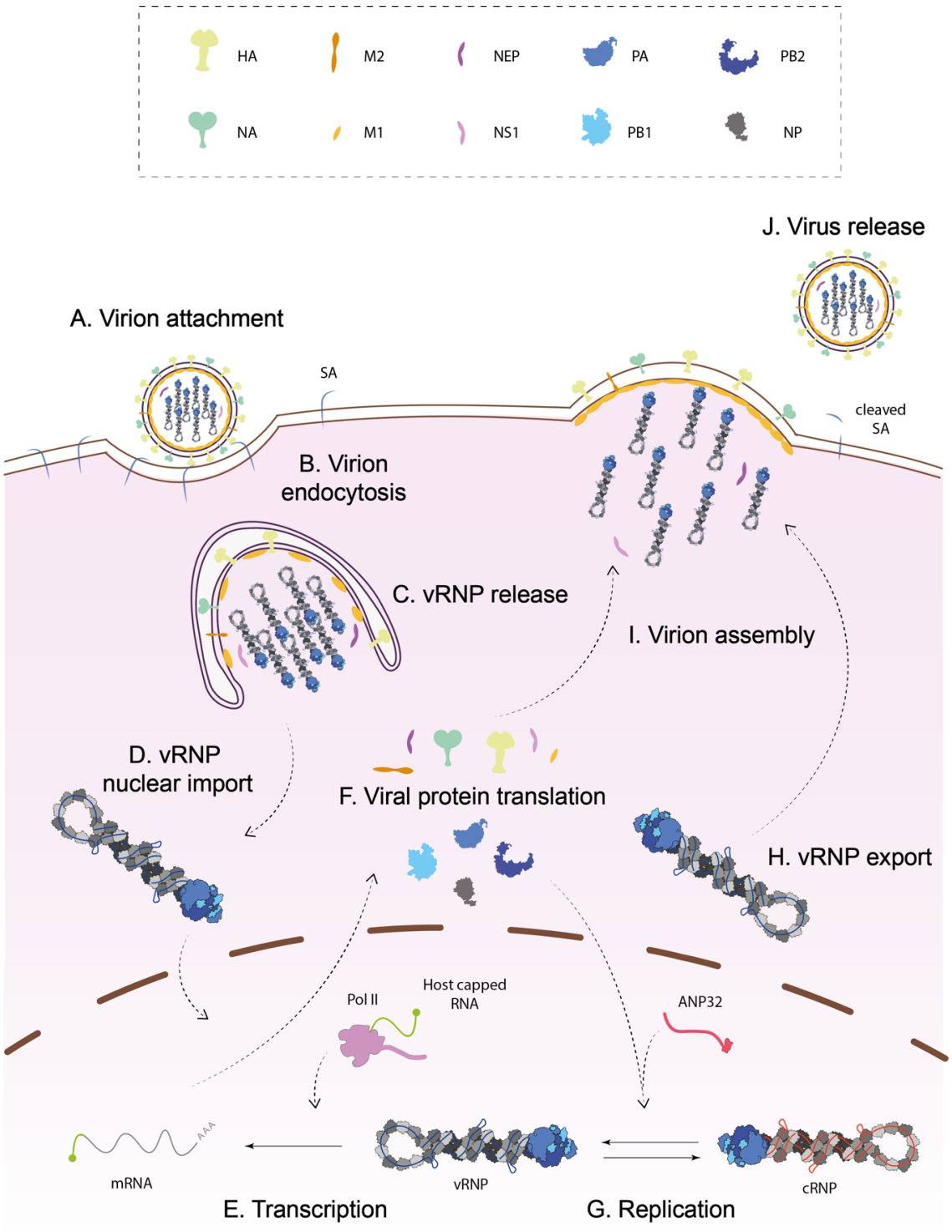


Figure 1.2 Influenza virus life cycle. *A. Influenza virion attaches to the host cell membrane by binding to SA-decorated cell surface receptors. B. Endocytosis facilitates viral entry. C. vRNPs are released into the cytosol upon endosomal and virion acidification. D. vRNPs are imported into the host nucleus through the nuclear pore complex. E. vRNPs carry out transcription to produce viral mRNA by associating with the host RNA polymerase II (Pol II) complex via its C-terminal domain (CTD) and cleaving a capped RNA primer for viral transcription initiation. F. Viral mRNA is transported into the host cell cytoplasm to instruct the production of viral proteins. G. Newly synthesised viral polymerase and NP are transported into the host nucleus and, together with host ANP32 protein, activate viral genome replication. The viral genome replicates via an intermediate cRNA which acts as template for vRNA synthesis. Newly synthesised cRNA and vRNA are assembled into cRNP and vRNP. H. Progeny vRNPs are exported out of the nucleus towards the cell membrane. I. Together, newly synthesised viral proteins and vRNPs are assembled into new virions at the cell membrane. J. Progeny influenza virion is released from the cell membrane and prevented from reattaching by cleaving SA off the cell surface receptors. (Viral proteins are labelled in the dotted-line box.) Adapted from Zhu et al., 2023.*

1.2 INFLUENZA VIRUS RIBONUCLEOPROTEIN COMPLEX

Influenza virus ribonucleoprotein complexes (RNPs) are central to the viral life cycle. Structurally, RNPs contains and protect non-capped and non-polyadenylated vRNA and cRNA; and functionally, they carry out transcription and replication of the viral genome (Eisfeld et al., 2015). Each vRNA segment is individually assembled into a vRNP, within which the single-stranded negative-sense vRNA forms a pseudo-circle with both termini, referred to as the 3' and 5' promoter, bound by a viral RdRp (Figure 1.3A). This vRNP-associated viral polymerase possesses RNA synthesis catalytic activity, which enables a vRNP to transcribe and replicate the resident viral genome segment to produce viral mRNA and cRNA, respectively. As

mentioned in earlier sections, the rest of the vRNA is coated with numerous NP. These NPs oligomerise into a flexible filament, which also forms a pseudo-circle, placing the two NPs closest to the two vRNA termini next to the resident viral polymerase. To form a pseudo-circle along the length of the vRNA, the NP filament creates a sharp turn at one end of the vRNP, creating a loop structure that contains three to eight NP molecules (Figure 1.3A). The NP filament within a vRNP twists into a highly-compact double helix inside a virion to efficiently package the viral genome; however, it can become highly dynamic during transcription and replication to release the RNA template into the active site without disrupting the overall RNP structure. Characterisation of cRNP reveals that it is structurally similarity to vRNP, but incapable of transcription (York et al., 2013).

Under an electron microscope, virion vRNPs appear as rod-like complexes of different lengths, with left-handed helical turns in the filament and an occasionally easily-identifiable NP loop at one end (Arranz et al., 2012; Coloma et al., 2020; Moeller et al., 2012) (Figure 1.3B). While the local resolution of the NP helical track has been increased to 7 Å, structural studies focusing on the polymerase region in an RNP are scarce and lack convincing results (Figure 1.3C-D).

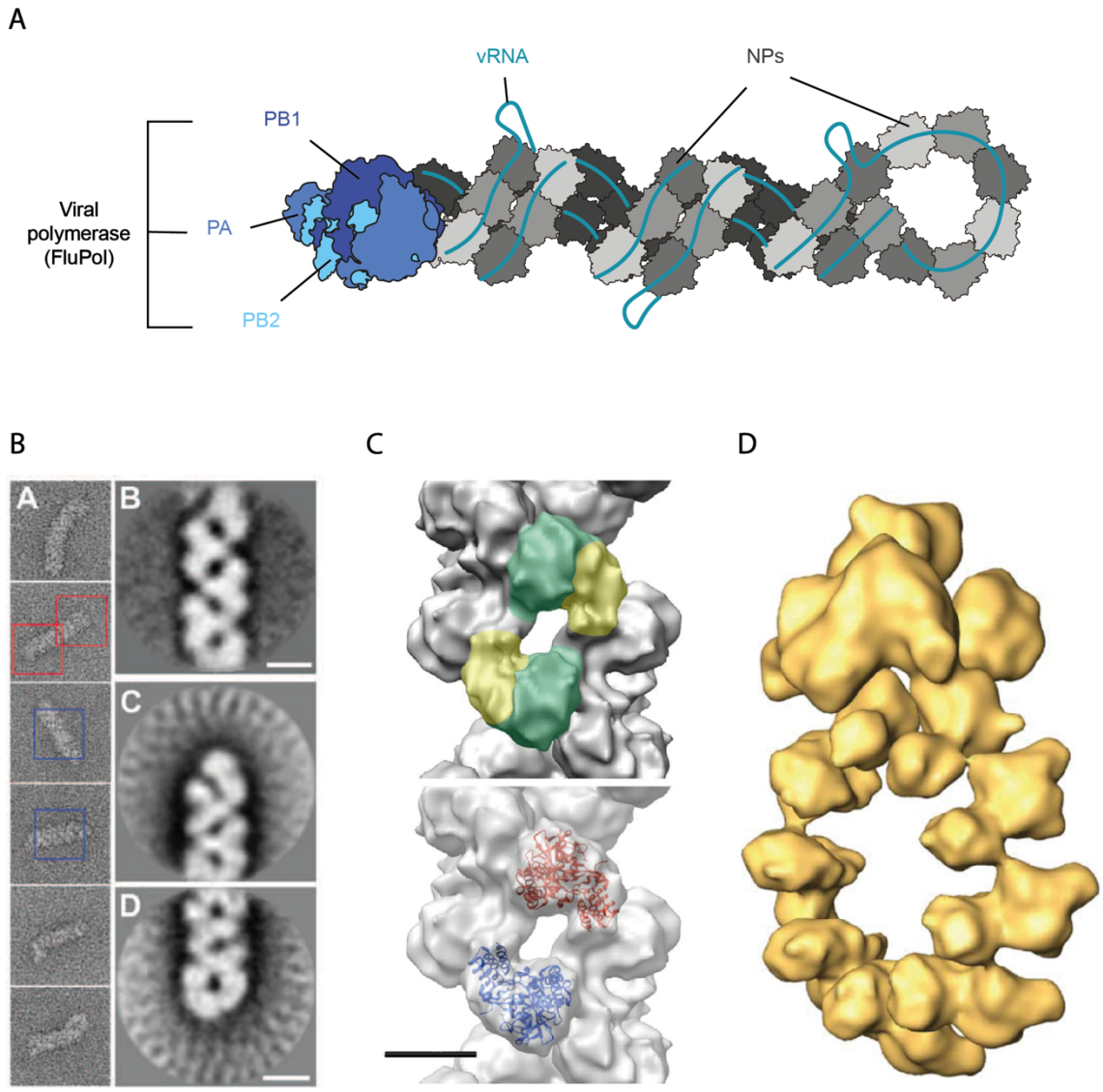


Figure 1.3 Structure of the influenza virus ribonucleoprotein complex. *A. Cartoon diagram of an influenza virus vRNP, with both vRNA termini being captured by a heterotrimeric viral polymerase and the rest of the vRNA bound by NP molecules. B. Electron micrograph of virion-derived vRNP showing the left-handed helical turns in the NP filament and the visible terminal NP loop. Adapted from Arranz et al., 2012. Scale bar: 100 Å. C. 3D reconstructed vRNP helical filament. Adapted from Coloma et al., 2020. Scale bar: 50 Å. D. 3D reconstructed mini-RNP. Adapted from Coloma et al., 2009.*

1.2.1 Influenza virus polymerase

More than 50% of the influenza viral genome encodes for one heterotrimeric protein complex — the viral polymerase, indicating the importance of this protein complex to the virus. Indeed, both viral transcription and replication depend on the RNA-synthesis activity of the viral polymerase. This heterotrimeric complex is composed of three similarly sized subunits: polymerase basic protein 1 (PB1), polymerase basic protein 2 (PB2), and polymerase acidic protein (PA, or P3 in influenza C and D viruses) (Te Velthuis & Fodor, 2016) (Figure 1.4A). Thanks to high-resolution structural maps generated using X-ray crystallography and cryo-electron microscopy (cryo-EM), many conformations of viral polymerases from all four types of influenza viruses have been revealed in the past decade, shedding light on the molecular mechanisms underlying transcription and replication by the influenza virus polymerase (Hengrung et al., 2015; Peng et al., 2019; Pflug et al., 2014; Reich et al., 2014; Serna Martin et al., 2018).

To understand the structure and function of a viral polymerase, one must visualize it not only as three linearized polypeptides, each with distinct functional domains, but also as one 3D entity with a rigid core and multiple flexible domains attached. These flexible domains can move relative to the core and to each other to form different polymerase conformations, which support the multifunctional nature of viral polymerase.

PB1 is known as the catalytic subunit, as it contains an RdRp domain ($PB1^{RdRp}$). The $PB1^{RdRp}$ is divided into six motifs, named motif A to F. Prior to catalysis, motifs D and F direct free nucleotides into the active site along a nucleotide entry channel lined with charged basic amino acids (Biswas & Nayak, 1994; Chu et al., 2012). The RNA synthesis active site is contained

within motif C. Two aspartate residues D445 and D446 (D446 and D447 in influenza C and D viruses) are found at the active site, each coordinating a magnesium ion, which catalyses the hydrolysis reaction between the 5' triphosphate group of the incoming nucleotide and the 3' hydroxyl group of the polymerase-bound nascent RNA primer, or RNA product (Chaturvedi & Shrivastava, 2005). Like other RdRps, the tertiary structure of the $PB1^{RdRp}$ displays a canonical right-hand fold with easily identified $PB1^{thumb}$, $PB1^{palm}$, $PB1^{fingers}$, and $PB1^{fingertips}$ subdomains (Venkataraman et al., 2018). In some polymerase conformations, a flexible β-ribbon located at the end of the $PB1^{thumb}$, called the priming loop, can be found protruding into the active site (Fan et al., 2019; Pflug et al., 2014; Reich et al., 2014). The influenza virus priming loop is important for replication and has homologues in other viral polymerases, such as hepatitis C virus and Dengue virus (Iglesias et al., 2011; Mosley et al., 2012; Te Velthuis et al., 2016) (Figure 1.4A).

In the quaternary structure of the viral polymerase, PB1 makes up the rigid core together with the PB2 N-terminus (PB2-N) and PA C-terminus (PA-C). To maintain an intimate association with PB2-N and PA-C in the core, $PB1^{palm}$ and $PB1^{thumb}$ form a strong interface with PA-C, and the PB1 C-terminal extension domain binds tightly to PB2-N. PB2-N contains a lid domain ($PB2^{lid}$) which plays an important role in separating the duplex of product and template RNA during the elongation step of RNA synthesis (Wandzik et al., 2020). On the other hand, PA-C serves as a large interface for binding of regulatory factors, such as host polymerase II (Pol II), via the Pol II C-terminal domain (CTD). Two CTD-binding sites on the influenza A, B and C viral PA-C have been identified, where basic residues on PA-C interact with the phosphorylated-serine-5 residues (pS5) of CTD (Lukarska et al., 2017; Serna Martin et al., 2018). This interaction is crucial for viral transcription initiation, especially cap-snatching, as it provides the viral polymerase with access to freshly-transcribed host-capped mRNA -

products of Pol II activity. Recent structural studies revealed that P3-C (PA-C equivalent in influenza C virus) is also involved in the formation of two distinct viral polymerase dimers and mediates the interaction between the viral polymerase and a host protein of the acidic nuclear phosphoprotein 32 (ANP32) family (Carrique et al., 2020; Fan et al., 2019) (Figure 1.4A-B).

The remainder of PA and PB2, namely the PA N-terminus (PA-N) and the PB2 C-terminus (PB2-C), make up the flexible domains that are difficult to resolve in cryo-EM maps. While PA-N consists only of the endonuclease domain (PA^{endo}) and a long flexible linker (PA^{linker}), PB2-C is made up of a series of globular domains linked flexibly to one another. These include the mid domain ($PB2^{mid}$), the cap-binding domain ($PB2^{cap}$), the linker ($PB2^{link}$), the 627 domain ($PB2^{627}$), and the nuclear localisation signal domain ($PB2^{NLS}$) (Te Velthuis & Fodor, 2016). Some domains have well-studied roles in polymerase function, but the characterization of functions of each flexible domain is still ongoing. PA^{linker} wraps around the viral polymerase core, joining up PA^{endo} and PA-C that sit on the opposite sides of PB1. PA^{endo} is structurally similar to other nucleases such as type II restriction endonucleases of *E. coli* (EcoRV) (Dias et al., 2009). As the name suggests, PA^{endo} possesses endonuclease activity and executes cap-snatching by cleaving a capped RNA primer off from host capped RNA. The active site of the endonuclease (to be distinguished from the PB1 RNA synthesis catalytic active site) contains a conserved D108 residue, in coordination with either a manganese or magnesium ion, that catalyses the endonucleolytic cleavage (Hara et al., 2006). A flexible loop (PA^{loop}) formed by PA residue 50-73 protrudes from the globular PA^{endo}. Equivalents of PA^{loop} are not found in influenza C virus, and its deletion in influenza A virus has no effect on the endonuclease activity but does inhibit replication (Nilsson-Payant et al., 2018) (Figure 1.4A-B).

PA^{endo} gains access to capped RNA through the help of $PB2^{cap}$, which binds specifically to the methyl-7-guanosine (m^7G) cap structure. Like in some host m^7G-binding proteins, two amino acids H357 and F404 in $PB2^{cap}$ form an aromatic sandwich with a π-bond in the middle that recognises the positively-charged nitrogen in m^7G (Tsurumura et al., 2013). $PB2^{627}$ received its name because it contains an amino acid residue, which is famously known to underlie host adaptation, at the 627 position. This particular amino acid underlies an important host-adaptation mechanism, in which a single mutation at the 627 position from glutamate (E) to lysine (K), could enable avian-adapted influenza viral polymerase to overcome the species barrier and replicate efficiently inside mammalian cells (Long et al., 2016). Finally, $PB2^{NLS}$, containing a bipartite nuclear localisation signal, is joined to $PB2^{627}$ via a long flexible linker, hence is often unresolved in cryo-EM structures (Figure 1.4A-B).

The mobility of the polymerase peripheral domains, granted by the flexible linkers between them, forms the structural basis of the polymerase's multifunctional nature. Recent structural studies have unveiled several distinct polymerase conformations, each characterized by unique positioning of flexible domains. Some of these conformations have been proposed to represent specific functional states, based on structural interpretations and supporting functional data (Zhu et al., 2023). Currently, structures with the most confident assigned functional states are transcriptase conformations. These conformations were observed in vRNA promoter-bound influenza A, B and C viral polymerases and are considered transcriptionally active because the positioning of the flexible domains is compatible with cap-snatching activity (Pflug et al., 2014; Reich et al., 2014; Serna Martin et al., 2018). Specifically, $PB2^{cap}$ is accessible for capturing host capped RNA and is positioned facing PA^{endo} to allow endonucleolysis. A similar conformation has also been observed in crystal structures, except with a 70-degree rotation in $PB2^{cap}$ (Reich et al., 2014). This slightly different conformation sheds light on the subsequent

stage of cap-snatching, where the cleaved capped RNA fragment is inserted into the polymerase RNA synthesis active site through the product exit channel (Walker & Fodor, 2019) (Figure 1.4C).

More recently, viral polymerases believed to be in the replicase and encapsidating conformations were observed in an influenza C virus asymmetric dimer structure (Carrique et al., 2020). These potential functional states were assigned because the product exit channel of the RNA-bound polymerase in the dimer structure aligns perfectly with the 5' cRNA promoter binding site of the apo polymerase. A typical RNA passage lined with basic amino acids was revealed, which is optimal for delivering product RNA. Notably, the vRNA-bound polymerase in this dimer structure differs significantly from the polymerase in transcriptase conformation, but appears structurally similar to a previously reported apo influenza C viral polymerase and a cRNA-bound influenza B viral polymerase (Hengrung et al., 2015; Reich et al., 2014; Thierry et al., 2016). Specifically, $PB2^{cap}$ is immobilized by packing closely against the $PB1^{palm}$, and $P3^{endo}$ is blocked because it is tightly packed against the $PB2^{NLS}$, rotated by 90°C relative to the transcriptase conformation. Similarly, $PB2^{627}$ is also rotated by 90°C, aligning with $PB2^{N2}$, $PB2^{mid}$, and $PB2^{link}$ to form a dimerisation interface with another viral polymerase (Figure 1.4D). However, since the 3' vRNA promoter is not resolved in the active site of the replicase, and there is no replication product in the structure, it remains undetermined whether this assigned replicase accurately represents a replicating polymerase.

In the assigned encapsidating polymerase of this dimer structure, there are marked differences in the arrangement of flexible domains compared to any previously observed conformations (Carrique et al., 2020). Most notably, several N-terminal domains of PB2 that make up the polymerase core are disordered, and the flexible PB2-C is packed closely against P3-C. This

conformation is likely to prevent either transcription or replication from occurring before a new RNP is fully assembled. It is important to note that these captured conformations represent snapshots of unknown stages during the dynamic RNA synthesis process, and further research is required to unambiguously assign specific functions (Figure 1.4E).

Polymerase subunits are synthesised individually in the cytosol and assembled into a heterotrimer with a 1:1:1 stoichiometry inside the nucleus. While PB2 is transported into the nucleus through an interaction between $PB2^{NLS}$ and importin-α, PB1 and PA are transported as a heterodimeric complex through binding to Ran binding protein 5 (RanBP5) via $PB1^{NLS}$ (Deng et al., 2006; Gabriel et al., 2011; Hutchinson et al., 2011; Pumroy et al., 2015; Tarendeau et al., 2007).

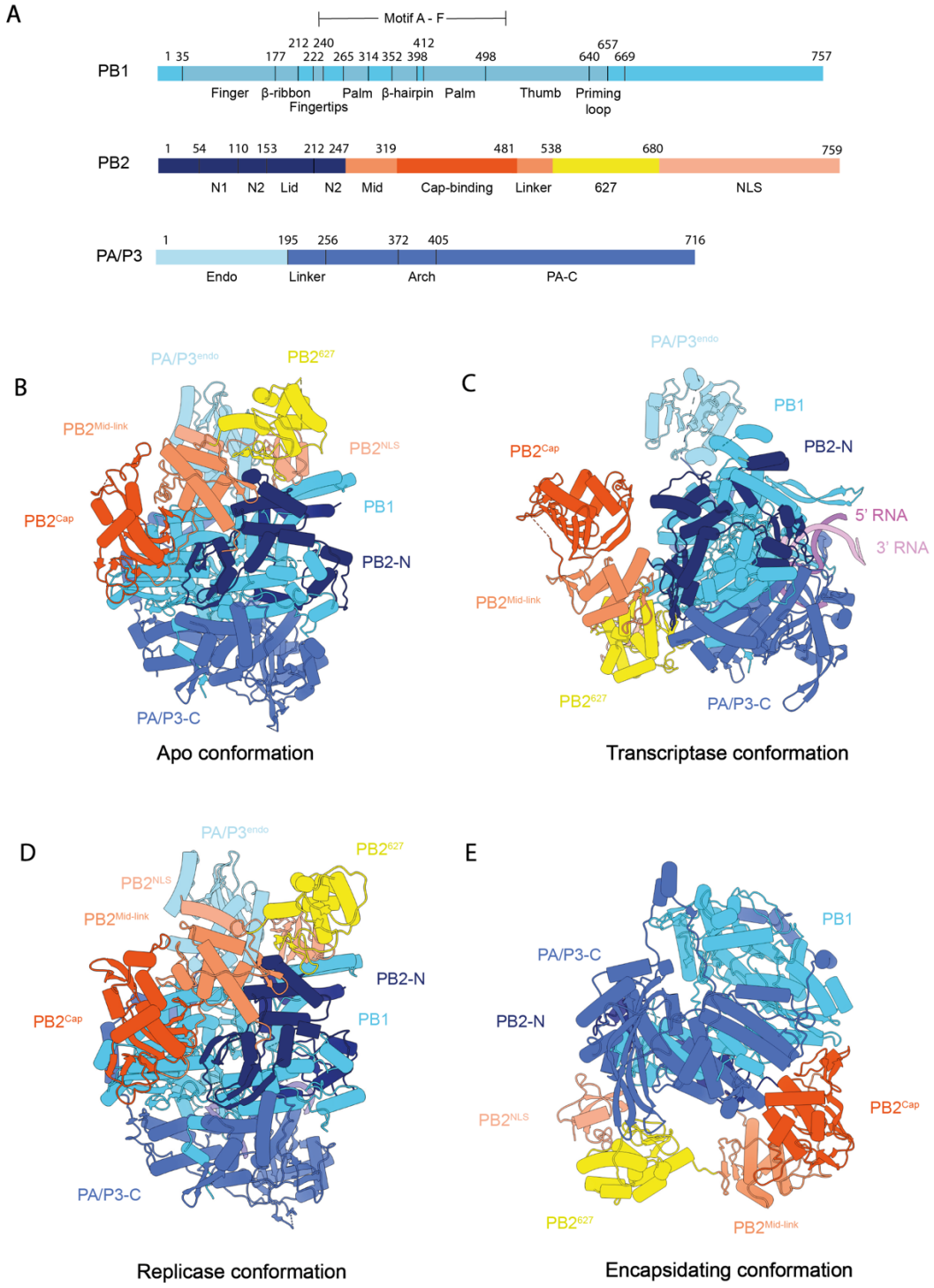


Figure 1.4 Structure and conformations of the influenza viral polymerase. *A. Schematic diagram illustrating the linear relationship of influenza A viral polymerase subunit domains. B. Cartoon structural models of influenza viral polymerase in apo conformation (PDB: 5D9A), transcriptase conformation (PDB: 4WSB), replicase conformation (PDB: 6XZR) and encapsidating conformation (PDB: 6XZR), showing the relative positioning of flexible domains in 3D space.*

1.2.2 Influenza virus nucleoprotein

Another major component of the RNP is NP, which makes up more than 90% of the mass of the RNP. Crystallographic studies of NP from influenza A, B and D viruses have unveiled that NP has a common crescent shape comprised of a head domain, a body domain and a tail loop (Donchet et al., 2019; Knight et al., 2021; Ng et al., 2012; Tang et al., 2021; Ye et al., 2006).

In the RNP filament, NP takes on an oligomeric form. This oligomerisation proceeds unidirectionally, with the tail-loop of one NP integrating into a groove in the body domain of the adjacent NP. In influenza A viruses, residue R416 within the tail-loop of one NP and residue E339 within the body domain of another NP have been identified to mediate the NP-NP interaction that leads to oligomerisation (Ye et al., 2006) (Figure 1.5A-B). NP oligomers containing different numbers of NP monomers have been seen in both crystal structures and under cryo-EM, and the orientations of NP relative to neighbouring NPs within an RNP varies due to the presence of a flexible linker between the tail-loop and the rest of the protein (Chenavier et al., 2023; Donchet et al., 2019; Knight et al., 2021; Ng et al., 2012; Ng et al., 2009). These flexibilities within an NP oligomer pose problems to structurally studying NP using any averaging methods for data processing, which is a common approach in current structural studies.

Purified NP naturally oligomerises in solution, suggesting that the conformational change in NP that accompanies NP oligomerisation is an energetically favourable process; however, during infection, newly synthesised NP is likely precluded from oligomerising before it is recruitment into new RNPs. The exact mechanisms underlying the regulation of NP's oligomeric state as well as its recruitment into RNP remain to be revealed. Post-translational

modifications (PTMs), such as phosphorylation, or the binding of NP to host chaperones and trafficking proteins such as importins and UAP56, likely play a role in keeping NP monomeric prior to RNP-assembly (Kawaguchi et al., 2011; Mondal et al., 2015; Turrell et al., 2015).

NP is thought to play a structural role as well as a regulatory role during transcription and replication. The primary function of NP is to provide a scaffold for the binding of vRNA and cRNA in an RNP, with an estimated ratio of 24 nt per one molecule of NP (Ye et al., 2006). The RNA binding ability of NP is granted by an RNA-binding groove nested between the head and body domains. This RNA-binding groove consists of two regions, G1 and G2, and is lined with basic amino acids. While G1 appears to be always accessible for RNA, structural studies suggest that G2 may be blocked by the C-terminal tail of NP when NP is in monomeric conformation (Chenavas et al., 2013; Knight et al., 2021) (Figure 1.5C).

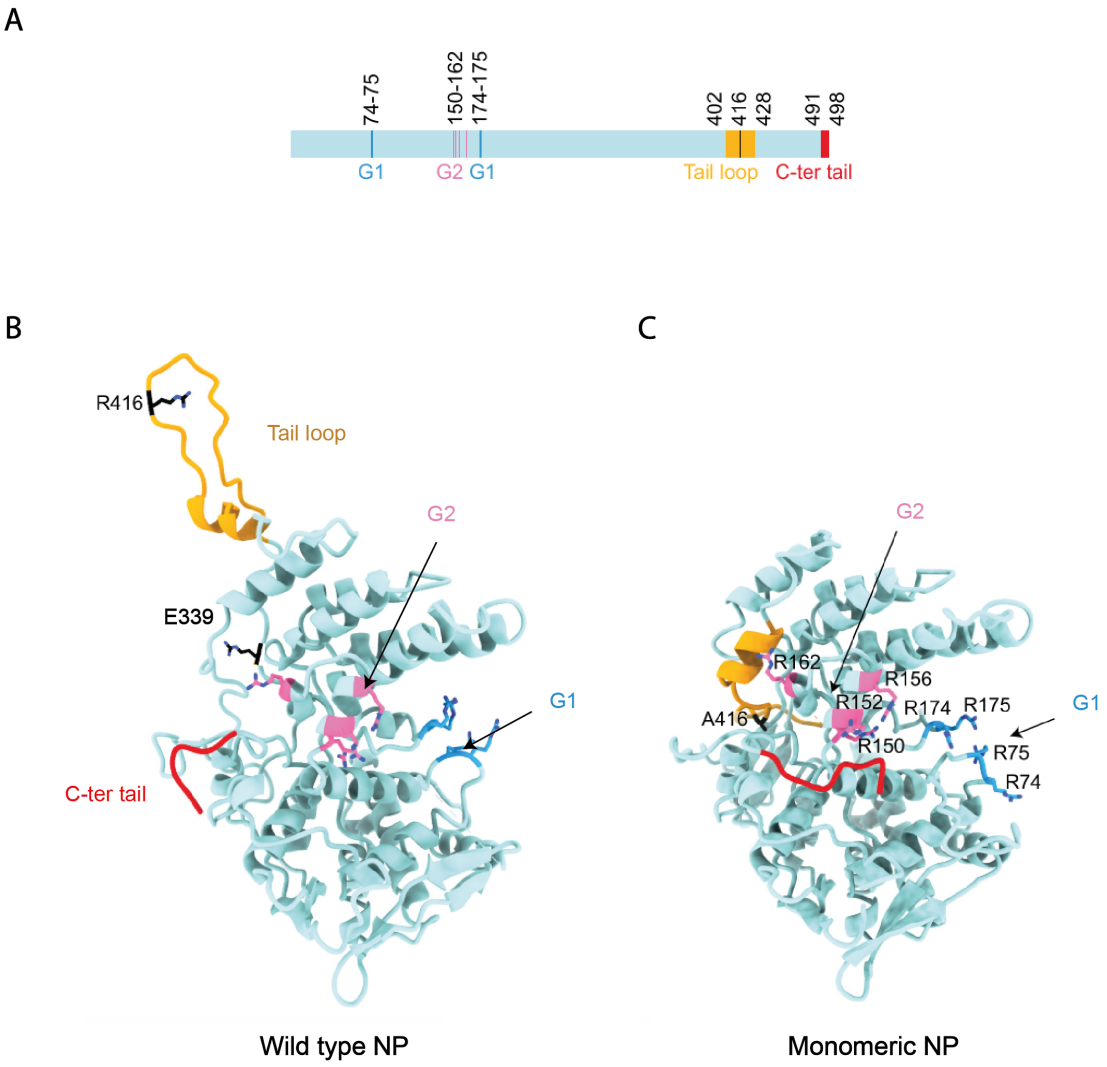


Figure 1.5 Structure of the influenza virus nucleoprotein. *A. Schematic diagram of the influenza A virus NP, showing the linear relationship between functionally important features in an NP molecule. B. Cartoon structural models of influenza A viral WT NP (PDB: 2Q06) and monomeric NP R416A (PDB: 3ZDP, mNPA), showing the relative positioning of tail loop, C-terminal tail and RNA binding grooves (G1 and G2) in 3D space. Key amino acids are shown in stick mode and labelled accordingly. Adapted from Wang et al., 2021.*

1.2.3 Intra-RNP interactions

1.2.3.1 Viral polymerase-NP interactions

High-resolution structural maps of viral polymerase and NP are being added to our databases at an astonishing speed, thanks to recent advancements in the production of high-quality recombinant proteins and improvements in the quality of raw data and processing power in crystallography and cryo-EM. However, intra-RNP and intermolecular interactions between the three RNP components, namely viral polymerase, NP, and resident RNA, are still not fully understood.

The interaction between viral polymerase and NP has been studied using various approaches, including immunoprecipitations in the context of cellular extracts or purified recombinant proteins, molecular docking based on known structures, surface plasmon resonance, and fluorescent microscopy. Several structural attempts have been made to resolve the viral polymerase's role within a functional RNP. Initially, a biologically active mini-RNP with a 248 nt vRNA and nine NP molecules was used for negative staining and cryo-EM (Coloma et al., 2009). Subsequently, virion-derived vRNPs and RNPs made from transfection were also imaged using cryo-EM (Arranz et al., 2012; Moeller et al., 2012). These studies claim to have identified the viral polymerase within the filament and suggested possible interactions between the polymerase PB2 subunit and the two proximal NPs. However, the resolution of the structures was low, being approximately 20 Å around the potential polymerase, hence it has been impossible to draw any firm conclusions. Additionally, the polymerase-NP interactions are likely to be distorted in the mini-RNP when compared to the naturally existing RNP due to the artificial constraints imposed on the two proximal NPs. Furthermore, in the two full-length

RNP structures, additional electron densities that could not be assigned to either viral polymerase or NP were present, casting further doubt on the identity of the potential polymerase.

Biochemical studies have revealed multiple interaction sites on both the viral polymerase and NP that mediate polymerase-NP interactions. Consistent with the RNP structural studies, PB2 has been recurrently identified as an interaction partner of NP, with some claims of $PB2^{627}$ playing a role in mediating this interaction (Biswas & Nayak, 1994; Biswas et al., 1998; Gui et al., 2014; Ng et al., 2012). However, a recent study also identified PA as an interaction partner of NP through residue 1-27 (Vidic et al., 2016). Several conserved residues on NP, such as D88, R204, W207, and R208, were identified as important for its association with viral polymerase through functional analysis (Gui et al., 2014; Marklund et al., 2012; Narkpuk et al., 2015; Tang et al., 2021). Nevertheless, the field has not produced any conclusive results on this matter, and the interaction between viral polymerase and NP still remains a mystery.

1.2.3.2 Polymerase-RNA interactions

RNA promoters refer to the terminal 14 to 15 nt at both ends of vRNA or cRNA, mostly bound by a viral polymerase in an RNP. Interactions between viral polymerase and its RNA promoter are important for maintaining the structural integrity of an RNP and are essential for polymerase activity (Cianci et al., 1995; Fodor et al., 1994; Pritlove et al., 1995). The 5' RNA promoter of vRNA and cRNA both form a hook-like structure by folding onto themselves via intramolecular base-pairing between residues 1-10. This 5' hook structure is tightly anchored in a pocket formed by the $PB1^{fingers}$ and covered by PA-C (Pflug et al., 2014). The binding of

the 5' vRNA promoter has been shown to activate an apo polymerase for its enzymatic activity, presumably by inducing the correct conformational rearrangement for the polymerase's active site. In contrast, 3' promoter of vRNA and cRNA are more dynamic, with a lower affinity compared to the 5' RNA promoter and multiple binding sites on the viral polymerase. Currently, there are three known 3' promoter binding sites: in the polymerase active site (Fan et al., 2019; Wandzik et al., 2020), on the polymerase surface close to the active site (Pflug et al., 2014), and in a groove formed at the junction of $PB1^{thumb}$, PA-C, and PB2-N (Fan et al., 2019; Peng et al., 2019). These positions likely represent the movement of the 3' promoter at different stages during RNA synthesis. The distal ends of the 3' and 5' promoter meet as they come out of the polymerase and form a short 5-8 base-pair RNA duplex which is stabilised by a PB1 ß-ribbon structure (Fodor et al., 1994; Kouba et al., 2019). Duplex formation occurs right after the 5' hook, from nucleotides 11 on the 5' promoter of both vRNA and cRNA; however, base-pairing starts from nucleotide number 10 on the 3' promoter of vRNA and number 12 on the 3' promoter of cRNA (Figure 1.6). This difference in base-pairing start sites between vRNA and cRNA 3' promoters is important because it leaves 9 nt of 3' vRNA but 11 nt of 3' cRNA inside the polymerase active site, positioning nucleotide number 2 and 3 of vRNA and nucleotide number 4 and 5 of cRNA opposite the catalytic ions when 3’ RNA enters the active site. An additional implication of this difference is that the long cRNA 3’ promoter is anchored within the viral polymerase, hence is less flexible compared to vRNA 3’ promoter (Robb et al., 2019). The variation in promoter positioning sheds light onto the different replication initiation processes on vRNA and cRNA templates and will be discussed further in later sections. Differences in promoter binding to viral polymerase, particularly of the 3' promoter, are likely to underlie the intrinsic identity differences between vRNP and cRNP, and are crucial for understanding the functional differences of vRNP and cRNP and how they might be distinguished by factors regulating RNP export.

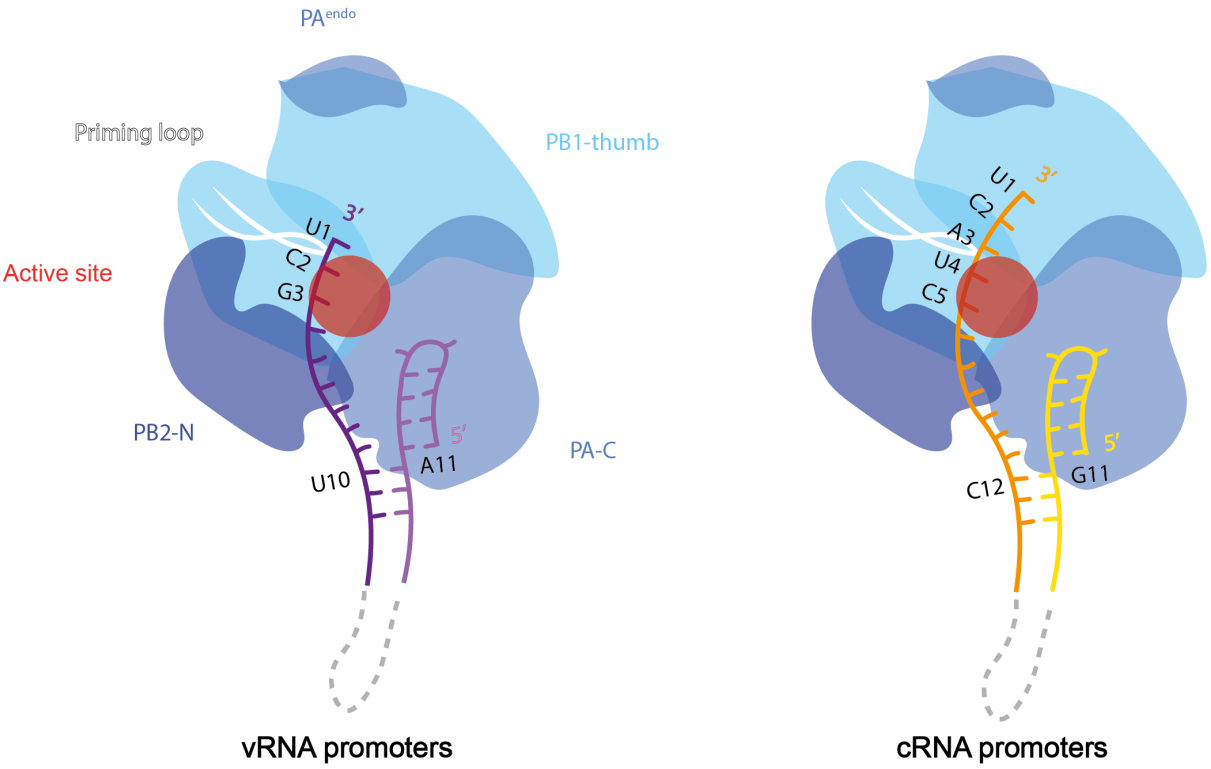


Figure 1.6 Structures of the vRNA and cRNA promoters. *Cartoon model of vRNA (left) and cRNA (right) 3' and 5' promoters, showing the difference in promoter base pairing and the positioning of the RNAs in the polymerase active site due to different 3' promoter lengths.*

1.2.3.3 NP-RNA interactions

The most extensive intermolecular interaction within an RNP occurs between NP and RNA. NP-RNA interactions also underlie the mechanism of vRNA and cRNA encapsidation during RNP assembly. Binding assays and sequencing data show that both G1 and G2 bind to the negatively charged phosphate backbone of RNA without displaying sequence specificity (Baudin et al., 1994). However, obtaining a high-resolution structure of NP binding to RNA by cryo-EM is difficult due to the small size of monomeric NP and the flexibility of oligomeric

NP. Indeed, to date, known high-resolution structures of purified NP contain no more than 3 fully resolved nucleotides, and in structures of RNPs or RNP-like particles where longer RNA was resolved, the local resolution of RNA density has never been below 4.5 Å (Chenavier et al., 2023; Ng et al., 2008; Ng et al., 2009; Ng et al., 2012a; Tang et al., 2021).

1.3 INFLUENZA VIRUS TRANSCRIPTION AND REPLICATION

The influenza viral polymerase is responsible for transcription and replication, producing 5' m^7G-capped and 3' polyadenylated viral mRNA, or 5' triphosphorylated and 3' unmodified cRNA and vRNA, respectively. This multifunctionality requires the polymerase to be capable of different enzymatic activities, despite only synthesising one type of product at a time. Kinetic studies on mRNA, cRNA, and vRNA accumulation reveal that transcription and replication are temporally segregated over the course of infection. Specifically, mRNA is produced early in infection, whereas cRNA and vRNA are only detectable after a sufficient level of viral protein accumulation is reached (Kawakami et al., 2011; Phan et al., 2021).

The following sections summarize our current molecular and structural understanding of: 1. how distinct enzymatic activities of the viral polymerase leads to the synthesis of mRNA, cRNA, and vRNA; 2. how viral and host factors are involved in regulating transcription and replication timing at different stages of infection.

1.3.1 Initiation

1.3.1.1 Primer-dependent initiation in mRNA synthesis

Unlike some other viruses that carry a capping enzyme, such as dengue and zika viruses, influenza viruses generate viral mRNA with 5' m^7G by stealing a cap-structure from host-capped RNA - a process known as cap-snatching (Bouloy et al., 1978; Dias et al., 2009; Plotch et al., 1981; Yuan et al., 2009). The majority of cap-donor RNA used by influenza virus transcription are mRNA products of the host DNA-dependent-RNA-polymerase II (Pol II). Pol II is regulated at various stages of transcription via phosphorylation on the Pol II C-terminal domain (CTD) (Jeronimo et al., 2016). Pol II CTD consists of 52 repeats of a heptapeptide, YSPTSPS, with 3 serine (S) residues at positions 2, 5, and 7 that can be phosphorylated at different time points during host mRNA synthesis. Serine-5 phosphorylation (S5p) occurs at the host transcription initiation phase and helps to recruit factors essential for the formation of a 5’ cap on the nascent mRNA product. On the other hand, serine-2 phosphorylation (S2p) is associated with the transcription elongation phase and helps to recruit cap-binding proteins after capping occurs (Lidschreiber et al., 2013). To ensure cap-snatching occurs during the short window between host-mRNA capping and elongation, the influenza virus polymerase specifically recognises and binds to S5p CTD of Pol II before viral transcription. Structures of viral polymerases associated with S5p Pol II CTD mimetic peptides show that polymerases of influenza virus A, B, and C all bind to S5p CTD, although they have slightly different CTD binding sites (Lukarska et al., 2017; Martínez-Alonso et al., 2016; Serna Martin et al., 2018). The binding of a Pol II CTD on a viral polymerase was found to be compatible with the polymerase adopting the transcriptase conformation. As described in an earlier section, this conformation allows $PB2^{cap}$ to bind to a nascent mRNA emerging from Pol II, and PA^{endo} to

cleave off a 10-15 nt capped RNA fragment (De Vlugt et al., 2019) (Figure 1.7A). The length of the cleaved RNA fragment is determined by the distance between $PB2^{cap}$ and PA^{endo}, and the variation is believed to be caused by the flexibility of their relative positions and a slight preference for cleavages after purines, particularly guanine (G). This enables the capped RNA fragment to form one or two Watson-Crick base pairs with the 3' vRNA promoter when the RNA primer is redirected into the polymerase active site by a 70-degree rotation of $PB2^{cap}$ around its flexible hinge (Reich et al., 2014). Transcription initiation occurs when the first guanine or cytosine (C) is added to the capped RNA primer, depending on whether its base pairing with the template happens at uridine 1 (U1) only or U1C2 (Figure 1.7B). As more nucleotides are added to the product mRNA, the template-product duplex becomes more stable. As a result, the vRNA promoter duplex is weakened and collapses through PB1 internal conformational changes, allowing elongation to proceed (Wandzik et al., 2020). As the template-product duplex grows, the priming loop rotates and unblocks the product exit channel and $PB2^{lid}$ helps to separate the duplex, which allows the single-stranded product RNA to leave through the product exit channel (Fan et al., 2019; Kouba et al., 2019).

1.3.1.2 Primer-independent terminal initiation in cRNA synthesis

In contrast to capped primer-dependent initiation, replication that produces 5' triphosphorylated cRNA and vRNA relies on primer-independent initiation, also known as *de novo* initiation. As described in an earlier section, due to the different positions where the promoter duplex forms on the 3' promoter of vRNA and cRNA, the lengths of the 3' vRNA and cRNA promoters that reside in the polymerase active site differ. The 3' vRNA promoter places 9 nt inside the active site, with C2 and G3 resting opposite the two catalytic residues. In

contrast, the slightly longer 3' cRNA promoter places 11 nt inside the active site, with U4 and C5 resting opposite the catalytic residues (Figure 1.6). The difference in the positioning of the 3' promoter in the polymerase active site results in two types of *de novo* initiations on vRNA and cRNA templates: terminal and internal initiations (Deng et al., 2006).

Adenosine triphosphate-guanosine (pppApG) dinucleotide, instead of other dinucleotides, typically forms during *de novo* initiation due to its structural stability and energetic favourability (Reich et al., 2017). As we established above, the 3' vRNA promoter normally positions C2 G3, instead of U1 C2, opposite the two catalytic residues D445 and D446. This structural arrangement suggests that energetically favourable pppApG formation on the 3' vRNA promoter relies on the repositioning of the 3' vRNA U1C2 next to the catalytic residues. Although a polymerase structure revealing the arrangement of the active site at the point of cRNA initiation has not yet been captured, biochemical studies and models of transcription initiation on vRNA templates suggest a potential role of the priming loop in accurately repositioning the 3' vRNA promoter at cRNA initiation (Fan et al., 2019; Oymans & te Velthuis, 2017; Pflug et al., 2014; Te Velthuis et al., 2016). These structures revealed that the priming loop tip residue P651 is in contact with the sugar-base of the 3' vRNA template, buttressing the 3' terminus towards the catalytic residues, hence providing a surface against which free pppA and pppG are stabilized and catalysed to form the first dinucleotide at the terminus of the 3' vRNA template. The requirement of the priming loop for efficient terminal initiation on the vRNA template is backed up by functional studies showing deficient cRNA synthesis upon mutations or truncation of the priming loop (Figure 1.7B).

1.3.1.3 Primer-independent internal initiation in vRNA synthesis

In contrast, since the 3' cRNA promoter has already placed 4U 5C opposite the catalytic aspartates, initial pppApG dinucleotide formation on the cRNA template naturally has high thermodynamic stability (Te Velthuis et al., 2016). However, this results in the initial dinucleotide being generated downstream of the 3' cRNA terminus, in a process called internal initiation. In agreement with this, the priming loop is shown to not be essential for efficient internal initiation of replication by a functional study (Te Velthuis et al., 2016). However, since initiation occurs internally on the 3' cRNA template, realignment of the initial dinucleotide to the template terminus prior to elongation is necessary to maintain genome integrity. In other words, realigning an internally initiated ApG dinucleotide to the 3' cRNA terminus is essential to produce a complete copy of cRNA. The priming loop, as well as a helix-loop-helix motif (HLH) of the polymerase PB1 subunit, have been shown to be important in this realignment process in several functional studies, where mutations at the tip of both the priming loop and in the HLH result in the production of failed-realignment products (Figure 1.7C) (Oymans & te Velthuis, 2017).

Similar to transcription, after more nucleotides are added to the product cRNA or vRNA, a product-template duplex is formed, which destabilizes the promoter duplex and allows elongation to proceed. Additionally, to make way for the extending RNA product, the priming loop must exit the active site, as seen in a previous polymerase structure where the priming loop extrudes into the solvent and projects towards the PB1 and PB2 interface (Kouba et al., 2019).

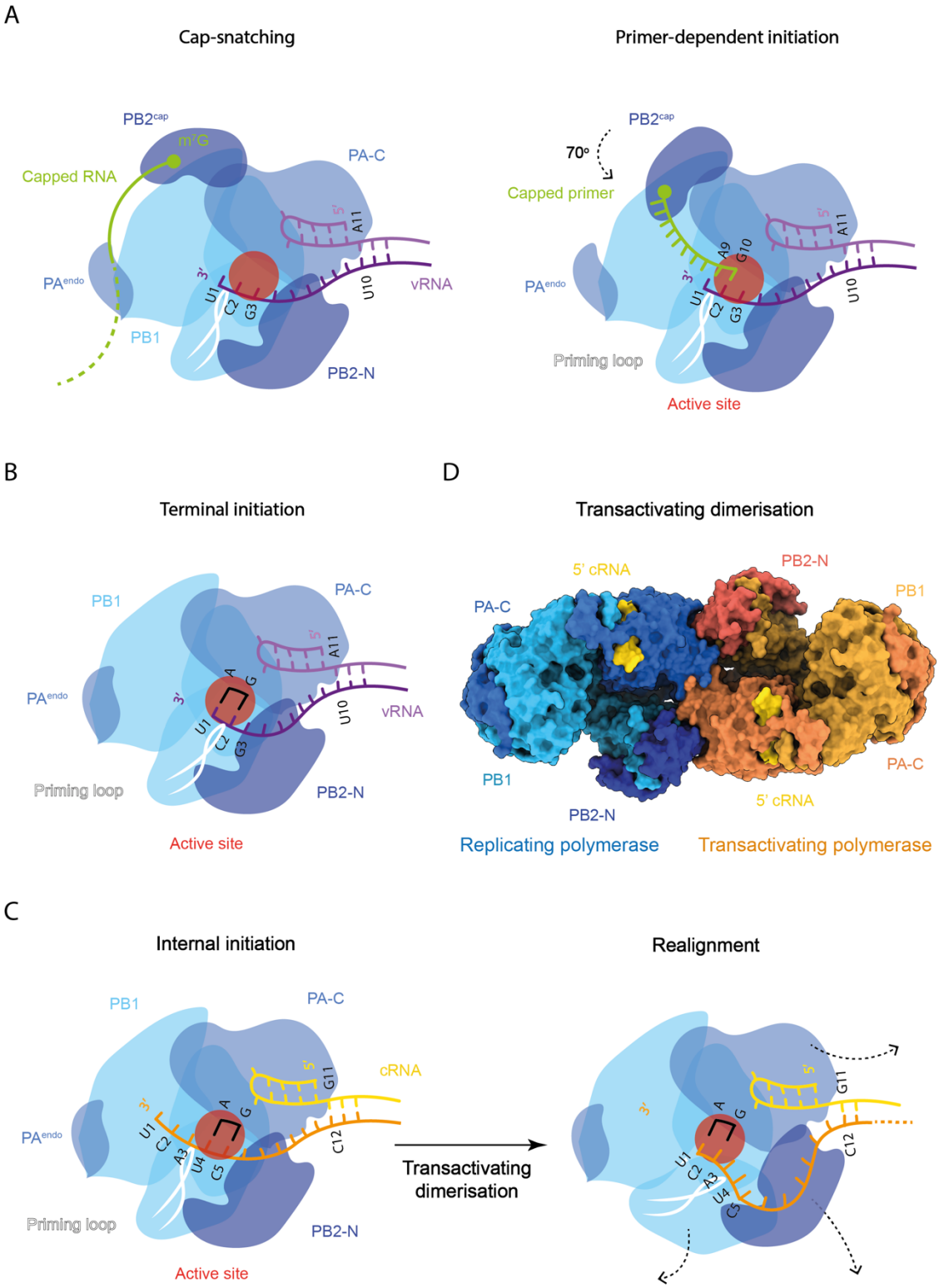

A
Cap-snatching
Primer-dependent initiation
PB2cap
m7G
Capped RNA
PA-C
PAendo
PB1
PB2-N
vRNA
70°
Capped primer
Priming loop
Active site
B
Terminal initiation
Priming loop
Active site
D
Transactivating dimerisation
5' cRNA
PB2-N
PB1
PA-C
Replicating polymerase
Transactivating polymerase
C
Internal initiation
Realignment
cRNA
Transactivating
dimerisation

Figure 1.7 Cartoon models of transcription and replication initiation. *A. Viral polymerase carries out cap-snatching of host capped RNA by binding to the RNA cap and cleaving downstream of the 3' terminus. Cleaved capped RNA primer is fed into the active site and initiates transcription. Dotted line shows the movement of the PB2 cap-binding domain. B. Terminal initiation occurs on vRNA template when the terminal U1C2 are positioned inside the active site. C. Internal initiation occurs on cRNA template and the ApG dinucleotide is realigned to the cRNA 3' terminus through action of the priming loop upon polymerase dimerisation with a transactivating polymerase resulting in a symmetric polymerase dimer. Dotted lines show the movements of flexible domains. D. Surface representation of a symmetric dimer of the influenza A virus polymerase (PDB: 6QX8). The three subunits of the replicating and transactivating polymerase are shown in different shades of blue and orange, respectively. B-D are adapted from Zhu et al.*

1.3.1.4 Polymerase symmetric dimerisation in replication

Viral polymerase dimerises in solution and has been observed in a number of crystallography or cryo-EM structures (Chang et al., 2015; Fan et al., 2019). A recent high resolution cryo-EM structure of the viral polymerase symmetric dimer has shed light on the potential mechanism of template realignment after internal initiation on the cRNA template. In this structure, two viral polymerases form a symmetric dimer through an interface that involves PA-C, $PB1^{thumb}$, and $PB2^{N1}$. Notably they are conformationally identical and both compatible with replication. By comparing the monomeric and dimeric structures of influenza A virus polymerase bound to cRNA promoter, the authors revealed a dimerisation-induced movement of a helical bundle in the polymerase core, formed by $PB1^{thumb}$ and $PB2^{N1}$. The authors proposed that this movement causes the priming loop of the polymerase to shift approximately 7 Å away from the active site, hence destabilising the 3' promoter and opening the 3' promoter binding site within the active site (Fan et al., 2019). Collectively, these structural rearrangements allow the destabilized 3' cRNA template to be translocated by the priming loop, realigning the 3' cRNA

terminus against the newly formed pppApG dinucleotide. Intervention in this symmetric dimerisation either through mutations or by blocking the dimerisation interface, has been shown to specifically impair the cRNA-to-vRNA replication step (Fan et al., 2019) (Figure 1.7D).

1.3.1.5 Polymerase asymmetric dimerisation in replication

During infection, cRNA and vRNA are detected after viral polymerase and NP accumulate to a threshold level, suggesting an involvement of viral polymerase and NP in both steps of replication. The most likely role of newly synthesised viral polymerase and NP in replication is to assemble nascent cRNA and vRNA into RNPs. Unlike viral mRNA, which does not trigger a host immune response with its 5' cap and 3' poly(A) structures, naked cRNA and vRNA can induce RNA-sensing innate immune pathways via pattern recognition receptors (PRRs), such as Toll-like receptors (TLRs) and retinoic acid-inducible gene I (RIG-I)-like receptors (RLRs) (Jensen & Thomsen, 2012). For this reason, newly produced cRNA and vRNA are believed to be assembled co-replicatively into RNPs, minimizing the exposure of the viral genome to host innate immune sensors. Additionally, host nuclear proteins belonging to the ANP32 family have been found to be important specifically for influenza viral replication and underlie the host-range specificity of the zoonotic influenza virus (Long et al., 2016).

In contrast to the symmetric dimer described above, a recent polymerase asymmetric dimer structure was resolved (Carrique et al., 2020). This structure leads to the idea of the formation of a replication platform for co-replicative assembly of RNP, which is required for both steps

of replication. Compared to symmetric dimer, asymmetric dimer has been observed in significantly lower abundance under cryo-EM and only in samples containing ANP32 proteins.

The influenza C virus polymerase asymmetric dimer observed in Carrique et al. reveals what is likely to be a platform at which viral genome replication occurs, providing valuable insight into the molecular mechanisms of replication, involving additional viral polymerase and host ANP32 family proteins. In this asymmetric dimer structure, one viral polymerase is bound to vRNA promoter, while the other polymerase lacks any RNA interactions. The product exit channel, found along PB1 finger and $PB2^{cap}$, of the RNA-associated polymerase aligns perfectly with the 5' cRNA binding site of the apo polymerase (Figure 1.8A). The two polymerases are connected by a passage of basic residues, which is optimal for RNA delivery (Figure 1.8B). Such ideal positioning implies the potential roles of each polymerase in the dimer: the RNA-bound polymerase acts as a replicase, whereas the RNA-free polymerase functions as an encapsidating polymerase. In agreement with this conclusion, the conformations of the RNA-bound polymerase and the apo polymerase are different from the transcriptase conformation, as described previously.

A copy of ANP32 family protein can be found in this structure, bridging the dimer by forming extensive contacts with both polymerases. ANP32 family proteins are found to specifically support influenza virus replication. Different isoforms of ANP32 family proteins exist, for instance, ANP32A and B in chicken (chA and chB) and ANP32A, B, and E in humans (huA, huB, and huE). However, the ability of specific viral polymerases to utilize different ANP32 family proteins varies. The most well-known example is the incapability of avian-adapted viral polymerase to utilize mammalian ANP32 proteins, which underlies the molecular basis of an important cross-species barrier of influenza virus as a zoonotic agent (Long et al., 2019;

Peacock et al., 2019). Intriguingly, a single amino acid mutation at PB2 residue 627 (residue 649 in influenza virus C) from glutamic acid into non-acidic lysine (E627K) was found to be enough to grant the avian viral polymerase the ability to utilize human ANP32 proteins (Almond, 1977; Long et al., 2019; Subbarao et al., 1993). This remarkable host-adaptation can be explained at the molecular level in the context of the asymmetric dimer structure. As observed in the structure, the N-terminus of the ANP32 leucine-rich region ($ANP32^{LRR}$) interacts with the replicating polsymerase at $PB2^{lid}$, $PB2^{627}$, and $PB2^{NLS}$, and its C-terminus interacts with the encapsidating polymerase at P3-C and PB1-N. The rest of ANP32, called the low complexity acidic region ($ANP32^{LCAR}$), is not fully resolved, but continuous electron density suggests that it forms contact with $PB2^{627}$ of the replicating polymerase (Figure 1.8A). In birds, $ANP32^{LCAR}$ has an extra 33-amino acid insertion compared to mammals. This insertion contains a mixture of acidic and basic residues and is shown to be capable of interacting with $PB2^{627}$ regardless of the acidity of residue 627 (residue 649 in influenza C virus). In contrast, mammalian ANP32 positions a series of highly-acidic residues, further downstream of $ANP32^{LCAR}$, opposite $PB2^{627}$. Consequently, huA and huB can only form a stable interaction with $PB2^{627}$ if the glutamic acid residue at 627 (649 in influenza C virus) is mutated to a less acidic residue.

While what constitutes a replication platform has been revealed, the sequence of events that lead to the formation of this structure remains a mystery. In the dataset that led to the reconstruction of the replicase-encapsidating polymerase dimer structure, a proportion of polymerase dimer was found to be missing the bridging ANP32 protein. Recently, our laboratory observed a single influenza A viral polymerase in an encapsidating conformation, bound to a molecule of ANP32 protein. In addition, the structure of a molecule of ANP32 protein binding only to a replicating polymerase in the absence of an encapsidating polymerase

has never been seen. The sequence of events leading to the formation of the replication platform could be deduced as follows. Firstly, a host ANP32 family protein captures a newly synthesised apo viral polymerase, locking the viral polymerase in an encapsidating conformation. Next, this encapsidating polymerase-ANP32 complex approaches an RNP-associated polymerase. A strong interaction between ANP32^{LCAR} and PB2^{627} of the RNP-associated polymerase stabilizes the asymmetric dimerisation between the RNP-associated polymerase and the apo polymerase. In this way, a replication platform is formed, causing the RNP-associated polymerase to assume a replicase conformation, which is compatible with primer-independent initiation of viral genome replication.

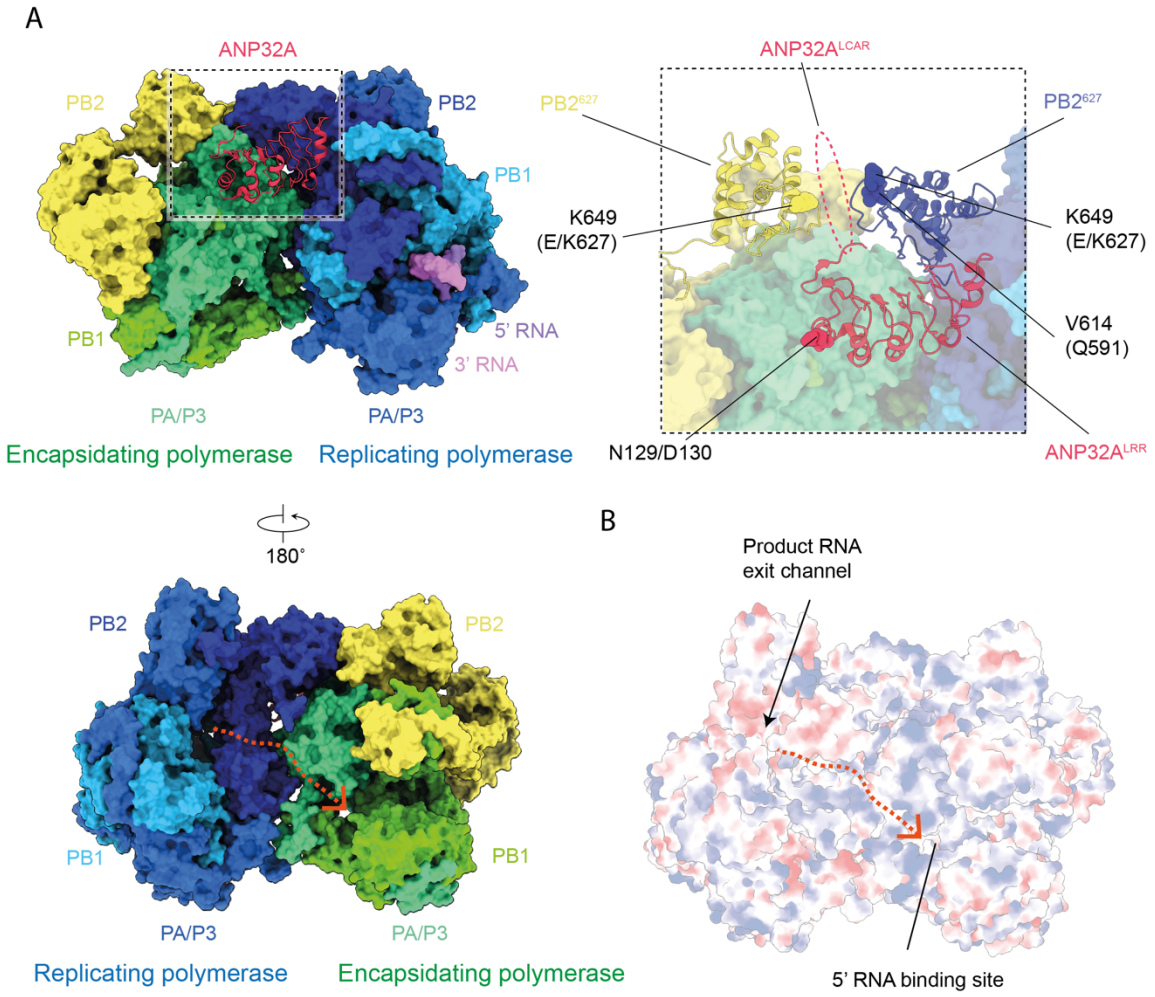


Figure 1.8 Structural model of an influenza virus genome replication platform. *A. Surface representation of an asymmetric dimer of the influenza C virus polymerase in complex with chicken ANP32A protein (PDB: 6XZR) (top left), with residues of interest between the replicating and encapsidating polymerase highlighted (top right). The 180-degree horizontally rotated model (bottom left) highlights the potential path for emerging product RNA. B. Electrostatic potential map of the replication platform showing the positively charged path for product RNA exit. Adapted from Zhu et al (2023).*

1.3.2 Elongation

After initiation, the viral polymerase must copy the rest of the template to produce mRNA, cRNA and vRNA. The elongation process is likely to be similar in viral genome transcription and replication. Regarding the viral polymerase, conformational rearrangements that favour subsequential elongation were observed after transcription initiation (Kouba et al., 2019; Wandzik et al., 2020). To grant viral polymerase access beyond 3' promoter, the RNA promoter duplex is disrupted. This disruption is caused by the growing template-product duplex, which destabilises the promoter duplex by rotating the stabilising PB1 β-ribbon. Meanwhile, as the $PB1^{thumb}$ opens up, the priming loop is extruded from the polymerase active site. This rearrangement frees up space for the template-product duplex to grow and unblocks the template exit channel. Thereafter, the template-product duplex strand separates through the action of stacking $PB2^{lid}$ on the latest duplex base pair (Kouba et al., 2019). Eventually, as elongation proceeds, the 3' end of the template RNA leaves the active site through the template exit channel and is sequestered at the 3' secondary binding site (mode B site, as mentioned in a previous section) (Walker et al., 2020; Wandzik et al., 2020). Consequently, during transcription and replication elongation, RNA template appears as two loops: a growing loop

and a shrinking loop, upon entering and exiting the polymerase, respectively. This configuration persists until termination occurs.

To copy the entire genome, the viral polymerases must gain access to the NP-coated RNA segment. Two possible models were proposed for this process: the first suggests that the template-associated NP filament completely disassembles and only reforms once RNA synthesis finishes. Alternatively, the other model hypothesized that the catalytic polymerase only causes local unwinding of the NP-RNA complex as it travels along the helical filament. Consequently, NP must release and recapture the template as RNA synthesis proceeds. Structural evidence is in line with the second model: polymerase structures show that the template entry and exit channels are located proximally on the polymerase, and RNP images reveal intact, NP-coated RNPs during transcription (Coloma et al., 2020; Kouba et al., 2019). Additionally, a recent structural study revealed extreme flexibility of NP-NP interactions within the RNP and proposed that NP filaments can slide via a mechanism called the 'processive helical track' in the parental RNP (Coloma et al., 2020). This high flexibility and proposed sliding movement between the two anti-parallel NP strands allow the RNP-bound polymerase to disrupt local NP-RNA structure and move within an RNP while staying bound to both ends. The model suggests that NP is temporarily released from template RNA as it enters the polymerase active site, and reassociates to template RNA as it emerges from the template exit site. This mechanism maintains RNP structure and integrity and allows for multiple rounds of transcription and replication per vRNP. Although this model is only shown for transcription, it is likely that the same mechanism happens in both steps of replication (Figure 1.9A).

To ensure multiple rounds of RNA synthesis on the same template, in particular mRNA synthesis, it is crucial that NP within parental RNP is conserved before and after transcription. In agreement to this proposal, additional NP is not needed for primary mRNA production in infected cells and *in vitro* in virion-purified vRNP. However, how template-associated NP is transiently displaced during RNA synthesis is poorly understood. Although there is no structural data for the underlying mechanism, biochemical data (as mentioned in an earlier section) indicate dynamic and transient interactions between NP and viral polymerase (Biswas & Nayak, 1994; Biswas et al., 1998; Gui et al., 2014; Marklund et al., 2012; Narkpuk et al., 2015; Tang et al., 2021; Vidic et al., 2016), which could underlie the molecular mechanism of how template-associated NP translocates from template entry to template exit during RNA synthesis.

The main difference between transcription and replication at the elongation step is whether new NP is recruited to product RNA. Naked RNA is at risk of being degraded or being recognised by cellular immunity pattern recognition receptor (PRRs). While *de novo* synthesised mRNA can be bound and protected by host mRNA binding proteins, cRNA and vRNA are believed to be assembled co-replicatively into RNPs (Carrique et al., 2020). After the 5' end of the nascent replication product is captured by an encapsidating polymerase, NP needs to be recruited onto the growing product RNA. Unlike some non-segmented negative-sense RNA viruses, which possess a viral phosphoprotein (P protein) for NP recruitment, influenza virus lacks a P protein equivalent (Curran et al., 1995; Leyrat et al., 2011; Masters & Banerjee, 1988; Mavrakis et al., 2006). Recent structural and functional studies have led to a model for NP-recruitment, mediated by ANP32^{LCAR}. As mentioned earlier, in the asymmetric polymerase dimer structure with ANP32 protein, ANP32^{LCAR} is found to be disordered, underscoring its flexibility within the complex (Carrique et al., 2020). Recent biochemical data

show that the highly acidic ANP32^{LCAR} can interact with multiple RNA-free NPs by binding to the RNA-binding grooves, G1 and G2 (Wang et al., 2022). Additionally, ANP32^{LCAR} is shown to be important for replication of a long template but not a short template in cell culture, presumably because long templates produce long RNA products that must be assembled into RNPs to be elongated (Wang et al., 2022). Collectively, these observations suggest a mechanism for NP recruitment in influenza virus: ANP32 family proteins in the replication platform act as the P protein equivalents to recruit NP to the replication product via their LCAR domain (Zhu et al., 2023). Specifically, the extended and highly flexible ANP32^{LCAR} sequesters newly synthesised RNA-free NP, enhancing its local concentration around the cRNA and vRNA product, thereby facilitating the integration of NP into the new RNP (Figure 1.9B).

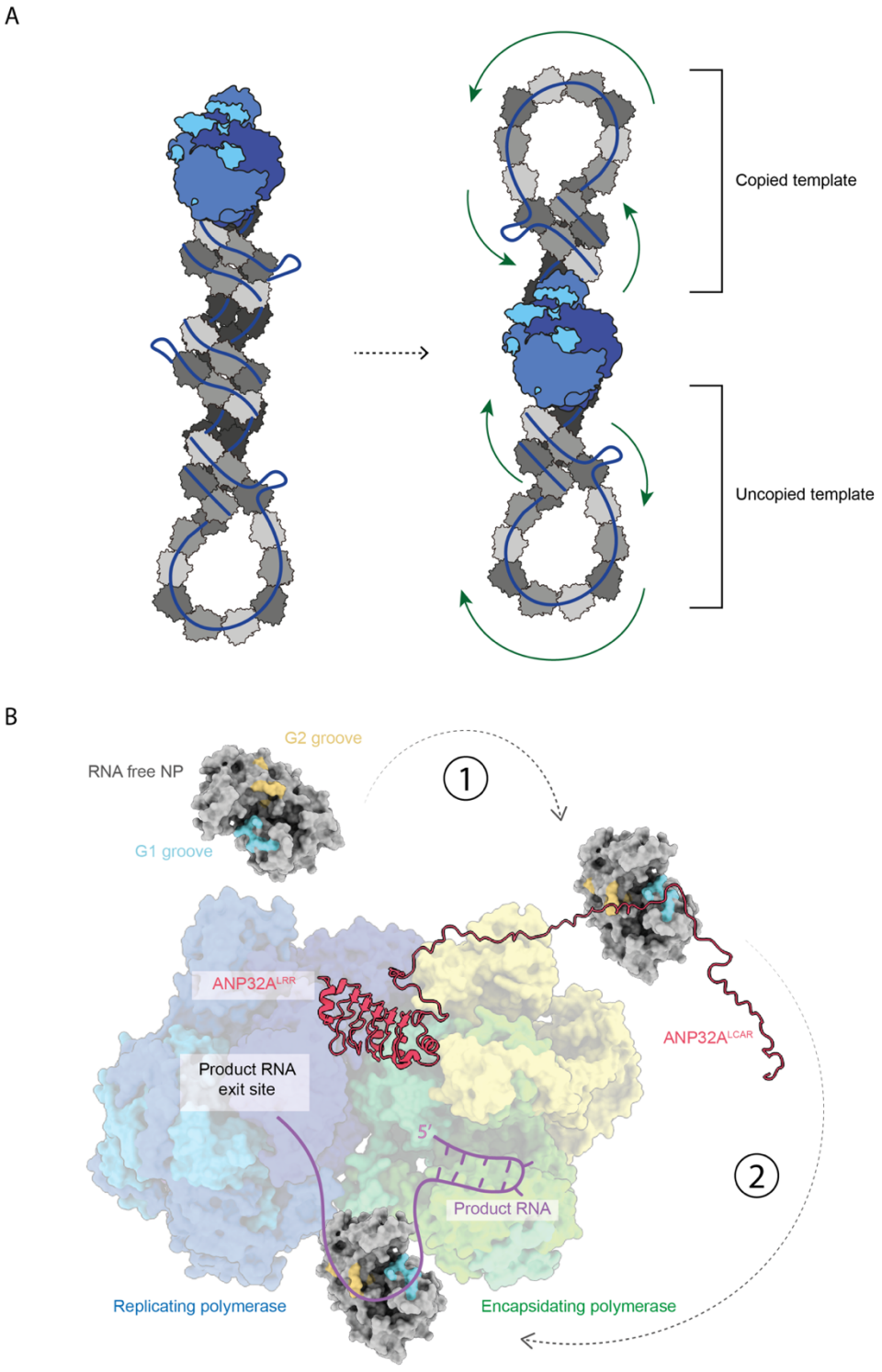


Figure 1.9 NP dynamics during replication. *A. Cartoon diagram of the 'processive helical track' model underlying how viral polymerase accesses NP-bound RNA during transcription and replication. B. Putative model of ANP32 dependent NP recruitment to the emerging nascent viral RNA. In step 1, ANP32^{LCAR} binds to the RNA free NP and increases the local concentration of NP around the replicating polymerase. In step 2, NP is displaced from ANP32^{LCAR} due to the higher affinity between NP and viral RNA. Adapted from Zhu et al. (2023).*

1.3.3 Termination

Transcription and replication terminate in different ways to produce products with different 3' termini. Transcription produces an mRNA transcript with a poly(A) tail, between 30-180 nt long, at its 3' terminus. This happens through an event called 'polymerase stuttering', in which the polymerase moves back and forth on an oligoU track close to the 5' end of the template. In detail, as transcription continues, the uncopied template RNA loop shrinks, until U17 (a residue 17 nt away from the 5' terminus) enters the active site. Further shrinkage becomes physically impossible due to the constraint caused by the anchored 5' hook. Base-pairing with the incoming nucleotide can still occur at U17, but the incorporation of an adenosine (A) opposite U17 stretches the template and breaks the relatively weak A-U bond, causing the template to retract and positioning U17 back into the active site (Poon et al., 1999; Pritlove et al., 1998; Robertson et al., 1981; Zheng et al., 1999). In this way, U17 enters and exits the active site repetitively, causing the polymerase to stutter at this position. After the poly(A) tail reaches a certain length, mRNA transcript is released from the viral polymerase (Davidritlove et al., 1998). Thereafter, the 3' promoter at the mode B site is allowed back to the primary promoter binding site (mode A site), reforming a promoter duplex with the 5' promoter (Figure 1.10).

Distinct from that of transcription, termination of cRNA synthesis is characterized by the release of the 5' vRNA promoter from its binding site and subsequent synthesis of a faithful copy of vRNA, without polyadenylation. Thereafter, in addition to relocating the template 3' terminus, the product 3' terminus is translocated into the 3' binding site of the encapsidating polymerase, marking the end of cRNA synthesis (Te Velthuis & Fodor, 2016). vRNA synthesis on cRNA terminates in a similar way to cRNA synthesis on vRNA. It is unclear whether the product 3' terminus needs any additional host or viral factors to be translocated into the new

polymerase. However, we can be confident that the precise positioning of the 3’ terminus of both template and product RNA is important for subsequent RNA synthesis by their respective RNPs (Figure 1.11).

The final mystery related to termination comes down to whether and how the unique mechanisms for transcription and replication initiation are linked to their respective termination processes. It is possible that this connection, determining the identity of the product RNA, involves mechanisms that result in the release of the 5' RNA promoter. For instance, it may involve the destabilization of 5' binding sites through the binding of replication cues, such as the new polymerase, ANP32 protein, and NP. However, the precise mechanism remains to be elucidated.

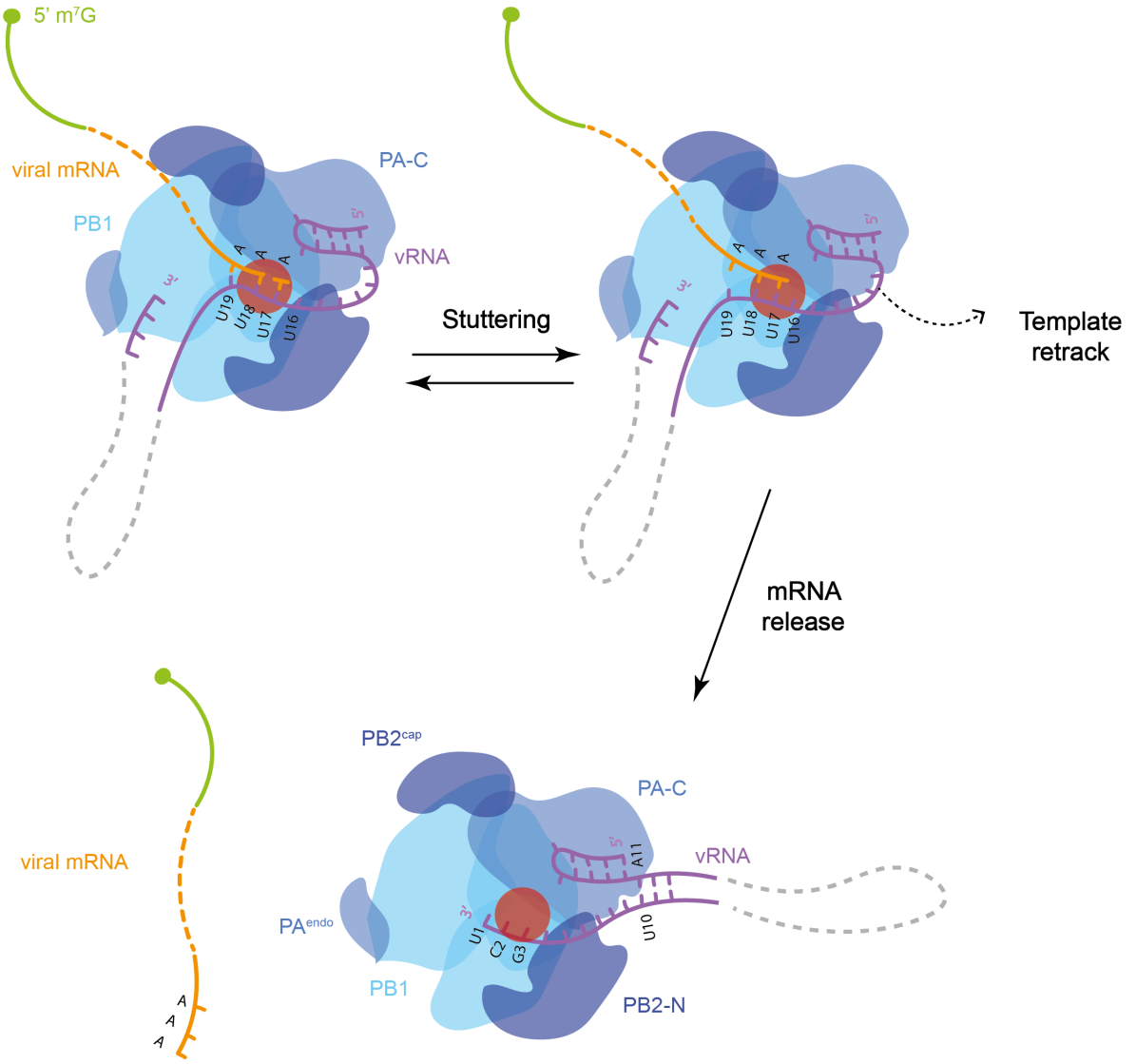


Figure 1.10 Cartoon model of transcription termination. *When transcription approaches termination, viral polymerase still binds tightly to the 5' vRNA promoter, creating a constraint in vRNA template as replication proceeds towards the end, until U17 enters the active site (top left). Further incorporation of adenosine opposite U17 creates a stuttering effect in the template RNA, resulting in a stretch of adenosine as U17 enters and exits the active site (top right). Stuttering terminates when the poly(A) tail reaches a certain length and the product mRNA dissociates from the viral polymerase (bottom). Thereafter, the 3' promoter of the template RNA is relocated back into the active site for the initiation of a new round of transcription.*

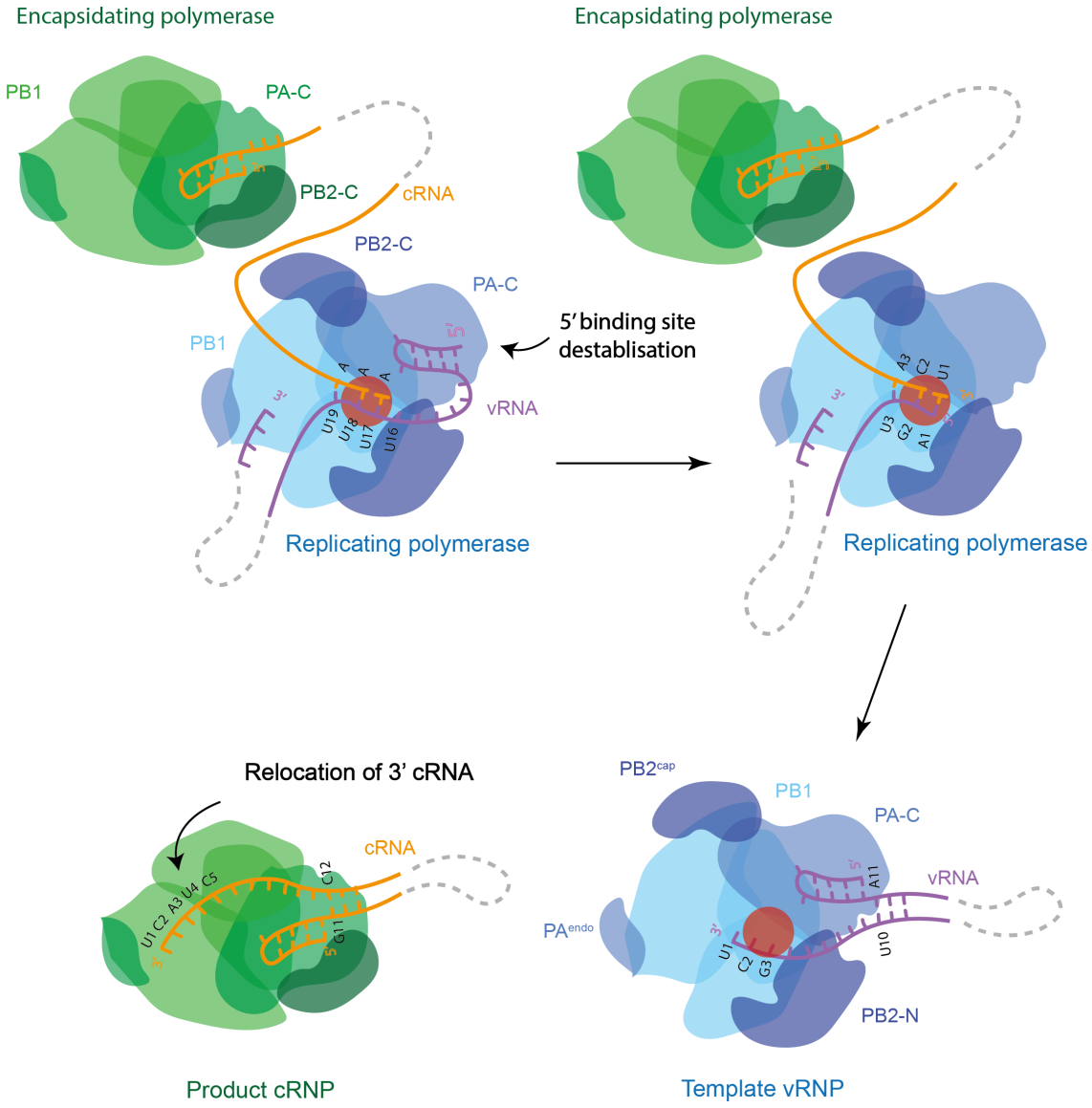


Figure 1.11 Cartoon model of replication termination. *The 5' promoter binding site is destabilised towards the end of viral genome replication (top left). The 5' promoter of the template RNA enters the active site and a faithful copy of the template (top right) is produced which dissociates from the replicating polymerase. Afterwards, the 3' promoter of the template RNA is relocated back into the active site, whilst the 3' promoter of the product RNA is positioned into the 3' binding site of the encapsidating polymerase (bottom).*

1.3.4 Molecular cues for transcription and replication

Accumulation profiles of mRNA, cRNA and vRNA during infection differ drastically from each other (Kawakami et al., 2011; Phan et al., 2021; Vreede et al., 2010). This observation raises the question how an RNP-associated polymerase changes conformations during infection to support specific enzymatic activities required for transcription or replication. It is likely that an interplay between viral and host factors governs the timing and activity of transcription and replication across various infection stages; however, the molecular mechanism is still not clear.

Previous studies suggest that vRNPs from incoming virions have a low level of cRNA synthesis activity even in the absence of viral protein synthesis (Deng et al., 2006; Turrell et al., 2013; Vreede et al., 2011). This gave rise to the 'stabilization model,' proposing that transcription and replication within vRNPs are stochastic and devoid of any molecular switch. The fact that cRNA only becomes detectable after viral protein production is due to the intrinsic instability of unmodified replication products, which are readily degraded by host RNA degradation processes. In other words, the produced cRNA only becomes detectable upon its stabilization and assembly into cRNPs, once free viral polymerase and NP become available.

While this model holds merit, previous studies indicate that interactions between viral and host proteins can regulate vRNP activity. The presence of transcription cues, such as host Pol II CTD, can enhance cap-snatching and transcriptional activity, and the presence of replication cues, such as free polymerase, ANP32 proteins and NEP, inhibits transcription and facilitates both steps of replication (Jorba et al., 2009; Krischuns et al., 2022; Robb et al., 2009; Serna Martin et al., 2018; York et al., 2013). Surprisingly, very recently, a study has also claimed that

Pol II CTD, which has classically been considered a transcription cue, can enhance replication activity in the presence of viral polymerase and ANP32 proteins (Krischuns et al., 2024). These findings raise questions about the precise roles of different factors in different modes of RNA synthesis. A simplified hypothesis suggests that while an RNP has intrinsic RNA synthesis activity, the availability of specific molecular cues at varying points of infection fine-tunes transcription and replication activities (Zhu et al., 2023). In agreement with this hypothesis, a common binding site on the viral polymerase PA-C has been identified and has been shown to mediate symmetric dimerisation, the binding of the encapsidating polymerase to ANP32^{LRR}, and the binding of the encapsidating polymerase to a replicating polymerase (Carrique et al., 2020; Keown et al., 2022; Krischuns et al., 2024) This suggests that competition between molecular interactions exists, and the outcome of such competition, as a result of binding affinity and molecular concentration, determines the activity of the viral polymerase. However, binding of a factor to either RNP-associated or newly synthesised polymerase may have different effects on overall viral RNA synthesis. Specifically, when a replication-promoting factor binds to a vRNP-associated polymerase, it increases vRNA synthesis; however, if it binds to a newly-synthesised viral polymerase, it would displace encapsidating conformation-inducing factor and prevent the newly synthesised polymerase from acting as an encapsidating polymerase, thereby inhibiting overall replication (Figure 1.12B).

It is worth noting that transcription and replication cues themselves could be further regulated, for example, by post-translational modifications (PTMs). Previous studies have identified that PTMs such as phosphorylation, acetylation, and ubiquitination, for the viral polymerase subunits and NP, which might influence transcription and replication (Giese et al., 2017; Günl et al., 2023; Hutchinson et al., 2012). For instance, phosphorylation in viral polymerase subunits has been reported to dissociate RNA from the polymerase and bias replication over

transcription, while phosphorylation of NP affects the NP oligomerisation state and RNP assembly (Chenavas et al., 2013; Dawson et al., 2020; Li et al., 2018; Mondal et al., 2015; Turrell et al., 2015). However, whether these PTMs are pro-viral, antiviral or neutral, and the extent to which protein functions are regulated by PTMs remains understudied.

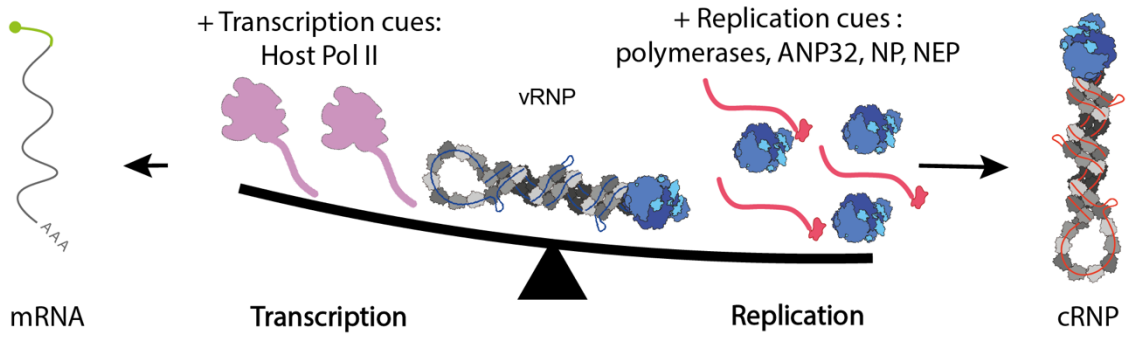


Figure 1.12 Molecular cues regulating the activity of the influenza virus polymerase. *The seesaw model of vRNP regulation. Interactions with transcription cues such as host RNA polymerase II drive vRNP towards mRNA production, whilst interactions with replication cues such as newly synthesised viral polymerase and ANP32 proteins drive vRNP towards cRNA production. Adapted from Zhu et al. (2023).*

1.4 NANOBODIES AND THEIR DERIVATIVES

Nanobodies, a type of engineered antibodies, have become valuable tools in therapeutics, diagnosis, and research over the past decade. The design of these artificial peptides takes advantage of a naturally existing variant of conventional heterotetrameric antibodies – heavy-chain-only homodimeric antibodies (HCAbs) that are produced alongside conventional antibodies in animals such as camelids and sharks. The variable domain of these HCAbs, which contains the antigen-binding site, can be cloned and expressed as single-domain antigen-

binding fragments, also known as nanobodies (Figure 1.13 A) (Bannas et al., 2017; Muyldermans, 2013; Pardon et al., 2014). These fragments consist of three hypervariable regions (HVs) flanked by conserved framework regions. The three HV loops form a continuous surface which is complementary to the surface of the epitope, hence are also called complementarity-determining regions (CDRs). Other than the high affinity and high specificity inherited from their derived antibodies, nanobodies offer several unique advantages as a tool in biotechnology and medicine, including small size, deep tissue penetration ability, as well as cheap and scalable production in prokaryotic expression systems. With these favourable biochemical properties, nanobodies have been a rising star in therapeutics, promising antivirals for passive immunization against, for instance, SARS-CoV-2 spike protein (Koenig 2021).

Not only have nanobodies been used as promising candidates for clinical development, they have also been involved in numerous research applications, in which the proof of principle has been successfully demonstrated. For example, nanobodies have been used to identify novel therapeutic targets and provide insight into mechanisms of protein functions at the molecular level (Hanke et al., 2016). In structural studies, nanobodies have been known to facilitate crystallization of proteins with flexible domains and imaging of conformationally dynamic proteins, such as membrane protein G protein-coupled receptors, by stabilizing protein complexes or fixing the protein in a certain conformation (McMahon et al., 2018). In addition, nanobodies have been grafted onto larger protein scaffolds to generate megabodies, with an increased molecular weight and unaffected binding affinity and specificity (Figure 1.13B) (Coupland et al., 2021; Uchański et al., 2021). Similarly, an increase in the molecular weight of a nanobody can also be achieved when the non-antigen-binding regions of a nanobody are engineered to either dimerise with another nanobody (gluebody) or to attract an antibody that has its flexible Fab hinge stabilized by another antibody (NabFabody) (Bloch et al., 2021; Ye

et al., 2022). These approaches produce nanobody derivatives that can bind specifically to a target, help to increase particle size, prevent preferential orientation, and provide a reference point in single-particle cryo-EM or cryo-ET reconstructions. In parallel, nanobodies fused to a fluorescent protein, called chromobodies or fluobodies, can be used to achieve single-molecule localisation when coupled with super-resolution imaging techniques (de Beer & Giepmans, 2020).

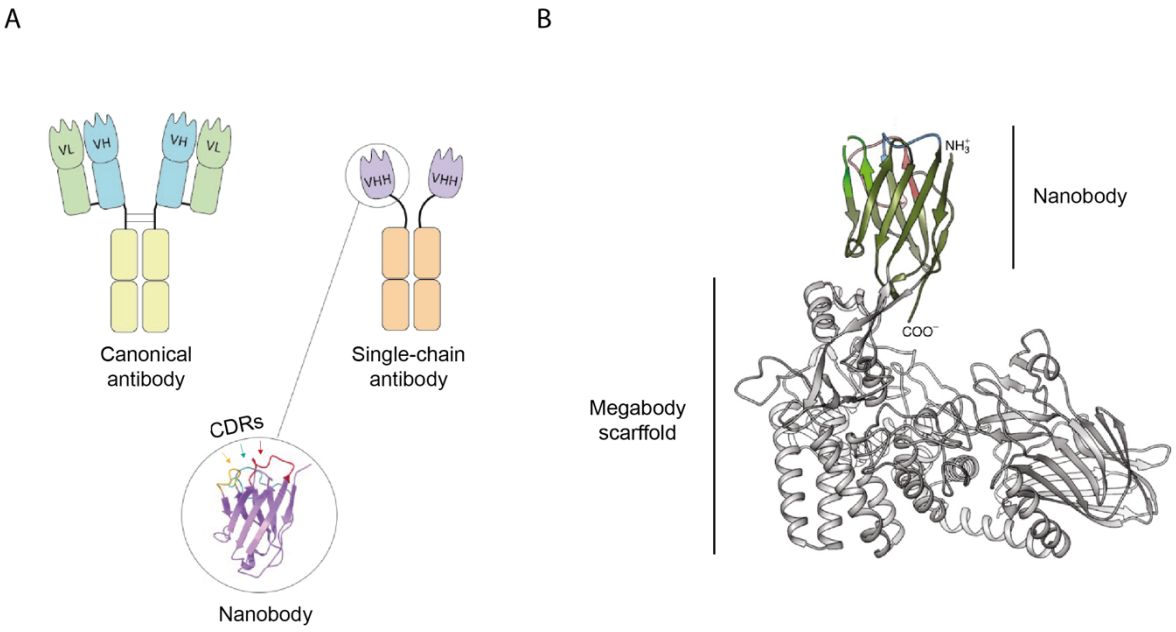


Figure 1.13 Nanobody and megabody structures. *A. Cartoon diagram of antibodies and structural model of nanobody. Antigen-binding CDR1-3 are highlighted in yellow, cyan and red. B. Structural model of megabody as a chimeric molecule, with an antigen specific nanobody and a scaffold protein. Adapted from Uchanski et al. (2021).*

1.5 OBJECTIVES

Influenza virus genome transcription and replication are central to the viral life cycle; however, there remain many unknowns about the structure of RNP, the viral transcription and replication complex, and the mechanisms that regulate RNP functions. The aim of this study is to gain deeper molecular and mechanistic insights into the influenza virus ribonucleoprotein complex using innovative approaches.

To investigate RNP functions in a simplified and highly controlled manner, I established an *in vitro* vRNP activity assay, enabling the reconstitution of viral transcription and a complete replication cycle within a test tube for the first time (Chapter 2). Using this assay, I defined the minimal components required for viral genome replication and dissected the roles of individual components (Chapter 3). Next, I took advantage of a panel of inhibitory nanobodies to investigate the viral polymerase, uncovering novel drug targets and gaining mechanistic insights into polymerase functions (Chapter 4). Lastly, to further understand how the polymerase and nucleoprotein interact during RNA synthesis, I attempted to create a polymerase and nucleoprotein complex, containing the minimal number of the component proteins, using recombinant proteins. I also developed megabodies as a structural tool, and attempted to carry out cryo-EM analysis of polymerase and nucleoprotein interactions in both resting and active RNPs (Chapter 5).

Chapter 2: Minimal Components to switch the activity of vRNP from transcription to replication

Most data from this chapter were published in:

***Zhu Z**, Fan H, Fodor E (2023) Defining the minimal components of the influenza A virus replication machinery via an in vitro reconstitution system. PLoS Biol 21(11): e3002370.*

2.1 INTRODUCTION

The influenza viral polymerase is responsible for both transcription and replication of the resident vRNA inside a vRNP. These two functions require distinct enzymatic activities and interactions with different viral and host factors. Within an infected cell, transcription depends on the interaction of viral polymerase with host Pol II, and relies on PA^{endo} and $PB2^{cap}$ to carry out cap-snatching on host capped RNA. This interaction grants the viral polymerase access to nascent host capped RNA, and the cap-snatching activity generates capped RNA fragments, which can subsequently be used as primers for transcription. Thereafter, the viral polymerase initiates transcription in a primer-dependent manner, copies the vRNA template, and terminates through stuttering prematurely at an oligo(U) sequence located near the vRNA 5' terminus. In contrast, cRNA synthesis utilizes primer-independent initiation and therefore requires *de novo* ApG synthesis activity. It also progresses through the entirety of the RNA template and

generates a complete replica of the vRNA. In addition, while mRNA can be produced without additional viral protein and can utilise host mRNA binding proteins, cRNA synthesis must be synchronized with the recruitment of newly produced viral polymerase and NP for cRNP assembly (Zhu et al., 2023).

Virion-isolated vRNP has been shown to be capable of synthesizing mRNA primarily, but also a limited amount of cRNA, without additional proteins (Vreede & Brownlee, 2007). This suggests that the vRNP-associated polymerase possesses both transcriptional and, at least to some extent, replicational activity. The two distinct modes of RNA synthesis within the same viral polymerase could be achieved by the polymerase transitioning between conformations that permit each enzymatic activity. However, during infection, newly synthesised viral polymerase is required for detecting viral cRNA synthesis (Vreede et al., 2004). Indeed, profiles of mRNA, cRNA and vRNA accumulation throughout infection reveal that mRNA is generated during the early stages, whereas cRNA and vRNA only become detectable once a critical level of viral proteins, in particular viral polymerase and NP, is reached. The requirement for new viral polymerase to transit from transcription to replication provides insight into possible molecular mechanisms underlying vRNP regulation. Three models have been proposed: according to the first 'trans-acting model', newly synthesised viral polymerase acts in-trans on the vRNA template, and is the main source of cRNA production during replication (Jorba et al., 2009). The second 'replication-switch' model states that newly produced viral polymerase act as a 'switch' to trigger the replicational activity of vRNP-associated polymerase (Gastaminza et al., 2003; Huarte et al., 2003). Lastly, the 'stabilisation model' suggests that there is no molecular switch that actively inhibits transcriptional activity and promotes replicational activity in vRNP-associated polymerase; instead, replication

products are stabilised by newly synthesised viral polymerase through RNP assembly (Vreede et al., 2004). However, it remains an enigmatic which, if any, of these models is correct.

Curiously, while influenza virus-infected cell lysate containing host and viral proteins is capable of stimulating cRNA synthesis in virion-derived vRNPs in a test tube *in vitro*, the addition of partially purified recombinant viral polymerase, expressed in mammalian cells, does not affect the transcriptional or replicational activity of purified vRNP (Beaton & Krug, 1986; Skorko et al., 1991; Sugiyama et al., 2015; Vreede & Brownlee, 2007). Therefore, viral polymerase alone is insufficient to promote replication. Numerous factors that affect vRNP functions have been identified from transfected or infected cells through methods such as pull-down and mass spectrometry; however, due to the complexity of the cellular environment in which these studies were carried out, the exact role of each factor is unclear. Perhaps the most well-known regulatory factors are the members of the host ANP32 protein family; these are essential factors for influenza virus genome replication and underlie viral host-specificity. In particular, the avian and mammalian isoforms of ANP32 family proteins differentially support replication of influenza viruses of avian and mammalian origin. Previous studies have shown that ANP32 family proteins contribute to replication in two ways: firstly, suggested by a cryo-EM structure, the ANP32^{LRR} bridges an asymmetric replicating-encapsidating dimer formed by the vRNP-associated polymerase and a newly synthesised viral polymerase, thereby creating a platform for viral genome replication; secondly, supported by biochemical data, ANP32^{LCAR} is involved in recruiting NP during vRNP assembly. However, our current knowledge is not sufficient to support that this structurally resolved ANP32-bound polymerase dimer represents the minimal requirement for viral genome replication. Thus in this chapter, I aimed to identify the minimal protein requirement for virion-purified vRNP to switch from transcription to translation *in vitro*.

2.2 RESULTS

2.2.1 Purification of active virion-derived vRNP

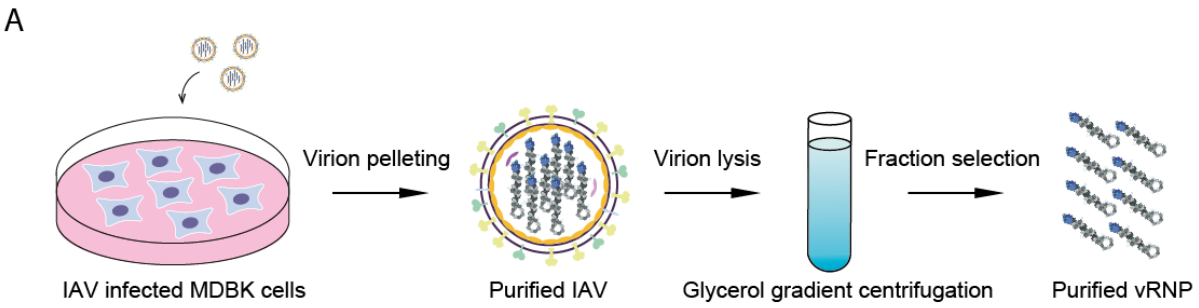
Early work from our group showed that vRNP isolated from virions is capable of producing mRNA and a limited amount of cRNA in the absence of any additional proteins. To create a system for studying the effects of regulatory factors on viral transcription and replication *in vitro*, I followed the classic protocol (Vreede & Brownlee, 2007) to purify virion-derived vRNPs from influenza A/WSN/33 (H1N1) viruses produced by infecting MDBK cells. In brief, I pelleted and lysed the virions, and separated vRNPs from other viral components using glycerol gradient centrifugation (Figure 2.1A). Fractions from glycerol gradient centrifugation were collected and analysed using SDS-PAGE followed by silver staining. Fractions containing viral polymerase and NP at approximately the expected stoichiometry were combined (Figure 2.1B). Virion-derived vRNPs were pelleted using ultracentrifugation, resuspended gently in storage buffer, and snap-frozen for mid-term (up to 6 months) storage at -80°C until their use in activity assays. Since the final vRNP sample contains RNA and a mixture of proteins, its concentration could not be accurately determined using Nanodrop. Therefore, I used two independent methods to determine the molarity of purified vRNP. Firstly, a known molarity of Nanogold beads was mixed with the vRNP sample at a certain measured concentration (X mg/ml), and the mixture was imaged using negative staining EM (results not shown). The ratio of Nanogold beads to vRNP was calculated by comparing their counted numbers across micrographs containing a total of more than 4000 Nanogold beads. The molarity of the vRNP sample at X mg/ml was calculated by dividing the known molarity of Nanogold beads by the Nanogold beads to vRNP ratio. In parallel, I estimated the molarity of purified vRNP by comparing the vRNP sample to a serial dilution of purified recombinant viral

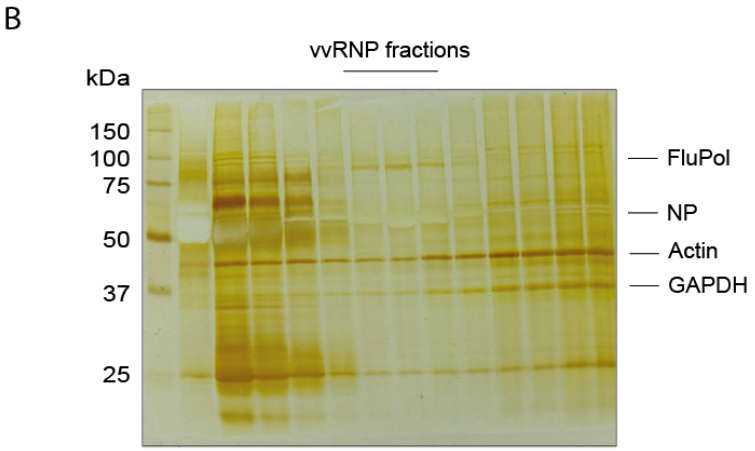
polymerase of known molar concentration, measured by Nanodrop, using SDS-PAGE followed by Coomassie staining (Figure 2.1C). The two methods resulted in similar estimated molar concentrations (with only 2X difference); therefore, an averaged value was used to calculate the molarity of vRNP used in my subsequent experiments.

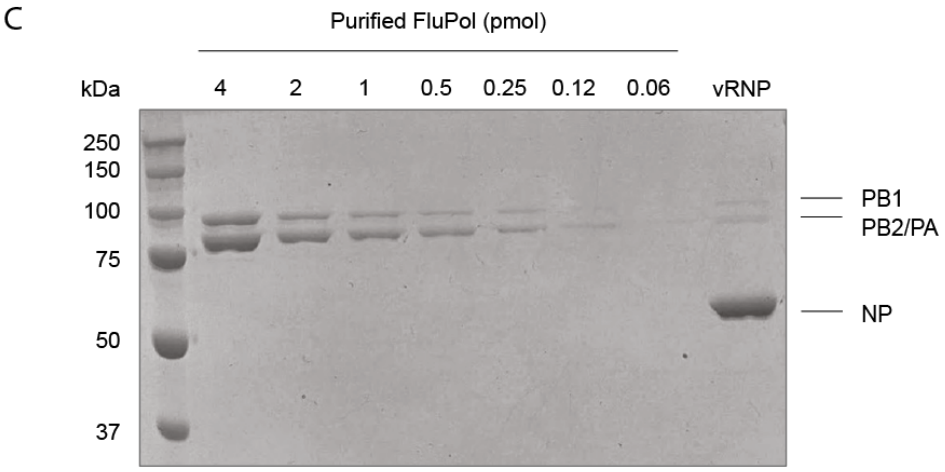
To confirm that the virion-derived vRNP possessed RNA synthesis activity, I performed vRNP activity assays as previously described in the presence or absence of either globin mRNA (a source of capped RNA primers) or ApG dinucleotides (mimic of pppApG dinucleotides formed during replication initiation) (Vreede & Brownlee, 2007). RNA extracted from the system was then analysed using primer extension, with radiolabelled NA segment-specific primers against either the positive or the negative strand. These primers generate cDNAs of mRNA, cRNA, and vRNA, which have different mobilities on a denaturing gel (Figure 2.1D). Although the mRNA and cRNA products were detected by the same primer (NA 160), they can be distinguished on a gel by their mobilities. This is because transcription initiates on a capped primer, therefore in the primer extension assay mRNA generates a longer cDNA product than cRNA. I also included a negative control lacking rNTPs to show that the vRNP preparation is cRNA and mRNA free, and only contains vRNA. In the presence of rNTPs, but without a primer, I observed a small amount of cRNA product. This signal for cRNA product was increased as a result of adding ApG primer to the reaction. Taken together, these observations confirm that my purified vRNP possesses cRNA synthesis activity, and that *de novo* synthesis of ApG at replication initiation presents the bottleneck of replication. Next, I observed that a notable amount of mRNA product was produced after capped globin mRNA was added, as a source of capped primer (Figure 2.1E). I also noticed a slight mobility difference between the signals for the cRNA product produced in the presence and absence of capped globin mRNA. Additionally, there was an increase in the signal intensity of the cRNA product when globin

mRNA is added. Further characterisation of these bands in Chapter 3 reveals that the band, with the mobility of cRNA, observed in the presence of newly-produced viral mRNA is likely derived from a mixture of cRNA and truncated viral mRNA through the endonuclease activity of RNP-associated viral polymerase.

Collectively, these results demonstrate that virion-purified vRNPs are transcriptionally and replicationally active, and that primer-dependent mRNA synthesis is preferred over primer-independent cRNA synthesis *in vitro* in the presence of capped mRNA.


A
Virion pelleting
Virion lysis
Fraction selection
IAV infected MDBK cells
Purified IAV
Glycerol gradient centrifugation
Purified vRNP


B
vvRNP fractions
kDa
150
100
75
50
37
25
FluPol
NP
Actin
GAPDH


C
Purified FluPol (pmol)
kDa
4
2
1
0.5
0.25
0.12
0.06
vRNP
250
150
100
75
50
37
PB1
PB2/PA
NP

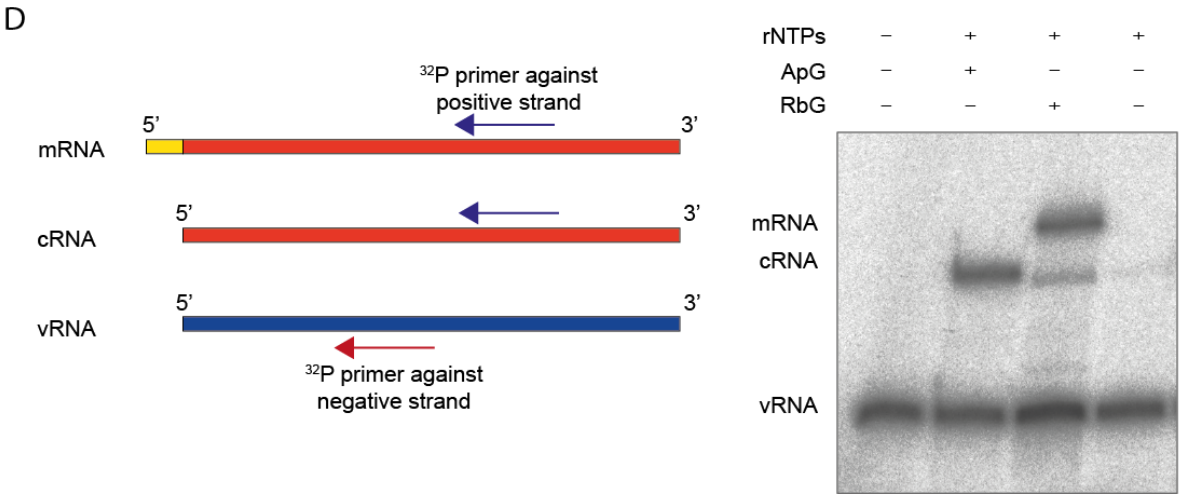

D
rNTPs
ApG
RbG
32P primer against positive strand
5'
3'
mRNA
cRNA
vRNA
32P primer against negative strand

Figure 2.1 Purification, semi-quantification and activity test of virion-derived vRNP. *A. Schematic diagram illustrating how WSN virion-derived vRNPs (vvRNPs) are isolated from virions amplified in MDBK cells using glycerol gradient. B. SDS-PAGE analysis followed by silver-staining of glycerol gradient fractions of vvRNP purification, from bottom (left) to top (right). C. SDS-PAGE analysis followed by Coomassie staining comparing the vvRNP sample to a serial dilution of recombinant NT60 viral polymerase (FluPol) of a known concentration. D. Schematic diagram illustrating the basis of RNA detection using the primer extension assay (left). vRNP activity was tested with or without nucleotides, and with or without ApG dinucleotides or rabbit globin mRNA (RbG) primer.*

2.2.2 Additional viral polymerase inhibits transcription

Next, I investigated the effect of free viral polymerase alone on vRNP activity using purified recombinant viral polymerase expressed in insect cells (a gift from Dr Haitian Fan) (Figure 2.2A). I added an increasing amount of viral polymerase, from 16X to 640X the molar concentration of input vRNP, to the reaction, both in the presence or absence of capped primer donor. I found that free viral polymerase only causes a minor increase in cRNA levels and has no effect on vRNA levels, but strongly inhibits transcriptional activity of vRNP in a dose-dependent manner (Figure 2.2B, C).

A surprisingly high concentration of viral polymerase is needed for effective viral replication. This is presumably because wild type viral polymerase naturally dimerises in solution to form symmetric regulatory dimer, whose dimer interface is in close proximity to that of the asymmetric replicating-encapsidating dimer on the encapsidating polymerase. This suggests that the majority of added viral polymerase, which is in the form of symmetric regulatory dimer, could not act as encapsidating polymerase due to steric hindrance. As a result, only a minor population of added viral polymerase, in monomeric form, could dimerise with the vRNP-

associated polymerase and affect vRNP activity. To test this hypothesis, I compared the effect of an increasing concentration of wild type (NT60-FluPol) or symmetric regulatory dimer mutant polymerase (NT60-$PA_{352\text{-}356A}$) (Fan et al., 2019), which has been shown to form fewer symmetric regulatory dimers in solution; I found that the symmetric regulatory dimer mutant polymerase causes similar transcription inhibition as the wild type polymerase but at a lower concentration (Figure 2.2D), supporting this hypothesis.

While viral polymerase could be reducing the mRNA levels by inhibiting transcription, it is possible that the diminished mRNA signal is due to the endonuclease activity of the added viral polymerase. To test this possibility, I compared the effect of an increase in wild type or endonuclease mutant polymerase (1918-PA_{D108A}) (Hara et al., 2006) which has been shown to be incapable of endonucleolysis. I found that the endonuclease mutant polymerase inhibits transcription to a similar level as the wild type polymerase (Figure 2.2E), opposing the possibility that mRNA is degraded by the endonuclease activity of the added polymerase. We cannot rule out the possibility that the endonuclease activity of the vRNP-associated polymerase is enhanced by an increase in free polymerase, which degrades mRNA. This however, seems unlikely from an evolutionary point of view.

Lastly, I compared the effect of wild type and active site mutant polymerase (NT60-$PB1_{D445A,D446A}$, or PB1a) (Biswas & Nayak, 1994). I found that catalytically inactive viral polymerase alone inhibits transcriptional activity of vRNP to the same extent as wild type viral polymerase alone (Figure 2.3A, B, lanes 4, 6).

Collectively, these results suggest that additional viral polymerase alone can inhibit the transcriptional activity of the vRNP-associated polymerase, but is insufficient to switch

transcription to replication. This inhibitory effect is likely to rely on the added polymerase being in a monomeric state, but does not depend on either the endonuclease activity or the catalytic activity of the added polymerase (Figure 2.2F).

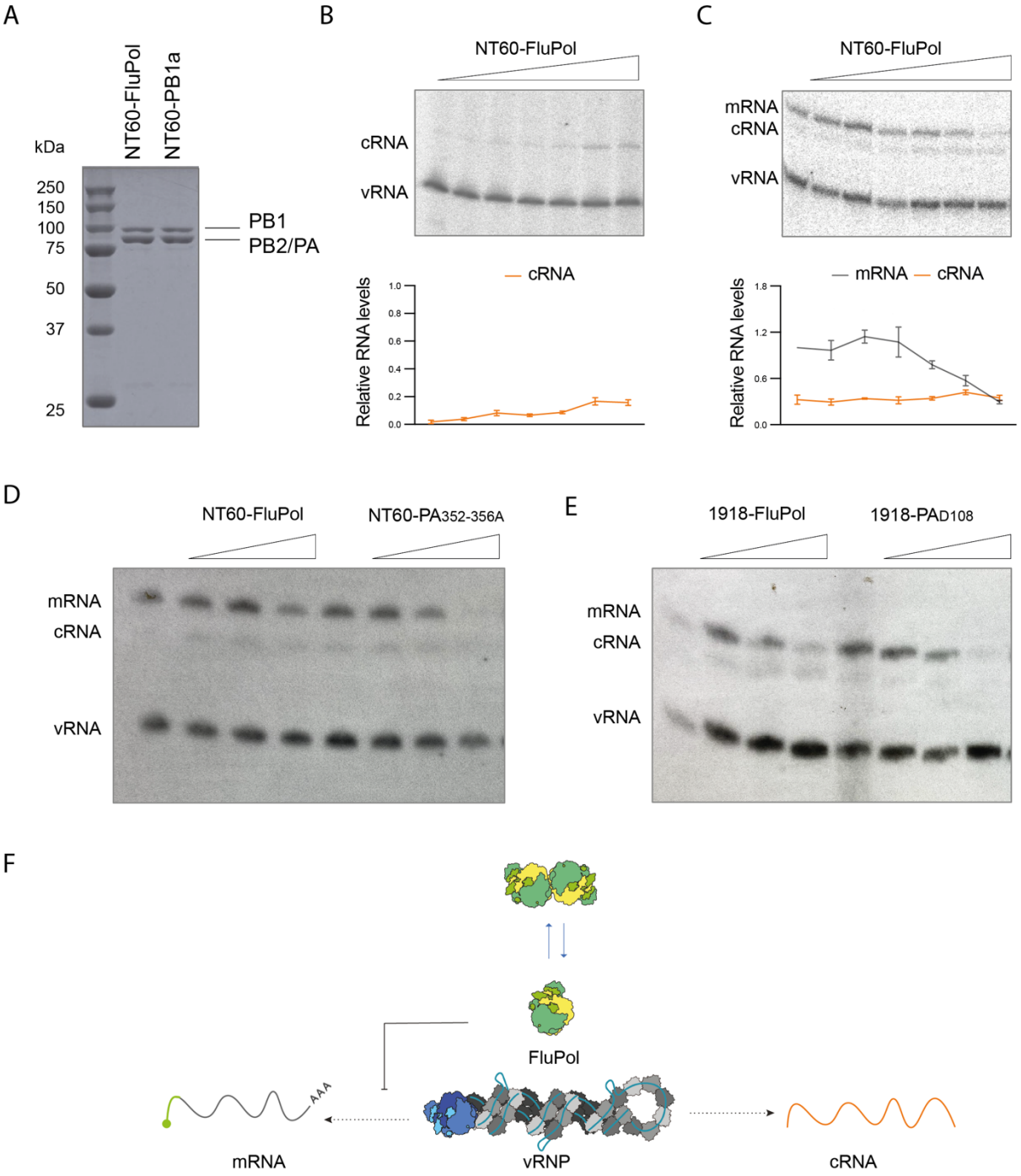


Figure 2.2 Free viral polymerase inhibits transcriptional activity of vRNP-associated polymerase. *A. SDS-PAGE analysis followed by Coomassie staining of wild type NT60 polymerase (NT60-FluPol) and catalytically inactive NT60 polymerase (NT60-PB1a) purified from insect cells. B-C. Effect of increasing amount of NT60-FluPol on vRNP activity, in the absence (B) or presence (C) of globin mRNA as a capped primer donor. A representative result of three independent RNA analyses by primer extension is shown (top). Relative RNA levels (bottom) were calculated by comparison to mRNA levels without added polymerase. Data present mean values ± s.e.m., based on n = 3 independent vRNP purifications and reactions. D. Effect of increasing amount of NT60-FluPol or symmetric regulatory dimer mutant NT60 polymerase (NT60-$PA_{352-356A}$). E. Effect of increasing amount of wild type 1918 polymerase (1918-FluPol) or endonuclease mutant 1918 polymerase (1918-PA_{D108A}). F. Schematic diagram showing the effect of viral polymerase alone on vRNP activity.*

2.2.3 Viral polymerase and ANP32 are the minimal components to switch the activity of vRNP from transcription to replication

The proposed replication complex consists of an asymmetric replicating-encapsidating dimer of viral polymerase and a copy of the ANP32 family protein (Carrique et al., 2020). This complex is likely to represent a replication platform, but there is no direct evidence that it contains all components required for efficient replication. Here, I studied the effect of viral polymerase and the ANP32 family protein on vRNP activity, thereby examining the functional relevance of this structurally resolved replication complex.

I found that the addition of ANP32B alone has no effect on RNA synthesis, either in the presence or absence of a capped primer donor. On the other hand, the addition of viral polymerase alone can inhibit mRNA synthesis in a globin mRNA-primed reaction but has little effect on cRNA synthesis, consistent with my earlier results. Excitingly, I observed a robust

increase in cRNA synthesis and a complete suppression of mRNA synthesis when ANP32B and viral polymerase were added together (Figure 2.3A-B).

Many studies have shown that cRNA accumulation only occurs after a critical amount of new viral polymerase is present, both in cells and as shown here *in vitro*. Therefore, a long-standing question is whether replication of the vRNA template in a vRNP is carried out by free viral polymerase newly synthesized during infection, which functions in a trans-acting manner. To answer this question, I compared the cRNA-promoting effect of ANP32B protein together with either wild type viral polymerase or PB1a, in the presence or absence of a capped primer donor. I found that PB1a, which lacks RNA synthesis activity, still causes a robust increase in cRNA synthesis, in both primed and unprimed conditions, to the same extent as the wild type viral polymerase (Figure 2.3A-B). Therefore, I confirmed that replication is carried out *in cis* by the vRNP-associated polymerase that also transcribes the vRNA template. By showing that the role of the additional viral polymerase is non-catalytic, I hypothesize that the free viral polymerase acts allosterically, promoting replicational activity by inducing a conformational change in the vRNP-associated polymerase.

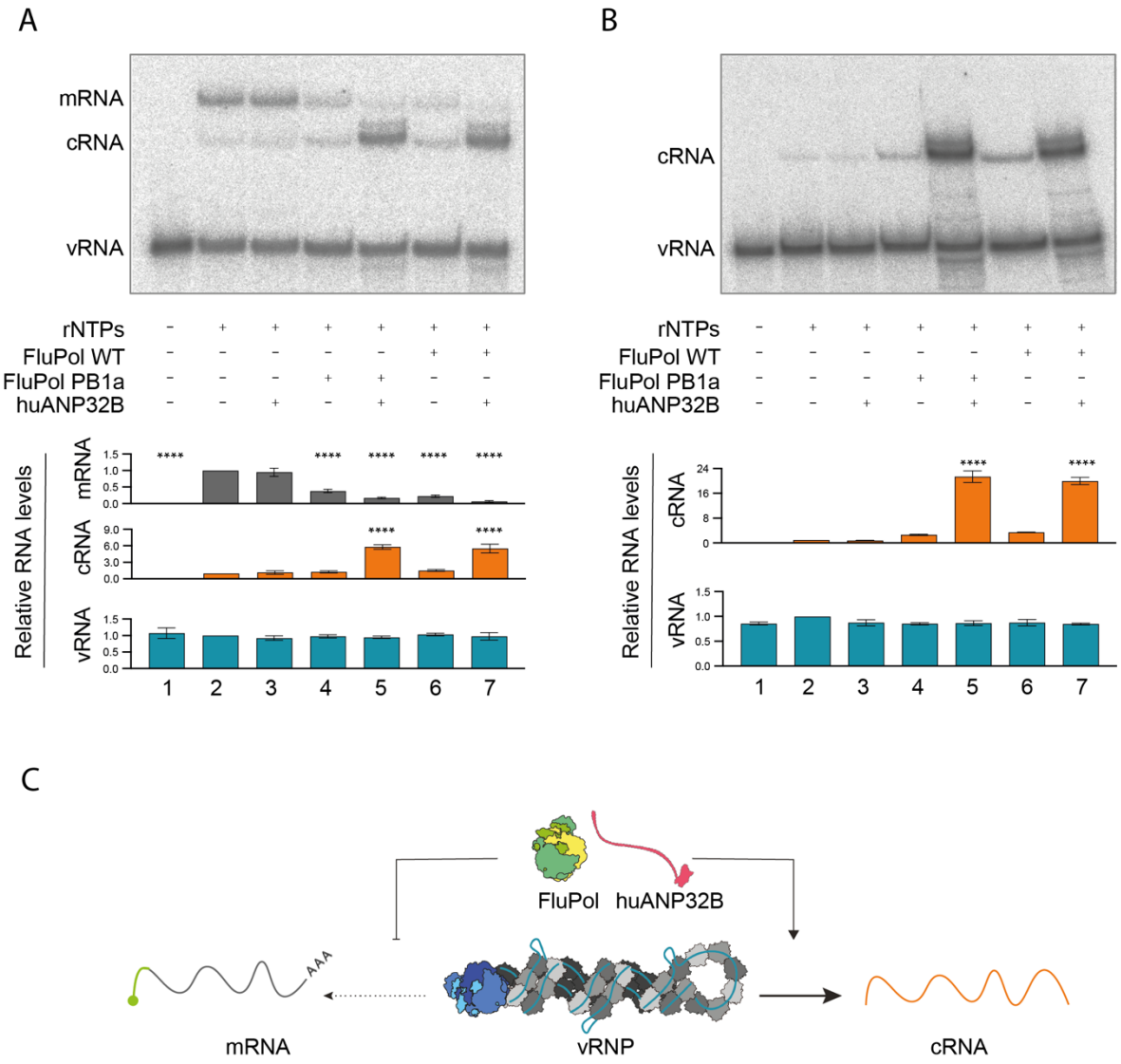


Figure 2.3 Viral polymerase and human ANP32B together promote vRNP replication. *A-B. Effect of active (NT60-FluPol) or inactive (NT60-PB1a) NT60 viral polymerase and human ANP32B (huANP32B) alone or together on vRNP activity, in the presence (A) or absence (B) of globin mRNA as capped primer donor. A representative result of three independent RNA analyses by primer extension is shown (top). Relative RNA levels (bottom) were calculated by comparison to the RNA levels obtained in the absence of FluPol and huANP32B. Ordinary one-way ANOVA was used to assess significance. Data represent mean values ± s.e.m., based on n = 3 independent vRNP purifications and reactions. ****, p<0.0001. C. Schematic showing the effect of viral polymerase and huANP32B on the transcriptional and replicational activity of vRNP.*

To test whether this observed effect of viral polymerase and ANP32B is dose-dependent, I titrated an increasing amount of viral polymerase and ANP32B, or ANP32B alone, into either an unprimed or primed system. First, I kept the ratio of free viral polymerase and ANP32B at 1:1. The result shows that in the primed condition, ANP32B alone has no effect on transcription or replication at all tested concentrations (Figure 2.4A). In contrast, titration of a mixture of viral polymerase and ANP32B, from 16X to 640X the molar concentration of input vRNP, causes a gradual decrease in mRNA signals to the baseline level and a trend-matching increase in cRNA levels until a plateau is reached (Figure 2.4B). Similarly, a dose-dependent increase in cRNA levels was also observed in the unprimed condition, except plateauing at a lower concentration of viral polymerase and ANP32B compared to in the primed condition, possibly because there is no competition from transcriptional activity (Figure 2.4C).

Next, I examined the dose-dependent effect of each component separately, by keeping either viral polymerase or ANP32 protein levels constant, while decreasing levels of the other component. My results show that at a constant level of ANP32 protein, around 320X that of input vRNP, decreasing viral polymerase gradually decreases cRNA levels and restoring mRNA levels. On the contrary, in the presence of constant viral polymerase, around 320X that of input vRNP, decreasing ANP32 does not significantly affect the reaction equilibrium until it is completely removed from the system (Figure 2.4D). These results collectively suggest that ANP32 protein is in massive excess when it is added at a 1:1 ratio to viral polymerase. This is possibly because only the monomeric population of the wild type viral polymerase can participate in the formation of the replication complex – an earlier proposed hypothesis. Therefore, when ANP32 and viral polymerase were added at 1:1 ratio, the number of ANP32 molecules that can participate in forming a replication complex is more than the viral polymerase that can act as an encapsidating polymerase (the effective viral polymerase). In

agreement with this hypothesis, I found that the influenza A virus monomeric mutant viral polymerase (mFluPolA), with mutations that disrupt the symmetric regulatory dimer formation (NT60-$PA_{352\text{-}356A}$), can promote replication to the same level as wild type viral polymerase (NT60-FluPol) at a lower concentration (Figure 2.4E) (Fan et al., 2019).

In primed conditions, as viral polymerase and ANP32B are added, the decrease in mRNA signals synchronizes perfectly with the increase in cRNA signals. This can be explained by a complete switch of transcriptional activity to replicational activity. Nevertheless, an alternative explanation is that mRNA is cleaved by the endonuclease activity of the added viral polymerase. To exclude this possibility, I demonstrated identical changes in RNA levels when a wild type viral polymerase is swapped for an endonuclease mutant polymerase (Figure 2.4F).

So far, my results allow me to define viral polymerase and ANP32 protein as the minimal requirement for the replicational activity of vRNP. This strongly suggests that the replication platform, made up of viral polymerases and ANP32 protein, proposed by Carrique et al., indeed underlies the molecular mechanism that promotes replication. According to the model, interactions between the replicating polymerase and the encapsidating polymerase, and between viral polymerases and $ANP32^{LCAR}$, are critical for stabilising this complex. To test this model, I first replaced influenza A viral polymerase with influenza C viral polymerase (FluPolC). Viral polymerases from different types of influenza viruses differ significantly at their dimerisation interface, especially in the $PB2^{627}$ domain, which interacts with the $ANP32^{LCAR}$ according to the model. Therefore, it is unlikely that dimerisation can occur between influenza A and influenza C viral polymerases. I found that in the presence of ANP32 proteins, influenza C viral polymerase cannot support replication in the influenza A viral polymerase, nor can it inhibit transcription (Figure 2.4G). To further test this model, I

compared the regulatory effects of the wild type viral polymerase (NT60-FluPol) and the polymerase with $PB2^{627}$ deleted (NT60-$\Delta PB2^{627}$) in the presence of ANP32B. I found that NT60-$\Delta PB2^{627}$, which lacks a major interacting domain responsible for complex formation according to Carrique et al., is unable to support replication under unprimed conditions (Figure 2.4H). My results thus support the replication platform model from our group, and indicate that interactions mediated by $PB2^{627}$ on the encapsidating polymerase are important for inhibiting transcriptional activity and promoting replicational activity in vRNP.

Collectively, these results demonstrate that RNA-free viral polymerase and host ANP32 proteins are the minimal requirements to switch vRNP function from transcription to replication, and that this regulatory event likely occurs in the context of the replication platform, through polymerase asymmetric dimerisation and interactions with ANP32 protein.

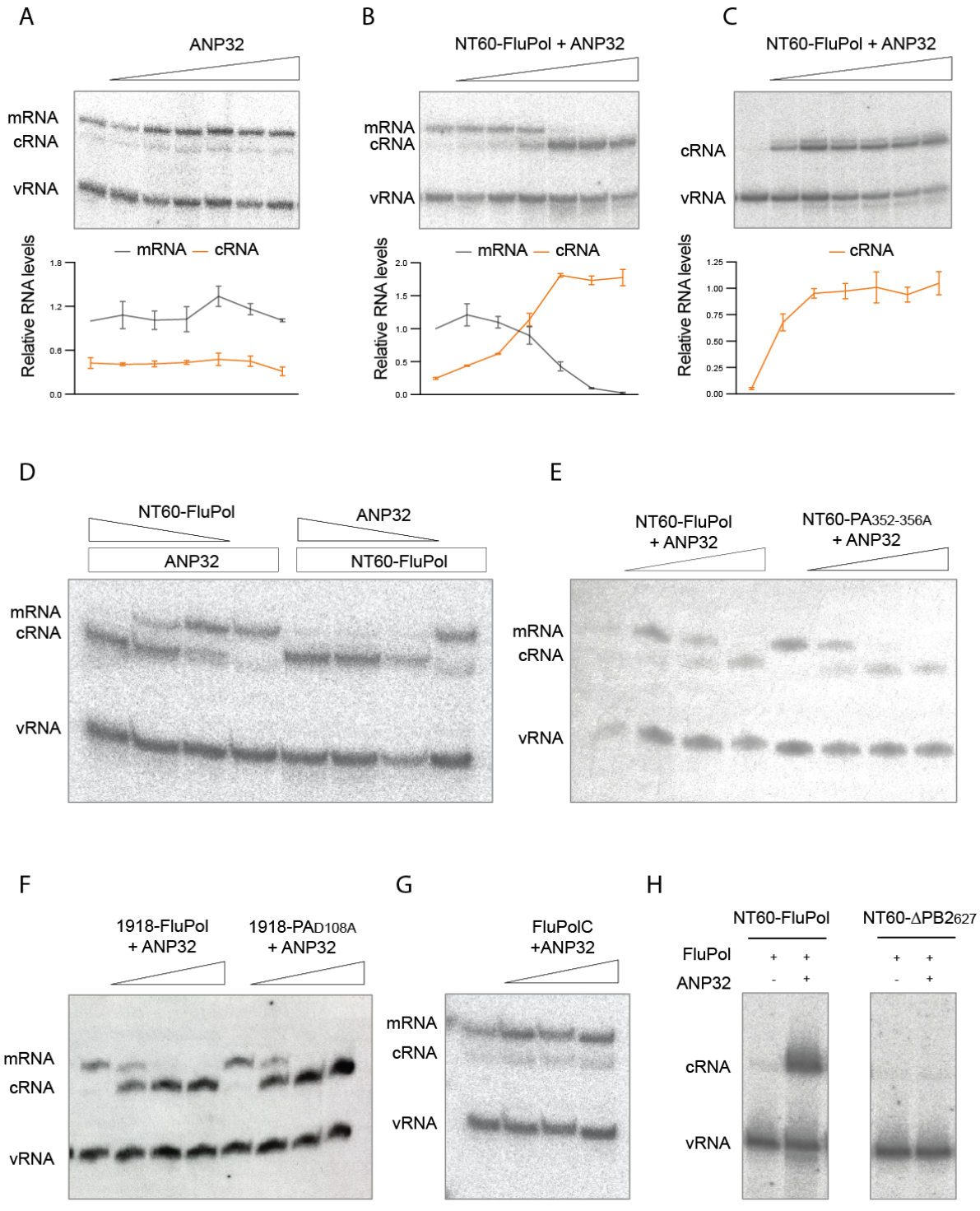


Figure 2.4 Further investigation into the requirements for the polymerase to promote replication. *A-C. Effect of increasing amount of ANP32 protein (A) or ANP32 protein and NT60-FluPol on vRNP activity (B-C), in the presence (A-B) or absence (C) of primer. A representative result of three independent RNA analyses by primer extension is shown (top). Relative RNA levels (bottom) were calculated by comparison to the mRNA levels without added proteins. Data present mean values ± s.e.m., based on n = 3 independent vRNP purifications and reactions.*

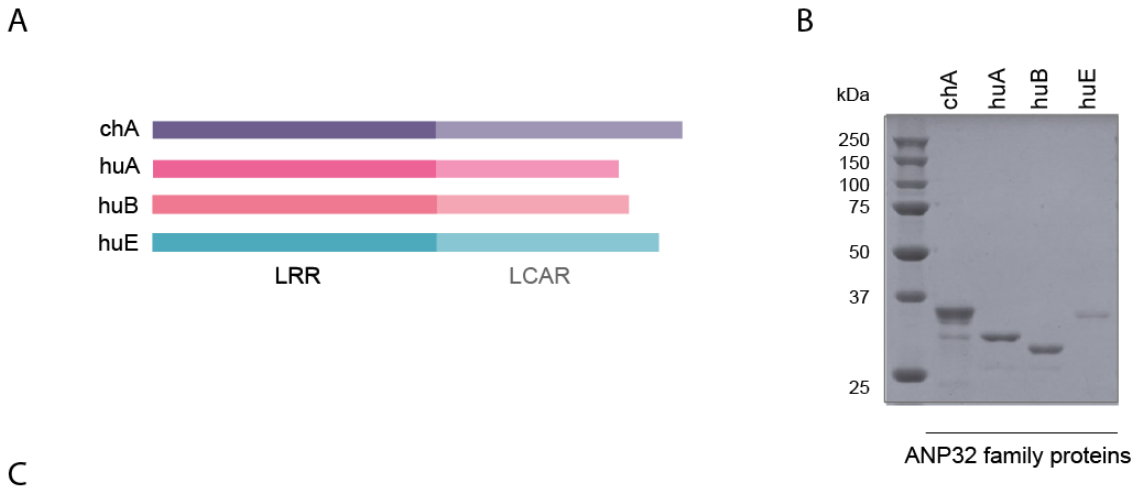
*D. Effect of decreasing amount of NT60-FluPol or ANP32 protein in the presence of constant ANP32 protein or NT60-FluPol, respectively. * putative cleaved mRNA. E. Effect of increasing amount of wild type (NT60-FluPol) or symmetric regulatory dimer mutant polymerase (NT60-PA$_{352-356A}$) in the presence of ANP32 protein. F. Effect of increasing amount of wild type 1918 polymerase (1918-FluPol) or endonuclease mutant 1918 polymerase (1918-PA$_{D108A}$) in the presence of ANP32 protein. G. Effect of increasing amount of influenza C virus polymerase (FluPolC) in the presence of ANP32 protein. H. Effect of wild type (NT60-FluPol) or PB2$_{627}$ deletion viral polymerase (NT60-ΔPB2$_{627}$) in the absence and presence of ANP32B.*

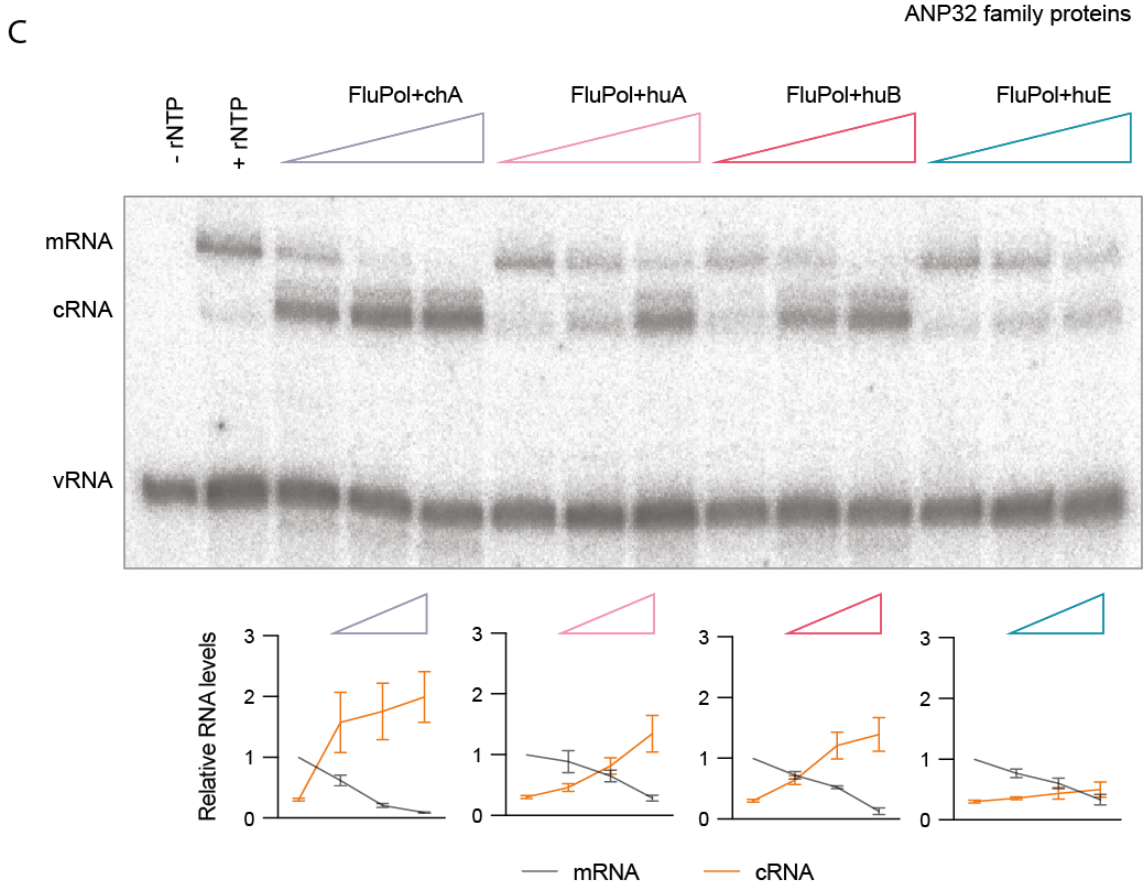
2.2.4 ANP32 isoforms each support replication to different levels and ANP32^{LCAR} is required for replication

Species-specific variation in ANP32 family proteins underlies the hindered replication of avian influenza A virus within mammalian hosts (Long et al., 2019). Specifically, the ANP32 proteins found in mammals are inadequate to support the efficient replication by avian influenza A viral polymerase. To gain further insight into the function of ANP32 family proteins, I carried out an investigation comparing the three human isoforms of ANP32 proteins, namely ANP32A, B, and E, in promoting replication in the presence of human-adapted viral polymerase (NT60-FluPol). Additionally, an avian ANP32 representative, chicken ANP32A, was included due to its distinct structure, namely a 33-amino acid insertion into ANP32^{LCAR}, compared to the human ANP32 proteins (Figure 2.5A-B). ANP32 proteins purified from insect cells were kindly gifted by Dr Haitian Fan. My results revealed that among the tested ANP32 proteins, chicken ANP32A exhibited the most robust support for replication. This observation is counterintuitive but expected, as it is consistent with previous findings showing the strongest affinity in chicken ANP32A for the human-adapted viral polymerase (Domingues & Hale, 2017). Activity results on human ANP32 proteins also agree with their relative affinity to viral

polymerase: ANP32B most strongly promotes replication, closely followed by ANP32A, while human ANP32E failed to support replication in the presence of viral polymerase. The transcription inhibitory effect of each ANP32 protein in the presence of viral polymerase is comparable to their replication-promoting ability, with chicken ANP32A appearing to be the most potent in transcription inhibition, and ANP32E showing an inhibition level that resembles that of polymerase alone. These outcomes agree with previous studies highlighting the inefficiency of human ANP32E in supporting viral replication in cells (Sheppard et al., 2023) (Figure 2.5C).

Interestingly, an avian-origin H5N1 influenza A virus isolated in Turkey from a turkey has recently been used in a study to investigate whether viruses could adapt to using ANP32E for replication in an ANP32A and ANP32B double-knockout mammalian cell line. A $PB1_{K577E}$, PA_{Q556R} double mutation was discovered upon the virus gaining the ability to replicate in the double-knockout cell line after serial passages (Sheppard et al., 2023). When we introduced these mutations into the viral polymerase (TurM-FluPol) and compared the replication-promoting ability of each type of ANP32 protein in the presence of this mutant polymerase, I found that human ANP32E can support replication to the same extent as other ANP32 proteins at the chosen concentration (Figure 2.5D). This result highlights the importance of these double mutations on viral-host adaptation.


A
chA
huA
huB
huE
LRR
LCAR
B
kDa
chA
huA
huB
huE
250
150
100
75
50
37
25
ANP32 family proteins


C
- rNTP
+ rNTP
FluPol+chA
FluPol+huA
FluPol+huB
FluPol+huE
mRNA
cRNA
vRNA
Relative RNA levels
0
1
2
3
mRNA
cRNA

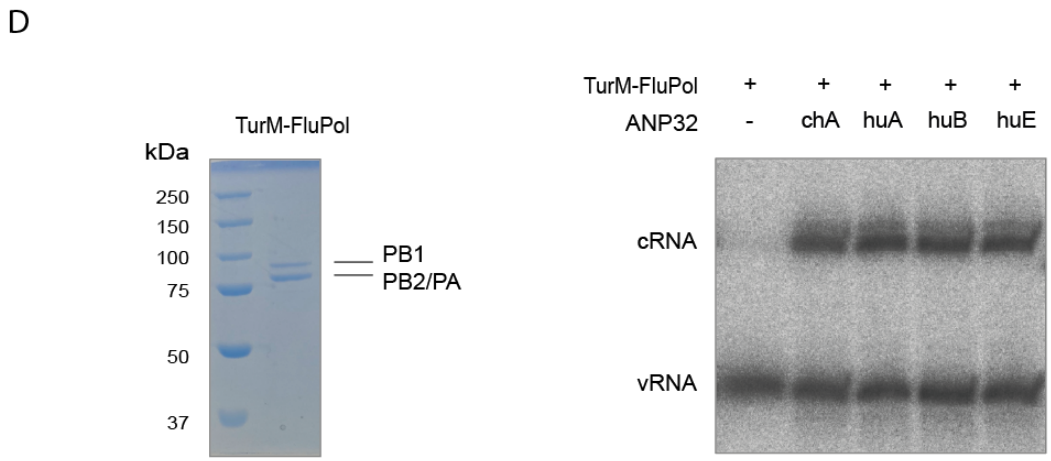
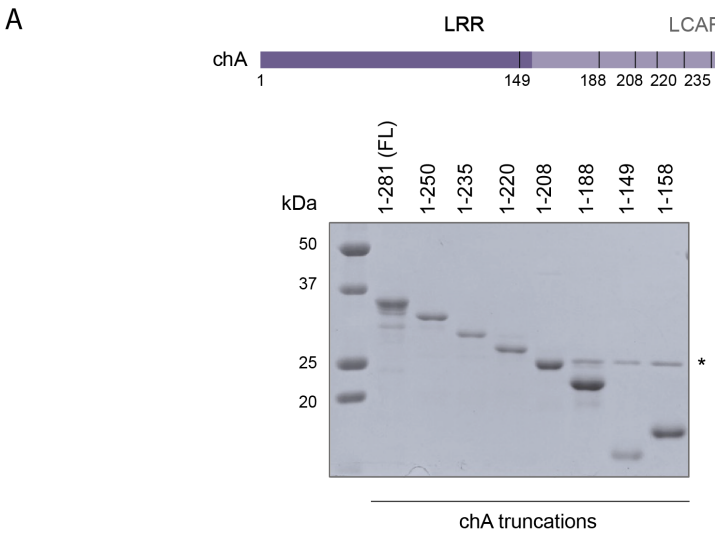

D
TurM-FluPol
kDa
250
150
100
75
50
37
PB1
PB2/PA
TurM-FluPol + + + + +
ANP32 - chA huA huB huE
cRNA
vRNA

Figure 2.5 Effect of ANP32 isoforms on vRNP replication. *A. Schematic diagram of chicken ANP32A (chA), human ANP32A (huA), human ANP32B (huB) and human ANP32E (huE) protein. LRR and LCAR are illustrated in dark and light colours, respectively. B. SDS-PAGE analysis followed by Coomassie-staining of ANP32 proteins purified from bacterial cells. C. Effect of increasing amount of NT60 viral polymerase (FluPol) and different ANP32 proteins on vRNP activity in the presence of globin mRNA as capped-primer donor. A representative result of three independent RNA analyses by primer extension is shown (top). Relative RNA levels (bottom) were calculated by comparison to mRNA levels without added FluPol and ANP32 protein. D. Effect of adding different ANP32 proteins together with Turkey* $PB1_{K577E}$, PA_{Q556R} *double mutant polymerase (TurM-FluPol) on vRNP activity in the presence of globin mRNA as capped-primer donor. A representative result of RNA analysis by primer extension is shown.*

Next, I investigated the role of $ANP32^{LCAR}$ in supporting replication. The known function of $ANP32^{LCAR}$ is to recruit NP for RNP assembly; however, since $ANP32^{LCAR}$ was not fully resolved in the ANP32-bound polymerase dimer structure, its involvement in the function of the replication complex is unclear (Carrique et al., 2020). To address this question, I swapped full-length chicken ANP32A for truncated chicken ANP32A with serial truncation of $ANP32^{LCAR}$ (a gift from Dr Haitian Fan) and tested their functions in the presence of viral polymerase in primed reactions (Figure 2.6A). My results show that deletions of up to 46 amino acid residues (1-235, 1-250) from the C-terminus of chicken ANP32A produced a similar effect in vRNP activity in the presence of viral polymerase, compared to full-length chicken ANP32A (1-281). However, as more amino acid residues were deleted, truncated ANP32A lost its ability to promote replication and inhibit transcription beyond that observed with adding polymerase alone (Figure 2.6B).

Taken together, these findings underscore the distinct impact of various ANP32 family proteins on regulating vRNP activity in the presence of viral polymerase with or without specific adaptative mutations. My results also provide direct evidence that $ANP32^{LCAR}$ plays

a previously uncharacterised role in facilitating viral genome replication, in addition to NP recruitment.



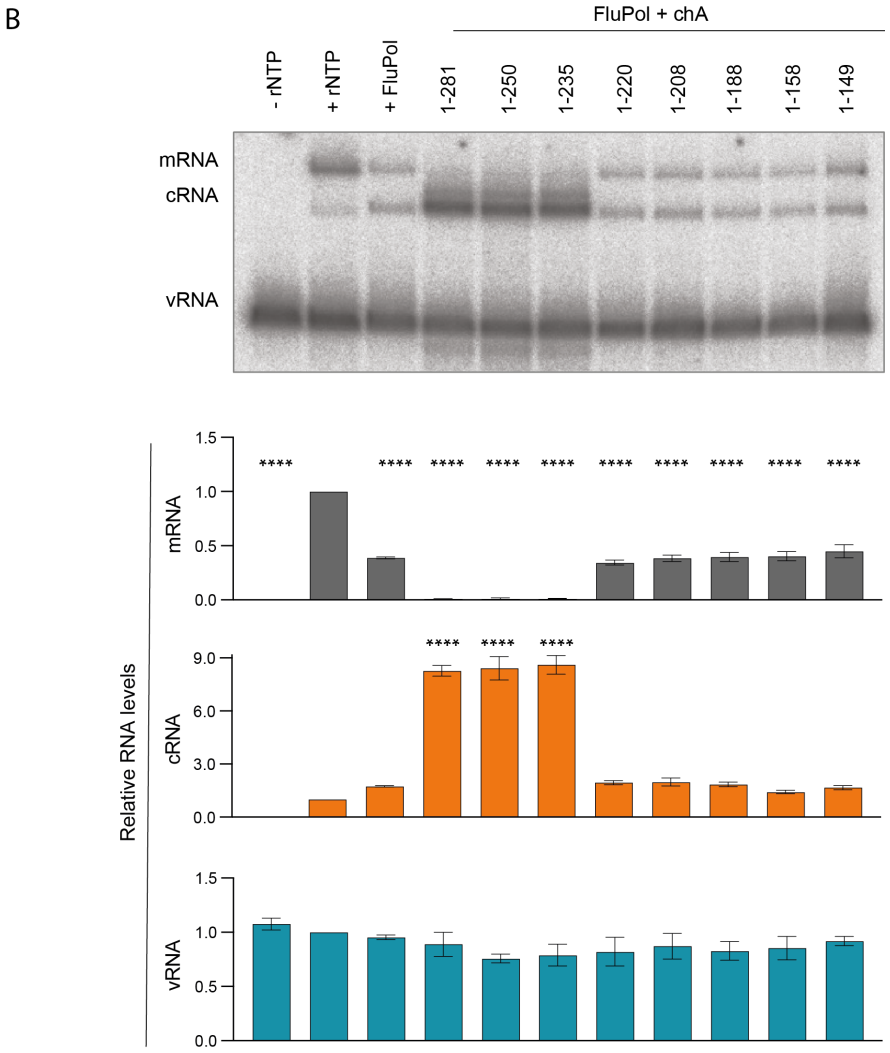


Figure 2.6 Effect of chicken ANP32^{LCAR} serial truncation on vRNP activity. *A. Schematic diagram of chicken ANP32A (chA) and truncated mutants lacking parts or the LCAR domain. LRR and LCAR are illustrated in dark and light purple respectively. The numbers under the schematic show the length of each mutant in amino acid (top). SDS-PAGE analysis followed by Coomassie-staining of chA full-length or truncated mutants purified from bacterial cells (bottom). * Shows the presence of 3C protease in some samples from post-purification tag cleavage. B. Effect of chA full length and truncated mutants together with NT60 viral polymerase (FluPol) on vRNP activity. A representative result of three independent RNA analyses by primer extension is shown (top). Relative RNA levels (bottom) were calculated by comparison to mRNA levels without added FluPol and chA. Ordinary one-way ANOVA was used to assess significance. In A-B, data are represented as mean values ± s.e.m., based on n = 3 independent vRNP purifications and reactions. ****, p<0.0001.*

2.2.5 Pol II CTD has different effects on vRNP-bound or RNA-free viral polymerase

Within cells, the level of transcription and replication is likely to be kept optimal at a given time by a balance between transcription-promoting cues and replication-promoting cues. Host Pol II, which binds to viral polymerase via its C-terminal domain (CTD), has long been believed to be a transcription cue because its interaction with viral polymerase is crucial for cap-snatching and primer-dependent transcription initiation (Engelhardt et al., 2005; Lamb & Choppin, 1977). Therefore, I sought to investigate the effect of Pol II CTD binding in the absence and presence of replication cues such as free viral polymerase and ANP32 proteins.

In mammals, Pol II CTD consists of 52 repeats of a 7-amino acid sequence, in which the phosphorylation on serine 5 residue has been shown to be important for the interaction with viral polymerase. When I added serine-5-phosphorylated (S5p) recombinant CTD mimic peptide (a gift from Alaa Baazaoui) with 26 repeats of the consensus sequence to primed

reactions, in the absence of viral polymerase or ANP32 protein, I observed an increase in transcription product. Surprisingly, I also observed an increase in replication product. On the other hand, addition of S5p CTD mimic peptide had no effect (Figure 2.8A). Similarly, both transcription and replication were enhanced by the phosphorylated CTD mimic peptides in the presence of additional viral polymerase (Figure 2.7A). However, when viral polymerase and ANP32 proteins are present simultaneously (at 320X the molar concentration of input vRNP), an increase in S5p CTD mimic peptide (from 16X to 640X the molar concentration of input vRNP) rescued the transcriptional activity but exerted inhibition in replication (Figure 2.8B).

Together, these results suggest that S5p CTD plays a general RNA-synthesis promoting role upon binding to vRNP. However, when bound to free viral polymerase in the presence of ANP32 protein, S5p CTD potentially hinders the free viral polymerase from interacting with ANP32 protein, therefore lowering the replication-promotion ability of free viral polymerase in the presence of ANP32 protein.

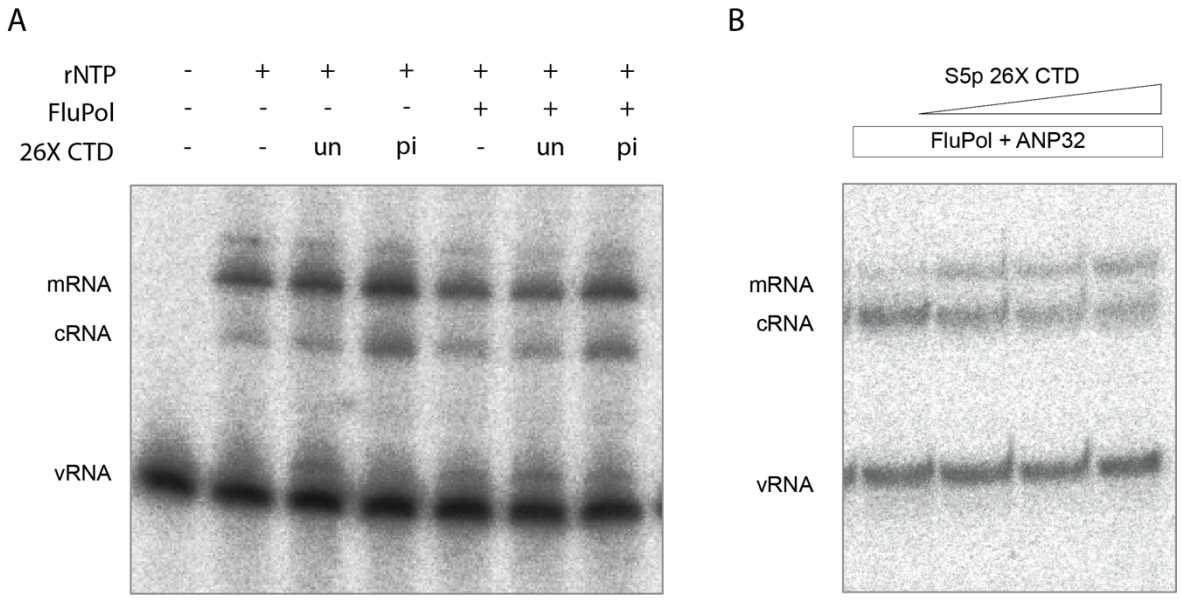


Figure 2.7 Effect of host Pol II CTD on vRNP activity. *A. Effect of unphosphorylated (un) or phosphorylated (pi) recombinant 26-repeat Pol II CTD mimic peptides (26X CTD) on vRNP activity in the presence of globin mRNA as capped-primer donor, with or without additional NT60 viral polymerase (FluPol). B. Effect of increasing amount of phosphorylated recombinant 26-repeat Pol II CTD mimic peptides (S5p 26X CTD) on vRNA activity in the present of FluPol, chicken ANP32A protein (ANP32) and globin mRNA as capped-primer donor. A representative result of RNA analysis by primer extension is shown.*

2.3 DISCUSSION

In this chapter, I aimed to understand how vRNP functions are regulated at the molecular level by identifying the minimal protein requirement for switching vRNP from transcription to replication *in vitro*. I first purified vRNPs from virions and confirmed that, in the presence of rNTPs and capped primer donor, these vRNPs preferably synthesise mRNA and have limited cRNA synthesis ability, as shown in previous studies (Vreede & Brownlee, 2007). Our observation is consistent with the order of events during influenza virus infection: the early onset of transcription and the delayed start of replication. At the molecular level, this implies that the virion-derived vRNP-associated polymerase is more likely to adopt a transcriptase conformation at the start of infection. I also found that, although in the absence of primer vRNP is lacking either transcriptional or replicational activity, addition of ApG dinucleotide significantly increases cRNA products. This finding shows that the initial dinucleotide formation event is likely to be the rate-limiting step for replication. In a vRNP, the position of the 3' promoter places C2G3 opposite the catalytic aspartates, and previous *in vitro* studies have reported cRNA products resulted from initiation from position 2 (Vreede & Brownlee, 2007). It is therefore possible that terminal dinucleotide *de novo* formation on the vRNA template is energetically unfavourable due to the suboptimal placement of the 3' promoter in the active site. As a result, vRNP has a preference for primer-dependent initiation of transcription.

Next, I demonstrated that additional RNA-free viral polymerase can inhibit the transcriptional activity of vRNP. This is consistent with the cryo-EM data showing that vRNA-associated viral polymerase bound to apo polymerase is in a transcriptionally-inactive conformation similar to RNA-free viral polymerase or cRNA-bound viral polymerase (Carrique et al., 2020; Hengrung

et al., 2015; Thierry et al., 2016). I also showed that additional viral polymerase alone is insufficient to support efficient replication. These results support a model in which RNA-free viral polymerase contributes to transcriptional inhibition by binding to and inducing conformational rearrangement that unfavours transcription in the vRNP-associated viral polymerase; however, to promote the rate-limiting step of replication, namely primer-independent terminal initiation, other mechanisms are required.

Excitingly, I then found that the addition of viral polymerase and ANP32 protein grants the vRNP-associated viral polymerase robust replicational activity. This finding strongly supports a recently proposed model of an influenza virus replication platform, made up of a viral polymerase asymmetric dimer and a copy of ANP32 protein, based on cryo-EM structures (Carrique et al., 2020). According to this model, interactions within the replication platform encourage the vRNP-associated viral polymerase to act as a replicase and the additional apo viral polymerase to act as an encapsidating polymerase. While previous studies have shown that this complex formation is necessary for replication, by mutating or blocking interaction surfaces, it was unclear whether this complex represents the minimal component to support efficient replication. Here, my data strongly support this model by showing that additional viral polymerase and ANP32 protein is sufficient to switch the enzymatic activity of vRNP from transcription to replication.

Since newly synthesised viral polymerase must be present for replication to occur, it has been under debate whether the additional apo polymerase operates *in trans* to replicate the vRNA template. Previous cRNA stabilization data show that the expression of an inactive viral polymerase can still support cRNA accumulation (Vreede et al., 2011). However, the caveat of these *in vivo* studies is that infections were carried out at a high multiplicity of infection (MOI),

meaning that a significant number of active viral polymerases may be present by dissociation from vRNPs; the proportion of active polymerase cannot therefore be known. In contrast, in my *in vitro* study, the input vRNP sample has been separated from dissociated viral polymerase by glycerol gradient, and a huge excess of inactive viral polymerase was added to ensure that any active viral polymerase dissociated from vRNP would constitute an insignificant proportion of the total free polymerase. In these conditions, I found that the additional viral polymerase would not require RNA synthesis activity to have transcription-inhibiting and replication-promoting effects. Additionally, I showed that the polymerase PB2^{627} domain, which mediates polymerase dimerisation and its interaction with host ANP32 proteins, is crucial for replication promotion. Together, my result disproved the 'trans-acting model' and provided support for the 'replication-switch model' for viral genome replication *in vitro*. In other words, it is the vRNP-associated viral polymerase that is responsible for synthesizing cRNA, and the additional viral polymerase only acts as an allosteric regulator to switch on the replicational activity in the vRNP-associated viral polymerase. It is worth noting that, during infection, the 'stabilisation model', in which viral polymerase and NP bind to and stabilises newly synthesised replication product, is likely to contribute additionally to the accumulation of cRNA due to the presence of RNA degradation enzymes in host cells.

Influenza A viral polymerase readily forms a symmetric regulatory dimer in solution (Fan et al., 2019). Interestingly, my results suggest that only free viral polymerase in a monomeric state can be involved in replication complex formation. Evidently, the dimerisation interfaces on the encapsidating viral polymerase in the asymmetric replicating-encapsidating dimer, and on both polymerases in the symmetric regulatory dimer, both involve PA-C. Therefore, a viral polymerase in dimeric form in solution will have no free interface to interact with ANP32 protein and the vRNP-associated viral polymerase. During influenza A viral infection, it is

possible that newly synthesised viral polymerases dimerise and form a polymerase reservoir, which can participate in replication upon dimer dissociation. Alternatively, unknown viral or host factors might be present that prevent viral polymerase symmetric dimerisation and allow monomeric viral polymerase to be incorporated into the replication platform. In either scenario, the timing of replication can be further regulated through controlling the polymerase oligomeric status.

Using this *in vitro* assay, I also confirmed that ANP32 is indispensable in supporting viral genome replication. Interestingly, I found that the majority of the ANP32^{LCAR} is needed to support *in vitro* vRNP replicational activity. My result and hypothesis are also consistent with a previous study showing that while ANP32^{LCAR}-truncated mutants can still support NP-independent replication, a reduced rate of replication has been observed at earlier time points during infection (Wang et al., 2022). Presumably, ANP32^{LCAR} is important for replication in a vRNP but can be somewhat compensated by other unknown mechanisms in cells during infection. Prior to the study, the involvement of ANP32^{LCAR} in NP-independent replication was not clear because ANP32^{LCAR} was not resolved in the cryo-EM structure of the replication complex and was only reported to be involved in RNP assembly. Therefore, my data suggest a previously-unknown role of ANP32^{LCAR} in viral genome replication.

Moreover, I found that avian and human ANP32 isoforms differ in their ability to promote genome replication together with viral polymerase containing human-adapted PB2-K627, as reported in previous studies (Long et al., 2016). Specifically, chicken ANP32A has the strongest effect, followed by human ANP32B, then human ANP32A, and human ANP32E has an undetectable effect. In other words, chicken ANP32A, human ANP32B and human ANP32A require from the lowest to the highest viral polymerase concentration to completely

switch vRNP transcription to replication. This observation could be explained by the decreased affinity of these ANP32 proteins for viral polymerase, as previously reported (Long et al., 2016; Zhang et al., 2020). Additionally, the degree of replicational support provided by different truncation of ANP32^{LCAR} aligns with the affinity of these LCAR-truncated ANP32 for viral polymerase in a pull-down assay (Domingues & Hale, 2017). The perfect match between the affinity of ANP32 protein for viral polymerase and its replication-promoting ability further reveals a possible role of ANP32 proteins prior to replication initiation: since my earlier data suggest that only viral polymerase in a monomeric state can participate in replication complex formation, it is possible that ANP32 proteins regulate the oligomerisation state of free viral polymerase through binding to and dissociating the readily formed polymerase dimer, and control the number of monomeric polymerase that can act as encapsidating polymerases. According to this hypothesis, an ANP32 protein with a higher affinity to viral polymerase is better at dissociating viral polymerase symmetric dimers. This hypothesis is backed up by structural data showing that the binding site of ANP32 protein on an encapsidating polymerase shares a similar surface as the dimerisation interface of the polymerases in the symmetric regulatory dimer (Carrique et al., 2020). The hypothesis is further supported by two very recent studies which independently reveal structures of ANP32 protein in complex with a monomeric viral polymerase, which is, unsurprisingly, in encapsidating conformation (Arragain et al., 2024; Staller et al., 2024).

Intriguingly, my results suggest that host Pol II CTD promotes transcription as well as replication but can also revert the replication-promoting effect of viral polymerase and ANP32 protein by inhibiting replication. This seemingly contradictory effect of CTD is in fact consistent at the molecular level because CTD could have a different impact on either vRNP-associated polymerase or free viral polymerase. The model supported by my data suggests that

upon binding to RNP-associated polymerase, CTD promotes RNA synthesis in general, i.e., both transcription and replication; however, upon binding to free viral polymerase, CTD prevents the polymerase from interacting with ANP32 proteins, thereby inhibiting the formation of the replication platform. Therefore, when there is only additional viral polymerase, CTD still increases the general activity of vRNP-associated viral polymerase; however, in the presence of ANP32 proteins, the addition of CTD diminishes the effect of ANP32 through binding competition. However, this is contradictory to a recent study suggesting that CTD promotes ANP32 binding to viral polymerase as well as polymerase dimerisation (Krischuns et al., 2024). In this study, the presence of CTD together with ANP32 protein appears to increase replicational activity of recombinant viral polymerase and allows production of full length cRNA on a 51 nt vRNA template. This discrepancy could be caused by differences between the two systems, one using vRNPs and the other using a synthetic RNA template in the absence of NP.

In summary, my data provide molecular insights into how transcription and replication are regulated in vRNP, which is crucial for a successful viral life cycle. Furthermore, I have provided strong evidence to support the recent structural model for a replication platform and revealed novel roles of ANP32^{LCAR} and CTD in viral genome replication, providing new possibilities for the development of new antivirals against the influenza virus by targeting the mechanisms that are crucial for regulating viral genome replication.

Chapter 3: *In vitro* reconstitution of a complete cycle of viral genome replication

Parts of the data from this chapter were published in:

***Zhu Z**, Fan H, Fodor E (2023) Defining the minimal components of the influenza A virus replication machinery via an in vitro reconstitution system. PLoS Biol 21(11): e3002370.*

3.1 INTRODUCTION

Influenza virus genome replication occurs through a two-step process. In the first step, a cRNA is produced using a viral vRNA template and is assembled into a cRNP together with a viral polymerase and an oligomer of NP. The second step involves the synthesis of vRNA from the cRNA template and is believed to take place within the context of newly assembled cRNP. During infection, viral genome replication only occurs in the presence of newly synthesised viral proteins, particularly viral polymerase and NP. This leads to a delay in viral genome replication compared to transcription. In agreement with this, virion-isolated vRNPs readily synthesise mRNA in the presence of capped-primer donor, but possess limited cRNA synthesis ability in both the primed and the unprimed conditions. In the previous chapter, I successfully identified the minimal protein requirement to switch on cRNA synthesis in purified vRNPs. This facilitates the reconstitution of the first step of replication inside a test tube. However,

vRNA levels remained unchanged. This means that the *de novo* synthesised cRNA product was unable to serve as a template for vRNA synthesis; in other words, the newly added viral polymerase remained inactive. To reconstitute both steps of replication *in vitro*, additional replication-promoting factors, such as NP, must be identified.

The role of NP has been a subject of investigation for many years. Initially, NP was believed to function as a molecular switch from transcription to replication (Beaton & Krug, 1986; Shapiro & Krug, 1988). However, this proposal was later challenged by studies indicating that NP appeared to have no effect on the initiation and termination of cRNA synthesis in a cellular context. These studies claimed that NP primarily plays a product-stabilizing role and supports replication through its RNA-binding and oligomerisation abilities in RNP assembly (Deng et al., 2006; Turrell et al., 2013; Vreede et al., 2011). One plausible explanation for the unchanged vRNA signals in Chapter 2 is that, due to the absence of NP, cRNA was not assembled into cRNP; therefore, vRNA synthesis could not happen in the context of cRNP. In this chapter, I aimed to define the minimal protein requirement for a full cycle of influenza genome replication inside a test tube and establish the first *in vitro* system that reconstitute both steps of viral genome replication. Furthermore, I aimed to utilize this system to dissect specific functions of viral replication-promoting factors.

3.2 RESULTS

3.2.1 NP stimulates replicational activity and enables the complete cycle of viral genome replication

In Chapter 2, I demonstrated that additional viral polymerase and ANP32 proteins robustly increase cRNA synthesis in purified vRNPs. However, vRNA levels remained constant, indicating that product cRNA could not serve as template for vRNA synthesis. This could be due to the lack of RNP assembly and/or the failure in synthesizing full-length cRNA, as primer extension can only detect the 5' terminus of cRNA products. Previous studies have shown that NP is involved in RNP assembly and is essential for the replication of full-length viral genome (Elton et al., 1999; Turrell et al., 2013; Wang et al., 2022) Therefore, I explored the effect of NP on virion-derived vRNP activity using recombinant NP from influenza virus A/NT/60/1968 (H3N2), purified from bacterial cells (a gift from Dr Haitian Fan).

To study the effect of NP, I first added recombinant NP, together with rNTPs, to purified vRNP, in unprimed or globin mRNA-primed conditions. I observed a barely noticeable increase in cRNA signals in unprimed conditions and a decrease in mRNA levels in primed conditions (Figure 3.1A, B, lane 2, 3). This observation indicates that, like free viral polymerase, NP alone does not significantly enhance replication, but might act as an allosteric activity modulator of the vRNP-associated viral polymerase, inhibiting its transcriptional activity. Alternatively, this could simply be a result of NP binding to, and depleting, the added globin mRNA. When NP was added in addition to either ANP32B or viral polymerase, vRNP activity was no different from adding either ANP32B or viral polymerase alone (Figure 3.1A, B, lane 4, 5, 8). Yet

excitingly, when NP was added together with both ANP32B and wild type viral polymerase, there was a significant increase in vRNA signals compared to input vRNA, or any conditions lacking NP (Figure 3.1A, B, lane 10). Remarkably, this increase in vRNA levels was accompanied by an additional increase in cRNA levels. These results suggest that cRNA produced by virion-derived vRNP is used as a template to produce new vRNA, which then could act as a template for more cRNA synthesis.

To test whether the increase in vRNA and cRNA levels is truly due to the catalytic activity of added viral polymerase, I replaced the wild type viral polymerase (NT60-FluPol) with a catalytically inactive viral polymerase (NT60-PB1a), which could not support further vRNA synthesis on a newly produced cRNA template. As expected, no increase in vRNA was observed if NT60-PB1a was used. However, intriguingly, the increase in cRNA level occurred regardless of whether active or inactive viral polymerase was used, suggesting that NP can act as an additional factor to directly promote cRNA synthesis in input vRNP (Figure 3.1A, B, lane 6, 7). Similar results were observed in primed or unprimed conditions, except that the degree of replication promotion in unprimed conditions is higher than that in primed conditions. This finding is consistent with my results from Chapter 2, where a lower concentration of viral polymerase and ANP32B is required to promote replication in unprimed conditions. In addition, when an avian ANP32 protein, chicken ANP32A, was used instead of human ANP32B, similar results were observed, apart from the degree of replication promotion being stronger with chicken ANP32A (results not shown), agreeing with my data from Chapter 2. To investigate whether the effects of NP are concentration-dependent, I titrated NP into the system with a constant concentration of viral polymerase and ANP32B (Figure 3.1C). I selected the constant concentration for viral polymerase and ANP32B from Figure 2.4B, as the lowest concentration sufficient to produce a plateau in cRNA synthesis in globin-mRNA primed conditions:

specifically, 160X that of input vRNP. To titrate NP into the system, I gradually increased the molar ratio of NP to viral polymerase and ANP32B, from 0.2X to 4X. This ratio seems rather low considering there is one viral polymerase but numerous NP in an RNP. However, it is worth remembering that, according to my previous data (Figure 2.4D), viral polymerase is the limiting factor when it is added at an equal molar ratio as ANP32B because only polymerase in monomeric state can participate in the formation of replication platform. I estimate that the effective viral polymerase (i.e. viral polymerase in monomeric state and can act as an encapsidating polymerase) accounts for roughly 1/20 of the total added polymerase in the presence of chANP32A. Taking this information into consideration, the estimated molar ratio of NP to the effective viral polymerase and human ANP32B would be more than 4X to 80X, covering the stoichiometry of NP and viral polymerase within an RNP. My results show that increasing NP in the presence of constant inactive viral polymerase and ANP32B steadily increases cRNA levels, and increasing NP in the presence of constant wild type viral polymerase and ANP32B leads to a robust, concentration-dependent elevation in both cRNA and vRNA levels (Figure 3.2C-D). However, it is not possible to assess whether additional cRNA was made when wild type viral polymerase was used instead of inactive viral polymerase, therefore, whether secondary cRNA synthesis occurred remains unclear. These results suggest that NP plays two independent roles in promoting vRNP replication: 1. by directly promoting cRNA synthesis and 2. by supporting vRNA synthesis on a cRNA template.

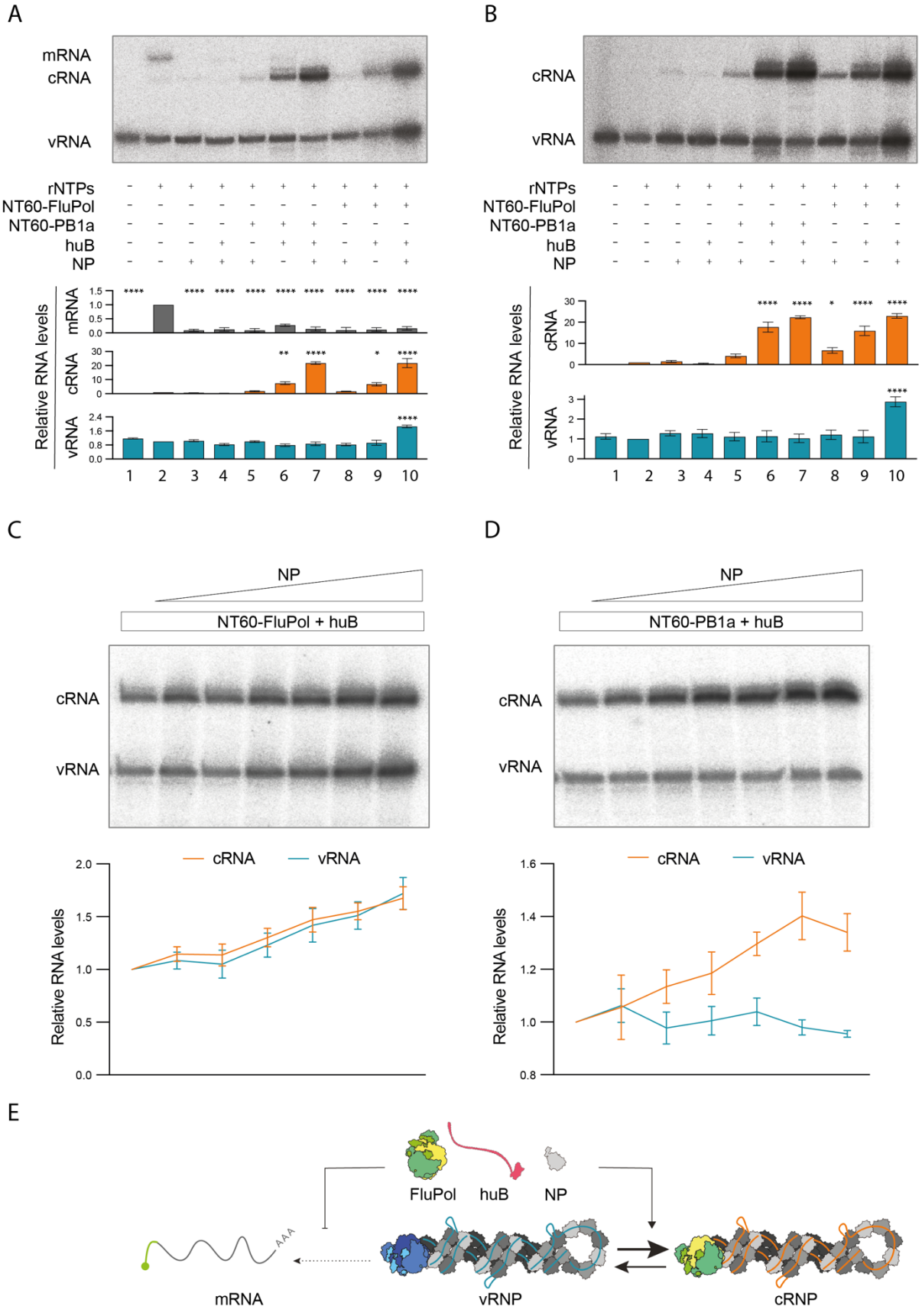

A
mRNA
cRNA
vRNA
rNTPs
NT60-FluPol
NT60-PB1a
huB
NP
Relative RNA levels
mRNA
cRNA
vRNA
1 2 3 4 5 6 7 8 9 10
B
cRNA
vRNA
rNTPs
NT60-FluPol
NT60-PB1a
huB
NP
Relative RNA levels
cRNA
vRNA
1 2 3 4 5 6 7 8 9 10
C
NP
NT60-FluPol + huB
cRNA
vRNA
cRNA
vRNA
Relative RNA levels
D
NP
NT60-PB1a + huB
cRNA
vRNA
cRNA
vRNA
Relative RNA levels
E
FluPol
huB
NP
AAA
mRNA
vRNP
cRNP

Figure 3.1 Viral NP promotes vRNP replication and enables a complete cycle of viral genome replication. *A-B. Effect of NP alone or in combination with inactive (NT60-PB1a) or active (NT60-FluPol) viral polymerase and human ANP32B (huB) on vRNP activity, in the presence (A) or absence (B) of globin mRNA as capped-primer donor. A representative result of three independent RNA analyses by primer extension is shown (top). Relative RNA levels (bottom) were calculated by comparing them to the RNA levels obtained in the absence of FluPol, huB and NP. Ordinary one-way ANOVA was used to calculate statistical significance. C-D. Effect of increasing amounts of NP in the presence of a constant level of huB and NT60-FluPol (C) or NT60-PB1a (D) on vRNP activity. A representative result of three independent RNA analyses by primer extension is shown (top). Relative RNA levels (bottom) were calculated by comparing them to the RNA levels without added NP. In A-D, data represent mean values ± s.e.m., based on n = 3 independent vRNP purifications and reactions. *, p<0.05; **, p<0.01; ***, p<0.001; ****, p<0.0001. E. Schematic diagram showing that FluPol, huB and NP are the minimal set of proteins required for a complete cycle of viral genome replication.*

We have so far only been focusing on one of the eight segments, namely the NA segment, using NA-specific primers, in our primer extension analysis. To investigate whether other viral segments are affected in the same way in this assay, I quantified mRNA, cRNA, and vRNA levels of each segment in the presence of viral polymerase, ANP32 protein and NP, using primers specific to each viral segment. Except for the HA segment, whose primers failed to work, all other segments showed similar changes in mRNA, cRNA, and vRNA levels as the NA segment when viral polymerase, ANP32B, and NP were added in different combinations. Therefore, it is likely that transcription and replication are regulated similarly in all segments. These results provide strong evidence that replication of all influenza viral segments has been reconstituted in a test tube.

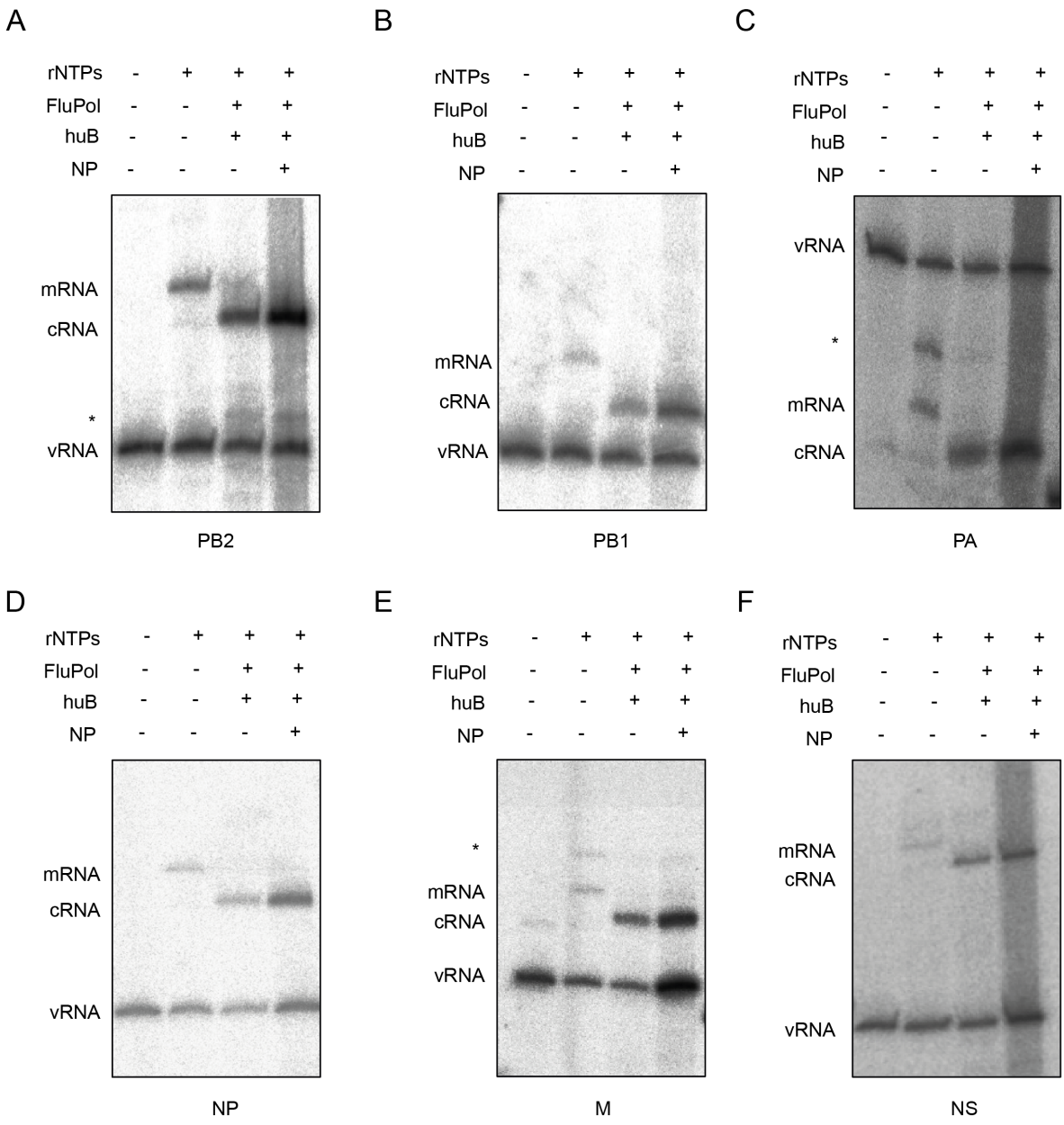


Figure 3.2 Viral polymerase, ANP32 protein and viral NP enable a full replication cycle in different influenza virus genome segments. *Effect of viral polymerase (FluPol), human ANP32B (huB) and NP alone or together on vRNP activity in segment 1 PB2 (A), segment 2 PB1 (B), segment 3 PA (C), segment 5 NP (D) segment 7 M (E) and segment 8 NS (F), in the presence of globin mRNA as capped-primer donor. A representative result of RNA analysis by primer extension is shown. *additional bands with unknown origins.*

I then investigated whether viral polymerases from different influenza viruses have activity in supporting replication. vRNP used in this study was isolated from A/WSN/33 (H1N1) virus,

but the recombinant viral polymerase and NP were cloned from A/NT/60/1968 (H3N2) virus because technical difficulties were encountered during the expression and purification of matching recombinant A/WSN/33 (H1N1) viral polymerase. Although viral polymerase of A/NT/60/1968 (H3N2) shares over 96% sequence identity and nearly 99% sequence similarity with its homologue from A/WSN/33 (H1N1), it remains a question whether a mismatch in subtypes between vRNP, free viral polymerase, and NP has masked any effect. To answer this question, I compared the effect of the same amount of influenza A viral polymerases from A/NT/60/1968 (H3N2), A/Brevig Mission/1/1918 (H1N1), A/WSN/33 (H1N1), and A/Fujian/01/2002 (H5N1) and a viral polymerase from influenza C virus (a gift from Dr Haitian Fan) (Figure 3.3A). I found that all of the tested influenza A viral polymerases show a similar effect on vRNP activity to the viral polymerase from A/NT/60/1968 (H3N2). Specifically, they inhibit transcription and strongly promote replication in the presence of ANP32B (Figure 3.3B). When NP from A/NT/60/1968 (H3N2) was added to the system, a notable further increase in cRNA levels and a less-obvious increase in vRNA levels was observed across all influenza A viral polymerases. In contrast, while influenza C viral polymerase can partially inhibit transcription, it cannot promote either step of replication (Figure 3.3B). I additionally compared the effect of wild type and inactive viral polymerase from A/NT/60/1968 (H3N2) and A/WSN/33 (H1N1) on WSN vRNP and showed that both active (WSN-WT) and inactive WSN polymerases (WSN-PB1a) resulted in RNA degradation in the system. This is likely caused by the presence of nuclease in purified WSN-WT and WSN-PB1a samples because WSN polymerases were not purified through size exclusion chromatography (SEC) due to their low yields. Therefore, while I showed that influenza A viral polymerase from can promote replication and inhibit transcription of WSN vRNP across tested viral subtypes, it was impossible to conclude whether WSN viral polymerase has an additional effect on WSN vRNP activity.

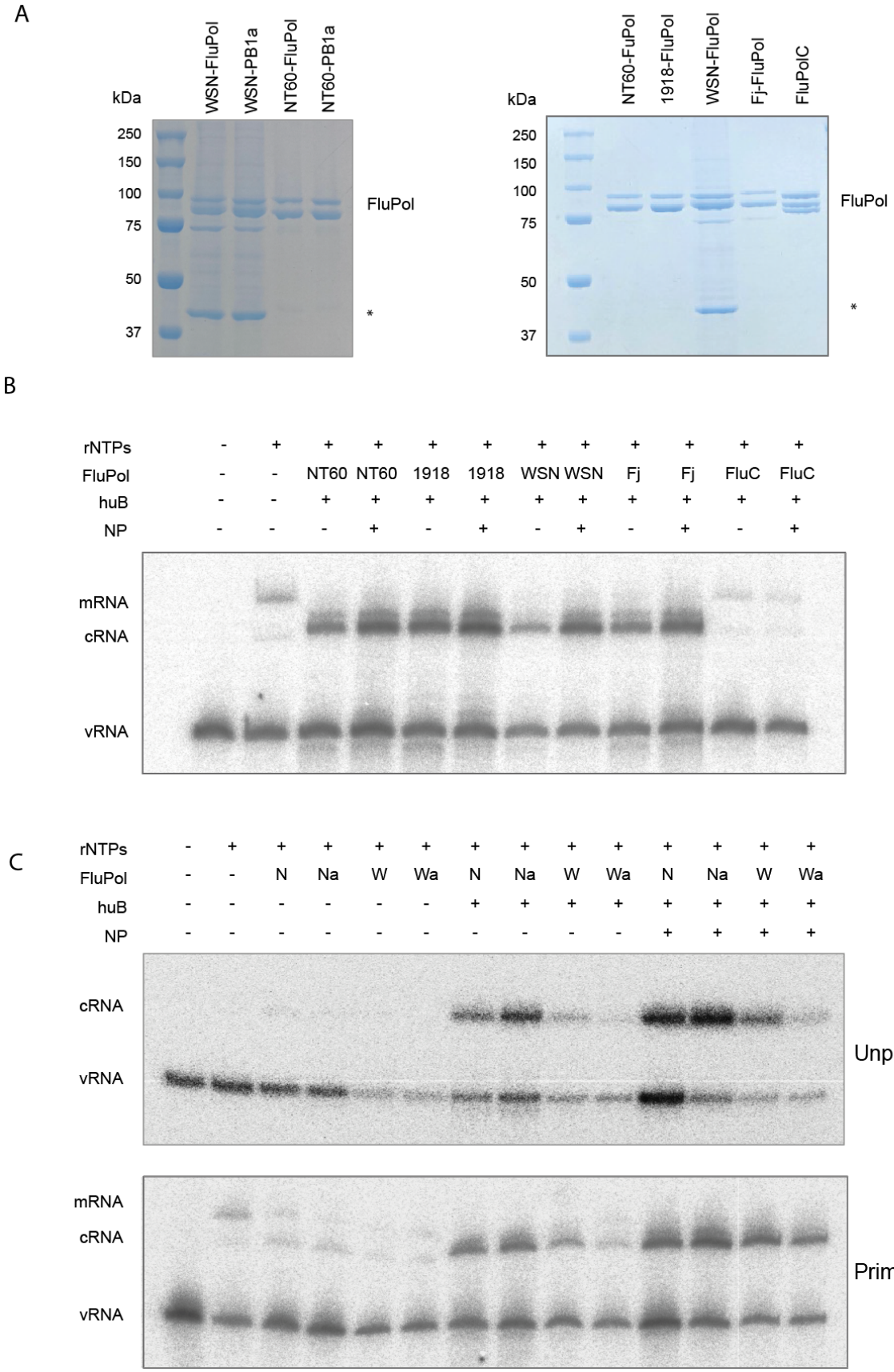

A
kDa
WSN-FluPol
WSN-PB1a
NT60-FluPol
NT60-PB1a
250
150
100
75
50
37
FluPol
*
kDa
NT60-FuPol
1918-FluPol
WSN-FluPol
Fj-FluPol
FluPolC
250
150
100
75
50
37
FluPol
*
B
rNTPs - + + + + + + + + + + +
FluPol - - NT60 NT60 1918 1918 WSN WSN Fj Fj FluC FluC
huB - - + + + + + + + + + +
NP - - - + - + - + - + - +
mRNA
cRNA
vRNA
C
rNTPs - + + + + + + + + + + + + +
FluPol - - N Na W Wa N Na W Wa N Na W Wa
huB - - - - - - + + + + + + + +
NP - - - - - - - - - - + + + +
cRNA
vRNA
Unprimed
mRNA
cRNA
vRNA
Primed

Figure 3.3 Effect of viral polymerase from different types and subtypes of influenza viruses on vRNP activity. *A. SDS-PAGE analysis followed by Coomassie staining of wild type (WSN-FluPol) and inactive (WSN-PB1a) A/WSN/33 (H1N1) viral polymerase, wild type (NT60-FluPol) and inactive (NT60-PB1a) A/NT/60/1968 (H3N2) viral polymerase, wild type A/Brevig Mission/1/1918 (H1N1) viral polymerase (1918-FluPol), wild type A/Fujian/01/2002 (H5N1) viral polymerase (Fj-FluPol), and wild type influenza C viral polymerase (FluPolC) purified from insect cells. * Indicates TEV protease contamination. B. Effect of different subtypes of influenza A viral polymerase (FluPol) or influenza C viral polymerase (FluC), and human ANP32B (huB), in the absence or presence of NP, on vRNP activity in globin-mRNA primed condition. C. Effect of active (N) or inactive (Na) NT60 viral polymerase, and active (W) or inactive (Wa) WSN viral polymerase on vRNP activity in unprimed or globin-mRNA primed conditions. For B-C, a representative result of RNA analysis by primer extension is shown.*

In this section, I demonstrated that vRNP, viral polymerase, ANP32 protein, and NP represent the minimal protein requirements to reconstitute a complete cycle of viral genome replication inside a test tube. Additionally, I identified a previously unknown regulatory role of NP in viral genome replication (Figure 3.1E).

3.2.2 Investigating the roles of NP in modulating vRNP activity

Results from the previous section indicate that NP plays more than one role in regulating RNP activity. It inhibits transcriptional activity, directly promotes cRNA synthesis together with viral polymerase and ANP32 proteins, and enables vRNA synthesis using newly produced cRNA as template. Here, I aimed to use the established *in vitro* assay to study the biochemical properties that enable NP to carry out these roles.

I first studied whether the RNA-binding and oligomerisation properties of NP are required for its functions. According to previous literature, these properties of NP enable it to act as a scaffold for genome packaging and allow the formation of the RNP filament. To study this, I took advantage of purified NP mutants expressed in bacterial cells (a gift from Dr Haitian Fan) (Figure 3.4A). Specifically, three NP mutants were used: 1. An RNA-binding mutant with eight arginine residues (R74A/R75A/R150A/R152A/R156A/R162A/R174A/R175A) in the RNA-binding G1 and G2 grooves replaced by alanine (NP8A); 2. A monomeric mutant with a single mutation (R416A) introduced in the tail-loop to disrupt the interaction with adjacent NP, hence cannot oligomerise (mNP); 3. A combination of NP8A and mNP, which is monomeric and deficient in RNA-binding (mNP8A). An electrophoretic mobility shift assay was kindly performed by Dr Haitian Fan to confirm that NP8A and mNP8A have impaired RNA binding ability (Figure 3.4B).

By replacing wild type NP with NP mutants in the *in vitro* assay, I found that while WT NP alone inhibits transcription in primed conditions, none of mNP, NP8A or mNP8A has comparable transcription inhibition activity at the tested molar concentration (320X that of input vRNP) (Figure 3.4C). I also found that mNP could promote both cRNA and vRNA synthesis to a comparable level as WT NP, whereas NP8A and mNP8A no longer support a complete cycle of replication, specifically, vRNA synthesis. Curiously, both NP8A and mNP8A partially retained their ability to stimulate cRNA synthesis by the input vRNP. Moreover, I observed that the further increase in cRNA levels is higher when NP capable of RNA-binding (WT NP or mNP) was added, compared to when NP incapable of RNA-binding was added (WT NP8A or mNP8A). This indicates that, the RNA binding ability of NP contributes to the synthesis of at least some cRNA, likely through enabling newly synthesised

vRNA to function as the template for secondary cRNA synthesis, in the context of vRNP (Figure 3.4D).

The NP monomeric mutant has previously been reported to have a lower RNA-binding ability because the G2 RNA-binding groove is blocked by its C-terminal tail (Chan et al., 2010; Chenavas et al., 2013; Elton et al., 1999). Therefore, the loss in transcription inhibition in mNP might be a side effect of NP losing its RNA-binding ability. To test this, I expressed and purified a previously characterised NP mutant, that contains an R416A mutation together with C-terminal tail deletion (mNPdC) (Wang et al., 2022). The deletion of the C-terminal tail unblocks G2 and reverts the RNA-binding ability of mNP. I performed NP titration in the presence of globin mRNA and showed that the mRNA inhibition ability of NP decreased massively in mNP, but only decreased slightly in mNPdC, agreeing with my hypothesis that the loss of transcription inhibition in the oligomerisation mutant is an indirect effect of the loss of RNA-binding ability due to G2 groove blockage (Figure 3.4E). I also performed titration of WT NP, mNP, or mNPdC, from 0.2X to 8X molar ratio at a constant level of viral polymerase and ANP32 proteins, in the presence of globin mRNA. I found that both mNP and mNPdC support cRNA synthesis and a full cycle of replication, consistent with my previous results (Figure 3.5F). Together, these results suggest that the RNA-binding ability, but not the oligomerisation ability of NP is required to inhibit transcription and to support a full cycle of replication. However, the direct cRNA synthesis-promoting role of NP requires neither ability.

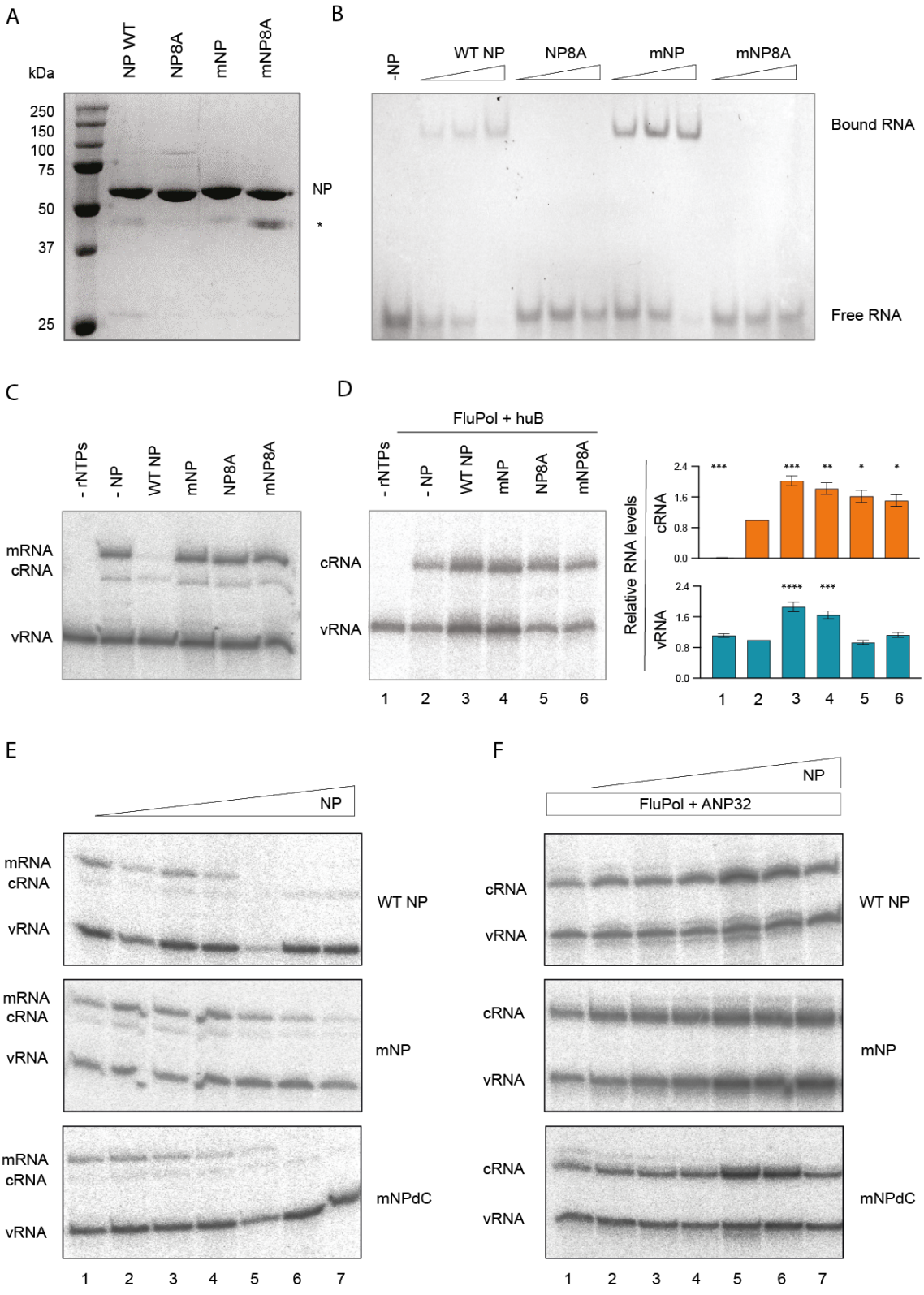


Figure 3.4 Effect of the oligomerisation ability and the RNA-binding ability of NP on its functions. *A. SDS-PAGE analysis followed by Coomassie staining of the wild type NP (WT NP), the RNA-binding deficient mutant NP (NP8A), the monomeric mutant NP (mNP) and the RNA-binding deficient monomeric NP (mNP8A) purified*

*from E. coli. * Indicates degraded NP. B. RNA-binding activity of WT and mutant NP assessed by electrophoretic mobility shift assay (EMSA) (Performed by Dr Haitian Fan). C. Effect of the oligomerisation and the RNA-binding ability on the transcription inhibition effect of NP in globin mRNA-primed conditions. D. Effect of the oligomerisation and the RNA-binding ability on the replication promotion effect of NP in the presence of viral polymerase (FluPol) and human ANP32B protein (huB) in globin mRNA-primed conditions. A representative result of three independent RNA analyses is shown (left). Relative RNA levels (right) were calculated by comparing them to the RNA levels obtained in the absence of NP. Ordinary one-way ANOVA was used to assess significance. Data represent mean values ± s.e.m., based on n = 3 independent vRNP purifications and reactions. *, p<0.05; **, p<0.01; ***, p<0.001; ****, p<0.0001. E-F. Effect of increasing amount of WT, M, and monomeric C-terminal deleted NP (MdC) in the absence (E) or presence (F) of a constant amount of wild type viral polymerase (FluPol) and ANP32 protein, in globin mRNA-primed conditions.*

Curiously, I observed that WT NP and mNPdC, but not mNP, inhibit viral genome replication at high concentrations (Figure 3.4E, F, lane 7). Since the RNA-binding grooves of NP have been reported to also bind to ANP32^{LCAR} through an interaction similar to the NP-RNA interaction, I hypothesized that excess NP inhibits replication by binding to ANP32^{LCAR} and preventing ANP32 proteins from interacting with viral polymerase to form the replication platform. To test this hypothesis, I decreased the amount of viral polymerase and ANP32 proteins that is kept constant in the NP titration experiment. By reducing the viral polymerase and ANP32 protein levels by 10X, from 160X to 16X of input vRNP, I theoretically decreased the amount of NP that is required to completely deplete ANP32 proteins. I left out globin mRNA for this experiment so that a significant level of replication can still happen at the chosen viral polymerase and ANP32 level, as shown in Figure 2.4C. I observed an earlier onset of replication inhibition at a lower viral polymerase and ANP32 level, starting from a proportionally lower NP concentration (Figure 3.4F, Figure 3.5A). This result is consistent with my hypothesis that excess NP inhibits replication by binding to and depleting ANP32

proteins. Therefore, I conclude that at a high relative molar concentration (from approximately 10X that of viral polymerase and ANP32 protein), NP inhibits replication.

In most of the experiments described above, I chose the lowest concentration of viral polymerase and ANP32 protein that produces a plateau in cRNA synthesis in the presence of globin-mRNA as capped primer donor. However, in unprimed conditions, the plateau is reached at a much lower viral polymerase and ANP32 protein concentration compared to the primed conditions, as shown in Figure 2.4B, C. Specifically, an increase in viral polymerase and ANP32 protein levels does not have an additional effect on cRNA synthesis after the level exceeds 32X that of input vRNP, in unprimed reactions; or 160X that of input vRNP, in the presence of globin mRNA (Figure 2.4B, C). This raises the question whether, in unprimed conditions, a lower level of viral polymerase and ANP32 can produce the same effect with NP, as in primed conditions. I hence examined the effect of viral polymerase and ANP32 levels, from 32X to 320X that of input vRNP, on vRNP activity in unprimed conditions. To avoid the replication inhibition effect of NP when it is at a high relative molar ratio to viral polymerase and ANP32 protein, I kept the molar ratio of NP to viral polymerase and ANP32 protein constant. I found that NP cannot promote replication efficiently in the presence of low levels of viral polymerase and ANP32 proteins, relative to input vRNP (Figure 3.5B). These results suggest that the effect of NP also depends on the concentration of viral polymerase and ANP32 protein.

Taken together, I showed that the roles of NP in influenza virus genome replication rely on its different properties, and also depend on the relative concentrations of viral polymerase and ANP32 proteins to input vRNP, as well as to NP.

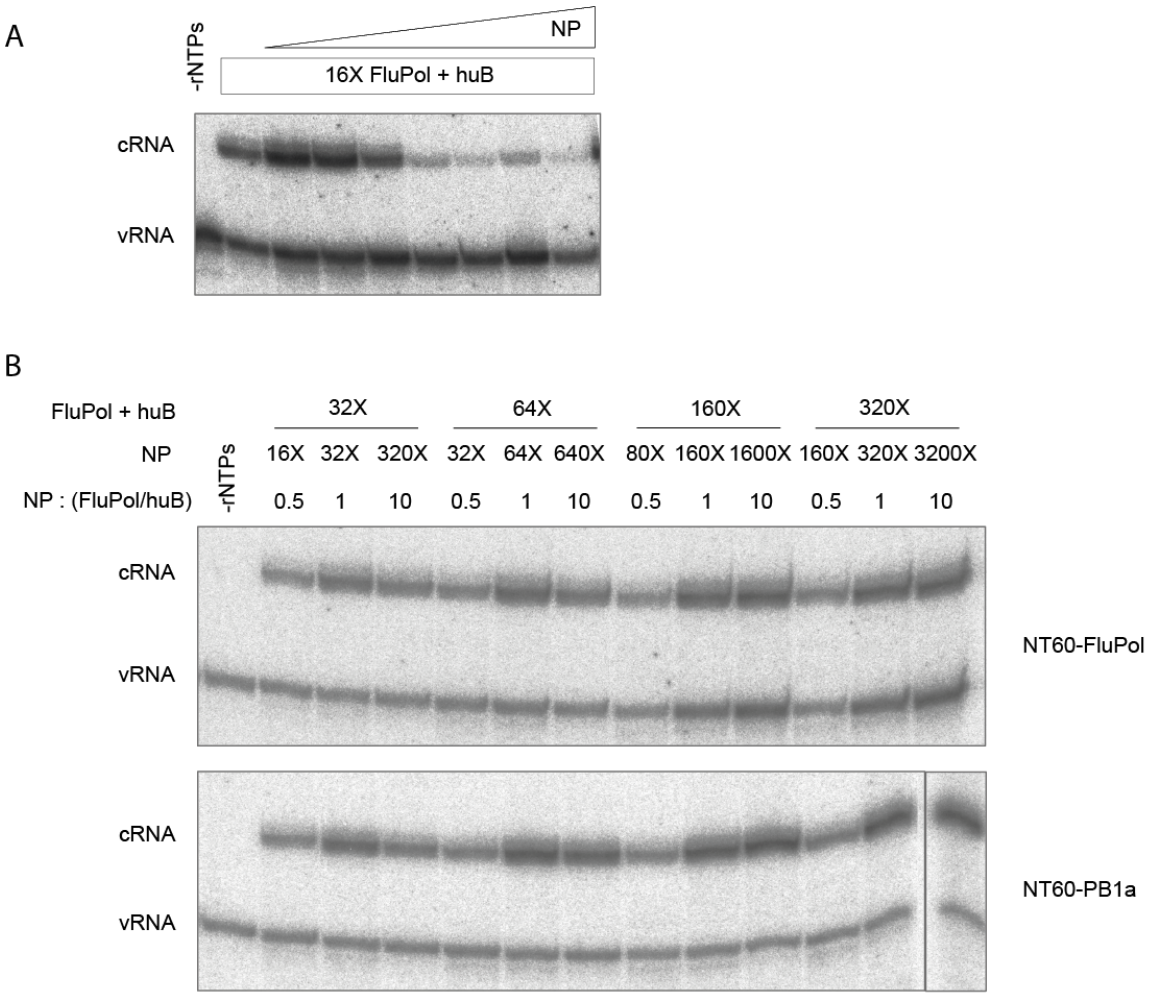


Figure 3.5 Relative amount of viral polymerase and ANP32 protein affect the replication promoting effect of NP. *A. Effect of increasing amount of NP on viral genome replication in the presence of a constant level of viral polymerase (FluPol) and human ANP32B (huB), in unprimed conditions. The levels of FluPol and huB are both kept at 32X that of input vRNP. B. Effect of increasing amount of wild type (NT60-FluPol) or inactive (NT60-PB1a) viral polymerase and huB on the replication promotion ability of NP, in unprimed conditions. For A-B, a representative result of RNA analysis is shown.*

3.2.3 Further characterisation of the RNA products from the in vitro assay

In previous sections, I studied mRNA, cRNA and vRNA synthesis by focusing on their major products detected in the primer extension assay. However, I observed many additional cDNA bands which represent products with different sizes. This observation suggests that each of the mRNA, cRNA or vRNA species might be heterogenous. In this section, I conducted further primer extension analysis on these non-major RNA products and investigated how different viral and host factors affect their production.

To begin with, I examined the input virion-derived vRNA using primer NA 1280 (expected size of cDNA from vRNA = 128 nt, Figure 3.6A) in a primer extension assay and resolved the products on a 6% polyacrylamide gel. Besides a strong cDNA signal at the expected size, I observed a series of discrete extra cDNA bands with faster mobility than the cDNA band produced from a full-length vRNA. These shorter cDNA were likely the reverse transcription products of vRNA with truncated 5' terminus. I also showed that, as I increase the input vRNP level, the intensity of these extra cDNA bands increases proportional to the intensity of the expected cDNA signal (Figure 3.6B), indicating that these extra bands are specific to input vRNA. To rule out the possibility that these extra bands were generated due to an intrinsic early-termination problem of the reverse transcriptase used in the primer extension assay, I compared the pattern of the extra bands between virion-derived vRNA (set 1) and vRNA derived from vRNP reconstitution in cells (set 2). I found that some, but not all, of the bands from set 1 have the same mobility as those from set 2 (Figure 3.6B). Therefore, I concluded that at least some of these non-major cDNA bands were specific to the virion-purified vRNA, hence my input vRNA likely contains internal deletions at the 5' terminus.

Next, I investigated the RNA synthesized in unprimed, ApG-primed, or globin mRNA-primed reactions, in the presence and absence of viral polymerase and human ANP32B. To characterize both positive- and negative-stranded RNAs, I carried out a primer extension assay using NA 160 (expected size of cDNA from cRNA = 160 nt, expected size of cDNA from mRNA > 160 nt) and NA 1280 (Figure 3.6A) and resolved the products on a 6% gel. In the tested conditions, I observed the presence or absence of cDNA signal resulted from full length mRNA (>160 nt), cRNA (160 nt), and vRNA (128 nt) consistent to my previous results. Nevertheless, when cRNA was synthesized in the presence of viral polymerase and human ANP32B, I also observed another series of discrete extra cDNA bands (set 3) (Figure 3.6C, lanes 11, 12, 13). Notably, in the presence of ApG, vRNP can produce a significant amount of cRNA without additional proteins; however, RNA products from this reaction do not generate extra cDNA bands with set 3 pattern (Figure 3.6C, lane 2, 5). When viral polymerase is added to the ApG-primed reaction, while the cDNA signal that represents cRNA remains roughly the same, extra cDNA bands that resemble set 3 appear. The pattern of the bands in set 3 is distinct from that of the bands in set 1 or set 2, and the band intensity is significantly higher. To investigate which primer set 3 bands were derived from, I performed primer extension assays with either primer NA 160 or NA 1280. The results show that set 3 bands were products of NA 160 (Figure 3.7D). Additionally, when I boosted transcriptional activity to a similar level as replication activity, I could not observe any extra bands that resemble set 3, indicating that set 3 bands are specific to cRNA (Figure 3.7C). These results indicate that some cRNA products synthesised *in vitro,* in the presence of additional viral polymerase, contain 5' termini with truncation or internal deletion.

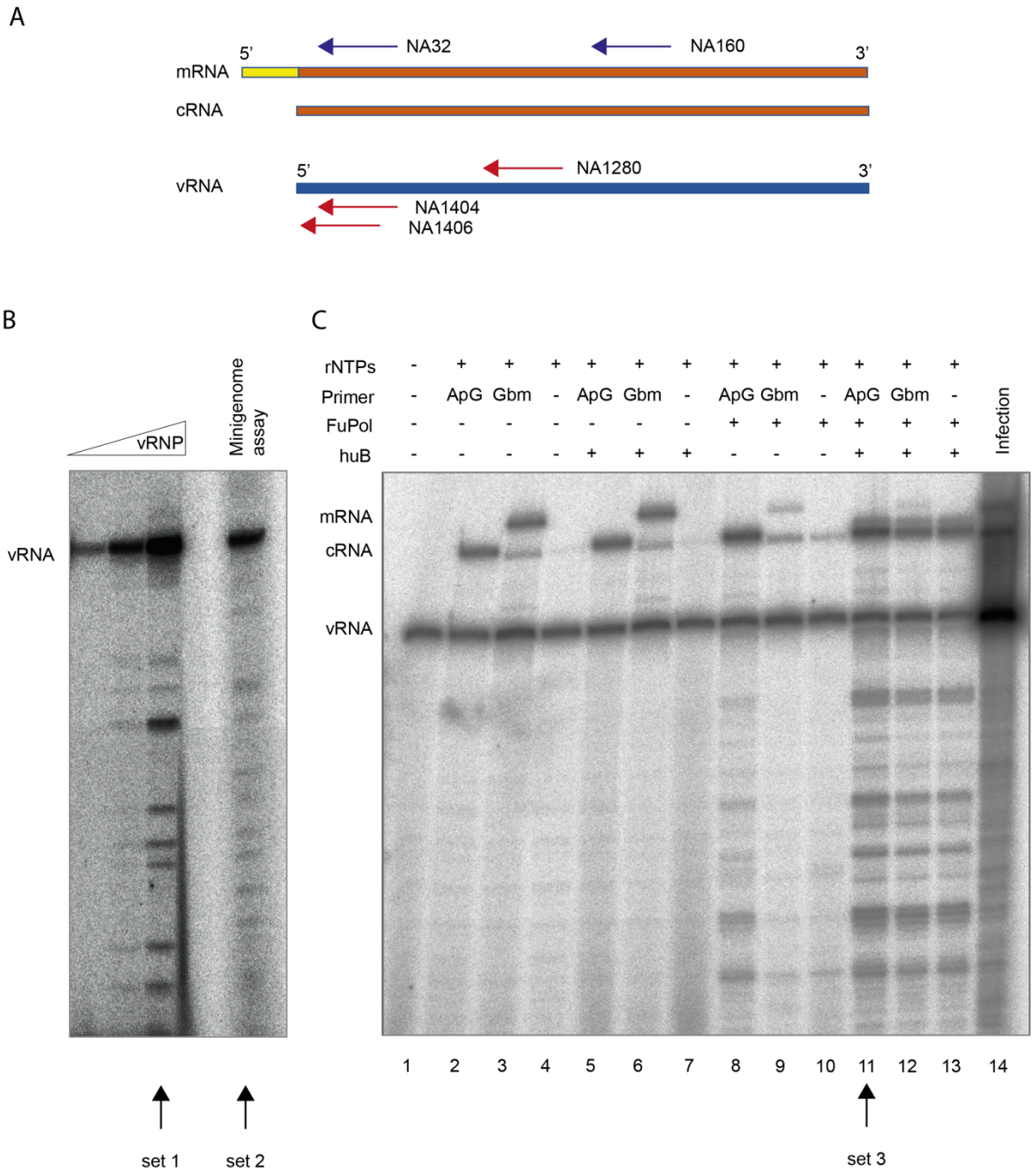
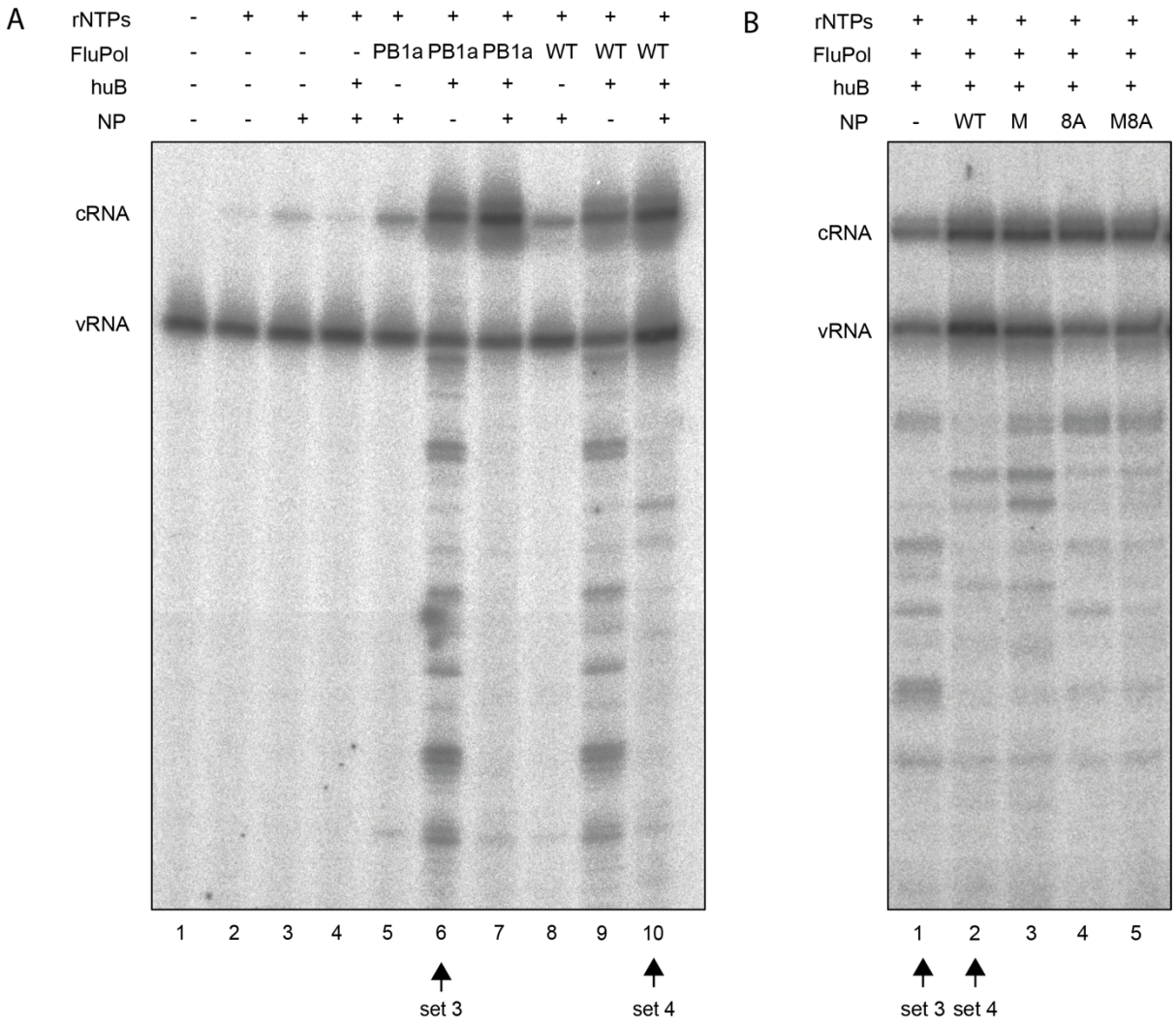


Figure 3.6 Input vRNA and cRNA produced in the presence of additional viral polymerase contains truncated 5' termini. *A. Schematic diagram illustrating the design of primers specific to the positive and the negative sense NA segment for the primer extension assay. B. Primer extension assay using NA 1280 showing the major vRNA band and the minor extra bands from virion-derived vRNP (set 1) and NA-based minigenome assay (set 2). C. Primer extension assay using NA 160 and NA 1280 showing the major mRNA, cRNA and vRNA bands and the extra minor bands from the in vitro vRNP assay in unprimed, ApG primed (ApG) or globin mRNA-primed (Gbm) conditions with or without additional viral polymerase (FluPol) and/or human ANP32B (huB), or from infection cells (Infection). The arrow highlights a lane containing extra bands (set 3) with a distinct pattern.*

Intriguingly, the introduction of NP in the presence of viral polymerase and human ANP32B leads to the elimination of set 3 bands (Figure 3.7A lane 6, 7, 9, 10). More interestingly, if the additional viral polymerase is active, a new pattern of extra bands (set 4) appears upon the addition of NP, coinciding with the disappearance of set 3 bands (Figure 3.7A lane 9, 10). To investigate which primer these set 4 bands were derived from, I performed the primer extension assay with either NA 160 or NA 1280. My results show that bands from set 4 were produced by a combination of NA 1280 and NA 160 (Figure 3.7E). The bands from set 4 derived from NA 160 (set 4 (NA 160)) share a somewhat similar pattern to set 3 bands, hence are likely the residual signal from set 3 bands. On the other hand, the bands from set 4 derived from NA 1280 (set 4 (NA 1280)) have a brand-new pattern. They are likely produced from *de novo* synthesised vRNA from the second step of replication. This suggests that some vRNA synthesised upon the addition of NP also contain truncated 5' terminus. I then investigated the effect of different NP mutants on synthesizing cRNA and vRNA with truncated 5' terminus through analysing the set 3 and set 4 (NA 1280) bands. My results showed that NP8A and mNP8A failed to completely eliminate set 3 bands and do not support the production of set 4 (NA 1280) bands (Figure 3.7B lane 4, 5). On the other hand, mNP struggled to eliminate set 3 bands but strongly induced production of set 4 (NA 1280) bands (Figure 3.7B lane3). Therefore, I conclude that NP can help to produce cRNA with an intact 5' terminus, and that both RNA-binding ability and oligomerisation ability of NP are important for this process. However, vRNA synthesised in the presence of NP contains truncations close to the 5' terminus, especially when NP oligomerisation is impaired. Additional mechanisms might be required to preserve vRNA integrity in infection.


A
rNTPs - + + + + + + + + +
FluPol - - - - PB1a PB1a PB1a WT WT WT
huB - - - + - + + - + +
NP - - + + + - + + - +
cRNA
vRNA
1 2 3 4 5 6 7 8 9 10
set 3
set 4
B
rNTPs + + + + +
FluPol + + + + +
huB + + + + +
NP - WT M 8A M8A
cRNA
vRNA
1 2 3 4 5
set 3 set 4

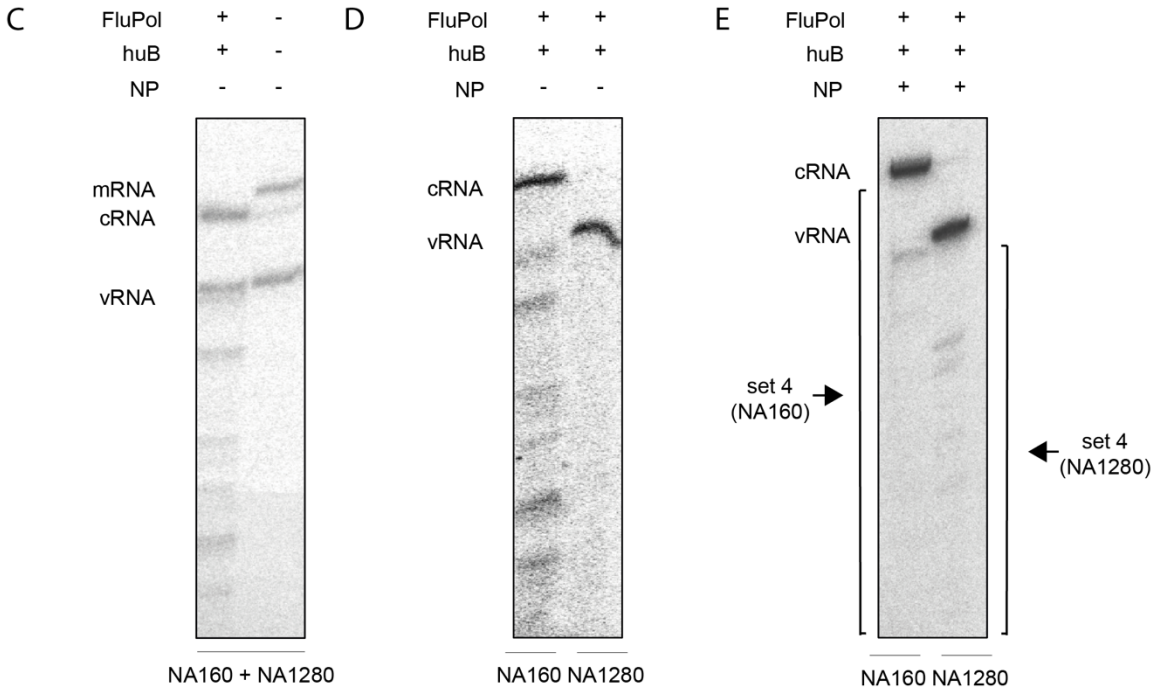

C
FluPol + -
huB + -
NP - -
mRNA
cRNA
vRNA
NA160 + NA1280
D
FluPol + +
huB + +
NP - -
cRNA
vRNA
NA160 NA1280
E
FluPol + +
huB + +
NP + +
cRNA
vRNA
set 4
(NA160)
set 4
(NA1280)
NA160 NA1280

Figure 3.7 The addition of viral and host factors results in the synthesis of aberrant replication products. *A. Effect of additional wild type (WT) or inactive (PB1a) NT60 viral polymerase (FluPol), human ANP32B (huB) and NP on the production of extra bands. B. Effect of additional WT, monomeric mutant (M), RNA-binding deficient mutant (8A) and RNA-binding deficient monomeric mutant (M8A) NP to a reaction with viral polymerase (FluPol) and human ANP32B (huB), on the production of extra bands. For A-B, arrows highlight lanes containing extra bands (set 3 and set 4) with a distinct pattern. C-E. Primer extension assay with either NA 160 or NA 1280, or both, on the RNA sample from globin mRNA-primed reactions, in the presence or absence of viral and host factors. Arrows highlight the distinct patterns of extra bands from set 4 that are specific to NA 160 or NA 1280.*

Notably, in both primed and unprimed conditions, when viral polymerase and human ANP32B are added, I observed an additional cDNA band just above the cDNA band derived from full length cRNA (Figure 3.6C lane 11-13). Moreover, the cDNA band at around 160 nt (the expected size of cDNA from cRNA) produced in mRNA-primed conditions in the absence of added proteins seems to have a slightly different mobility from that produced in unprimed (or ApG primed) conditions, or that produced when replication-promoting proteins are added. However, it is difficult to distinguish cDNA products that differ in length by only a few nucleotides on a 6% gel, which I have been using up until now. Therefore, I carried out primer extension assays on different RNA samples using 20% gels which can resolve cDNA at a higher resolution. Higher percentage gels are optimised for the analysis of smaller cDNA fragments. Therefore, I designed primers closer to the 5' termini of the positive- or the negative-sense influenza NA segment: NA 32 (expected size of cDNA from cRNA = 32 nt), NA 1406 (expected size of cDNA from vRNA = 22 nt) and NA 1404 (expected size of cDNA from vRNA = 25 nt). I used these primers on the RNA samples obtained from infected cell debris, minigenome assay, unprimed or primed *in vitro* vRNP assay in the absence or presence of host and viral factors, and resolved the cDNA products on a 20% gel alongside partially degraded radioactive primers as markers. Using the positive-sense primer NA 32 I showed that,

the RNA sample from infected cells produces a smear of cDNA that represent mRNA with varied lengths and a discrete cDNA signal at 32 nt that represent full length cRNA (Figure 3.8A lane 1). The different lengths of mRNA are likely the result of viral polymerase differentially cleaving capped host mRNA and generating transcription primers at different lengths, or capped-primer priming at different positions on the template. A cDNA band 2 nt shorter than the major cDNA band, at 30 nt, was also observed in the infected sample (Figure 3.8A lane 1). In contrast, the RNA sample from the globin mRNA-primed reaction produces a more discrete cDNA band that is derived from mRNA, as there is only one species of capped primer donor, and it contains one preferential cleavage site (Figure 3.8A lane 6). In addition, cDNA of the RNA sample from the mRNA-primed reaction contains a quadruplet of bands around the expected cDNA size of a full length cRNA, between 30 and 33 nt. The major band at 30 nt possibly represents cleaved viral mRNA by excess viral polymerase with endonuclease activity, because it is also observed in the cDNA of the infected RNA sample - the only other condition where viral mRNA is present. A cDNA doublet with similar signal intensity, at 33 nt and 32 nt, is also observed in unprimed conditions, in presence and absence of viral polymerase, and in ApG-primed condition (Figure 3.8A lane 2-7). This cDNA doublet are potentially derived from the correct cRNA product (cRNA 5'+0) and a cRNA product with an additional nucleotide added to the 5' terminus (cRNA 5'+1). The addition of viral polymerase, human ANP32B and NP produces cDNA with additional bands between 34 nt and 38 nt (Figure 3.8A lane 3-4). Consistent with my primer extension data on cRNA, my results from using the negative-sense primer NA 1406 and NA 1404 showed that, although most of the vRNA formed in the presence of viral polymerase, ANP32B, and NP have an intact 5' terminus (vRNA 5'+0), some vRNA products also have 1-3 additional nucleotides at the 5' termini (vRNA 5' +1, +2, or +3), with a decreasing abundancy (Figure 3.8B lane 4 and 8). I also revealed that both input vRNA and the RNA sample from minigenome assay contain vRNA with intact 5' terminus

(vRNA 5' +0) vRNA, as well as vRNA with one additional nucleotide (vRNA 5' +1) (Figure 3.8B lane 2-3, 6-7). Together, my results suggest that although full-length cRNA and vRNA are likely to have formed in my *in vitro* vRNP assay, RNA products containing truncated or elongated termini are also present.

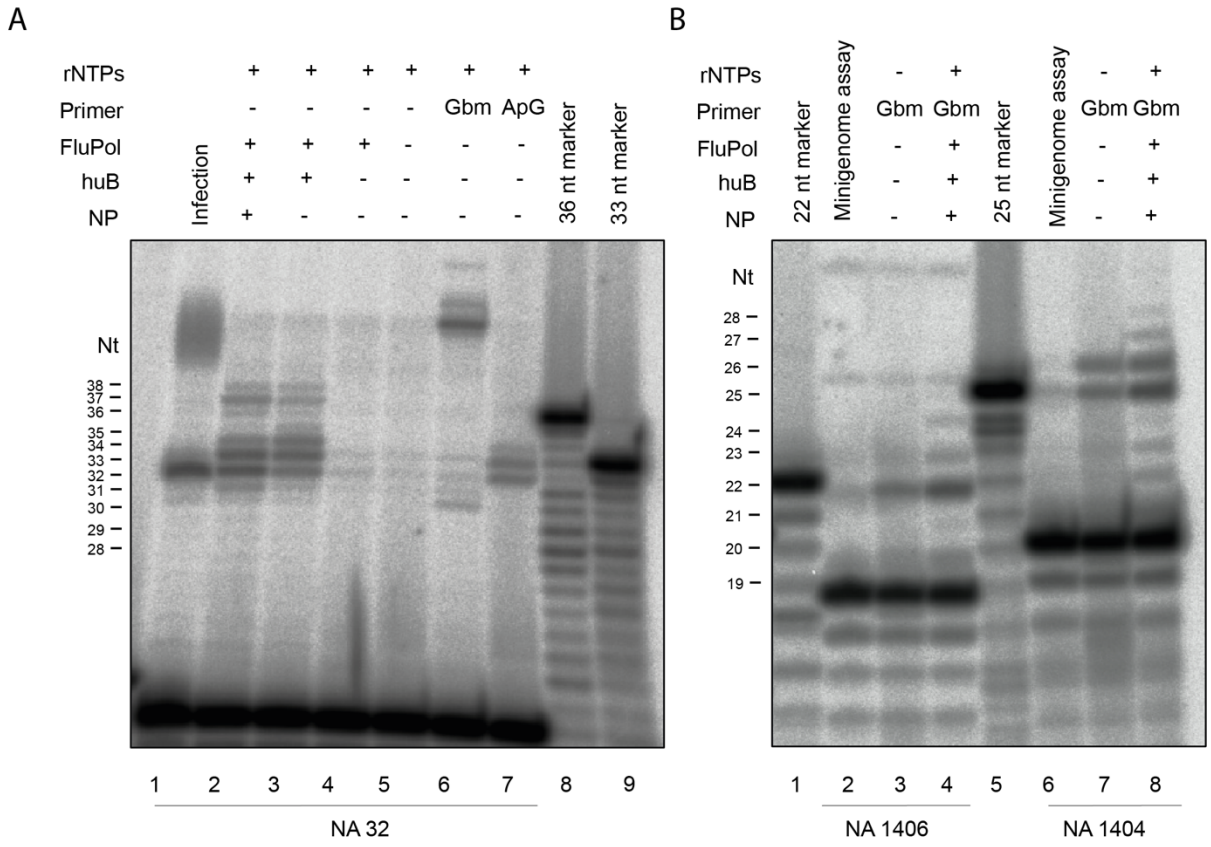


Figure 3.8 Viral RNA in the in vitro vRNP replication reconstitution assay contains heterogeneous 5' termini*. A. Primer extension assay with NA 32 primer (expected product size = 32 nt) on the RNA sample from infected cell lysate (Infection), in vitro vRNP assay with or without ApG or Globin mRNA primer (Gbm), in the presence or absence of viral polymerase (FluPol), human ANP32B (huB) and NP. Partially degraded radioactive primers, with lengths 36 nt or 33 nt, are used as size markers. B. Primer extension assay with NA 1406 (expected product size = 22 nt) or NA 1404 (expected product size = 25 nt) primer on RNA sample from infected cell lysate, in vitro vRNP assay with or without ApG or Gbm primer, in the presence or absence of FluPol, huB and NP. Degraded radioactive primers, with lengths at 22 nt or 25 nt, are used as size markers.*

3.3 DISCUSSION

In this chapter, I successfully reconstituted a full cycle of influenza viral genome replication, a major step in the viral life cycle, inside a test tube. This result enables me to define the minimal protein requirement for a complete cycle of vRNP replication as viral polymerase, ANP32 protein, and NP. This piece of work also presents the first successfully assembled functional RNP from recombinant proteins, as well as a truly cell-free vRNP replication system that faithfully replicates both steps of replication.

As discussed in Chapter 2, the addition of viral polymerase and ANP32 proteins to virion-isolated vRNPs allows the formation of a replication complex and the switch from transcription to replication in vRNPs. Here, the system was taken further by adding recombinant NP. The addition of NP, along with viral polymerase and ANP32 proteins, enables product cRNA to act as template for subsequent vRNA synthesis, most likely through assembly of functional cRNP which is capable of replicating. In a similar way to vRNP replicational activation, this functional cRNP is possibly activated for replication by the presence of additional viral polymerase and ANP32 protein. The role of NP supported by these data is consistent with previous studies in cellular systems, where the accumulation of NP is necessary for efficient replication of a full-length genome (Turrell et al., 2013; Wang et al., 2022). The combined results from Chapters 2 and Chapter 3 hence lead to the proposal of a minimal model for completing the viral genome replication cycle, in which non-vRNP-associated polymerase, ANP32 proteins, and NP are key components (Figure 3.9).

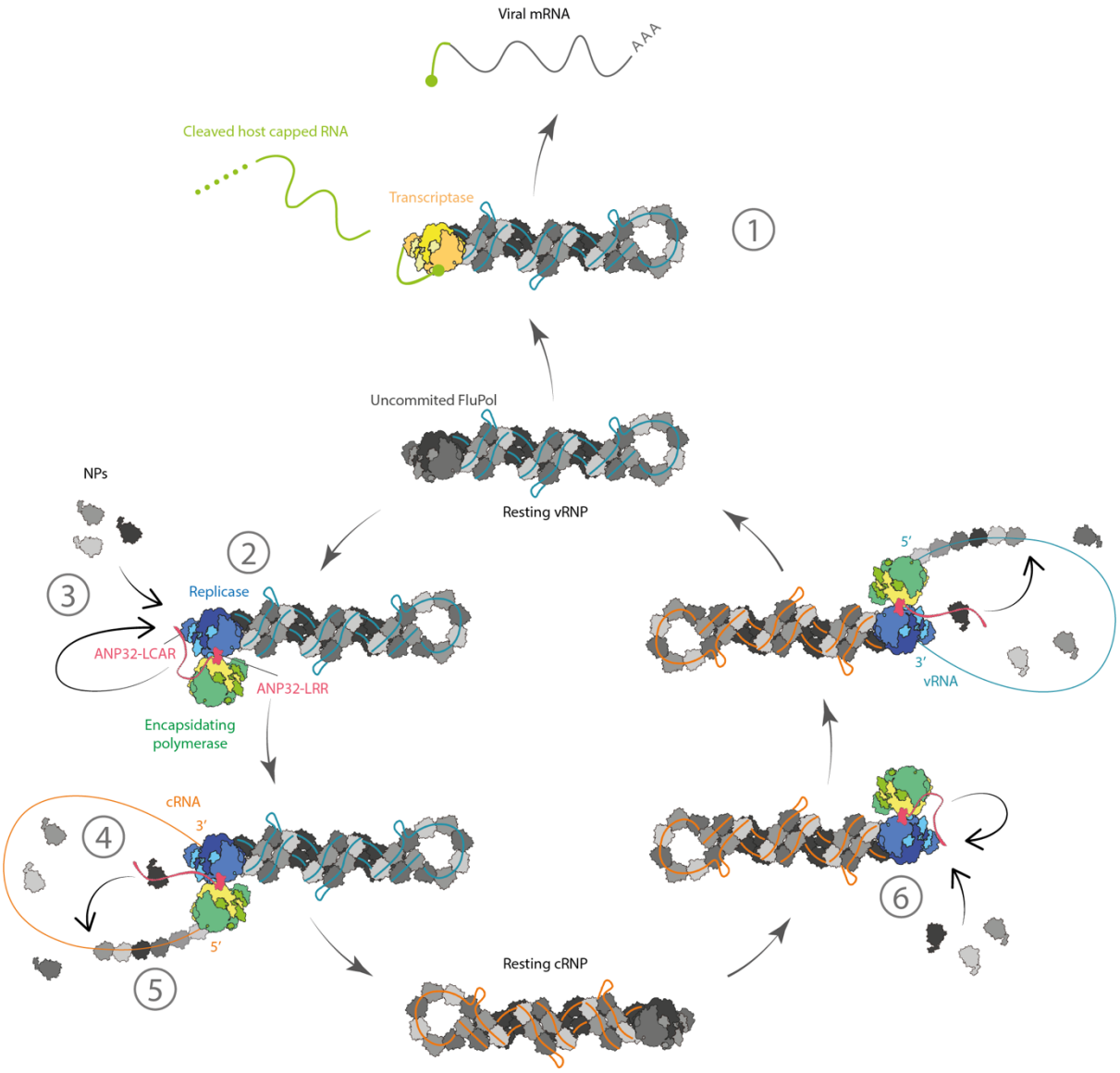


Figure 3.9 Model summarising the modulation of vRNP activity by viral and host factors in the in vitro vRNP replication reconstitution assay. *Virion-derived vRNPs preferentially synthesise mRNA in the presence of a capped RNA primer donor with the vRNP-resident polymerase acting as transcriptase (1). Additional viral polymerase and ANP32 bind to the vRNP-resident polymerase and switch the vRNP-resident polymerase by allosteric modification from a transcriptase to a replicase that initiates cRNA synthesis in a primer-independent manner (2). The ANP32^{LRR} stabilises the polymerase dimer while the ANP32^{LCAR} plays a role in promoting genome replication through a yet to be discovered mechanism. NP further promotes replication in an RNA-binding independent manner (3). The 5' end of cRNA product is captured by the encapsidating polymerase and the ANP32^{LCAR} promotes the co-replicational assembly of cRNA into cRNP by binding to NP and increasing its local density (4). The RNA-binding ability of NP is required for the assembly of functional cRNP that can carry out the second step of replication (5). Additional polymerase and ANP32 bind to the resident polymerase of newly*

assembled cRNP, converting the uncommitted polymerase to a replicase to perform primer-independent initiation of vRNA synthesis (6). The synthesis of vRNA and its assembly into vRNP proceed in a manner similar to that described for cRNA above.

For many years, the role of NP in regulating vRNP function has remained a topic of debate. Early studies suggested that NP is the critical factor responsible for switching vRNP from transcription to replication (Beaton & Krug, 1986; Shapiro & Krug, 1988). However, later studies indicated that the primary function of NP is to stabilize replication products in cells, by binding to and protecting cRNA and vRNA, and has no direct role in regulating replicational activity of RNP (Deng et al., 2006; Turrell et al., 2013; Vreede et al., 2011; Wang et al., 2022). Here, I present compelling evidence that NP plays multiple roles in replication, both structural and regulatory. My data suggest that, structurally, NP helps to form a functional RNP that can copy its resident RNA, as well as protect newly synthesised cRNA from degradation in a cellular environment where RNAses are present. These roles rely on the ability of NP to bind to RNA. Excitingly, my data also unveiled a previously unknown regulatory role of NP in directly promoting replication, which is independent of either the RNA binding or the oligomerisation ability of NP. To isolate the role of NP in each of the two distinct steps of viral replication, I took advantage of an inactive viral polymerase, which could not support further replication upon RNP assembly. In this way, I demonstrated that NP plays a direct regulatory role in RNP functions and acts an additional cue that promotes replication. It is highly likely that NP further promotes replication through interactions with the replication platform, however, the exact molecular mechanism remains to be revealed. A recent low resolution cryo-EM density reconstruction from our lab (unpublished data from Dr Haitian Fan) reveals electron density resembling a molecule of NP that is attached to PA^{endo} of a viral polymerase in a replicase conformation. Particularly, this NP-like structure appears to stabilize the

interaction between PA^{endo} and $PB2^{NLS}$ in the replicase conformation. This structure sheds light on a potential mechanism underlying the direct regulatory role of NP: stabilising the replicase conformation of the RNP-associated polymerase. This use of NP as a replication promoting factor is particularly advantageous because NP is abundant, and its association to the replication platform could create a reservoir of NP for recruitment by ANP32 proteins during subsequential RNP assembly.

In addition, I demonstrated that NP, like free viral polymerase, inhibits transcription and slightly promotes replication in vRNP without other factors. Like with free viral polymerase, it is possible that NP acts allosterically by altering the conformation of the RNP-associated polymerase; however, it is also possible that NP binds to and depletes capped-primer donor mRNA, thereby inhibiting transcription. Surprisingly, I also showed that excess NP can inhibit replication. Through further investigating this observation, I conclude that the exact role of NP depends on the relative concentration of viral polymerase and ANP32 proteins to NP, and to input vRNP. This versatile role of NP in regulating RNP functions can be particularly useful to fine-tune the activity of RNPs over the course of infection, when the requirement for vRNP activity changes. Together, the dependency of vRNP-associated viral polymerase on the presence of free viral polymerase and NP, as well as host ANP32 family proteins, provides the influenza virus with an elegant mechanism to temporally regulate the process of genome replication: Primarily, viral genome replication only occurs when a sufficient level of viral polymerase and NP is reached. This can prevent the production of naked cRNA and vRNA, which can activate the innate immune response via the RIG-I pathway (Pichlmair et al., 2006). Additionally, when excess NP is made, viral replication is down regulated in time to proceed onto virion assembly.

It is worth noting that the most severe cases of influenza virus infection are due to over activation of innate immunity, and defective viral RNAs are primary inducers of innate immunity (Te Velthuis et al., 2018). The *in vitro* vRNP replication system developed in this chapter is well-suited for assessing the fidelity of the viral polymerase, namely its potential to generate defective viral RNA, in different conditions. Although the primer extension assay has limited ability to analyse RNA products due to its lack of coverage over the region upstream of the primer, I manage to show that cRNA with intact 3' and 5' termini must have been synthesised in my system in the presence of viral polymerase, ANP32 protein, and NP because some vRNA with an intact 5' end has been produced, which relies on the cRNA template having an intact 3' terminus. Nevertheless, whether full-length cRNA is the major product remains to be determined. My further analysis of the RNA products showed that some cRNA produced in my system is truncated close to the 5' terminus, and that these truncations result from the presence of viral polymerase. A possible explanation is that the endonuclease activity of excess additional viral polymerase, or remaining RNAse A from the purification process, cleaves product cRNA at specific sites, therefore produces truncated cRNA of discrete lengths. This is further supported by the fact that the addition of NP, which is capable of binding to RNA, decreases the amount of truncation, presumably because it protects cRNA from the polymerase endonuclease activity. Additionally, I observed extra nucleotides at the 5' end in product cRNA and vRNA, which is consistent with observations from some previous viral polymerase functional studies (Barr & Fearns, 2010; Deng et al., 2006; Zhang et al., 2010). Since viral genome segment termini are particularly vulnerable to degradation, this process could be a possible mechanism to protect and restore damaged viral genome termini. Taken together, my *in vitro* system has been shown to be a valuable tool for evaluating the role of different viral and host factors in affecting the production of defective viral RNA, which is important in innate immune activation and, ultimately, in viral virulence and pathogenesis.

Co-infection of influenza viruses has been a captivating research topic due to its impact on disease severity and its contribution to the emergence of new viral strains through genetic reassortment (Shao et al., 2017). In this chapter, I showed that the viral polymerases from a range of influenza A viruses can promote replication in A/WSN/33 (H1N1), raising a question whether this could be important in elevating disease severity during co-infection and promoting reassortment.

While there are numerous advantages to studying influenza virus vRNP functions using a simplified *in vitro* system, it is crucial to recognise the inherent differences between such an *in vitro* system and the actual infection conditions. RNA synthesis by vRNP happens in a drastically different environment during cellular infection compared to our *in vitro* reconstituted system. Firstly, the relative and absolute concentrations of vRNP, as well as host and viral factors, differ significantly in the *in vitro* system and in an infected cell. Specifically, our *in vitro* system involves adding a substantial excess of cap donor mRNA, as well as viral and host factors, to a concentrated purified vRNP sample, which does not accurately reflect in vivo infected conditions. In an infected cell, host capped mRNA available for cap-snatching, and host ANP32 proteins are likely to at much lower concentrations, and viral polymerase and NP are only synthesized later during infection. Furthermore, an infected cell contains numerous other host and viral factors that directly or indirectly influence vRNP functions or regulate transcription and replication. For example, additional stimulatory and inhibitory factors, such as viral NEP, likely contribute to the regulation of vRNP functions in infection. Enzymes that degrade viral and host proteins as well as unprotected viral RNAs are also present. Although wild type viral polymerase and NP readily oligomerize in solution, various mechanisms involving different binding partners are proposed to maintain them in a monomeric state within cells (Kawaguchi et al., 2011; Mondal et al., 2015; Turrell et al., 2015). Compensatory

mechanisms for pathways crucial for vRNP functions are also likely to exist in an infected cell to ensure successful viral replication. These mechanisms possibly use factors not typically engaged in regulating vRNP functions and could explain some discrepancies observed in my *in vitro* system versus cellular systems. For instance, the ANP32^{LCAR} truncation mutant still supports NP-independent replication in cells but is unable to support replication in the absence of NP *in vitro*. Cellular processes could also indirectly affect vRNP functions by competing with the vRNP for substrates such as nucleotides. Electron tomography has revealed that the interior of a cell is densely packed with dissolved solutes including ions, metabolites, proteins, and RNA, with an estimated average of 30-40% of cellular volume occupied by proteins and RNA (Medalia et al., 2002). Therefore, it is not impossible that the processes of transcription and replication, and their regulation, is different in a crowded cellular environment. Although attempts to recreate this crowded environment by adding PEG8000 to our *in vitro* system have not shown any effect on vRNP functions (data not shown), we have no sufficient evidence to exclude this possibility. A cell also has various functional compartments created by either membranes or phase separation (Diekmann, Y., & Pereira-Leal, 2013; Wheeler & Hyman, 2018). The precise location of transcription and replication inside the cell nucleus remains unknown. It is conceivable that, unlike in our *in vitro* uncompartmentalized system, vRNP transcribes vRNA in proximity to host Pol II to produce capped mRNA but replicates at a different location in the nucleus. Another major consideration is the difference between native and recombinant proteins. In this study, we did not account for the PTMs of viral and host factors; however, modifications such as phosphorylation, acetylation, and ubiquitination are known to affect various aspects of viral replication in cells (Chenavas et al., 2013; Coon, et al., 2020; Giese et al., 2017; Günl et al., 2023; Hutchinson et al., 2012; Li et al., 2018; Mondal et al., 2015; Turrell et al., 2015). Additionally, the host Pol II CTD mimic protein used in our study contains at most 26 heptad repeats of the C-terminal domain. A complete Pol II complex,

with 52 heptad repeats and many more subunits, is likely to affect vRNP function differently in infection. Our *in vitro* assay reconstituted a simple system involving only vRNPs from one influenza virus type (WSN virus); however, during co-infection and superinfection, where vRNP and viral factor from different types of viruses are present simultaneously, further complications could arise. Although the simplified *in vitro* system allows the detection of molecular events that might be masked in a complex in vivo environment, caution is warranted when inferring from these observations. For instance, we observed cleaved viral mRNA product likely resulting from the endonuclease activity of vRNP-associated viral polymerase, and aberrant RNA products from replication. However, whether these products are generated during infection and whether they reach a biologically significant level remain unclear. Finally, a cellular system is dynamic and the protein and RNA levels are constantly changing. This variance from our *in vitro* system has several implications. Firstly, reaction kinetics in a cellular environment differ significantly from those in a test tube because products such as mRNA and vRNP are removed from the system when they are exported out of the nucleus. Secondly, the levels of host and viral factors involved in vRNP function can not only decrease but also increase because *de novo* protein production occurs in the cell until host shutdown. In contrast, in the *in vitro* system, no additional factor is added once the reaction starts, hence the concentrations of viral polymerase, ANP32 proteins, and NP can only decrease over time. Lastly, since viral mRNA produces viral proteins such as viral polymerase and NP during infection, a feedback loop is present in vivo; however, in our *in vitro* system, mRNA levels and viral protein concentration are not linked.

In summary, my work marks the first successful reconstitution of a complete influenza virus genome replication cycle in a truly *in vitro* experiment. This robust system offers a highly controlled and simplified system for decoding the molecular details of influenza virus

replication and dissecting the roles of each regulatory factor. Using this novel *in vitro* system, we have determined the minimal protein requirements for influenza viral genome replication, and uncovered new and intriguing roles for these protein factors. Taken together, my Chapter 2 and Chapter 3 allow me to present a more complete model for the molecular mechanism underlying influenza viral genome replication (Figure 3.9). The outcome of the study also presents exciting opportunities to further investigate influenza viral genome replication. In the context of the aim for my thesis, this assay has given rise to the possibility of using cryo-electron microscopy techniques to image isolated vRNPs during transcription and replication. Regarding the broad interest of the field, this approach of studying influenza viral genome replication holds the potential to be extended for and applied to other negative-sense RNA viruses.

"What I cannot create, I do not understand' – by Richard Feynman.

Chapter 4: Inhibition of 1918 influenza virus polymerase using nanobodies

Most data from this chapter were published in:

Keown JR, **Zhu Z***, Carrique L*, Fan H*, Walker AP, Serna Martin I, Pardon E, Steyaert J, Fodor E, & Grimes JM (2022) Mapping inhibitory sites on the RNA polymerase of the 1918 pandemic influenza virus using nanobodies. Nature Communications, 13(1), 251.*

** Joint authorship*

4.1 INTRODUCTION

Influenza viruses are causal agents of seasonal epidemics and pandemics, imposing significant burdens on the global healthcare system (Carrat et al., 2008). However, currently available influenza vaccines need to be updated annually and only offer partial protection. When prevention fails, and severe symptoms of influenza virus infection presents, antivirals targeting different stages of a viral life cycle are the next line of defence. Current antivirals include drugs against viral polymerase, which reduce RNA synthesis, thereby inhibiting viral replication (Liu et al., 2021). The viral polymerase is a promising candidate for antiviral targets because of its indispensable function during the viral life cycle and its relatively conserved amino acid

sequence across different viral subtypes. However, the emergence of drug-resistant strains poses a challenge to the effectiveness of current antivirals. This arms race has led to the need to develop new therapeutic approaches that inhibit key targets, such as the viral polymerase, in novel ways.

Nanobodies, derived from the variable domain of single-chain camelid antibodies, are recently popularised tools for protein structural and functional studies. They have gained significant attention in antiviral research due to their unique properties, including small size, high stability and specificity, and strong binding-affinity to antigens. For instance, nanobodies have shown promise in therapeutic development against SARS-CoV-2 and norovirus, by inhibiting the interaction between SARS-CoV-2 spike protein and the human cell surface receptor angiotensin-converting enzyme 2 (ACE2), or binding to and disassembling intact norovirus viral particles (Koromyslova & Hansman, 2015; Pymm et al., 2021). Additionally, recent studies on nanobodies against influenza viral protein components have demonstrated the use of nanobodies to study protein function and to identify antiviral targets (Fan et al., 2019; Hanke et al., 2016; Schmidt et al., 2016).

Presently, all circulating influenza A viruses in humans are believed to have evolved from an ancestor virus, A/Brevig Mission/1/1918 (H1N1), which triggered the worldwide 1918 influenza pandemic, and resulted in between 50 to 100 million deaths globally. While the reason behind the exceptional pathogenicity of this virus remains uncertain, various viral factors, including the viral surface proteins and the viral polymerase, have been reported to contribute to its high lethality. Prior to my study, single-chain antibodies against the apo or the promoter-bound recombinant viral polymerase from A/Fujian/01/2002 (H5N1) were raised in llamas at the VIB research institute by the Steyaert lab. After affinity-selection by phage-

display and verification via enzyme-linked immunosorbent assays, 24 nanobodies were generated using standard methods (Figure 4.1A) (Pardon et al., 2014). Thereafter, Dr Jeremy Keown revealed the polymerase binding sites of these nanobodies on the 1918 viral polymerase using cryo-EM and crystallography, and showed that these binding sites are relatively conserved across influenza A viral subtypes (Figure 4.1B) (Keown et al., 2022).

In this chapter, I aim to functionally evaluate the inhibitory activity of our nanobodies against the 1918 viral polymerase, and to identify new ways to inhibit polymerase activity.

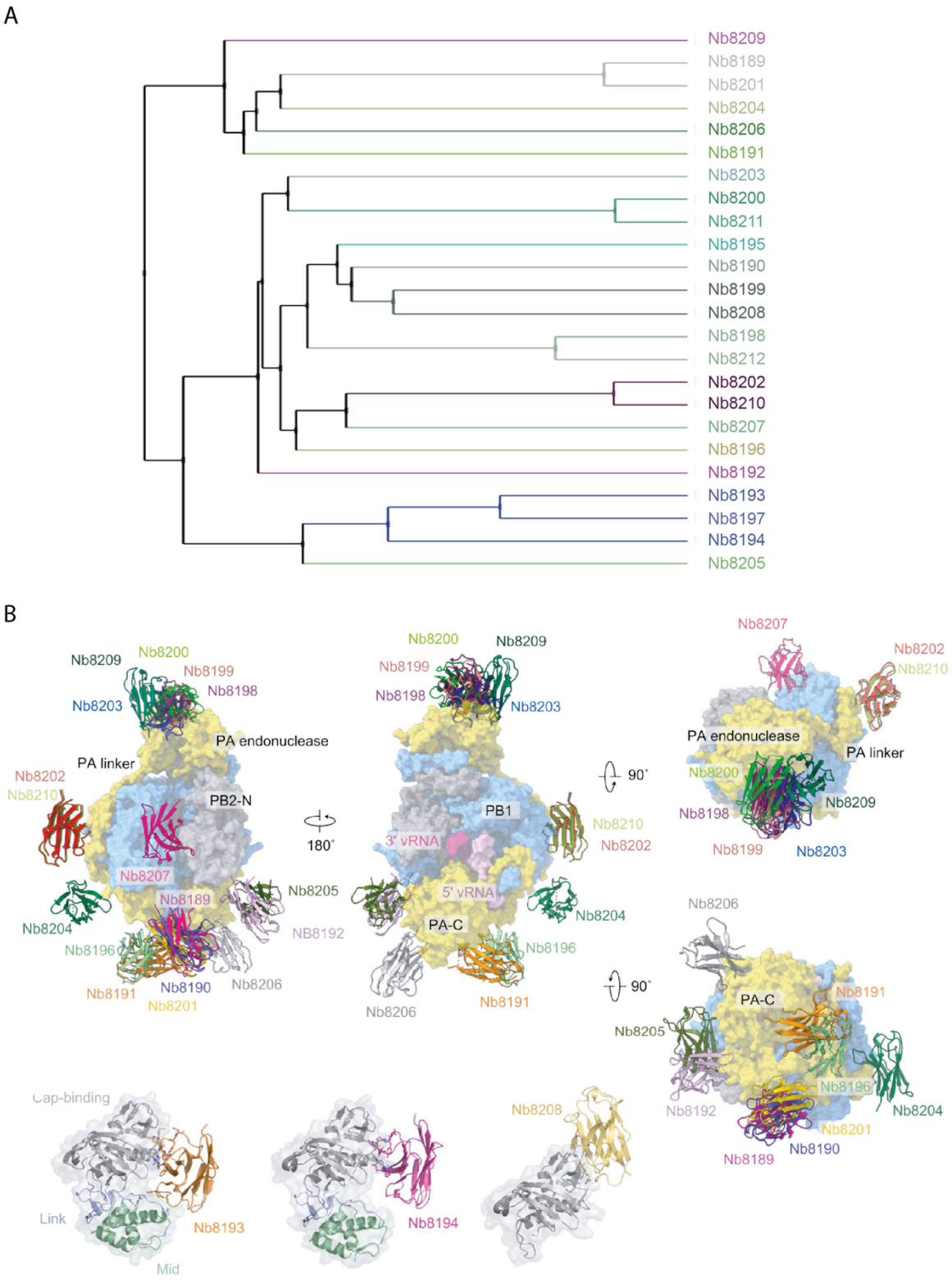


Figure 4.1 Nanobody phylogenetics and binding site information. *A. Phylogenetic tree showing the relationships of the 24 nanobodies, constructed based on protein sequences. B. Structural model showing the binding site of each nanobody on the 1918 viral polymerase. The figure is adapted from Keown, Zhu, Carrique, Fan 2022.*

4.2 RESULTS

4.2.1 The purified 1918 viral polymerase is functionally active

I first examined whether the 1918 viral polymerase, used to generate the nanobody binding map, is active in RNA synthesis. To assess the replicational activity of the purified 1918 polymerase (a gift from Dr Jeremy Keown) (Figure 4.2A), I performed *in vitro* polymerase activity assays with purified 1918 polymerase bound to vRNA or cRNA promoter, in the absence of primers. I supplied the polymerase with rNTPs containing radiolabelled GTP, and showed that purified 1918 polymerase can synthesise cRNA or vRNA products using vRNA or cRNA short-templates, respectively (Figure 4.2B). Addition of ApG dinucleotides to the reactions increased the signals for both cRNA and vRNA products (Figure 4.2C). These results suggest that purified 1918 polymerase is capable of replication, and that the primer-independent initiation is a rate-limiting step for both cRNA and vRNA synthesis. Furthermore, I tested whether purified 1918 polymerase is transcriptionally active when bound to vRNA promoter and supplied with either an 11 nt long capped RNA primer or a 20 nt long cap donor RNA. The 20 nt long capped RNA is cleaved before being used for transcription, hence can be used to assess the cap-snatching activity of viral polymerase. I observed the formation of transcription products and the simultaneous disappearance of capped RNA in both cases (Figure 4.2D). This result indicates that purified 1918 polymerase is also transcriptionally active, and possesses endonuclease activity. Consistent with my observations in Chapter 3, by-products 1 nt longer than the major products are observed in both transcription and replication (Figure 4.2D), suggesting that a nucleotide addition process is commonly seen. Taken together, these results demonstrate that the 1918 polymerase used in the structural study is functionally active.

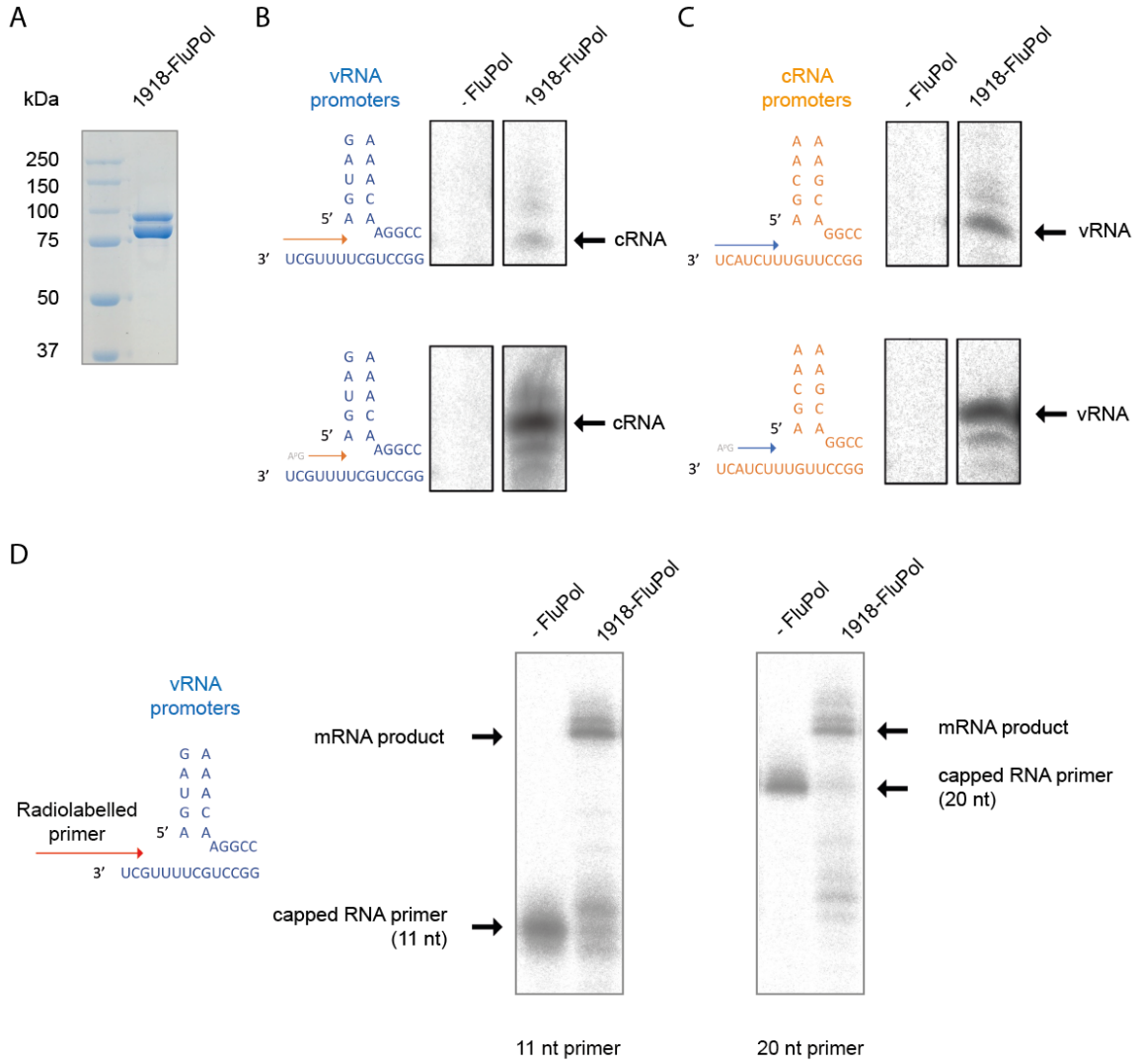


Figure 4.2 Activity analysis of the purified 1918 viral polymerase. *A. SDS-PAGE analysis followed by Coomassie staining of wild type 1918 influenza viral polymerase purified from insect cells. B-D. In vitro polymerase activity assays showing that the purified recombinant 1918 viral polymerase (1918-FluPol) can synthesise cRNA on a short vRNA template (B), vRNA on a short cRNA template (C) in the absence (top) and presence (bottom) of ApG primer, and can synthesise mRNA upon being supplied an 11 nt or a 20 nt capped primer donor. The 20 nt primer donor requires cleavage by the endonuclease domain of the viral polymerase prior to being used in transcription initiation. A representative gel of 3 independent reactions is shown.*

4.2.2 Nanobodies inhibit the 1918 viral polymerase activity in vitro

The inhibitory effects of nanobodies on 1918 polymerase activity were investigated next. I first studied this *in vitro* within the context of purified 1918-RNP. Given that purifying vRNP from virions containing 1918-vRNP would not comply with my local biosafety regulations, I cloned a TAP-tag onto 1918-PB2 and transfected it together with 1918-PA, PB1 and NP to reconstitute 1918-RNP in HEK293T cells. A Pol I promoter-controlled NA gene was transfected simultaneously to produce vRNA for RNP reconstitution. In this way, 1918-RNP with a TAP-tag was produced in cell culture and was purified through affinity purification. (Figure 4.3A). I tested the activity of purified RNP and showed that purified 1918-RNP was capable of producing mRNA and a small amount of cRNA upon the addition of rNTPs in the presence of globin mRNA (which needs to be cleaved by endonuclease activity to be used for transcription) or a 12 nt capped primer as capped-primer donor (Figure 4.3B). Unfortunately, RNP purified from cells contained a mixture of cRNP and vRNP, as indicated by the presence of cRNA and vRNA signals in the input sample (no rNTP condition). However, purified cRNP did not show significant activity, as the vRNA levels remained unchanged upon the addition of rNTPs (Figure 4.3C-D).

Meanwhile, I purified a panel of viral polymerase-binding nanobodies from the periplasm of WK6 *E. coli* using osmotic lysis, followed by Ni-affinity purification (see Figure 4.3E). By adding an equal amount of purified nanobodies to the *in vitro* reaction, I observed that a subgroup of nanobodies, such as Nb8190 and Nb8207, significantly inhibited 1918-RNP activity under both globin mRNA-primed or 12 nt capped primer-primed conditions (Figure 4.3C-D). Intriguingly, I found that nanobodies which either bound at similar locations, or phylogenetically closely-related did not necessarily have similar inhibitory potency. For

example, Nb8189 and Nb8190, both binding to PA-C at the same site, and being closely linked phylogenetically, had drastically different inhibitory effects on mRNA or cRNA synthesis. Additionally, only one of the five nanobodies binding to PA^{endo}, Nb8200, showed a significant inhibitory effect on transcription in the presence of globin mRNA, and only two of the three nanobodies binding to $PB2^{cap}$, Nb8193 and 8194, significantly inhibited transcription in the presence of either capped primer donors (see Figure 4.3C-D). The difference in inhibitory activities could, to a certain degree, be attributed to different binding affinities of each nanobody to the 1918 polymerase. In summary, my results demonstrate that a subset of polymerase nanobodies can inhibit 1918-RNP activity *in vitro.*

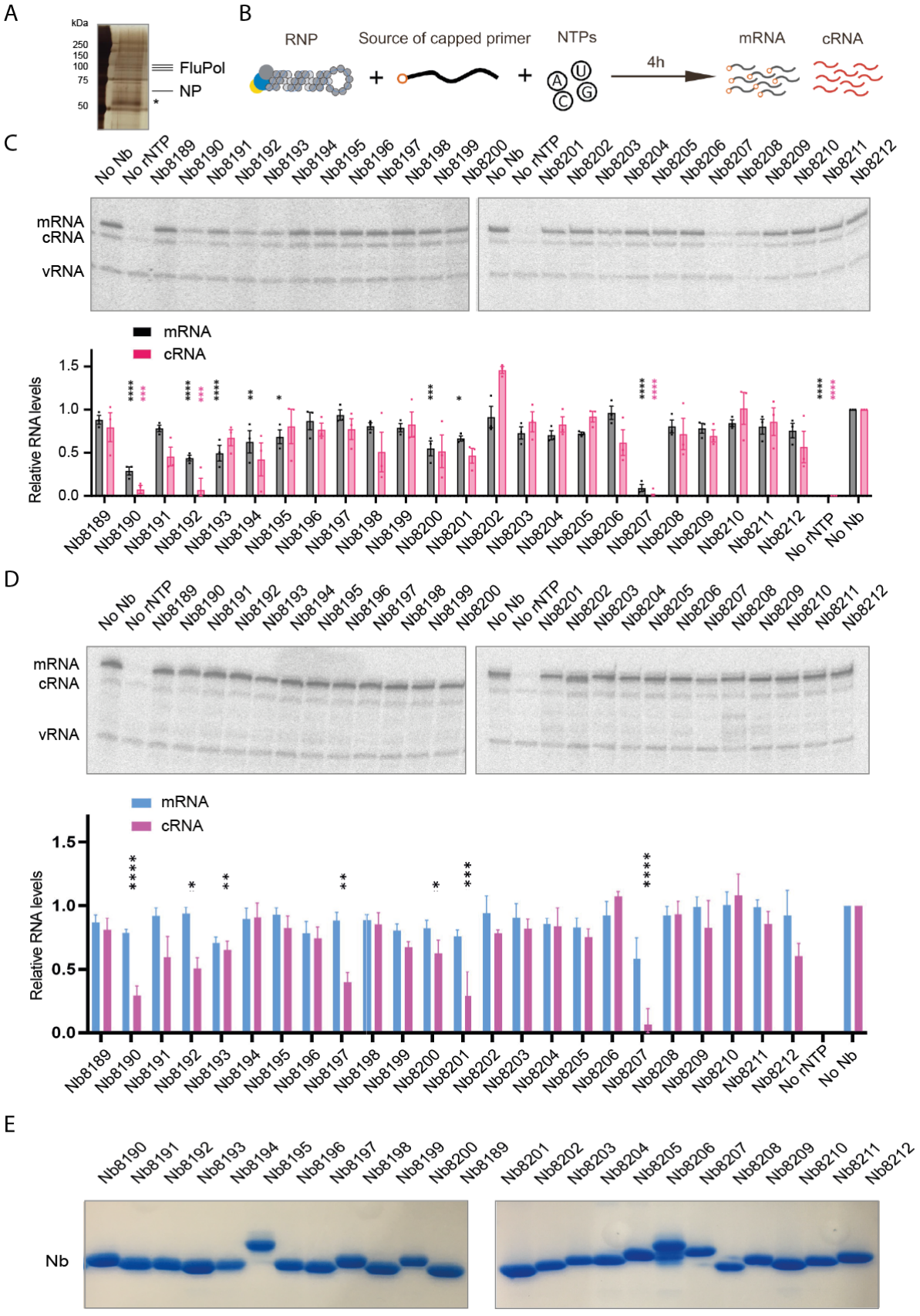

A
kDa
250
150
100
75
50
FluPol
NP
*
B
RNP
Source of capped primer
NTPs
4h
mRNA
cRNA
C
No Nb
No rNTP
Nb8189
Nb8190
Nb8191
Nb8192
Nb8193
Nb8194
Nb8195
Nb8196
Nb8197
Nb8198
Nb8199
Nb8200
No Nb
No rNTP
Nb8201
Nb8202
Nb8203
Nb8204
Nb8205
Nb8206
Nb8207
Nb8208
Nb8209
Nb8210
Nb8211
Nb8212
mRNA
cRNA
vRNA
mRNA
cRNA
Relative RNA levels
1.5
1.0
0.5
0.0
No rNTP
No Nb
D
mRNA
cRNA
vRNA
mRNA
cRNA
Relative RNA levels
1.5
1.0
0.5
0.0
E
Nb

Figure 4.3 Nanobodies inhibit the 1918 viral polymerase activity in vitro. *A. SDS-PAGE analysis followed by silver staining of the PB2-TAP-purified recombinant 1918-RNP. The sample lane was contaminated by the overloaded marker in the adjacent lane. * shows the co-purified IgG heavy chain. B. Schematic diagram showing the set up for the in vitro recombinant 1918-RNP activity assay. C-D. Effect of nanobodies on the activity of recombinant 1918-RNP in the presence of globin mRNA (C) or a 12 nt capped primer (D) as capped-primer donor. A representative result of 3 independent RNA analyses by primer extension is shown (top). Relative RNA levels (bottom) were calculated by comparing with the RNA levels obtained in the absence of nanobodies. Data are mean ± s.e.m. n = 3 independent RNP purifications and reactions. Ordinary one-way ANOVA was used to compare mRNA and cRNA levels in the absence (no Nb) and presence of the indicated nanobody*, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$. E. SDS-PAGE analysis followed by Coomassie staining of the same amount of nanobodies (Nb) purified from E. coli.*

4.2.3 Nanobodies inhibit the 1918 viral polymerase activity in cell culture

It is advantageous to assess nanobody inhibition on polymerase activity *in vitro,* because the amount of nanobody and RNP can be controlled across samples; however, this approach overlooks many other processes during which the activity of a viral polymerase could be affected during infection, including nuclear import, polymerase heterotrimer formation, RNP assembly, polymerase interaction with host and viral factors, and RNP membrane export. Therefore, I next studied the effect of these nanobodies on 1918 polymerase activity in cells. To reconstitute 1918-RNP activity in cells, I used a minigenome assay, where I transfected plasmids encoding 1918-RNP components (1918-PA, PB1, PB2, and NP), along with a Pol I promoter-controlled vRNA-like gene, into HEK293T cells. Initially, a luciferase-reporter gene was transfected as vRNA, which produces luciferase protein upon transcription and translation in cells. Product luciferase protein is an enzyme which catalyses an oxidising reaction that produces light when its substrate, luciferin protein, is added. The light intensity is measured by

a luminometer, as the readout of luciferase protein expression, which reflects the overall RNP activity. PB2 can be left out during transfection to create a negative control, providing a baseline luciferase level. To screen for the effect of these nanobodies, I additionally transfected plasmids encoding each individual nanobody, and calculated the relative luciferase intensity compared to positive control, where the nanobody plasmid is swapped for an empty vector.

Previous work from our lab by Dr Alex Walker showed that Nb8205 is highly inhibitory toward WSN viral polymerase in the minigenome assay (Fan et al., 2019). However, my initial screening results suggest that Nb8205 had no effect on 1918 polymerase activity, which is inconsistent with the previous result (Figure 4.4A). Since nanobodies were raised using avian A/Fujian/01/2002 (H5N1) viral polymerase, it is possible that nanobodies show different activity levels when used to inhibit polymerase from different viral strains. To assess whether this inconsistency was strain-dependent, I tested two nanobodies, Nb8205 and Nb8208 (the latter being the most inhibitory nanobody in my initial screening minigenome assay) on the activity of polymerases from five different strains of influenza A virus: A/NT/60/1968 (NT60), A/Fujian/01/2002 (Fj), A/Puerto Rico/8/34 (PR8), A/WSN/33 (WSN), and A/Brevig Mission/1/1918 (1918). I found that different viral strains can have different sensitivities towards nanobody inhibition. In brief, while NT60, Fj and WSN viral polymerases are sensitive to both nanobodies, PR8 viral polymerase is only sensitive to Nb8205 inhibition but insensitive to Nb8208 inhibition, and 1918 viral polymerase showed reduced sensitivity to both nanobodies (Figure 4.4B).

While the insensitivity towards Nb8208 in PR8 viral polymerase is likely to be due to strain-dependent differences, it is possible that the difference in sensitivity toward Nb8205 and Nb8208 inhibition in 1918 polymerase is because 1918-RNP components were encoded using

a different plasmid backbone. Specifically, in my minigenome assay, RNP components were all encoded in pcDNA plasmids, except for 1918-RNP components, which were encoded in pCAGGS plasmids. Both pCAGGS and pcDNA plasmids are typical transfection plasmids used to express exogenous proteins; however, it is known that genes encoded in pCAGGS plasmids have a higher protein expression level compared to genes encoded in pcDNA plasmids. Therefore, I compared the sensitivity of the same viral polymerase encoded in different vectors (pCAGGS, pcDNA, or pHW2000) towards Nb8205 and Nb8208 inhibition (Figure 4.3C). pHW2000 is a bidirectional plasmid used for viral reverse genetic engineering (Hoffmann et al., 2000). This plasmid produces a positive-sense viral gene segment that can be used to translate viral proteins, as well as to template negative-sense vRNA, and as such can also be used for the minigenome assay. My results showed that RNP encoded by pCAGGS plasmids were generally much less sensitive to nanobody inhibition in cells compared to RNP encoded by pcDNA or pHW2000 plasmids. A possible explanation of this observation is that upon transfection of the same amount of plasmid, pCAGGS plasmids produce significantly more viral polymerase than pcDNA or pHW2000 plasmids. Therefore, when RNP components were expressed in pCAGGS plasmids, the nanobody to viral protein ratio is substaintially lower, resulting in a lower percentage of RNP-associated viral polymerase being bound and inhibited by the transfected nanobody.

To test this hypothesis, I examined how nanobody inhibition is affected if the nanobody to viral protein ratio is increased, by raising the plasmids ratio between the nanobody and the RNP components (in pCAGGS plasmid) (Figure 4.3D). I found that an increase in nanobody to polymerase ratio drastically decreases luciferase intensity, i.e. increases the level of nanobody inhibition. I also examined how the luciferase intensity is affected if the ratio between the plasmids encoding vRNA and RNP component proteins is altered. I found that the luciferase

intensity decreases when I reduce the level of luciferase vRNA, but not when I reduce the levels of RNP component proteins (Figure 4.3E). Taken together, these results suggest that there was an excess of RNP component proteins in my initial minigenome screening assay, compared to both nanobody protein and the luciferase vRNA, therefore a large proportion of viral polymerase is not associated with an RNP. This results in a low occupancy of nanobody on RNP-associated viral polymerase and led to the inconsistency compared to the previous results from our lab.

In order to increase the occupancy of nanobody on RNP-associated viral polymerase, I adjusted the plasmid ratio according to my titrations and repeated the minigenome screening assay. My results revealed that more than half of the nanobodies inhibited 1918 viral RNP in cells, including those identified as inhibitory from the previous *in vitro* RNP activity assay, as expected (Figure 4.5A). Excitingly, some nanobodies that showed only minor effect, if at all, in the *in vitro* assay appeared to be highly inhibitory in cells. These include Nb8208, which binds to $PB2^{cap}$, and Nb8203 and Nb8209, which bind to PA^{endo} (Figure 4.5A). The difference in nanobody inhibition *in vitro* and *in cellulo* suggests that surfaces bound by Nb8208, Nb8203 and Nb8209 are possibly involved in functions that are not required for RNP *in vitro* activity but are essential for RNP activity in a more complex cellular environment.

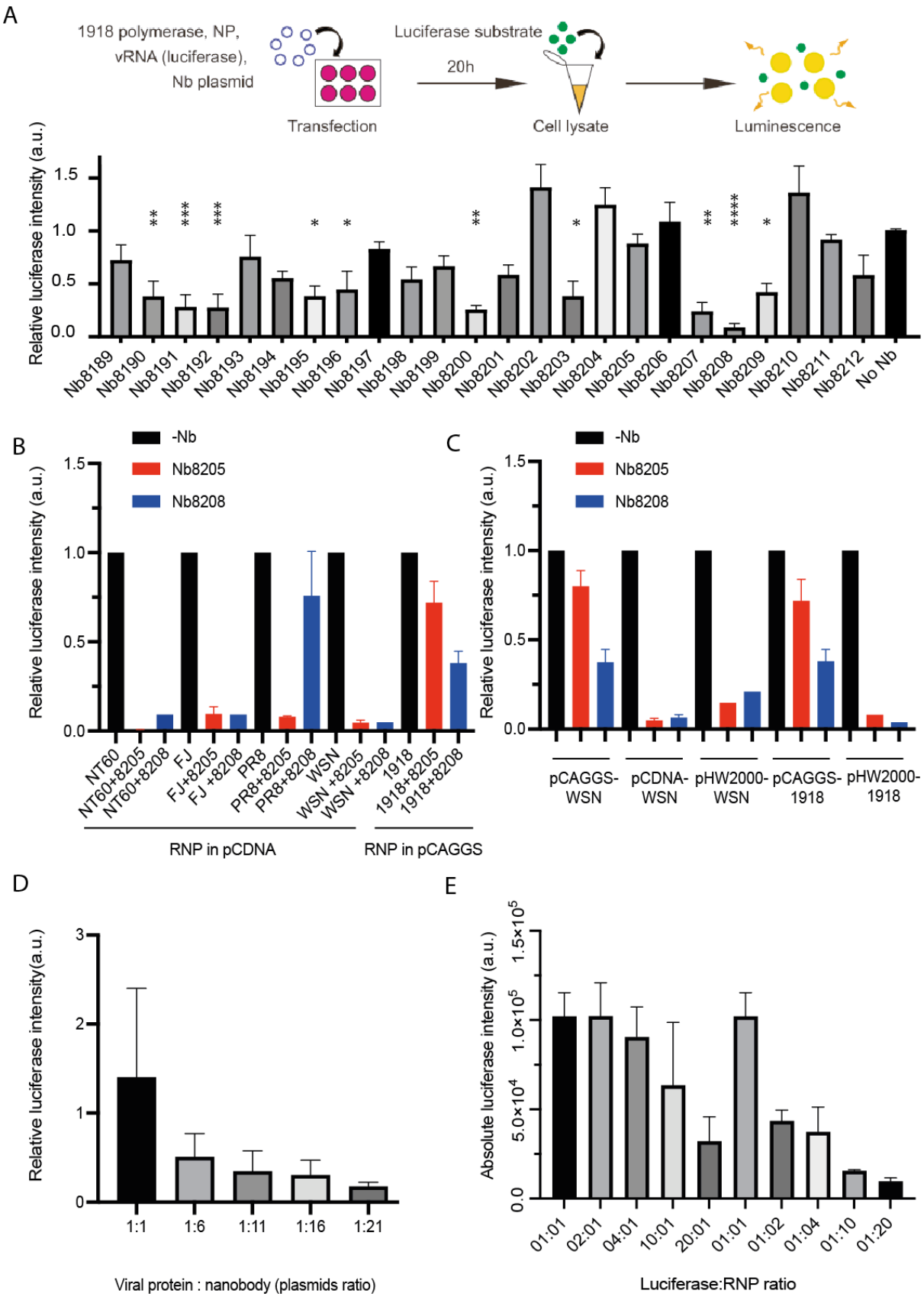

A
1918 polymerase, NP,
vRNA (luciferase),
Nb plasmid
Transfection
20h
Luciferase substrate
Cell lysate
Luminescence
Relative luciferase intensity (a.u.)
Nb8189
Nb8190
Nb8191
Nb8192
Nb8193
Nb8194
Nb8195
Nb8196
Nb8197
Nb8198
Nb8199
Nb8200
Nb8201
Nb8202
Nb8203
Nb8204
Nb8205
Nb8206
Nb8207
Nb8208
Nb8209
Nb8210
Nb8211
Nb8212
No Nb
B
-Nb
Nb8205
Nb8208
Relative luciferase intensity (a.u.)
NT60
NT60+8205
NT60+8208
FJ
FJ+8205
FJ +8208
PR8
PR8+8205
PR8+8208
WSN
WSN +8205
WSN +8208
1918
1918+8205
1918+8208
RNP in pCDNA
RNP in pCAGGS
C
-Nb
Nb8205
Nb8208
Relative luciferase intensity (a.u.)
pCAGGS-WSN
pCDNA-WSN
pHW2000-WSN
pCAGGS-1918
pHW2000-1918
D
Relative luciferase intensity(a.u.)
1:1
1:6
1:11
1:16
1:21
Viral protein : nanobody (plasmids ratio)
E
Absolute luciferase intensity (a.u.)
01:01
02:01
04:01
10:01
20:01
01:01
01:02
01:04
01:10
01:20
Luciferase:RNP ratio

Figure 4.4 Viral subtypes and plasmid backbones affect nanobody inhibition of polymerase activity in cell culture. *A. Schematic diagram of the luciferase-reporter minigenome assay (top). Effect of nanobodies on 1918 viral polymerase activity measured using a luciferase-reporter minigenome assay, in which same amount of each of pcDNA-nanobody, pPOLI-luciferase, pCAGGS-PA, pCAGGS-PB1, pCAGGS-PB2 and pCAGGS-NP were transfected (bottom). Data are mean ± s.e.m. n = 3 independent transfections with n =2 technical replicates. Ordinary one-way ANOVA was used to compare the relative luciferase intensity in the presence and absence of nanobodies. *, p<0.05; **, p<0.01; ***, p<0.001; ****, p<0.0001. B. Effect of Nb8205 and Nb8208 on viral polymerases from different viral subtypes measured using a luciferase-reporter minigenome assay. Data are mean ± range. n = 1 to 3 independent transfections with n =2 technical replicates. C. Viral polymerase activity in different expression systems measured using a luciferase reporter minigenome assay. Data are mean ± range. N = 1 to 3 independent transfections with n =2 technical replicates. D. Effect of decreasing RNP-component to nanobody plasmid ratio on the inhibition of viral polymerase activity by nanobodies measured using a luciferase reporter minigenome assay. Data are mean ± range. n = 2 independent transfections with n =2 technical replicates. E. Effect of change in pPOLI-Luciferase to RNP-component plasmid ratio on the inhibition of viral polymerase activity by nanobodies measured using a luciferase reporter minigenome assay. Data are mean ± range. n = 2 or 4 independent transfections with n =2 technical replicates.*

The luciferase reporter assay cannot distinguish transcriptional and replicational activity because it produces a single readout of light intensity as an overall representation of the RNP activity. Therefore, to assess the levels of individual viral RNA species, I performed the minigenome assay with NA segment as vRNA using pPOLI-NA plasmid, and quantified mRNA, cRNA and vRNA levels using the primer extension assay. A significant increase in mRNA, cRNA, and vRNA levels in conditions where all RNP components were transfected compared to the negative control lacking PB2 subunit confirmed that 1918 viral RNP was transcriptionally and replicationally active in HEK293T cells (Figure 4.5B). My primer extension results revealed the inhibitory effect of nanobodies on mRNA, cRNA and vRNA,

and showed that the inhibitory effect of nanobodies on viral RNA levels is largely consistent with that on viral protein production (Figure 4.5B).

An inherent drawback of the minigenome assay over the *in vitro* RNP activity assay is that the protein expression levels of different nanobodies could be different, resulting in apparent differences in inhibition. Therefore, I attempted to compare nanobody expression levels using monoclonal anti-llama antibodies. Unfortunately, I found that the detection ability of anti-llama antibodies against these nanobodies is nanobody-specific. Similarly, I also showed that different nanobodies are detected to differing levels by silver staining. These results suggest that the apparently highly-conserved nanobodies might be surprisingly different in their biochemical properties (Figure 4.5C). According to a previous Fodor and Grimes Lab member, Dr Itziar Serna Martin, these closely related nanobodies were unexpectedly different in their expression levels. Hence it is difficult to conclude whether nanobodies that did not show inhibition in cells were truly non-inhibitory or were expressed at low levels. Despite this, my results of the minigenome assay revealed certain highly inhibitory nanobodies that could be used in subsequent studies.

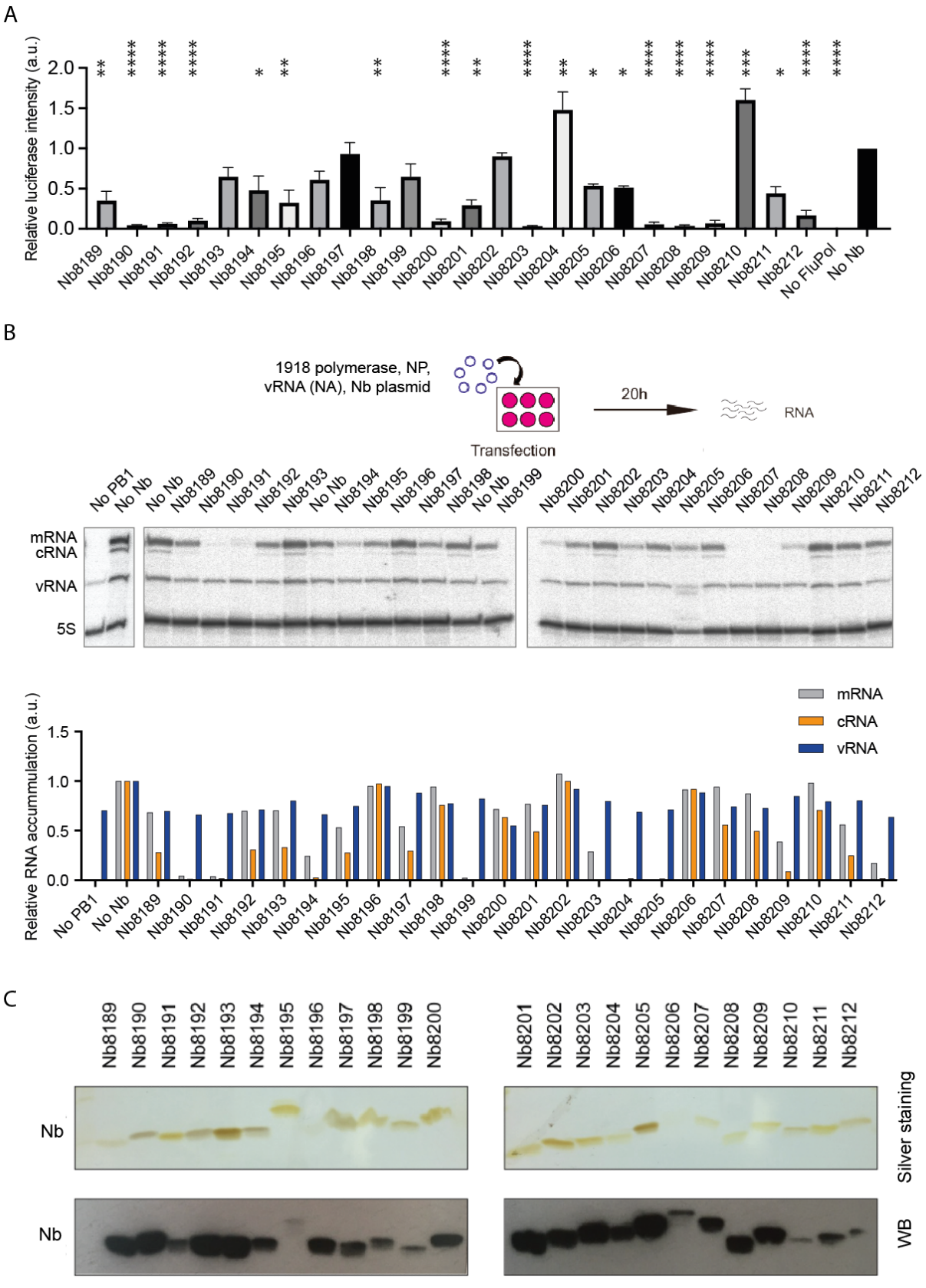

A
Relative luciferase intensity (a.u.)
2.0
1.5
1.0
0.5
0.0
Nb8189 Nb8190 Nb8191 Nb8192 Nb8193 Nb8194 Nb8195 Nb8196 Nb8197 Nb8198 Nb8199 Nb8200 Nb8201 Nb8202 Nb8203 Nb8204 Nb8205 Nb8206 Nb8207 Nb8208 Nb8209 Nb8210 Nb8211 Nb8212 No FluPol No Nb
B
1918 polymerase, NP, vRNA (NA), Nb plasmid
Transfection
20h
RNA
No PB1 No Nb No Nb Nb8189 Nb8190 Nb8191 Nb8192 Nb8193 No Nb Nb8194 Nb8195 Nb8196 Nb8197 Nb8198 No Nb Nb8199 Nb8200 Nb8201 Nb8202 Nb8203 Nb8204 Nb8205 Nb8206 Nb8207 Nb8208 Nb8209 Nb8210 Nb8211 Nb8212
mRNA cRNA
vRNA
5S
Relative RNA accummulation (a.u.)
mRNA
cRNA
vRNA
1.5
1.0
0.5
0.0
No PB1 No Nb Nb8189 Nb8190 Nb8191 Nb8192 Nb8193 Nb8194 Nb8195 Nb8196 Nb8197 Nb8198 Nb8199 Nb8200 Nb8201 Nb8202 Nb8203 Nb8204 Nb8205 Nb8206 Nb8207 Nb8208 Nb8209 Nb8210 Nb8211 Nb8212
C
Nb8189 Nb8190 Nb8191 Nb8192 Nb8193 Nb8194 Nb8195 Nb8196 Nb8197 Nb8198 Nb8199 Nb8200
Nb8201 Nb8202 Nb8203 Nb8204 Nb8205 Nb8206 Nb8207 Nb8208 Nb8209 Nb8210 Nb8211 Nb8212
Nb
Nb
Silver staining
WB

Figure 4.5 Nanobodies inhibit the 1918 viral polymerase activity in cell culture. *A. Effect of nanobodies on 1918 viral polymerase activity measured using a luciferase-reporter minigenome assay. Data are mean ± s.e.m. n = 3 independent transfections with n =2 technical replicates. Ordinary one-way ANOVA was used to compare the relative luciferase intensity in the presence and absence of nanobodies. *, p<0.05; **, p<0.01; ***, p<0.001; ****, p<0.0001. B. Effect of nanobodies on 1918 viral polymerase activity measured using an NA-based minigenome assay (bottom). RNA levels were determined by primer extension. Schematic diagram showing the NA-based minigenome assay (top). The gel of primer extension assay is shown (middle) with quantification (bottom). n = 1 transfection. C. SDS-PAGE analysis followed by silver staining (top) or Western Blotting (WB, bottom) of purified nanobodies (Nb) loaded at the same amount.*

4.2.4 Nanobodies inhibit viral growth

Up until now, I have been assessing nanobody inhibition of the RNA synthesis function of the 1918 viral polymerase. Next, I addressed whether nanobody inhibition could affect viral infection. I selected a subset of inhibitory nanobodies with unique binding sites on 1918 viral polymerase, namely Nb8190, Nb8191, Nb8192, Nb8193, Nb8206, Nb8207, Nb8208 and Nb8209. These binding sites are shown to be highly conserved across different influenza A viruses (Keown et al., 2022). Additionally, I included a non-inhibitory nanobody that binds to PA^{linker}, Nb8210, as a control.

To assess viral infection in the context of the 1918 viral polymerase, I generated reassortant influenza A/WSN/33 viruses containing either 1918 viral polymerase (WSN-$1918_{polymerase}$) or 1918 viral polymerase and NP (WSN-1918_{RNP}). I determined the viral titres using plaque assay (Figure 4.6A). Notably, the titre of the WSN-$1918_{polymerase}$ reassortant virus was four logs lower than that of the WSN or WSN-1918_{RNP} reassortant viruses, potentially suggesting an incompatibility between the WSN-NP and the 1918 polymerase (results not shown).

I carried out a viral growth assay with the WSN-1918$_{RNP}$ reassortant virus on HEK293T cells that had been transfected with plasmids encoding one of the selected nanobodies prior to infection. Viral titres in the cell culture media at 16 h, 24 h, 32 h, and 48 h post-infection (hpi) were determined by plaque assay. Additionally, viral mRNA, cRNA and vRNA levels in infected cells at each time point were determined by primer extension assay (Figure 4.7). An FHS (Final Honour Scheme) student, Grace Wheeler, performed the same assays in parallel with an additional three nanobodies: Nb8193, Nb8206, and Nb8209, along with the control nanobody Nb8210. Together, we observed that Nb8190, Nb8191, Nb8192, Nb8206, Nb8207, and Nb8209 strongly inhibited viral replication at the tested time points. Nb8208 strongly inhibited viral growth at 24 and 36 hpi, but the viral titre recovered compared to the controls, at 48 hpi (Figure 4.6C). Consistent with previous studies, the control nanobody Nb8210 had an insignificant effect on viral infection, showing that the observed decrease in viral titre was due to the presence of inhibitory nanobodies. I also showed that the inhibition in viral RNA levels matches the inhibition in viral titres (Figure 4.7). This result suggests that the restriction in viral growth is associated to nanobody inhibition in polymerase activity.

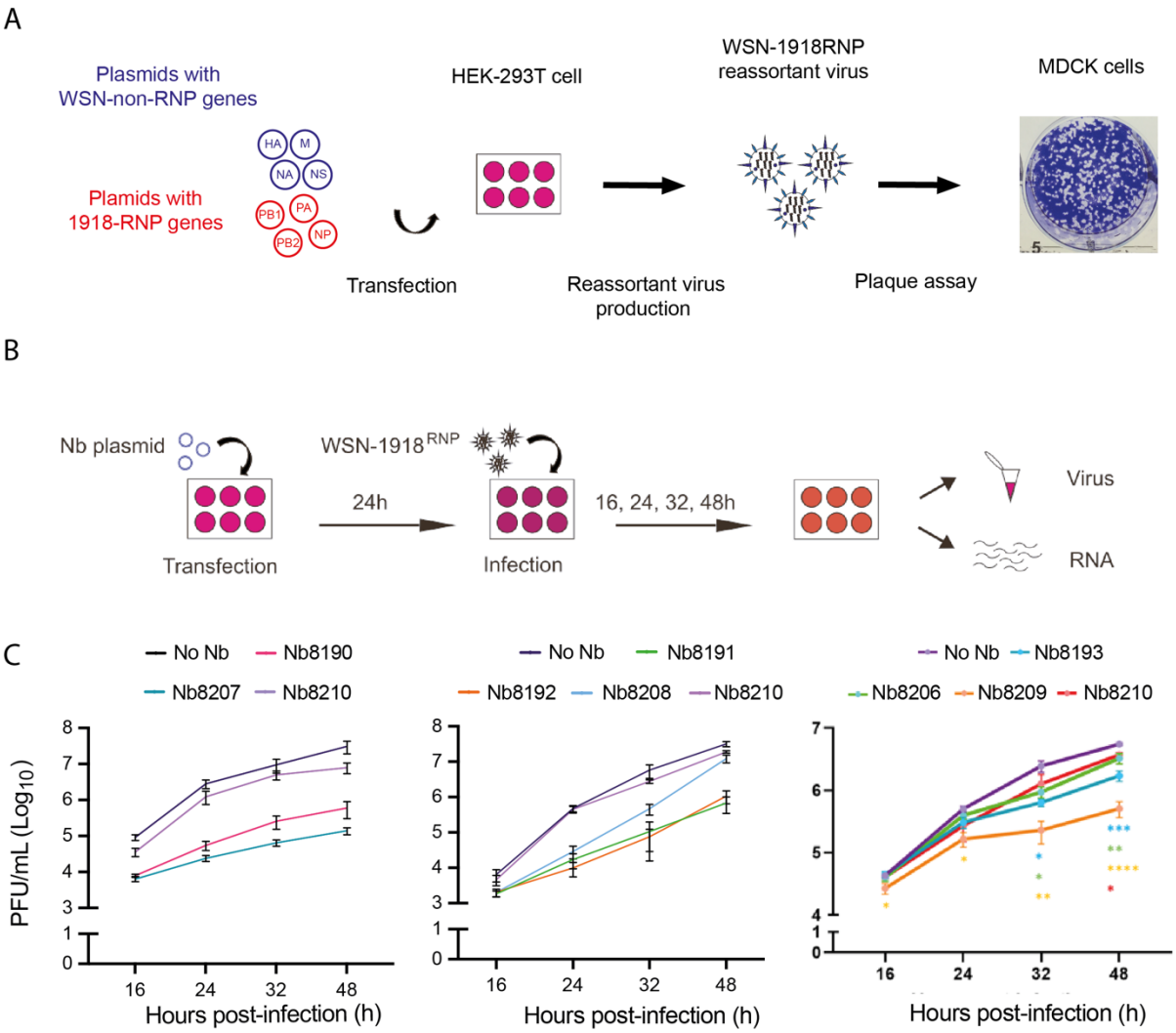


Figure 4.6 Nanobodies reduce viral growth in infection. *A. Schematic diagram showing the rescue and titre determination of WSN-1918$_{RNP}$ reassortant virus. B Schematic diagram of the viral growth assay, where viral titres in the cell media and viral RNA in the cell debris were analysed separately at 16, 24, 32 and 28 h post infection (hpi). C. Effect of three sets of nanobodies on viral growth of a WSN-1918$_{RNP}$ reassortant virus in HEK293T cells at specific time points. Data are mean ± s.e.m. n = 3 independent transfections and infections for each set of nanobodies. PFU stands for plaque-forming units. The right-most panel in C is adapted from Grace Wheeler's FHS report.*

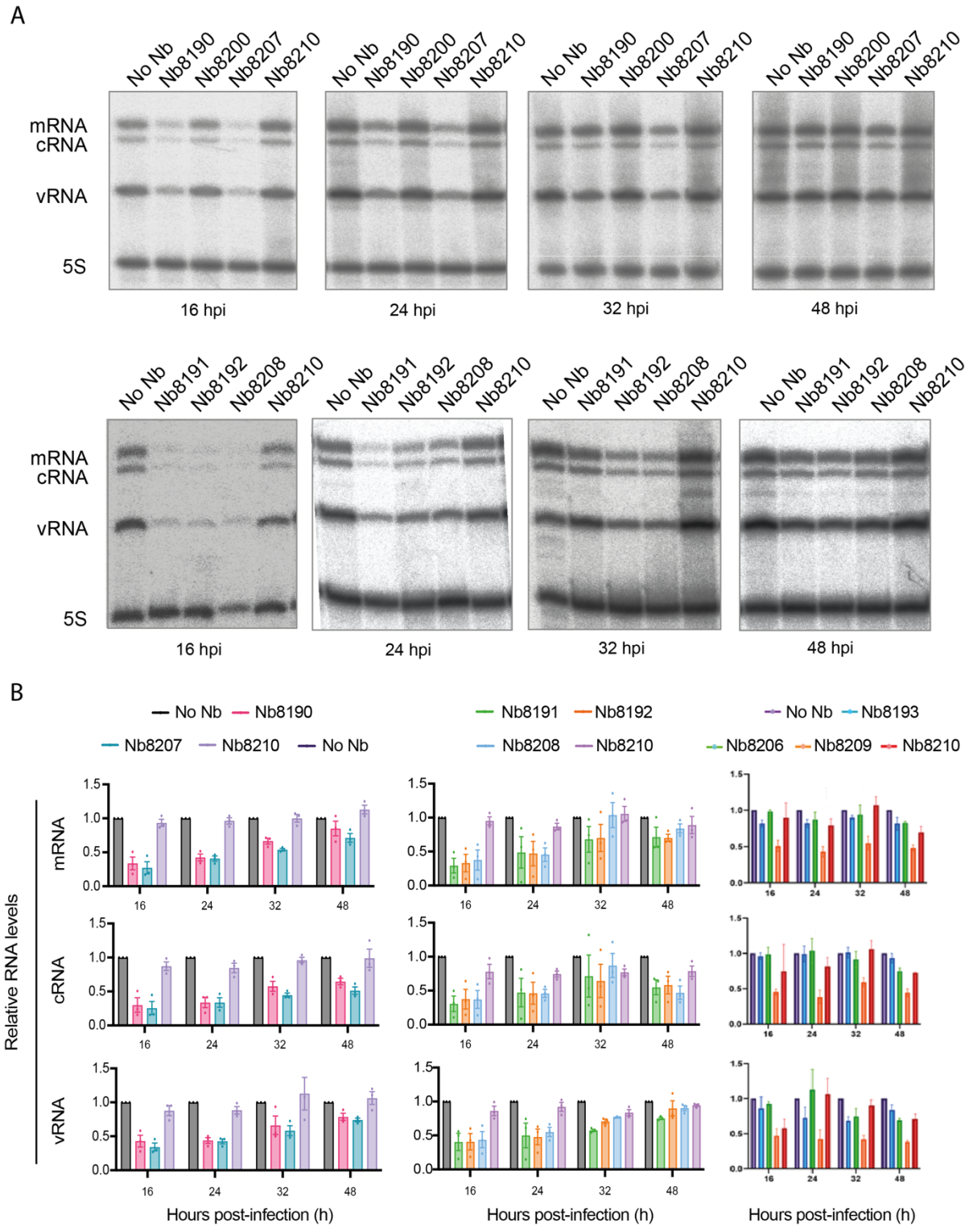


Figure 4.7 Nanobodies inhibit viral polymerase activity in infection. *A-B. Effect of three sets of nanobodies on the polymerase activity of a WSN-1918$_{RNP}$ reassortant virus in HEK293T cells at specific time points. A representative gel of 3 independent primer extension analyses of two of the three sets of nanobodies is shown (A). Quantification of data is shown as mean ± s.e.m. n = 3 independent transfections and infections for each set of nanobodies (B). The right-most panels in B are adapted from Grace Wheeler's FHS report.*

It is worth noting that nanobodies raised against an avian H5N1 viral polymerase exhibited such strong inhibitory effects on an H1N1 pandemic viral polymerase. This suggests that it might be possible to develop pan-inhibitory drugs against influenza A virus based on these nanobodies. To assess whether these nanobodies are inhibitory across different influenza A viral subtypes, I conducted a minigenome luciferase reporter assay using viral polymerase and NP from H3N2, H5N1, and H1N1 viral subtypes. I found that all selected nanobodies had a conserved inhibitory effect across the tested influenza virus subtypes (Figure 4.8A). However, I previously demonstrated that Nb8208 had decreased inhibitory effects against A/Puerto Rico/8/34 RNP. Together, these results suggest that the inhibitory effects of nanobodies are mostly conserved across different influenza A viral subtypes, but some subtypes might be resistant to certain nanobodies, possibly due to mutations at the interaction interface.

To demonstrate that cell viability was not affected by the expression of nanobodies, I performed a cell viability assay on the nanobodies I selected and showed that transfection of these nanobodies had no effect on cell viability 24 hpi (Figure 4.8B). I also verified nanobody expression prior to infection using Western blot and showed that all nanobodies I tested were expressed prior to infection (Figure 4.8C). However, as mentioned earlier, the expression levels between different nanobodies were not comparable using Western blot due to the difference in antibody detecting ability. Nevertheless, at least for the more readily detected nanobodies, I showed that nanobody protein levels remained unchanged at the last time point (Figure 4.8D).

Taken together, these results suggest that nanobodies expressed in cells can reduce viral growth by inhibiting the function of the viral polymerase.

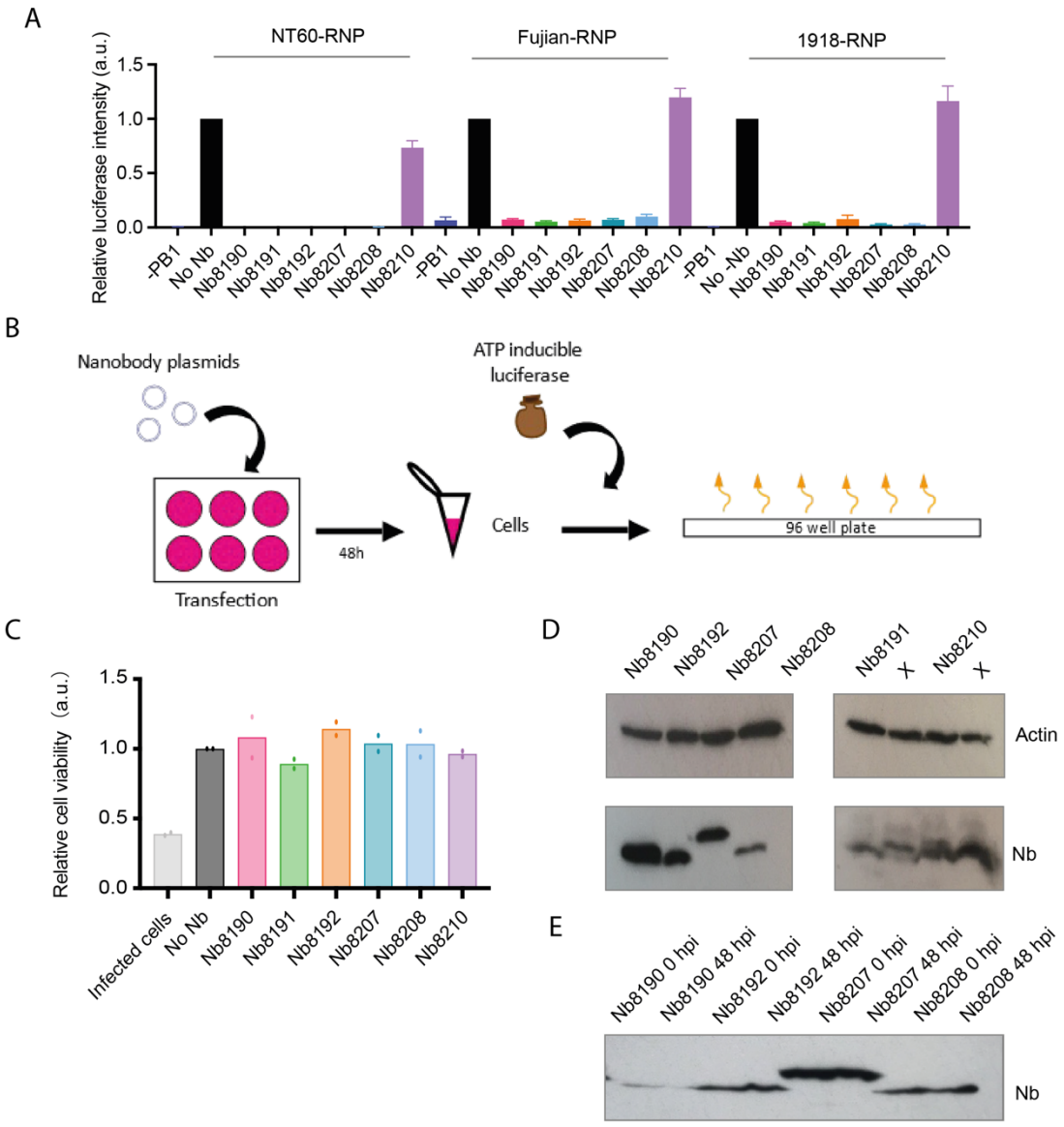


Figure 4.8 Further investigation on the selected inhibitory nanobodies. *A. Effect of the selected nanobodies on H3N2 (NT60), H5N1 (Fujian) and H1N1 (1918) influenza A virus polymerase activity measured using a luciferase reporter minigenome assay. Data are mean ± s.e.m. n = 3 independent transfections with n = 2 technical replicates. Ordinary one-way ANOVA was used to compare the relative luciferase intensity in the presence and absence of nanobodies for each influenza A viral subtype. B. Schematic diagram of the luciferase-based cell viability assay. C. Luciferase-based cell viability assay on HEK293T cells transfected with plasmids expressing selected nanobodies. Data are mean ± range. n = 2 independent transfections each with n = 2 technical replicates. D. Expression of selected nanobodies (Nb) in HEK293T cells 24 h post transfection analysed by Western Blot using anti-alpaca polyclonal antibody. β-actin (actin) was used as a loading control. E. Expression of well-detected nanobodies 16, 24, 32 and 48 h post infection (hpi), was analysed by Western Blot using anti-alpaca polyclonal antibody. A representative blot is shown for D-E.*

4.2.5 Inhibitory mechanisms of the nanobodies

Next, I was interested in understanding how these inhibitory nanobodies affect 1918 polymerase function at the molecular level. To shed light on the underlying inhibitory mechanisms, I conducted several assays to assess different aspects of the polymerase function independently.

To investigate the effect of nanobodies on viral transcription, I first examined their effect on the 1918 viral polymerase interacting with host Pol II CTD. This interaction is crucial during infection as it enables the viral polymerase to access nascent capped RNA produced by the host RNA polymerase, which is important for transcription initiation (Engelhardt et al., 2005; Lukarska et al., 2017). To mimic this process *in vitro*, I purified recombinant 1918 viral polymerase and performed a pull-down assay to measure its binding to a Pol II CTD mimic peptide with four heptad repeats (4X CTD), each with a phosphorylated serine-5 (pS5). A 4X CTD with phosphorylated serine-2 (pS2) was used as a negative control. The effect of purified nanobodies on CTD binding was assessed by adding excess nanobodies to the pull-down assay and measuring the change in 1918 viral polymerase that has co-immunoprecipitated with immobilized 4XCTD. My results showed that both Nb8191 and Nb8192 significantly reduced the binding of 1918 viral polymerase to pS5 4X CTD (Figure 4.9A).

I then addressed whether nanobodies had an effect on primary transcription, specifically, mRNA synthesised by incoming virion-derived vRNP. To answer this question, I performed a primary transcription assay by adding cycloheximide at the onset of viral infection of HEK293T cells expressing one of the selected nanobodies. Cycloheximide blocks ribosomal protein translation, thus prevents viral protein production at the beginning of infection. In this

way, cRNA and vRNA synthesis could not occur due to the lack of viral protein translation, while nanobodies transfected prior to infection would have been synthesised. Since the mRNA level is always affected by the vRNA level due to secondary transcription, in infection or in minigenome assay, the primary transcription assay allowed for the dissection of effects on viral transcription alone. My primary transcription results showed that Nb8190, Nb8191, Nb8192, and Nb8207 directly inhibited viral mRNA synthesis (Figure 4.9B).

The next step was to dissect the inhibitory effect of nanobodies on each of the two steps of viral replication, in other words, to separate cRNA synthesis inhibition from vRNA synthesis inhibition. To do so, I performed a cRNA stabilization assay, in which a catalytically inactive viral polymerase (1918-PB1a) and 1918-NP were expressed along with a nanobody prior to infection and cycloheximide treatment. In this way, cRNA could accumulate upon being stabilized by 1918-PB1a and NP, by being assembled into cRNP, but it could not be further replicated due to the inactive 1918-PB1a. I quantified cRNA levels in the presence of each nanobody and showed that all selected nanobodies, except for the control nanobody, inhibited cRNA synthesis (Figure 4.9C).

Finally, I assessed the second step of replication, namely vRNA synthesis on a cRNA template, by replacing 1918-PB1a used in the cRNA stabilization assay with an active 1918 polymerase (1918-FluPol). This allowed replication to proceed after cRNA synthesis, producing more vRNA using *de novo* cRNA as a template. These results showed that all selected nanobodies, except for the control nanobody, strongly inhibited vRNA synthesis (Figure 4.9D).

Taken together, the results from this section allow us to speculate the inhibitory mechanisms of a panel of nanobodies on 1918 polymerase functions. However, because individual

nanobody binding site on a viral polymerase could be involved in multiple molecular events during transcription and replication, the exact inhibitory mechanism is not completely clear.

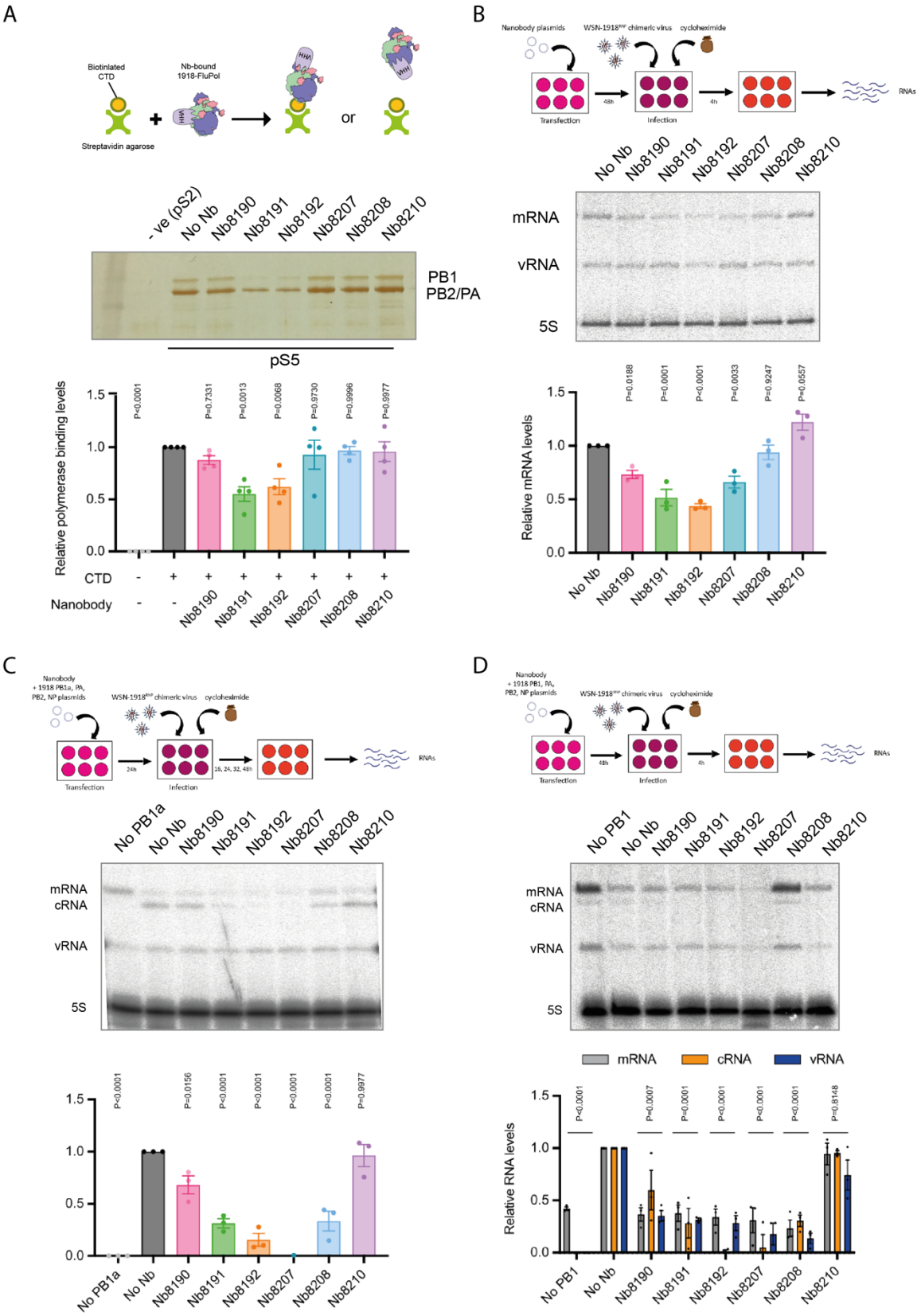


Figure 4.9 Investigating the inhibitory mechanisms of selected nanobodies. *A. Schematic diagram of the CTD-binding assay (top). Effect of nanobodies on the binding of purified recombinant 1918 viral polymerase (1918-FluPol) to serine-5-phosphorylated (pS5) Pol II CTD mimic peptide (CTD). Serine-2-phosphorylated (pS2) Pol II CTD mimic peptide was used as a negative control (-ve). A representative silver-stained gel from 3 independent analyses is shown (middle). Quantification of gels from n =4 independent binding assays is shown (bottom). Data are mean ± s.e.m. Ordinary one-way ANOVA was used to compare the relative polymerase binding in the presence and absence of nanobodies. P < 0.05 is considered significant. B. Schematic diagram of the primary transcription assay (top). Effect of nanobodies on primary transcription by a WSN-1918$_{RNP}$ reassortant virus in HEK293T cells expressing the indicated nanobodies prior to infection. A representative gel of 3 independent primer extension assays is shown (middle). Quantification of gels were shown as mean ± s.e.m. n = 3 independent transfections and infections. Ordinary one-way ANOVA was used to compare the relative mRNA levels in the presence and absence of nanobodies. P < 0.05 is considered significant. C. Schematic diagram of the cRNA stabilisation assay (top). Effect of nanobodies on cRNA levels produced by a WSN-1918$_{RNP}$ reassortant virus in HEK293T cells expressing the indicated nanobodies and influenza virus NP and polymerase subunits, including a polymerase active site mutant PB1 (PB1a), prior to infection. Plasmids to express PB1a (no PB1a) was omitted as negative control (no PB1a). A representative gel of 3 independent primer extension assays is shown (middle). Quantification of gels from n = 3 independent transfections and infections (bottom). Data are mean ± s.e.m. Ordinary one-way ANOVA was used to compare the relative cRNA levels in the presence and absence of nanobodies. P < 0.05 is considered significant. D. Schematic diagram of the replication assay (top). Effect of nanobodies on mRNA, cRNA and vRNA levels produced by a WSN-1918$_{RNP}$ reassortant virus in HEK293T cells expressing the indicated nanobodies and influenza virus NP and wild type polymerase subunits prior to infection. Plasmid to express PB1 was omitted as negative control (no PB1). A representative gel of 3 independent primer extension assays is shown (middle). Quantification of gels from n = 3 independent transfections and infections (bottom). Data are mean ± s.e.m. Ordinary one-way ANOVA was used to compare the relative RNA levels in the presence and absence of nanobodies. P < 0.05 is considered significant.*

4.3 DISCUSSION

In this study, I explored the inhibitory effects of a panel of nanobodies, with well-defined binding sites on the 1918 influenza virus polymerase, in various contexts. Through a combination of *in vitro* and cell-based assays, I demonstrated that a number of nanobodies can significantly impede 1918 viral polymerase activity. Expanding these findings into infection, I showed that a selected subset of nanobodies restrict viral growth by inhibiting RNP functions during infection. Coupled with the cryo-EM data generated by Dr Jeremy Keown from our lab, which reveal the nanobody binding sites on the 1918 viral polymerase (summarised in Table 4.2), my functional results uncovered novel inhibitory sites on the surface of the 1918 influenza virus polymerase. These sites are sensitive to nanobody binding and hold potential as targets for the development of new therapeutics against different strains of influenza A viruses (Figure 4.10). However, there are problems to overcome before these nanobodies can be transformed into therapeutics. Firstly, nanobodies typically struggle to penetrate cell membranes, presenting challenges when targeting viral polymerase, which is only accessible in the cytoplasm. Therefore, these nanobodies can only provide a mechanistic basis to design small molecular peptides that can enter the cells. Secondly, despite their small size, nanobodies are still larger than small molecule compounds and can lead to inhibition by causing steric hindrance. Therefore, to ensure similar inhibitory functions to be retained, extra considerations are needed when designing small molecule drugs based on nanobodies.

Additionally, I propose three promising novel inhibitory mechanisms against influenza viral polymerase for future drug development:

1. Conformational blockage: Nanobodies may inhibit viral polymerase by obstructing specific conformations essential for transcription, replication, or encapsidation. Specifically, each functional conformation of the viral polymerase relies on a distinct arrangement of flexible domains to facilitate its function. Therefore, nanobodies binding to particular sites on the viral polymerase might lock a polymerase in a certain conformation, or prevent it from adopting a specific conformation. For example, Nb8190 binds to a site against which $PB2^{627}$ is packed in the transcriptase conformation; therefore, it inhibits polymerase function by preventing the viral polymerase from adopting the transcriptase conformation, decreasing mRNA levels without affecting polymerase binding to the host Pol II CTD. Similarly, the binding of Nb8207 to the viral polymerase has been shown to be incompatible with the packing of PB2-C in either the transcriptase or the replicase conformation. As a result, my data show that Nb8207 prevents mRNA, cRNA, and vRNA synthesis without affecting CTD binding. Likewise, the binding of Nb8208 prevents the polymerase from assuming an encapsidating conformation by blocking the packing of $PB2^{cap}$, thereby inhibiting both steps of replication.

2. Disruption of host and viral factor interactions: Recent structural studies have shown that PA-C provides a binding hotspot for both viral and host factors, such as additional viral polymerase, ANP32 family proteins, and host Pol II CTD (Carrique et al., 2020; Fan et al., 2019; Lukarska et al., 2017; Serna Martin et al., 2018) (summarised in Table 4.1). Nb8191 and Nb8192 bind proximally on PA-C and have both been found to inhibit CTD binding to the 1918 polymerase, primary transcription, and both steps of replication. My results, therefore, support the model that PA-C contains a competition site for multiple regulatory factors, providing a surface for Pol II CTD binding during transcription and for polymerase dimerisation during replication. Therefore, occupying PA-C with a high affinity binding

partner prevents the binding of viral or host factors needed for transcription or replication, thereby inhibiting polymerase function.

3. Product Exit Channel Blockage: Some nanobodies, such as Nb8207, inhibit transcription or replication by blocking the product exit channel. This mechanism does not affect CTD binding but prevents RNA synthesis in a straightforward manner.

Notably, Nb8209 was not included in further mechanistic analysis. However, like Nb8208, it was not inhibitory *in vitro* but highly inhibitory *in cellulo*, suggesting its involvement in a process that was not reconstituted in my *in vitro* recombinant RNP activity assay. As mentioned in an earlier section, a recent cryo-EM structure from our lab has revealed an interaction between viral polymerase and NP at PA^{endo}. This leads to my speculation that Nb8209 possibly inhibits NP interaction with viral polymerase, thus preventing NP from directly promoting RNA synthesis at the replication platform; however, further analysis is needed to prove or disprove this hypothesis.

In summary, in this chapter I have unveiled novel inhibitory sites on the surface of the 1918 influenza virus polymerase using nanobodies. These inhibitory sites appear to be conserved across influenza A virus subtypes, offering promising targets for the development of new universal anti-influenza therapeutics. However, overcoming the challenges associated with nanobody delivery into cells while retaining the effect of steric hinderance will be crucial for translating these findings into practical antiviral treatments. Furthermore, I have proposed three distinct inhibitory mechanisms against influenza viral polymerase, which not only provide insights for antiviral design but also shed light on fundamental aspects of influenza virus biology.

Interfaces on viral polymerase	*Domains involved*	*Grouping*
Symmetric dimer interface	*PA-C, PB1^{thumb} and PB2^{N1}*	*Interface 1*
Pol II CTD binding interface	*PA-C*	*Interface 1*
Interface with replicase on encapsidase	*PA-C, PB2^{627}, PB2^{NLS}*	*Interface 1*
Interface with encapsidase on replicases	*PB2^{627}, PB2^{N2}, PB2$^{mid\text{-}link}$*	*Interface 2*
Interface with ANP32^{LRR} N-term on replicase	*PB2^{lid}, PB2^{627}, PB2^{NLS}*	*Interface 2*
Interface with ANP32^{LRR} C-term on encaspidase	*PA-C and PB1-N*	*Interface 1*

Table 4.1 Summary of the interaction interfaces on influenza A viral polymerase important for transcription and replication.

Nanobodies	*Biding site NO.*	*Binding site domains and residues*	*Effect on viral polymerase*
Nb8189	*Site 4b*	*PA-C (615-630), PB1-N (10-15)*	*Blocks PB2 packing in transcriptase conformation*
Nb8190	*Site 4b*	PA-C (615-630), PB1-N (10-15)	*Blocks PB2 packing in transcriptase conformation*
Nb8191	*Site 4c*	*PA-C (624-626, 551-557, 531-534, 405-420), PB1 (1-2)*	*Blocks 'interface 1', in particular the Pol II CTD binding interface*
Nb8192	*Site 4a*	*PA-C (426-445)*	*Blocks 'interface 1', in particular the symmetric dimer interface*
Nb8193	*Site 5a*	*PB2 (264, 388-390, 482, 524)*	*Blocks PB2^{cap} packing in encapsidating conformation*
Nb8194	*Site 5a*	*PB2 (482, 388-390)*	*Blocks PB2^{cap} packing in encapsidating conformation*
Nb8195	*NA*	*ND*	*NA*
Nb8196	*Site 4c*	*PA-C (705-712, 262-268, 624-627)*	*Blocks 'interface 1', in particular the Pol II CTD binding interface*
Nb8197	*NA*	*ND*	*NA*

Nb8198	*Site 1*	*PA^{Endo} (1-10, 170-185)*	*Inhibit endonuclease function*
Nb8199	*Site 1*	*P^{Endo} (1-10, 170-185)*	*Inhibit endonuclease function*
Nb8200	*Site 1*	*P^{Endo} (1-10, 170-185)*	*Inhibit endonuclease function*
Nb8201	*Site 4b*	*PA-C (615-630), PB1-N (10-15)*	*Blocks 'interface 1'*
Nb8202	*Site 2*	*PA^{Linker} (PA 220-240)*	*No effect*
Nb8203	*Site 1*	*PA^{Endo} (1-10, 170-185)*	*Inhibit endonuclease function*
Nb8204	*Site 2*	*PA^{Linker} (PA 250-265)*	*No effect*
Nb8205	*Site 4a*	*PA-C (490-496, 293-296, 426-445), PB1(565-585)*	*Blocks 'interface 1', in particular the symmetric dimer interface*
Nb8206	*Site 4d*	*PA-C (309-321, 555-558)*	*Blocks 'interface 1', in particular the symmetric dimer interface*
Nb8207	*Site 3*	*PB1 (280-295, 99-103, 265-268), PB2 (139-146)*	*Blocks product exist channel*
Nb8208	*Site 5b*	*PB2 (392, 466-468, 472)*	*Blocks $PB2^{cap}$ packing in encapsidating conformation*
Nb8209	*Site 1*	*PA^{Endo} (1-10, 170-185)*	*Inhibit endonuclease function*
Nb8210	*Site 2*	*PA^{Linker} (PA 220-240)*	*No effect*
Nb8211	*NA*	*ND*	*NA*
Nb8212	*NA*	*ND*	*NA*

Table 4.2 Summary of the nanobody binding sites on the 1918 viral polymerase and the expected effects on viral polymerase based on binding information. *NA: Not Applicable. ND: Not Determined.*

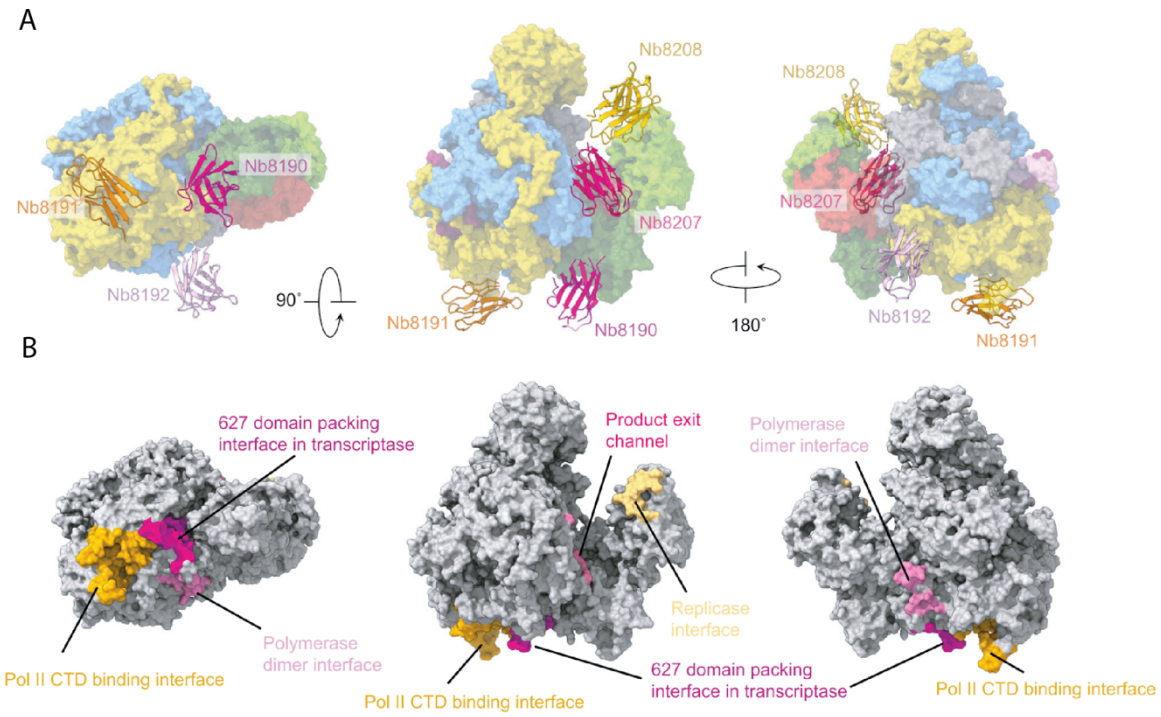


Figure 4.10 Inhibitory sites and potential inhibitory mechanisms of selected nanobodies on the 1918 viral polymerase. *A. Binding position of inhibitory nanobodies mapped on to the transcriptase conformation of the polymerase (PDB ID: 7NHX). Note that the binding of Nb8190 and Nb8207 is for illustration purposes only, as the binding of these nanobodies is incompatible with the transcriptase conformation shown. vRNA promoter and key polymerase subunit or domains are highlighted in different colours: 5' vRNA (pink), 3' vRNA (violet), PA (yellow), PB1 (blue), $PB2^{cap}$ (light green), $PB2^{627}$ (dark green), $PB2^{Mid\text{-}link}$ (red), PB2-N (grey). The location of each nanobody interaction site annotated with proposed biological function. Adapted from Keown, Zhu, Carrique, Fan et al. 2022.*

Chapter 5: New structural approaches to study intra-RNP interactions

5.1 INTRODUCTION

The influenza virus has a segmented RNA genome, organised into vRNP complexes. Each single-stranded vRNA segment forms a closed loop, with both ends captured by a single heterotrimeric RNA-dependent RNA polymerase, decorated with NPs that oligomerise, loop around, and twist into an antiparallel helix (Compans et al., 1972). Recent advancements in crystallography and cryo-EM techniques have facilitated the acquisition of high-resolution structures of viral polymerases of all four types of influenza viruses, and of NP of influenza A, B, and D viruses (Donchet et al., 2019; Hengrung et al., 2015; Ng et al., 2012a; Ng et al., 2009; Peng et al., 2019; Pflug et al., 2014; Reich et al., 2014; Serna Martin et al., 2018). Despite the success in resolving the structures of viral polymerase and NP, high-resolution structural studies on intra-RNP interactions, specifically between viral polymerase and NP, and between NP and RNA, are missing.

A recent cryo-EM study on RNPs provided structural insight into the NP filament of an RNP, by resolving the helical part of vRNP to a resolution between 7 and 25 Å (Coloma et al., 2020). Nevertheless, structural studies on the RNP-associated viral polymerase, and its surrounding NPs, remain scarce. The struggle in resolving the interactions between viral polymerase and

NP is caused by a combination of factors. For instance, distinguishing viral polymerase from NP in cryo-EM data is challenging because the morphology of a viral polymerase appears similar to multiple NPs to human eyes, as well as to particle picking software. Moreover, each electron micrograph of an RNP sample reveals a low number of RNP-associated viral polymerases because there is only one viral polymerase per RNP. Consequently, a substantial number of electron micrographs is required to acquire a meaningful dataset for the RNP-associated viral polymerase. Lastly, the interaction between the viral polymerase and NP is likely to be conformationally heterogeneous, especially during RNA synthesis. Flexibility within a protein complex poses challenges in structural biology, due to the requirement to computationally average many thousands of particles with near identical three-dimentional structures during data processing to obtain a density map with sufficient structural information to produce a model of the protein complex. Moreover, it has been unsuccessful to resolve more than 3 nt of RNA in the RNA-binding grooves of NP bound to various length of RNA fragments (Chenavier et al., 2023; Tang et al., 2021). This obstacle is rooted in the fact that the wild type NP engages in oligomer formation through a flexible head-to-tail interaction, generating inherent flexibility which is unfavourable for structural analyses. However, employing a monomeric NP does not resolve the issue, as each NP monomer (56 kDa) is excessively small, which makes image alignment infeasible.

All of the above challenges are common to structural biologists, and the field has been actively exploring strategies to overcome them. One approach is to use antibody derivatives such as megabodies, or NabFab antibodies, to form stable and rigid complexes with the target protein. These antibody derivatives can increase the size of target proteins and can serve as references to define features for accurate image alignment (Bloch et al., 2021; Coupland et al., 2021; Uchański et al., 2021).

The goal of this chapter is to develop methods to study intra-RNP interactions. First, I generated megabodies as structural tools, using previously characterized nanobodies against viral polymerase or NP. I then attempted to form a protein-RNA complex comprising one molecule of each of viral polymerase, RNA, and NP to simplify the system for cryo-EM imaging. Finally, I investigated whether megabody-bound vRNPs could be studied using electron microscopy.

5.2 RESULTS

5.2.1 Cloning, expression and purification of megabodies

To facilitate the study of viral polymerase-NP and NP-RNA interactions, I generated megabodies by grafting nanobodies onto a megabody protein scaffold using a published method (Uchański et al., 2021). In summary, 17 megabody constructs were created by cloning nanobody chain B-G, containing CDR1-CDR3, into an *E. coli* expression vector (pMESP23-NO) with a nanobody chain A and a megabody scaffold sequence (Figure 5.1A). The larger of the two megabody scaffolds (namely cYgjKNO) was selected from Uchanski et al. 2021, with a size of 87 kDa and a distinctive question-mark shape under cryo-EM. Among these constructs, 12 megabodies were cloned from nanobodies against viral polymerase (used in chapter 4), and five megabodies were cloned from nanobodies targeting influenza A virus NP (characterized in Hanke et al., 2016 (Hanke et al., 2016) (See Table 1 summary). I purified all 17 megabody constructs at various yields but will focus on four constructs in the remaining part of my thesis: NP megabody MbNP1, and viral polymerase megabodies Mb8192, Mb8206, and Mb8207 (Table 1).

All megabodies have a PeLB signal that facilitates their secretion into the bacterial periplasm, allowing for purification via osmotic shock. To purify megabody proteins, plasmids containing megabody genes were transformed into WK6 cells — an *E. coli* strain that can recognise the PeLB signal and secrete the PeLB-megabody into the periplasm. Various optimisation steps were carried out for the megabody expression and purification process (results not shown). I found that cultures derived from the glycerol stock of the transformed bacteria exhibited significantly reduced megabody expression. Two other *E. coli* strains (BL21 or C41) that also

recognise the PeLB signal were tested but were shown to be less efficient at producing megabodies in my set up. I also investigated the effect of induction temperature on megabody expression and purification. I found that although less megabody was expressed post-induction at 18°C compared to at 28°C, similar yields of megabody proteins could be obtained after purification. This suggests that megabodies synthesised at 18°C have better solubility or stability than that synthesised at 28°C. As a control, plasmids containing only the megabody scaffold were transformed in parallel. A comparison of the pre- and post-IPTG induced bacterial samples using SDS-PAGE analysis revealed a good expression level of the megabody scaffold but highly varied expression levels across different megabody constructs (Figure 5.1B). Osmotic lysis of the bacterial periplasm could produce megabodies with less contamination; however, in order to increase yield, I fully lysed bacterial cells using lysozyme, which released megabodies from both the cytoplasm and periplasm. Ni-NTA agarose was used to specifically bind to and immobilise His-tagged megabodies. These bound megabodies were eluted, concentrated and buffer-exchanged using a 10 kDa cut-off concentrator and loaded onto a Superdex 200 SEC column in the megabody buffer (Figure 5.1C). The peak eluting at around 13.5 ml represents my purified megabody (Figure 5.1D), which was concentrated and stored at -80°C. However, some megabodies aggregate in solution, resulting in multiple peaks in SEC analysis and a massively decreased yield (Table 1). Lastly, intact mass spectrometry was performed on the purified megabodies by Dr Rod Chalk at the Mass Spectrometry facility to confirm the identity of each megabody according to their predicted mass (results not shown).

Intriguingly, I observed significant variations in the yields of megabodies among the different constructs. The yield of each megabody construct, which was reflected by the UV absorbance levels during SEC analysis, roughly corresponds to the quantity of full-length megabody present in the cell lysate, as demonstrated in the SDS-PAGE analysis. In certain constructs, a

reduction in the signal of full-length megabody is accompanied by an increase in proteins smaller than the expected size of a megabody (Figure 5.1B-C). These findings collectively suggest that the differences in megabody yields may be attributed to varying degrees of protein degradation experienced by each megabody construct.

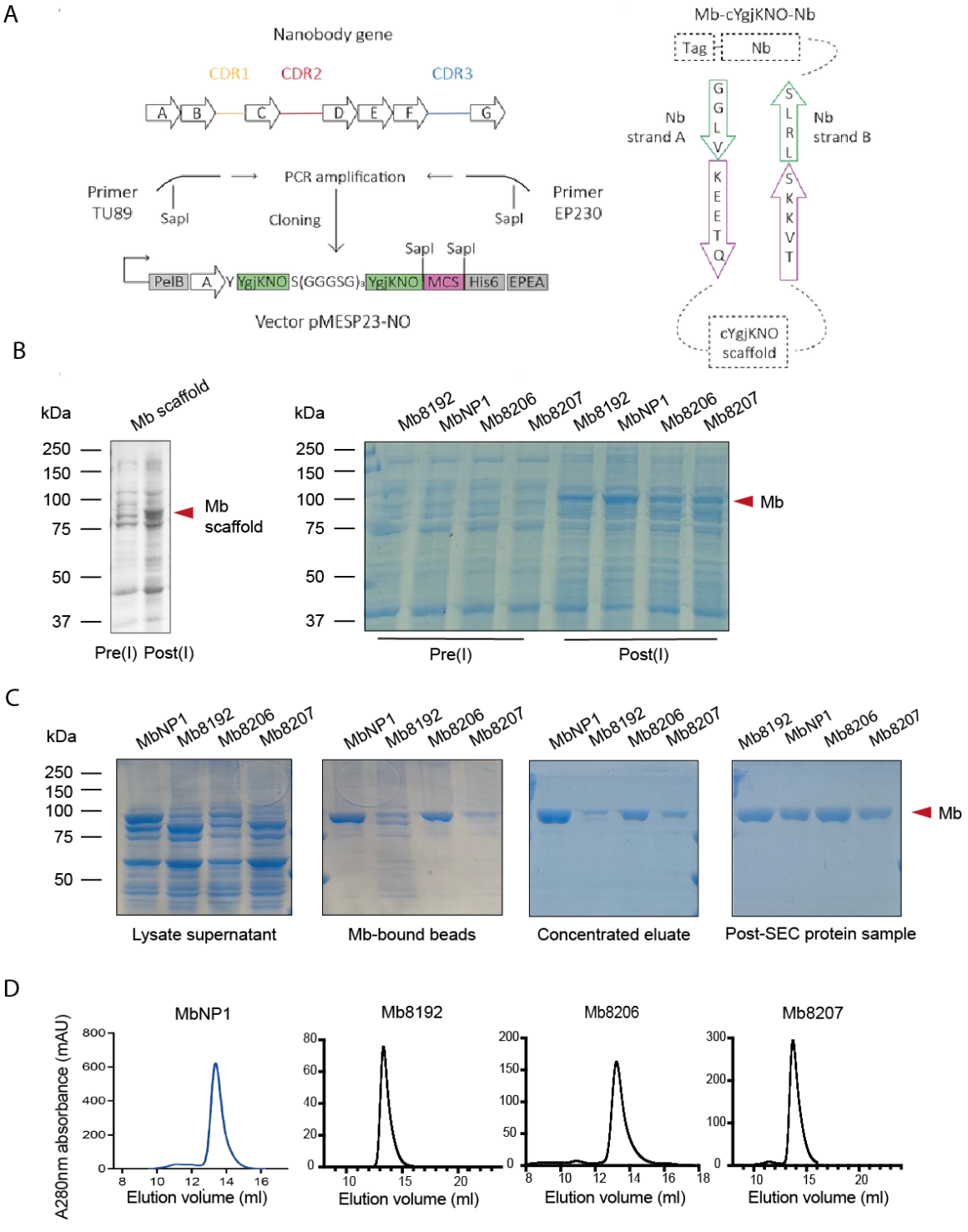


Figure 5.1 Cloning, expression and purification of megabodies. *A. Schematic diagram illustrating the process of grafting nanobody (Nb) chain B-G onto megabody (Mb) scaffold cYgjKNO in vector pMESP23-NO by restriction enzyme-based cloning (left). The amino acid sequence at the nanobody-megabody scaffold junction showing the two antiparallel ß-strands (right). Adapted from Uchanski et al. 2021. B. SDS-PAGE analysis of the pre (Pre(I))- and post- (Post(I)) induction WK6 cells expressing megabody (Mb) scaffold (87 kDa) (left) and MbNP1, Mb8192, Mb8206 and Mb8209 (appx. 101 kDa) (right). C. SDS-PAGE analysis of samples taken from different stages of megabody purification: supernatant of lysosomal cell lysate after centrifugation (lysate supernatant), megabody bound Ni-NTA beads (Mb-bound beads), concentrated and buffer exchanged eluted megabody sample from Ni-NTA beads (concentrated eluate) and post-SEC purification megabodies adjusted to the same concentration (post-SEC protein sample). D. SEC profiles of purified megabodies on a Superdex200 column. Megabodies elute at around 13.5ml.*

Megabody	*Target*	*Purification*	*Major SEC peak size (mAU)*	*Binding to ligand*
Mb8189	FluPol	Good	150	Yes
Mb8191	FluPol	Not good	20	Not tested
Mb8192	FluPol	Very good	400	Yes
Mb8193	FluPol	none	0	N/A
Mb8194	FluPol	Good	100	Yes
Mb8201	FluPol	Not purified	NA	N/A
Mb8202	FluPol	Bad, aggregate	10	Unclear
Mb8204	FluPol	Bad, aggregate	10	Unclear
Mb8205	FluPol	Not purified	NA	N/A
Mb8206	FluPol	Good	150	Yes
Mb8207	FluPol	Very good	300	Yes
Mb8210	FluPol	Bad, aggregate	30	Unclear
MbNP1	NP	Ok	50	Yes
MbKV052	NP	Very good	600	Yes
Mb151	NP	Good	150	Yes
Mb170	NP	Good	120	Yes
Mb495	NP	Good	100	Yes

Table 5.1 Summary of megabody constructs generated in this study. *Information on ligand for the nanobody part, purification, yield per1 L of bacterial culture in UV absorbance (mAU) on Superdex 200 column, and binding properties to ligands.*

5.2.2 Binding properties of megabodies

Following the purification of the megabodies, I investigated their ability to bind to the antigens for which their nanobody components were raised (See Table 1 summary). To assess this, I incubated selected megabodies with either purified NP or viral polymerase on ice, and then examined complex formation using SEC analysis (Figure 5.2A).

For testing the NP megabody MbNP1, I used an influenza A virus R416A monomeric mutant NP (mNP) because wild type NP forms oligomers in solution, which hinders both SEC analysis and cryo-EM study. A readily available short RNA, namely the 3' vRNA promoter, was added to test whether megabody-bound NP retains its ability to bind to RNA. I showed that purified MbNP1 alone and RNA-bound mNP each elute as a single peak in SEC, at around 13.5 ml and 14.5 ml, respectively. Mixing MbNP1 with RNA-bound mNP at an approximately 1:1.2 ratio resulted in the diminishment of both the MbNP1 peak and the RNA-bound mNP peak, produced a new peak at around 12 ml (Figure 5.2B). SDS-PAGE analysis of different SEC fractions of the MbNP1 + mNP + 14 nt RNA sample revealed the co-presence of megabody and mNP in the newly appeared peak (Figure 5.2C). I then concentrated the newly appeared peak and measured the 260/280 ratio of the concentrated sample using Nanodrop. The increase in 260/280 ratio from around 0.6 (average of purified MbNP1 and mNP) to above 0.9 indicated the presence of RNA in the newly appeared peak. These results suggest that MbNP1 binds to RNA-bound mNP in solution. Additional tests using apo mNP, and C-terminal tail-deleted monomeric NP (mNPdC) with or without RNA, demonstrated that MbNP1 binds to NP in all tested cases (results not shown).

Similarly, I tested the binding properties of the polymerase megabodies Mb8192, Mb8206 and Mb8207. Using SEC analysis, I showed that the polymerase megabodies each elute as a single peak at 13.5 ml and that vRNA promoter-bound 1918 viral polymerase alone appears as two peaks at around 9.5 ml and 12 ml, representing the tetrameric polymerase peak and the highly overlapped dimeric and monomeric polymerase peaks, respectively (Figure 5.1 D, Figure 5.3A). I then demonstrated that mixing Mb8192 and Mb8207 with the promoter-bound viral polymerase at a 2:1 ratio resulted in two new peaks, eluting at around 8.5 ml and 10 ml; but only one new peak at around 10 ml in the case of Mb8206. The appearance of new peaks was accompanied by the simultaneous disappearance of the viral polymerase and the megabody peaks (Figure 5.3A-D). I then analysed the selected SEC fractions by SDS-PAGE analysis. As the megabody runs at a similar position as the PB1 subunit on an SDS-PAGE gel, it was challenging to resolve them on an SDS-PAGE gel. However, I observed that in the fractions that correspond to the newly appeared peaks, the intensity of the top band increased to a similar level as the bottom band in a viral polymerase complex duplex. This can serve as an indicator of the successful formation of the megabody-polymerase complex for all three polymerase megabodies.

In summary, my results confirmed that MbNP1, Mb8192, Mb8206, and Mb8207 form complexes with their ligands in solution under the tested buffer conditions.

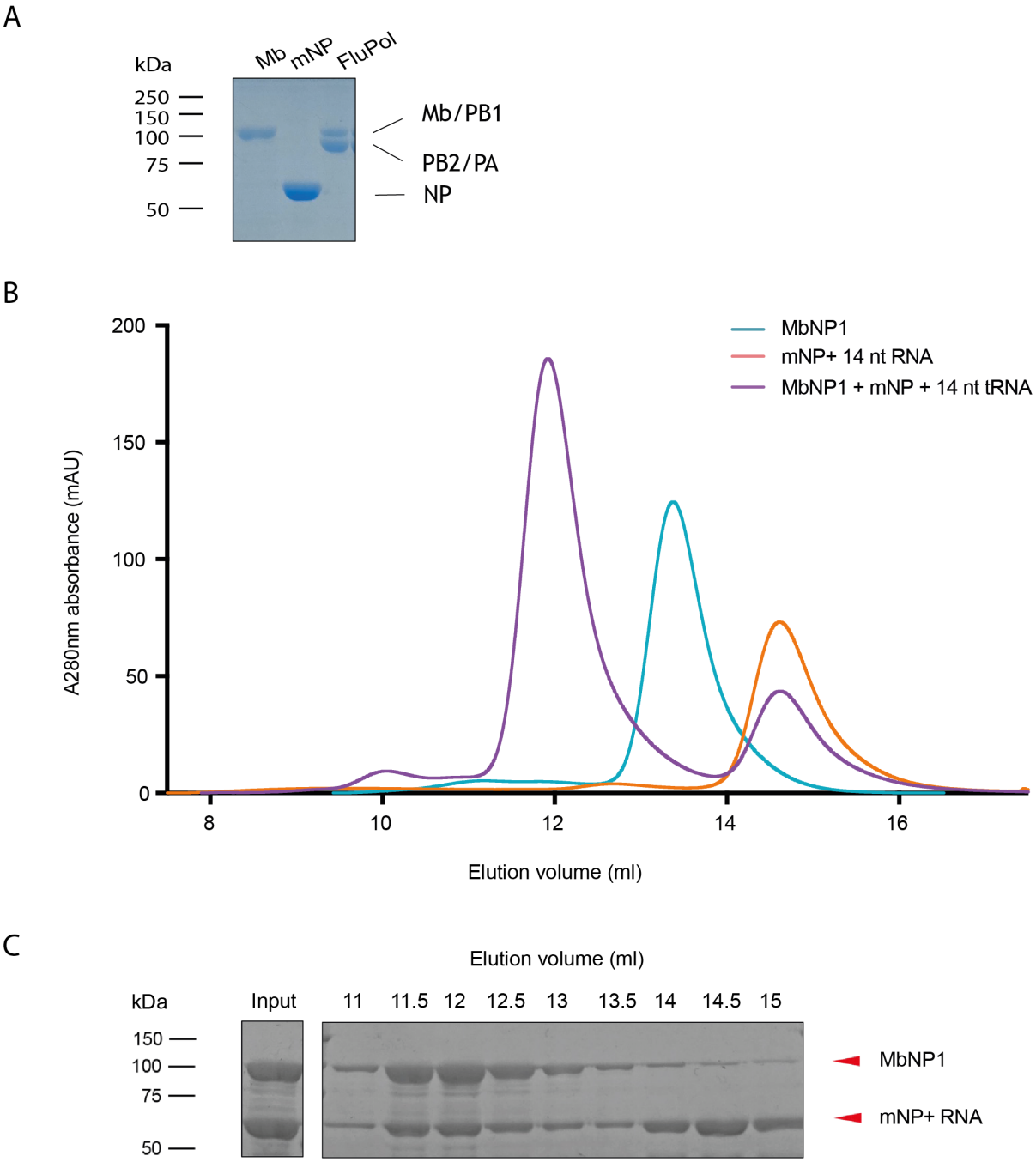


Figure 5.2 Binding of megabodies to NP. *A. SDS-PAGE analysis of purified megabodies (Mb), influenza A virus monomeric NP mutant (mNP) and influenza A virus polymerase (FluPol) used in subsequent complex formation. B. Elution profiles of SEC analysis on mNP with RNA (mNP + 14 nt RNA, orange), NP megabody (MbNP1, cyan), or RNA-bound mNP in complex with NP megabody (MbNP1 + mNP + 14 nt RNA, purple) on a Superdex 200 column. C. SDS-PAGE analysis of the SEC fractions of the MbNP1 + mNP + 14 nt RNA sample.*

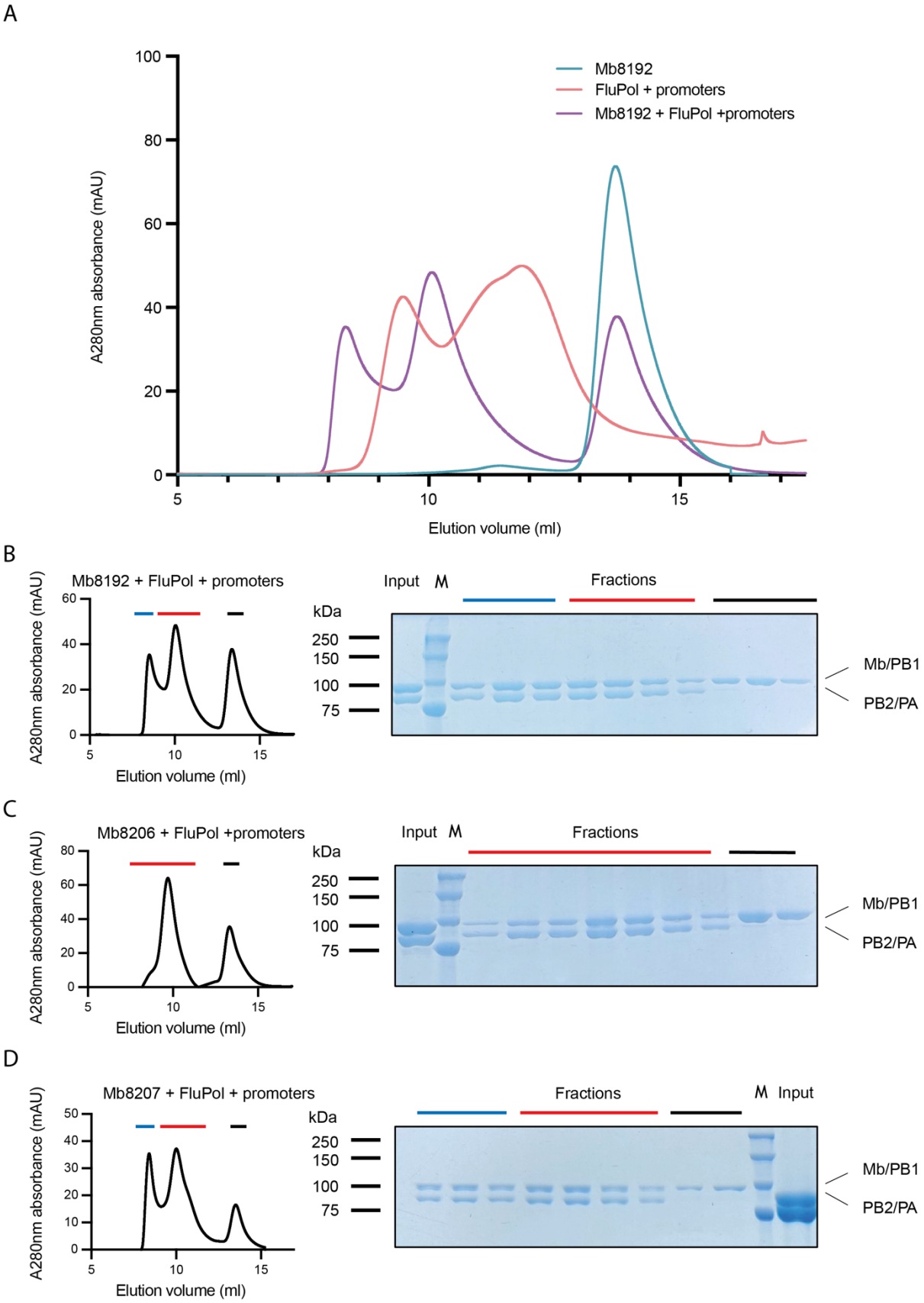

A
Mb8192
FluPol + promoters
Mb8192 + FluPol +promoters
A280nm absorbance (mAU)
Elution volume (ml)
B
Mb8192 + FluPol + promoters
Input
M
Fractions
kDa
250
150
100
75
Mb/PB1
PB2/PA
C
Mb8206 + FluPol +promoters
D
Mb8207 + FluPol + promoters

Figure 5.3 Binding of megabodies to viral polymerase. *A. Elution profiles of SEC analysis on influenza A viral polymerase with vRNA promoter (FluPol + promoter, orange), FluPol megabody (Mb8192, cyan), or promoter-bound viral polymerase in complex with FluPol megabody (Mb8192 + FluPol + promoter, purple) on a Superdex200 column. B-D. Elution profile of SEC analysis of promoter-bound FluPol in complex with Mb8192 (B), Mb8206 (C) or Mb8207 (D) (left). SDS-PAGE analysis of the indicated SEC fractions from the elution profiles (blue, red and black lines) (right).*

5.2.3 Cryo-EM analysis of the megabody-polymerase and the megabody-NP complexes

After successfully forming the megabody-polymerase and the megabody-NP complexes in solution, I applied the samples to EM grids to prepare for resolving structures using cryo-EM. EM data collection was graciously performed by Dr Haitian Fan, Dr Jeremy Keown, or Dr Loïc Carrique at Division of Structural Biology (STRUBI). In the 2D classification of the MbNP1-mNP sample, I observed kidney-shaped particles, comprising a smaller lobe and a larger lobe. Using cryoSPARC, I generated an *ab initio* model based on the 2D classes. After further modification by heterogeneous-refinement and non-uniform refinement, I obtained an Coulomb potential map (Figure 5.4A). By fitting the models of a monomeric NP (56 kDa), a megabody (100 kDa), or a megabody-mNP complex (156 kDa) into the map, I found that the larger lobe fits with the megabody scaffold and the smaller lobe can accommodate a nanobody. NP could not be fitted into the map. Therefore, the particles observed in 2D classification are most likely MbNP1 alone (Figure 5.4B). I also tested other NP megabodies, MbKV052 and Mb495, which were also shown to bind to RNA-bound mNPA in solution; however, again, only the megabody alone was detected in all tested cases (results not shown). It is likely that the megabody-mNP complex dissociates during cryo-EM sample preparation, resulting in

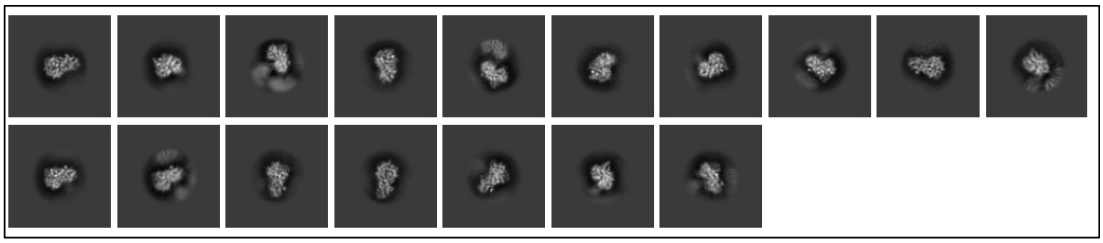
separated megabodies and mNP particles, and the dissociated-mNP are not large enough to be resolved. To combat this, I attempted to use pegylation or gradient fixation (Grafix) to stabilize the Mb-mNP complex (Kastner et al., 2008; Roberts et al., 2012). However, pegylation led to the loss of NP in the post-pegylation sample, as shown in the SDS-PAGE analysis, and the 2D classification of the post-Grafix sample showed no significant improvement (Figure 5.4C-E).

Next, I investigated the megabody-viral polymerase complexes in the presence of vRNA promoter using cryo-EM, with data collected by Dr Haitian Fan, Dr Jeremy Keown, or Dr Loïc Carrique at STRUBI. Notably, while approximately a third of the viral polymerase bound by Mb8192 and Mb8207 were in dimeric state, nearly all viral polymerase bound by Mb8206 was in monomeric state. This can be explained by Mb8206 binding to the viral polymerase dimerisation interface and blocking symmetric dimerisation which occurs readily in solution. It also aligns with my SEC analysis result (Figure 5.3C), where Mb8206-bound viral polymerase appears as a single monomeric peak at around 10 ml. On the micrograph after motion correction, I observed clear viral polymerase particles with good contrast and density. Excitingly, I also observed small and dense dots associated to the polymerase particles, in zoomed-in micrographs. These black dots were not observed in viral polymerase only samples, and were likely attached megabodies (Figure 5.5A). Indeed, 2D classification revealed that more than 50% of viral polymerase particles in a monomeric state have an extra question-mark shape density attached, resulting in a chicken-like shape. Based on the 2D classes of the monomeric viral polymerase with extra density, I used cryoSPARC to carry out *ab initio*, heterogeneous-refinement and non-uniform refinement and obtained an average electron density map for each of the complexes. To fit the viral polymerase and megabody into the Coulomb potential maps, I generated models of the viral polymerase with each of the three polymerase megabodies attached, based on the nanobody binding information. The megabody-

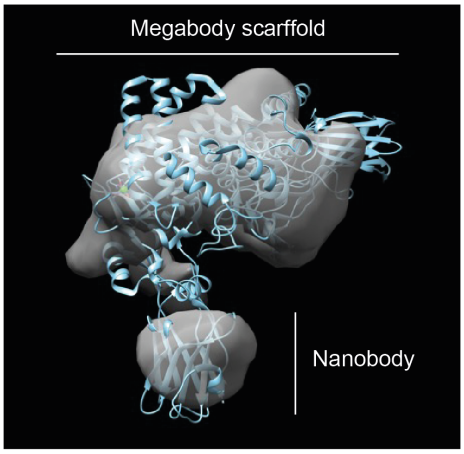
viral polymerase model could be fit into their respective Coulomb potential maps, with the chicken head being the megabody scaffold, the chicken neck being the nanobody joint, and the chicken body being the viral polymerase (Figure 5.5B-D).

In conclusion, I demonstrated that megabodies and viral polymerase forms a stable complex in solution and under cryo-EM. However, while RNA-bound mNP forms a stable complex with megabodies in solution, the complex is unstable under cryo-EM.

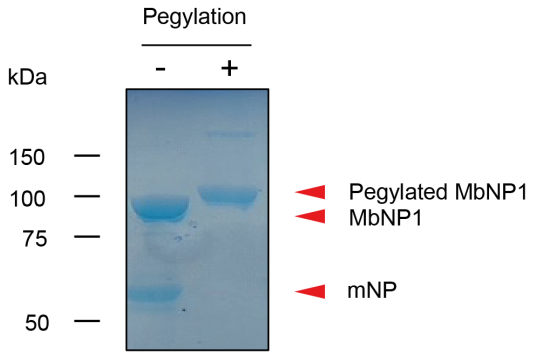
A



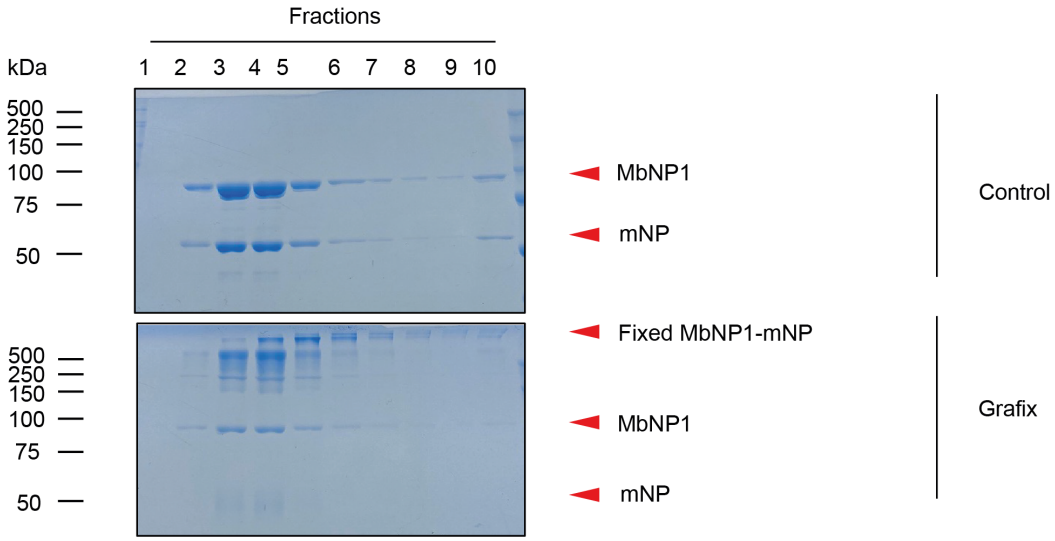
B


Megabody scarffold
Nanobody

C


Pegylation
-
+
kDa
150
100
75
50
Pegylated MbNP1
MbNP1
mNP

D


Fractions
kDa
1 2 3 4 5 6 7 8 9 10
500
250
150
100
75
50
MbNP1
mNP
Control
Fixed MbNP1-mNP
MbNP1
mNP
Grafix

E

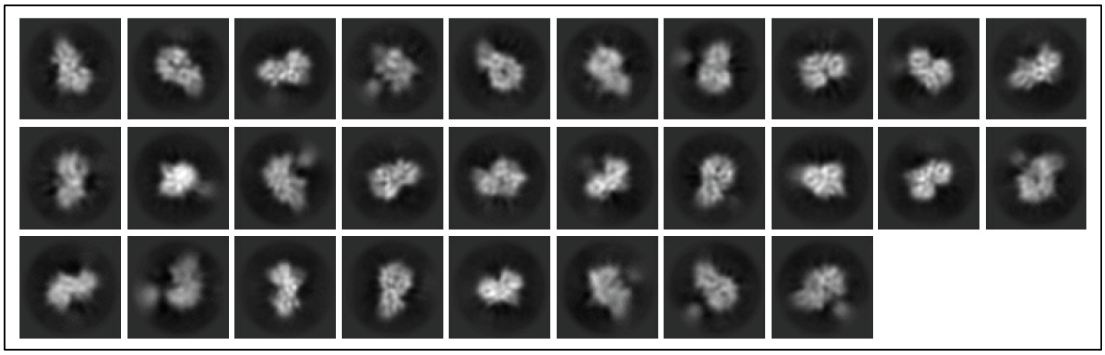
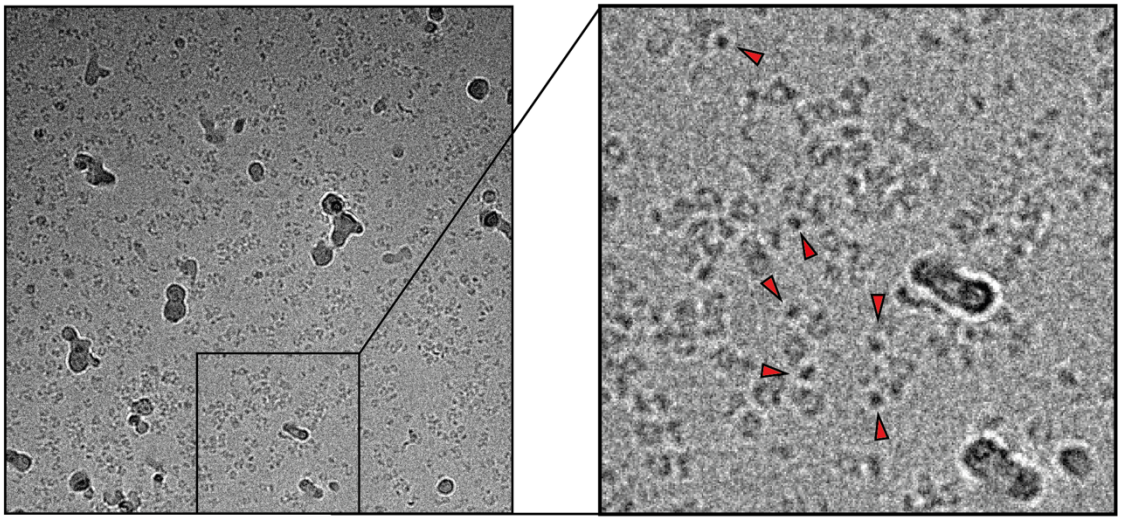


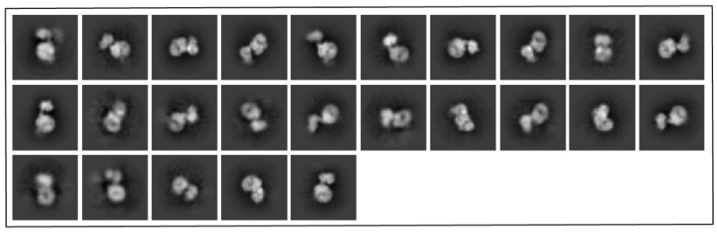
Figure 5.4 Cryo-EM micrographs and initial structural model of the megabody-NP samples. *A. Representative 2D class averages of the cryo-EM data set of RNA-bound mNP in complex with MbNP1. B. Modelled structure of megabody alone was fitted into the electron density map obtained from the 2D class after 3D reconstruction, with the top major density being the megabody scaffold and the bottom minor density being the grafted nanobody. C. SDS-PAGE analysis of RNA-bound mNP in complex with MbNP1 before and after pegylation. D. SDS-PAGE analysis of different fractions from Grafix on RNA-bound mNP in complex with MbNP1 in the absence (control) and presence of glutaraldehyde (Grafix). E. Representative 2D class averages of the cryo-EM data set of post-Grafix sample.*

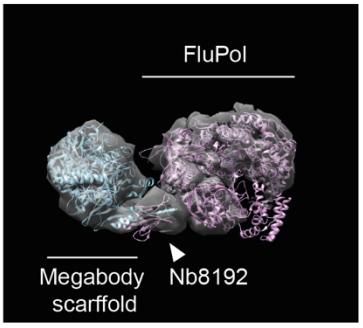
A



B

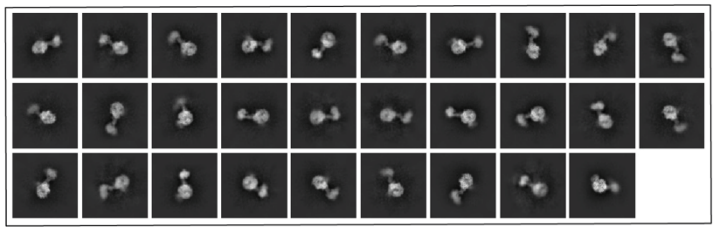
Mb8192 + FluPol + promoters

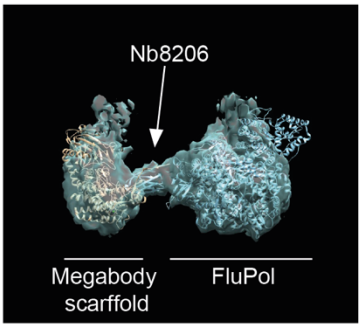




C

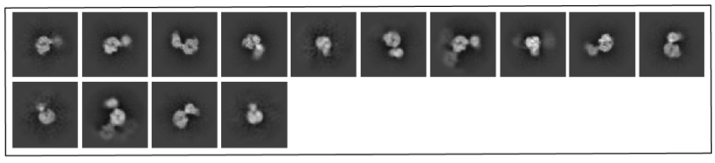
Mb8206 + FluPol + promoters





D

Mb8207 + FluPol + promoters



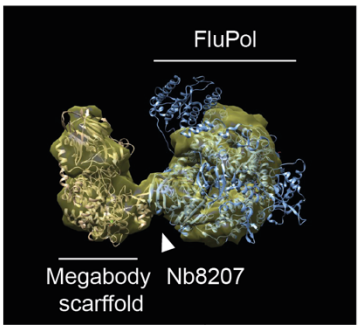


Figure 5.5 Cryo-EM micrographs and structural models of the megabody-viral polymerase samples. *A. Representative motion corrected micrograph of the vRNA promoter-bound viral polymerase in complex with FluPol megabody (left), and its zoomed-in view of the highlighted box (right). Red arrows point towards some of the individual FluPol particles with megabody (dense black dot) attached. B-D. Representative 2D class averages of the cryo-EM data set of the promoter-bound viral polymerase in complex with FluPol megabody Mb8192 (B), Mb8206 (C) and Mb8207 (D) (left). Modelled structures of FluPol bound by each megabody at the specific site fitted into the electron density map obtained from the 2D class after 3D reconstruction, with the larger density being FluPol and the smaller density being the megabody scaffold and the connecting density being the nanobody.*

5.2.4 In vitro assembly of a polymerase-NP complex

During the structural study of viral polymerase and its proximal NP within an RNP, I initially aimed to create a simpler system by assembling a complex composed of a viral polymerase and a single nucleoprotein in the presence of vRNA promoter. I designed a 39 nt 5' vRNA promoter mimic RNA, with a 14 nt promoter region and an additional 25 nt that allows for NP association. To prevent NP oligomerisation during complex assembly, I used the monomeric influenza A virus NP (mNP), which appears as a single monomeric peak in SEC analysis.

To determine the conditions under which the 39 nt RNA binds to mNP, I conducted SEC analysis of mNP in the presence of RNA in either 150 mM NaCl or 500 mM NaCl buffers. An increase in salt level typically disrupts intermolecular interactions. In both salt conditions, two peaks were observed in the elution profile. I analysed the protein composition in individual peak using SDS-PAGE and concentrated the sample to check for the presence of RNA by measuring the 260/280 ratio using Nanodrop. The results demonstrated that in the 150 mM NaCl buffer, both peaks, eluted at P1 = 12 ml and P2 = 13.5 ml, contained NP and RNA (Figure 5.6C red line, Figure 5.6G). However, in the 500 mM NaCl buffer, mNP and RNA eluted

separately at P3 = 14.5 ml and P4 = 16 ml, respectively, on a Superdex 200 column (Figure 5.6C blue line, Figure 5.6G). Given that in an RNP, a single NP associates with an average of 24 nt (Ye et al., 2006), I speculate that P2 represents a complex of a single mNP and RNA, while P1 represents a complex of two mNP associated to one piece of 39 nt RNA. Unfortunately, wild type influenza A virus polymerase is unstable in the buffer condition required for NP-RNA interaction, i.e. 150 mM NaCl (Figure 5.6D). This is possibly because the viral polymerase readily forms a dimer, and aggregates and precipitates out of solution in low salt buffer.

While wild type influenza A viral polymerase (FluPolA) is unstable in 150 mM NaCl buffer in the presence of vRNA promoter, I discovered that promoter-bound influenza A virus monomeric mutant polymerase (mFluPolA) (introduced in Chapter 2) is relatively stable under this condition. I showed that the addition of RNA stabilizes mFluPolA in low salt but also causes mFluPolA to dimerise, forming two peaks which elute at 10 ml and 11 ml (Figure 5.6F purple and yellow lines). These two peaks correspond to the dimer and the monomer fractions of mFluPolA on the Superdex 200 column, respectively. Furthermore, the addition of both 5' and 3' vRNA promoter to NP also leads to the oligomerisation of mNP (Figure 5.6F cyan line). Next, I incubated mFluPolA with RNA promoter before adding an excess of mNP and performing SEC analysis. I observed three major peaks in the complex formation elution profile, eluting at P1 = 10 ml, P2 = 11 ml, and P3 = 15 ml, respectively (Figure 5.6F, red line). SDS-PAGE analysis revealed that both P1 and P2 contained mFluPolA, promoter, and a small but detectable amount of NP (Figure 5.6H). In contrast, P3 eluted at the same volume as the apo mNP and only contained mNP (Figure 5.6F red and yellow lines). Promoter-bound mFluPolA alone and RNA-bound NP oligomer alone eluted at the same elution volume as P2. Therefore, the presence of proteins and RNA in P2 could be the result of the co-elution of

promoter-bound mFluPolA and RNA-bound mNP rather than successful complex formation. It is possible, yet less likely, that the presence of proteins and RNA in P1 is also due to co-elution because RNA-bound NP oligomer does not elute at 10 ml. Therefore, it is possible that mFluPolA-mNP complex has formed in the presence of promoter in P1. Unfortunately, the concentrated sample was not of sufficient quality for cryo-EM analysis.

I also attempted to form FluPol-NP complex using influenza C viral proteins because influenza C viral polymerase (FluPolC) is relatively stable in 150 mM NaCl and appears as monomer in solution. To form the complex, I generated monomeric influenza C viral NP (mNPC) by introducing R427A and E354A mutations, equivalent to R416A and E339A mutations in influenza A viral NP based on sequence alignment. In the absence of RNA, the mixture of FluPolC and mNPC resulted in two peaks eluting at the same volumes as apo FluPolC and mNPC, 11.5 ml and 13.5 ml, respectively. Although the elution volumes of the two peaks slightly overlapped, SDS-PAGE analysis revealed that these peaks mainly comprised viral polymerase and NP, respectively (Figure 5.6E yellow line, Figure 5.6H). I then demonstrated that the addition of the 39 nt 5' vRNA promoter to influenza C viral polymerase resulted in an additional peak of FluPolC dimer, which elutes around 9.5 ml (Figure 5.6E purple). On the other hand, the addition of the 5' vRNA promoter to mNPC resulted in two peaks at 12 ml and 9.5 ml, representing either one or two NP molecules associated with one molecule of RNA, respectively (Figure 5.6E cyan line). Finally, I incubated FluPolC with the 5' vRNA promoter before adding an excess of mNPC and analysing by SEC. The elution profile of the complex formation sample contains three major peaks: P1 = 9.5 ml, P2 = 11.5 ml, and P3 = 13.5 ml. Further analysis revealed that P1 and P2 both contained all three necessary components for complex formation, with viral polymerase and NP at different ratios (Figure 5.6H). However, since RNA-bound FluPolC and RNA-bound mNPC alone also elute at these two positions, I

cannot rule out the possibility that P1 is simply a result of co-elution of FluPolC-RNA and mNPC-RNA.

In conclusion, it remains unclear whether a complex of the viral polymerase, NP, and the RNA promoter was formed. This is primarily because viral polymerase and NP both oligomerise in the presence of promoter RNA, making it challenging to perform complex analysis by SEC.

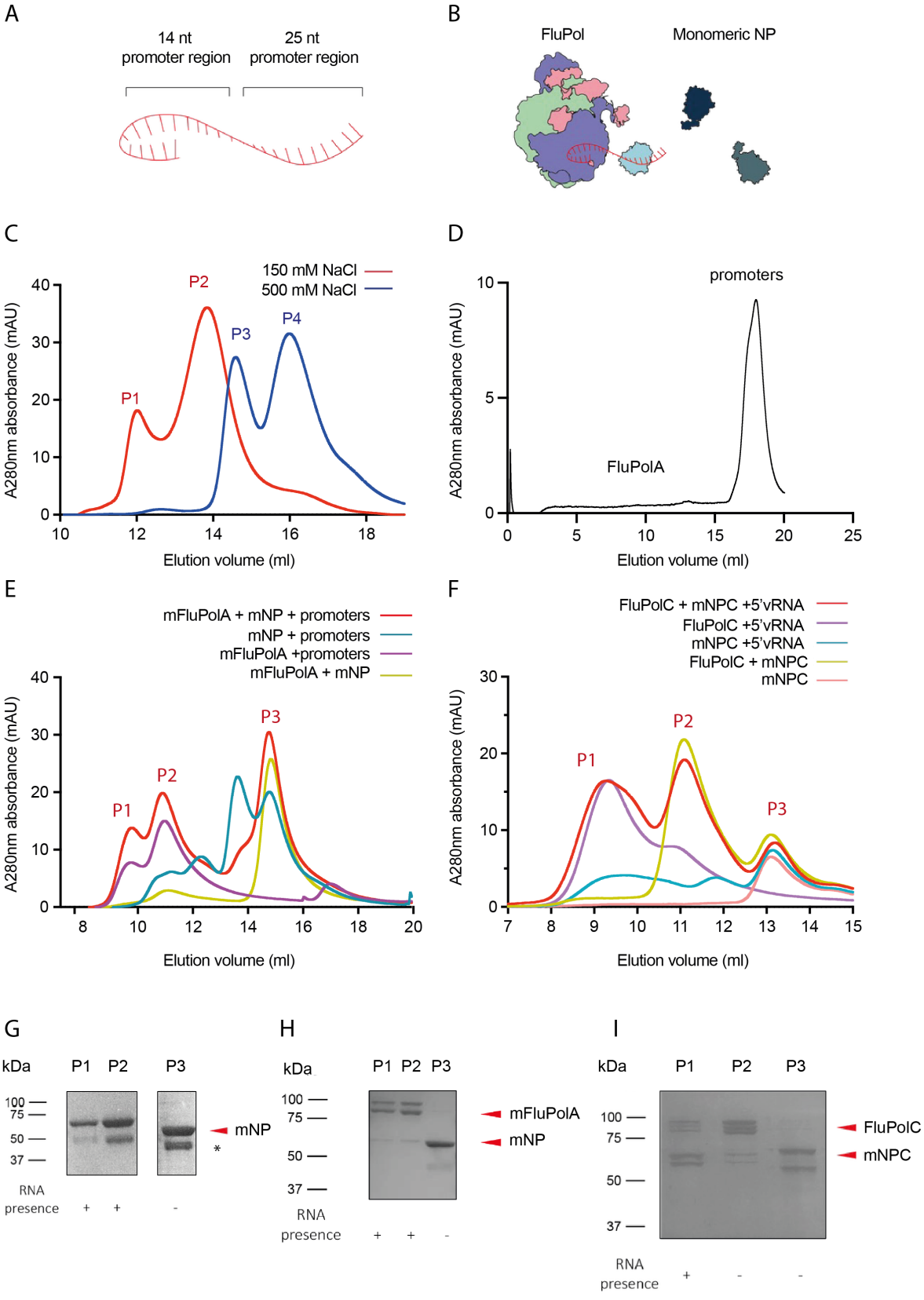

A
14 nt
promoter region
25 nt
promoter region
B
FluPol
Monomeric NP
C
150 mM NaCl
500 mM NaCl
P1
P2
P3
P4
A280nm absorbance (mAU)
Elution volume (ml)
D
promoters
FluPolA
A280nm absorbance (mAU)
Elution volume (ml)
E
mFluPolA + mNP + promoters
mNP + promoters
mFluPolA +promoters
mFluPolA + mNP
P1
P2
P3
A280nm absorbance (mAU)
Elution volume (ml)
F
FluPolC + mNPC +5'vRNA
FluPolC +5'vRNA
mNPC +5'vRNA
FluPolC + mNPC
mNPC
P1
P2
P3
A280nm absorbance (mAU)
Elution volume (ml)
G
kDa
P1
P2
P3
100
75
50
37
mNP
*
RNA
presence
+
+
-
H
kDa
P1
P2
P3
100
75
50
37
mFluPolA
mNP
RNA
presence
+
+
-
I
kDa
P1
P2
P3
100
75
50
37
FluPolC
mNPC
RNA
presence
+
-
-

Figure 5.6 In vitro assembly of a viral polymerase-NP complex for structural analysis. *A. Schematic diagram of the 39 nt 5' vRNA promoter designed for the complex formation. B. Schematic diagram of the complex formation using mNP. C. SEC elution profiles of the monomeric mutant influenza A viral NP (mNP) mixed with equal molar ratio of 39 nt 5' vRNA promoter in buffers with either 150 mM NaCl (red) or 500 mM NaCl (blue). D. SEC elution profiles of wild type influenza A viral polymerase (FluPolA) in the presence of excess vRNA promoter in a buffer with 150 mM NaCl. There was no peak for FluPolA due to it being unstable in 150 mM NaCl. E. SEC elution profiles of monomeric influenza A viral NP (mNP) + monomeric influenza A viral polymerase (mFluPolA) (yellow), mNPA+ vRNA promoter (blue), mFluPolA + vRNA promoter (purple) and mFluPolA + vRNA promoter + NP (red). F. SEC elution profiles of monomeric influenza C viral NP (mNPC) (beige), mNPC + influenza C viral polymerase (FluPolC) (yellow), mNPC + 39 nt 5' vRNA promoter (blue), FluPolC + 39 nt 5' vRNA promoter (purple) and FluPolC + mNPC + 39nt 5' vRNA promoter (red). G. SDS-PAGE analysis of indicated peaks in Figure 5.6C and the presence of RNA based on the 260/280 values. H. SDS-PAGE analysis of indicated peaks (Figure 5.6E, red) and the presence of RNA based on the 260/280 values. I. SDS-PAGE analysis of indicated peaks (Figure 5.6F, red) and the presence of RNA based on the 260/280 values. *Possible NP degradation*

5.2.5 Electron microscopy analysis of megabody-bound vRNP

Given the challenges in developing a simplified system to study viral polymerase-NP and NP-RNA interactions, I shifted my focus to a more complex but biologically relevant system: virion-purified vRNPs. As previously observed, NP megabodies tend to dissociate from NP during cryo-EM sample preparation, while megabody-viral polymerase complexes can be successfully imaged under cryo-EM. I therefore started by labelling vRNPs with viral polymerase megabodies. To enhance the occupancy of megabodies on RNP-associated viral polymerase, I incubated virion lysate with an excess amount of three polymerase megabodies before separating vRNP from other viral proteins using a glycerol gradient (Figure 5.7A).

I analysed the glycerol gradient fractions and showed that, without the addition of megabodies, vRNP alone eluted in fraction 7 to fraction 10. This was shown on SDS-PAGE gel as a doublet of viral polymerase between 75 and 100 kDa, and a significant excess of NP at around 56 kDa (Figure 5.7B). The viral polymerase doublet typically appears as a top PB1 band and a bottom PA and PB2 band, at 1:2 ratio. Since megabody has the same mobility as the top PB1 band of the viral polymerase doublet, the addition of one megabody transformed the typical polymerase doublet into two bands, in a 1:1 ratio (Figure 5.7C). With the addition of three different megabodies, the top band thickness increased further, reflecting the elevated RNP-associated megabody to viral polymerase ratio. Western Blot confirmed the presence of his-tagged megabodies in fraction 7 to fraction 10 (Figure 5.7D). I therefore combined these fractions for structural studies.

I first imaged the megabody-bound vRNP sample, with three different megabodies added, using negative stain EM. Grid preparation and data collection were kindly performed by Dr Errin Johnson at the EM facility in the Sir William Dunn school of pathology. Here, I coupled the His-tagged megabodies with 5 nm-diameter Ni-NTA-Nanogold beads to facilitate visualization. In the micrographs, I observed that vRNPs were evenly distributed with occasional aggregations (Figure 5.8A). A majority (89.9%) of Nanogolds co-localised with a vRNP, indicating a robust association (Figure 5.8B). I then examined the number of viral polymerases, as the number of Nanogold hotspots, along each vRNP. I revealed that 85.2 % of vRNPs had one Nanogold hotspot, 10.7% lacked Nanogold association, 4.6% had two hotspots, and none had more than two (Figure 5.8C). This observation suggested that most vRNP likely contains only a single viral polymerase and each vRNP has a maximum of two viral polymerases. Notably, 90.3% of Nanogold hotspots were located at the very end of a vRNP (Figure 5.8D), suggesting the viral polymerase mostly resides at the RNP terminus. Lastly, I

counted the number of Nanogolds at each hotspot and found that 59.8%, 24.6%, 10.7%, 2.8%, 1.3% and 0.8% of hotspots contain one to five, or six or more Nanogolds (Figure 5.8E). The hotspots with more than three megabodies possibly represent megabody-bound polymerase dimers.

Cryo-EM grids were prepared using the same megabody-bound vRNP sample but without Nanogold, because Nanogold can create ripple effects around the bound protein by causing electron scattering and affect the image resolution in cryo-EM data processing. While helical filaments representing vRNPs were visible, megabodies were less obvious compared to the megabody-bound viral polymerase samples (Figure 5.8F). Small, dense dots, possibly representing megabodies, were scarcely visible (Figure 5.8G). Cryo-EM dataset processing encountered challenges due to the lack of obvious megabodies. Despite these hurdles, I have laid the groundwork by attaching 1–3 megabodies onto vRNP-associated viral polymerase at a high occupancy. The sample has been successfully imaged under negative-staining EM, and a large cryo-EM dataset has been collected. Although it remains challenging, further work will be carried out by my colleagues with advanced skills in cryo-EM data processing.

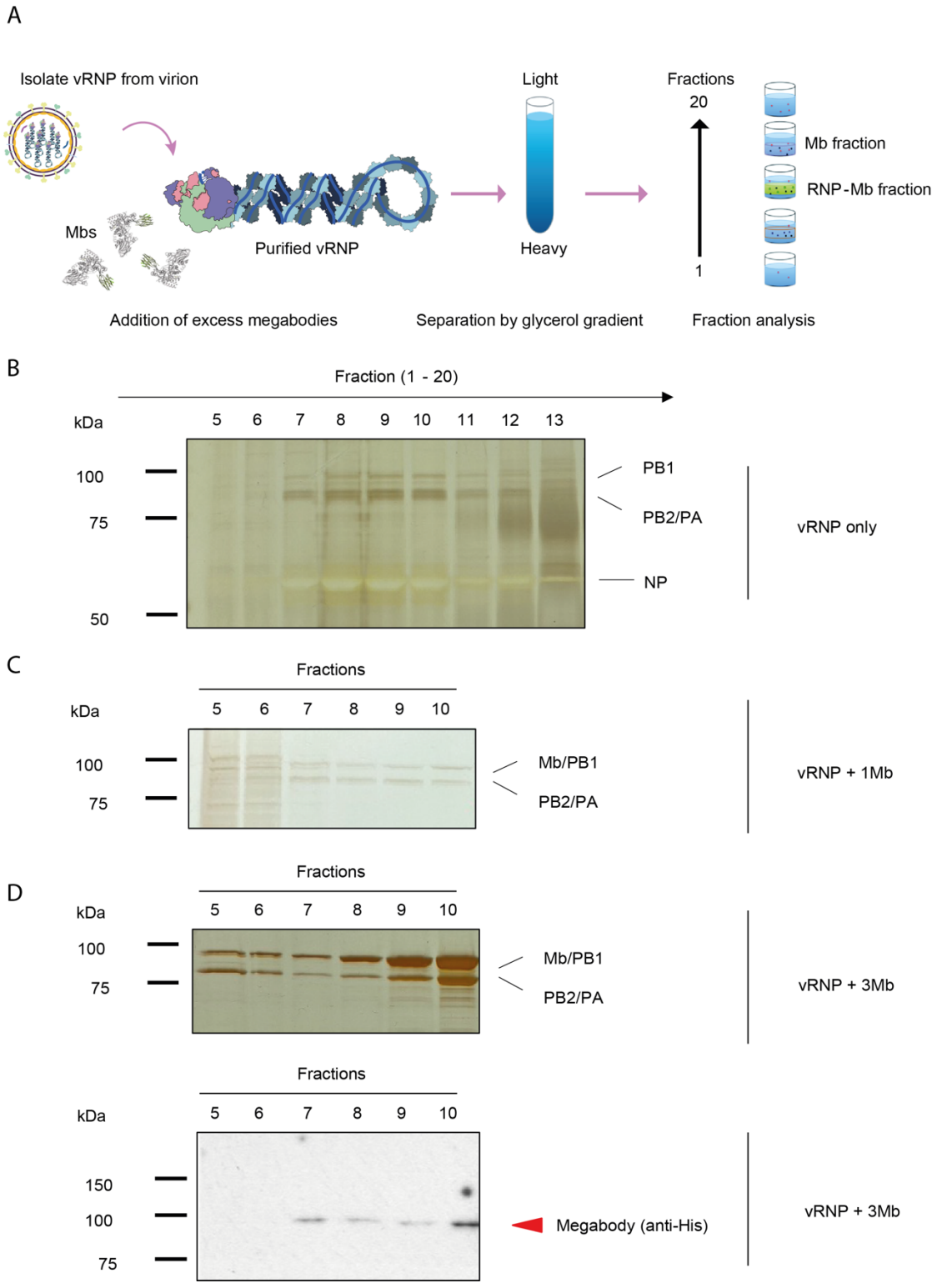


Figure 5.7 Binding of megabodies to virion-purified vRNP. *A. Schematic diagram showing the process of binding three different viral polymerase megabodies to virion purified vRNP before glycerol gradient centrifugation. B-D. SDS-PAGE analysis followed by either silver staining (B-D, top) or anti-His antibody Western Blot (D, bottom) of different fractions collected from glycerol gradient centrifugation of virion-purified vRNP only (B), vRNP with one megabody (C), or vRNP with three different megabodies (D).*

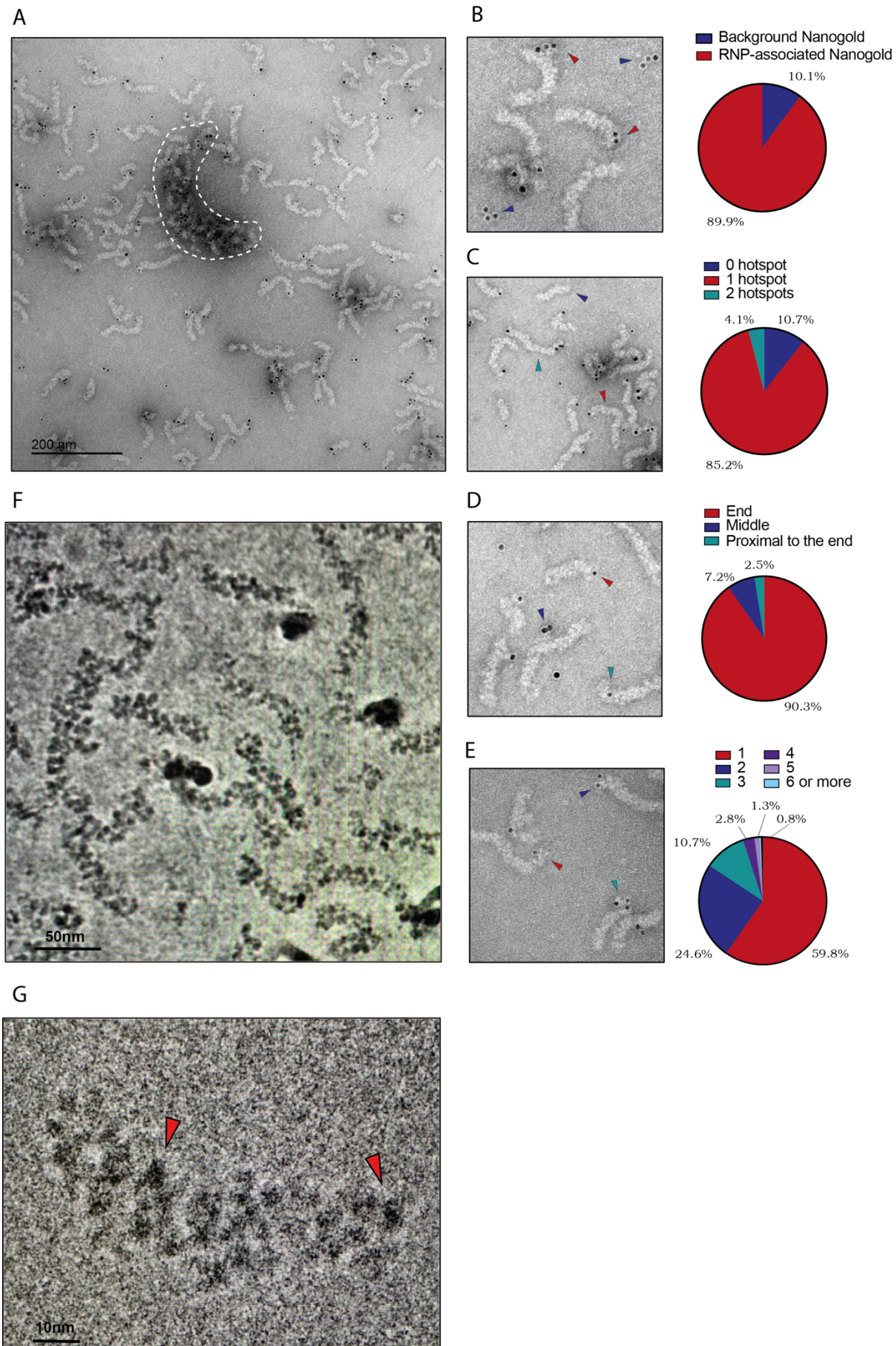

A
200 nm
B
Background Nanogold
RNP-associated Nanogold
10.1%
89.9%
C
0 hotspot
1 hotspot
2 hotspots
4.1%
10.7%
85.2%
F
50nm
D
End
Middle
Proximal to the end
2.5%
7.2%
90.3%
E
1
2
3
4
5
6 or more
1.3%
2.8%
0.8%
10.7%
24.6%
59.8%
G
10nm

Figure 5.8 Electron micrographs and analysis of the megabody-bound vRNP samples. *A. A representative negative staining electron microscopic micrograph of virion-purified vRNP with three megabodies added. White dotted line highlights vRNP aggregation which is not used in subsequential quantifications. B. Quantification of number of free (blue) or RNP-associated (red) Nanogold beads (right). N= 4384. Representative micrograph showing free (blue arrow) or RNA-associated (red arrow) Nanogold beads (left). C. Quantification of vRNP with 0 (blue), 1 (red), or 2 (cyan) Nanogold hotspots (right). N= 2271. Representative micrograph showing vRNP with 0 (blue arrow), 1 (red arrow), or 2 (cyan arrow) Nanogold hotspots (left). D. Quantification of Nanogold hotpots located at the terminus of a vRNP (red), in the middle of a vRNP (blue), or proximal to the end of a vRNP (cyan) (right). N= 2121. Representative micrograph showing Nanogold hotpots located at the termini of a vRNP (red arrow), in the middle of a vRNP (blue arrow), or proximal to the end of a vRNP (cyan arrow) (left). E. Quantification Nanogold hotpots with 1 (red), 2 (blue), 3 (cyan), 4 (purple), 5 (light purple), or 6 or more (light blue) Nanogold beads (right). N= 2121. Representative micrograph showing Nanogold hotpots with 1 (red arrow), 2 (blue arrow), 3 (cyan arrow) Nanogold beads (left). F. A representative motion corrected cryo-EM micrograph of virion-purified vRNP with 3 megabodies added. G. A representative motion corrected cryo-EM micrograph showing a single vRNP with potentially two megabodies attached (red arrows).*

5.3 DISCUSSION

The dynamic interactions within a vRNP underlie its conformational changes during transcription and replication. This chapter focuses on developing methods for studying intra-RNP interactions that are critical for vRNP functions, specifically, the viral polymerase-NP interaction and the NP-RNA interaction.

The challenges in resolving polymerase-NP interaction primarily stem from difficulties in recognizing and locating the viral polymerase within a vRNP using cryo-EM. On the other hand, studying NP-RNA interactions faces obstacles due to the wild type NP oligomers being structurally flexible and the NP monomers being too small. To address these issues, I attempted to use nanobody-derivative tools, such as megabodies, which theoretically provide distinct features for particle picking and can increase particle sizes for cryo-EM analysis. Although the structural determination of vRNP in resting states or during transcription and replication has not succeed by the time this thesis was written, I present functional and structural data showing the development of megabodies as tools against viral polymerase or NP. My work in this chapter has paved the way for structurally studying viral polymerase and NP in the context of vRNP in the future.

The development of functional viral polmerase or NP megabodies has been a lengthy process. The yields of the 17 megabody constructs generated in this study ranged from practically 0 to 8 mg per 1 L of bacterial culture. Some megabodies faced issues such as degradation (e.g., Mb8192 and Mb8207) or aggregation (e.g., Mb8202, Mb8204, and Mb8210). However, several megabodies against viral polymerase or NP were successfully purified, and were shown to retain ligand-binding specificity in solution. Thereafter, more challenges arose when

analysing megabody-ligand complexes using cryo-EM. While high-resolution maps for megabody-bound viral polymerases were generated, only megabodies were detected in the megabody-bound NP samples. Since no electron density was detected around the nanobody CDRs in the cryo-EM map, we concluded that the megabody-NP complex dissociated during cryo-EM sample preparation, which is not an uncommon problem in cryo-EM studies. Attempts to stabilize such an interaction by pegylation or glutaraldehyde cross-linking using Grafix have not been successful. Further improvements are needed to increase the stability of such complex to allow visualization of megabody-NP using cryo-EM, in order to study NP-RNA interactions.

In the course of developing strategies for studying viral polymerase-NP and NP-RNA interactions, I sought to simplify the system by forming a complex with one viral polymerase, vRNA promoter, and one NP molecule *in vitro*. However, SEC analysis presented difficulties, particularly because the association of RNA and NP requires low salt conditions, yet viral polymerases were unstable at such salt concentrations. Further investigation to find a condition that satisfies both is needed. Additionally, viral polymerase or mNP both form homo-oligomers in the presence of vRNA promoter used in my study. This hinders the interpretation of the SEC elution profiles of complex formation. In particular, it is difficult to tell successful complex formation from the co-elution of homo-oligomers of viral polymerase and NP if the elution volumes of the complex and the homo-oligomers are close together. Lastly, samples from complex formation have not been of high enough quality to be suitable for cryo-EM analysis. Unfortunately, during my time in the lab, I was unable to optimise the conditions for *in vitro* formation of a viral polymerase-NP complex that is suitable for cryo-EM study. This approach, nevertheless, remains a possibility for future research.

Finally, viral polymerase-NP interactions were studied in the context of virion-purified vRNPs, a biologically relevant setting, using polymerase megabodies. Negative stain microscopy data showed that purified vRNPs used in my study are of good quality and appear with a similar morphology as in previous studies (Coloma et al., 2020; Compans et al., 1972; Moeller et al., 2012). Using Nanogold labelling on anti-His tagged megabodies, I was able to confirm that in the presence of an excess amount of three different megabodies, more than 95% of vRNPs have at least one megabody-Nanogold attached. Additionally, since the location of megabodies represents that of vRNP-associated viral polymerase, I confirmed that viral polymerase mostly locates at one of the termini of a vRNP filament, as megabody-Nanogold was found at the terminus of a vRNP in more than 90% of Nanogold-labelled RNPs. My observations agree with previous studies, in particular, the recent negative staining results from Coloma et al. (Coloma et al., 2020). This study showed the terminal localisation of influenza viral polymerase in resting vRNP and the re-localisation of viral polymerase to the central region of an active vRNP, which agrees with their proposed progressive helical track model. In this model, viral polymerase moves along the RNA template during transcription, as the antiparallel NP strands slide against each other. A growing loop of RNA comes out of the product exit channel and is re-captured by NP, resulting in the apparent relocation of viral polymerases along the vRNP. This process ends and the vRNP enters a resting mode again when the uncopied RNA loop disappears, re-placing the viral polymerase at the terminus of the vRNP.

We collected a high-quality cryo-EM dataset on my three-polymerase megabody-bound vRNP sample using a ThermoFisher FEI Krios at the Electron Bio-Imaging Centre (eBIC). One of the most challenging parts of studying viral polymerase-NP interaction in vRNP using cryo-EM has been to identify the viral polymerase in the sample. Unfortunately, even with megabodies attached, picking viral polymerase in a vRNP sample remains challenging, as

megabodies do not provide significant features for particle picking, either by manual particle picking or by using various cryo-EM particle picking algorithms. The blob picker function could not be used as NP is in massive abundance compared to viral polymerase, therefore will overwhelm the polymerase signals. It is possible to generate back-projected viral polymerase structures with one to three megabodies attached and use them for reverse model-based picking. Either template picking or machine learning-based picking algorithms that train software with a reversely-generated model could be tried. Alternatively, since more than 90% of the viral polymerase locates at the end of a vRNP, end picking presents a possibility because if both ends of a vRNP are selected, theoretically, over 45% of the picks should contain a viral polymerase. Although time-consuming, this can technically be done manually. Instead, machine learning-based end-picking could be employed by using the deep picker function in cryoSPARC or using crYOLO.

In summary, while in this study, I was unable to reveal viral polymerase-NP or NP-RNA interactions, I am confident that we are now closer than ever to being able to, for the first time, visualize these intra-RNP interactions using cryo-EM techniques. Megabodies generated in this study can be further optimized and aid with future structural studies. Due to my limited time in the lab, the follow-up work will be kindly taken over by my colleagues.

Conclusions and Perspectives

In this thesis, I embarked on the journey of elucidating the mechanistic principles that drive influenza virus vRNP functions. As I outlined at the start, our overarching objective was to visualize vRNPs throughout key functional processes at the molecular level in real-time. This would reveal how dynamic intra-RNP interactions lead to conformational changes that underpin different vRNP functional states. However, accomplishing this goal faces two primary challenges: the difficulty in imaging vRNP functions during infection, in the intricate cellular environment; and the struggle in precisely locating viral polymerase or viral RNA within the context of vRNP, let alone resolving viral polymerase-NP or NP-RNA interactions. My work presented here addresses these challenges and contributes to a deeper understanding of vRNP functions.

An exciting breakthrough in this thesis is the establishment of a minimal *in vitro* system for vRNP functions, which reconstitutes transcription and both steps of replication in a test tube. Previous studies showed that virion-isolated vRNPs readily transcribe in a test tube. In this thesis, I was able to, for the first time, switch vRNP from transcription to replication, by adding in purified viral polymerase and ANP32 proteins. This success not only validates the recently proposed replication platform model based on a cryo-EM structure, but also provides insights into the molecular regulation of transcription and replication within a vRNP. By further exploring this *in vitro*-reconstituted switching event, I studied the roles of the viral polymerases and the ANP32 protein. I concluded that vRNP-associated viral polymerase carries out both

transcription and replication. Furthermore, the additional viral polymerase acts as an allosteric regulator to suppress transcriptional activity and, in the presence of an ANP32 protein, to promote replication. In addition, I identified a novel role for ANP32^{LCAR} in promoting replication.

Taking this system further, I added purified NP and enabled the second step of replication inside a test tube. This marks the first successful *in vitro* reconstitution of the entire genome replication cycle of a negative-sense virus. Through establishing this *in vitro* system, I defined the minimal protein requirements for influenza virus genome replication. I also revealed a previously unknown role for NP in directly promoting viral genome replication. Taken together, my work has led to a more complete model for influenza virus vRNP functions described in Figure 3.10.

At the time of my DPhil, a molecular tool that was gaining popularity in both structural and functional studies was the camelid derived nanobody. In this thesis, I studied viral polymerase functions by employing a panel of nanobodies against influenza viral polymerase, with known binding sites on the polymerase surface. These nanobodies revealed novel inhibitory sites on the polymerase surface, which can be used as potential antiviral targets. I also investigated their inhibitory mechanisms and proposed new approaches to inhibit influenza viral polymerase activity. Notably, a subset of nanobodies were found to inhibit transcription or replication by locking the polymerase in a specific conformation or preventing the polymerase from assuming a certain conformation. With this function, they offer valuable tools for enriching vRNPs in distinct functional states during future imaging.

I also attempted to address the issue with localising viral polymerase in a vRNP in structural studies. Despite the successful generation of megabodies that can bind to purified viral polymerase or vRNP, and be resolved by cryo-EM analysis when in complexes with viral polymerase; however, the identification of viral polymerase in a vRNP with the help of megabodies remains a challenge.

Regrettably, there was insufficient time during my DPhil to start structurally resolving the viral polymerase within the context of resting, transcribing, or replicating vRNP. Nevertheless, with the minimal *in vitro* system for vRNP functions, megabodies that facilitate the localisation of viral polymerase, and nanobodies that constrain the polymerase in specific conformations, we are currently at the closest point to successfully imaging vRNP during its functions. However, substantial challenges persist, as the addition of an excess of viral and host proteins in the current *in vitro* system may impede the imaging process. Addressing this requires either a reduction in the amount of added protein or the purification of vRNPs post-reaction. Additionally, the identification of megabody-bound polymerase in a vRNP using cryo-EM remains challenging. Therefore, methods that utilize megabodies to pinpoint vRNP-associated viral polymerase need to be developed. In short, to achieve ultimate success, optimizations are still required for both the *in vitro* system and the structural tools.

The establishment of the minimal *in vitro* system for vRNP functions brings forth extensive possibilities for future development. It provides a new way to study vRNP functions in a simplified and high-controlled environment, and has the potential to provide answers to many unanswered questions on influenza virus transcription and replication. Primarily, this system can be used to inspect the direct effects and dissect the roles of different viral and host factors in vRNP functions. Additionally, it can be used to study the effect of specific genetic mutations

in viral or host proteins on vRNP activity. While introducing mutations into recombinant proteins added to this system is relatively straightforward, doing so for proteins within the vRNP can be difficult. This is due to the potential impediment of viral replication which results in the production of insufficient virions for vRNP purification. A potential approach involves binding nanobodies or small molecular drugs to vRNPs to study protein changes on the vRNP side. Alternatively, recombinant RNPs can be purified from RNP reconstitution in cells; however, the indifferent purification of both vRNP and cRNP via this method can complicate subsequent analysis. A solution to this could be to further purify RNPs based on their RNA sequences, which enables vRNP and cRNP to be isolated separately. An advantage of this approach is that the successful purification of cRNP can then enable the establishment of a minimal *in vitro* system for cRNP functions, which can be used for comparison with vRNP functions. Another potential way to introduce mutations into vRNPs is to generate vRNPs from scratch, using recombinant proteins and synthetic vRNA. This approach also allows the sequence of vRNA and the length of vRNP to be manipulated, which can be useful in studying how vRNA sequence affects vRNP functions and cryo-EM analysis of homogenous vRNPs. Furthermore, this system offers a platform to study RNA products generated during transcription and replication. How do different host and viral factors strike a balance between faithful RNA synthesis and the production of subgenomic RNA? Considering the tight link between aberrant RNA products and innate immunity, applying RNA products of this system to various immunoactivation assays can elucidate how viral and host factors affect innate immunity induction by influencing viral RNA synthesis. Another intriguing question that could be answered by leveraging this system is how post-translational modifications of newly synthesized viral polymerase subunits and NP affect transcription and replication. By introducing and removing PTMs, or PTM mimics, on the recombinant proteins, we can gain a better understanding of the precise tuning of vRNP functions via post-translational

modification. Moreover, this system can be used to develop high-throughput screening methods to identify drugs against influenza virus vRNP functions. An additional corollary of this involves the study of how human genetic variations in host proteins, such as ANP32 family proteins, affect replication efficiency. Furthermore, if we manage to isolate individual vRNP undergoing transcription and replication from this system, we can create a highly-controlled low-noise environment to study viral polymerase dynamics using single-molecule approaches. In a broader context, the success in reconstituting influenza virus genome transcription and replication *in vitro* sets an encouraging example for the development of similar techniques in other negative-sense RNA viruses.

In conclusion, this thesis represents a significant stride toward the ultimate goal of imaging vRNPs in all functional states. The successful bottom-up reconstitution of viral transcription and replication substantiates our current comprehension of influenza virus biology. Through the creation of the *in vitro* system for vRNP functions, we not only deepened our understanding of influenza virus transcription and replication, but also unveiled numerous avenues for further advancements in influenza virus research.

Chapter 6: Materials and methods

6.1 MATERIALS

6.1.1 Plasmids

Plasmids used in this study were from previous Fodor Lab members or generated in this study by molecular cloning.

Name	Original source
pcDNA3	Fodor et al., 2002 (Ervin Fodor et al., 2002)
pcDNA-WSN-PB1	Fodor et al., 2002 (Ervin Fodor et al., 2002)
pcDNA-WSN-PB1a	Vreede et al.,2004 (Vreede et al., 2004)
pcDNA-WSN-PB2	Fodor et al., 2002 (Ervin Fodor et al., 2002)
pcDNA-WSN-PA	Fodor et al., 2002 (Ervin Fodor et al., 2002)
pcDNA-WSN-NP	Fodor et al., 2002 (Ervin Fodor et al., 2002)
pPOLI-NA	Fodor et al., 1999 (Ervin Fodor et al., 1999)
pPOLI-Luc	Manz et al., 2016 (Mänz et al., 2016)
pcDNA-Nb8189	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8190	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8191	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)

pcDNA-Nb8192	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8193	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8194	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8195	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8196	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8197	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8198	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8199	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8200	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8201	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8202	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8203	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8204	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8205	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8206	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8207	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8208	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8209	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8210	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8211	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcDNA-Nb8212	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pCAGGS-1918-PB1	Tumpey et al., 2005 (Tumpey et al., 2005)
pCAGGS-1918-PB1a	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pcCAGGS-1918-PB2	Tumpey et al., 2005 (Tumpey et al., 2005)
pcCAGGS-1918-PB2-TAP	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)

pCAGGS-1918-PA	Tumpey et al., 2005 (Tumpey et al., 2005)
pCAGGS-1918-NP	Tumpey et al., 2005 (Tumpey et al., 2005)
pcDNA-Fj-PB1	Kashiwagi et al., 2009 (Kashiwagi et al., 2009)
pcDNA-Fj-PB2	Kashiwagi et al., 2009 (Kashiwagi et al., 2009)
pcDNA-Fj-PA	Kashiwagi et al., 2009 (Kashiwagi et al., 2009)
pcDNA-Fj-NP	Kashiwagi et al., 2009 (Kashiwagi et al., 2009)
pcDNA-PR8-PB1	Wanitchang et al., 2011 (Wanitchang et al., 2011)
pcDNA-PR8-PB2	Wanitchang et al., 2011 (Wanitchang et al., 2011)
pcDNA-PR8-PA	Wanitchang et al., 2011 (Wanitchang et al., 2011)
pcDNA-PR8-NP	Subbarao et al., (K. Subbarao et al., 2003)
pcDNA-NT60-PB1	Kashiwagi et al., 2009 (Kashiwagi et al., 2009)
pcDNA-NT60-PB2	Kashiwagi et al., 2009 (Kashiwagi et al., 2009)
pcDNA-NT60-PA	Kashiwagi et al., 2009 (Kashiwagi et al., 2009)
pHW2000-WSN-PB1	Hoffmann et al., 2000 (Hoffmann et al., 2000)
pHW2000-WSN-PB2	Hoffmann et al., 2000 (Hoffmann et al., 2000)
pHW2000-WSN-PA	Hoffmann et al., 2000 (Hoffmann et al., 2000)
pHW2000-WSN-NP	Hoffmann et al., 2000 (Hoffmann et al., 2000)
pHW2000-1918-PB1	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pHW2000-1918-PB2	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pHW2000-1918-PA	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pHW2000-1918-NP	Keown, Zhu, Carrique, Fan et al., 2022 (Keown et al., 2022)
pMESy4-Nb8189	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8190	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8191	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8192	Fan et al., 2019 (Fan et al., 2019)

pMESy4-Nb8193	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8194	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8195	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8196	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8197	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8198	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8199	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8200	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8201	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8202	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8203	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8204	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8205	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8206	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8207	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8208	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8209	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8210	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8211	Fan et al., 2019 (Fan et al., 2019)
pMESy4-Nb8212	Fan et al., 2019 (Fan et al., 2019)
pMESP23NO-Mb8189	Generated in this study in Chapter 5.
pMESP23NO-Mb8191	Generated in this study in Chapter 5.
pMESP23NO-Mb8192	Generated in this study in Chapter 5.
pMESP23NO-Mb8193	Generated in this study in Chapter 5.
pMESP23NO-Mb8194	Generated in this study in Chapter 5.

pMESP23NO-Mb8201	Generated in this study in Chapter 5.
pMESP23NO-Mb8202	Generated in this study in Chapter 5.
pMESP23NO-Mb8204	Generated in this study in Chapter 5.
pMESP23NO-Mb8205	Generated in this study in Chapter 5.
pMESP23NO-Mb8206	Generated in this study in Chapter 5.
pMESP23NO-Mb8207	Generated in this study in Chapter 5.
pMESP23NO-Mb8208	Generated in this study in Chapter 5.
pMESP23NO-Mb8210	Generated in this study in Chapter 5.
pHEN6-VHH52-LPETG-His	Schmidt et al., 2016 (Schmidt et al., 2016)
pHEN6-VHH-KV052-LPETG-His	Schmidt et al., 2016 (Schmidt et al., 2016)
pHEN6-KV151-LPETG-His	Schmidt et al., 2016 (Schmidt et al., 2016)
pHEN6-KV170-LPETG-His	Schmidt et al., 2016 (Schmidt et al., 2016)
pHEN6-KV495-LPETG-His	Schmidt et al., 2016 (Schmidt et al., 2016)
pMESPNO23-MbNP1	Generated in this study in Chapter 5.
pMESPNO23-MbKV052	Generated in this study in Chapter 5.
pMESPNO23-Mb151	Generated in this study in Chapter 5.
pMESPNO23-Mb170	Generated in this study in Chapter 5.
pMESPNO23-Mb495	Generated in this study in Chapter 5.
pGEX-6P-1-NPA	Wang et al., 2022 (Wang et al., 2022)
pGEX-6P-1-mNP	Wang et al., 2022 (Wang et al., 2022)

pGEX-6P-1-mNPC	Generated in this study in Chapter 5.
pGEX-6P-1-chANP32A	Wang et al., 2022 (Wang et al., 2022)
pGEX-6P-1-huANP32B	Wang et al., 2022 (Wang et al., 2022)

Table 6.1. List and descriptions of plasmids used in this study.

6.1.2 DNA oligonucleotides

DNA oligonucleotides used in this study were purchased from Sigma-Aldrich. DNA oligonucleotides were dissolved and diluted in nucleotide-free water (Invitrogen) to 100 mM.

Name	**Sequence (5' – 3')**
pCAGGS_1918_PB1_PF	CCTGAACACAATGACCAAAG
pCAGGS_1918_PB1_PR	CATGATCTCAGCGAACTCCTC
pCAGGS-PB1a Rev	GTCACACTCGAGCATGCCCGGGGGAGGTATATCTTTCG
pCAGGS-PB1a Fwd	AGCAAGTCTAGAGCCTCTGCTAACCATGTTCATGCCTTC
TU89	GTCCTCTCAGAGTCCCTGCTTCTCGAGTTCC
EP230	TGGGTCCAGTGGCAGAGGTCACCTTCTCGTCAGGA
Vector_F (PB2-TAP)	GAATTCGGATGGCCATCAATGCGGCCGCTATG
Vector_R (PB2-TAP)	CTTAGTTCTTTTATTCTTTCCATGGTGGATCCGAG
Fragment_F (PB2-TAP)	GTACCGAGCTCGGATCCACCATGGAAAGAATA
Fragment_R (PB2-TAP)	CTCTTTTCCATAGCGGCCGCATTGATGGCCAT
NA 160 cloning	AGGACTGGAATTCCTCCAGTATGGTTTTGATTTCCG
NA 1280 cloning	CCTTGACATATGCTGGACTAGTGGGAGCATCAT

22nt primer marker	CATTTAAACTCCTGCTTTTGCT
25nt primer marker	ATTCATTTAAACTCCTGCTTTTGCT
NA-160	TCCAGTATGGTTTTGATTTCCG
NA-1280	TGGACTAGTGGGAGCATCAT
NP-160	TGATTTCAGTGGCATTCTGG
NP-1438	TGATGGAAAGTGCAAGACCA
PB1-146	TCCATGGTGTATCCTGTTCC
PB1-2208	TGATTTCGAATCTGGAAGGA
PB2-182	GCCATCATCCATTTCATCCT
PB2-2198	TGCTAATTGGGCAAGGAGAC
M1-162	AGCCATTCCATGAGAACCTC
M-890	GAAAAGAGGGCCTTCTACGG
NS1-226	CGCTCCACTATTTGCTTTCC
NS-734	TGATTGAAGAAGTGAGACACAG
PA-146	TGAGTGCATATTGCTGCAAAT
PA-2021	TTCTTATCGTTCAGGCTCTT
5S 100	TCCCAGGCGGTCTCCCATCC
NA 32	GGTTTGGATTCATTTAAACT
NA 50	CAATGGTTATTATTTTCTGG
NA 22	CATTTAAACTCCTGCTTTT
NA 25	ATTCATTTAAACTCCTGC
NA 1404	GTTCAAAAAACTCCTTGTTT
NA 1406	CAAAAAACTCCTTGTTTCT
pMESP23-NO BB 1	ATGGTTGCTTTGACGTATGCGG
pMESP23-NO BB 2	TGGTGTAAAGTGTGGTCGAACC

pMESP23-NO BB 3	AGCAGGATTGGTTAAAGCGGCA
pMESP23-NO BB 4	TTAATGAATCGGCCAACGCGCG
pMESP23-NO BB 5	TATCTGCGCTCTGCTGAAGCCA
pMESP23-NO BB 6	GGATCTTACCGCTGTTGAGATC
pMESP23-NO Ins F	GGTGGCAGAGATGTACCCTAAA
Lac promoter F	TTTACACTTTATGCTTCCGG
pMESy4_Nb_P3	GCCTTATCCGGTAACTATCG
pMESy4_Nb_P4	CCGCTTACAGACAAGCTGTG
pMESy4_Nb_P5	CCCTCATAGTTAGCGTAACG
pMESy4_Nb_P6	CGGTGACGGTGATAATTCACC
pMESy4_Nb_P7	CCGGTTCCCAACGATCAAGG
pCAGGS_P1	CTCTAGAGCCTCTGCTAACC
pcDNA seq	GGCAACTAGAAGGCACAGTCG
M13 fwd_complementary	ACTGGCCGTCGTTTTAC

Table 6.2 List and sequences of DNA oligonucleotides used in this study.

6.1.3 RNA oligonucleotides

RNA oligonucleotides used in this study were purchased from Horizon Discovery. RNA oligonucleotides were dissolved and diluted in nucleotide-free water (Invitrogen) to 4 mM.

Name	Sequence (5' – 3')
5' IAV vRNA	AGUAGAAACAAGGCC
3' IAV vRNA	GGCCUGCUUUUGCU
5' IAV cRNA	pAGCAAAAGCAGGCC
3' IAV cRNA	GGCCUUGUUUCUACU
FluA_vRNA_5'_29mer	AGUAGAAACAAGGCCGUAUAUGAGACAGA
FluA_vRNA_5'_39mer	AGUAGAAACAAGGCCGUAGAAUGAUGUAUAUGAGACAGA
11-mer RNA	ppGAAUACUCAAG
20-mer RNA	ppAAUCUAUAAUAGCAUUAUCC
12-mer capped RNA	m7GpppAAUCUAUAAUAG

Table 6.3. List and sequences of RNA oligonucleotides used in this study.

6.1.4 Host RNA polymerase II CTD mimic peptides

CTD mimic peptides used in this study containing four repeats of the heptapeptide consensus sequence (YSPTSPS) with specified PTMs were designed by Monica Martínez-Alonso and

synthesised by Cambridge Peptides Ltd. via solid-phase peptide synthesis. Peptides were dissolved and diluted in DMSO to 10 mg/ml.

Name	Sequence (5' – 3')
Unphosphorylated CTD (UnP)	Biotin - PEG - YSPTSPSYSPTSPSYSPTSPSYSPTSPS Biotin
Serine 2 phosphorylated CTD (S2P)	Biotin - PEG - YS*PTSPSYS*PTSPSYS*PTSPSYS*PTSPS
Serine 5 phosphorylated CTD (S5P)	Biotin - PEG - YSPTS*PSYSPTS*PSYSPTS*PSYSPTS*PS

Table 6.4. List and sequences of the synthetic DNA-dependent RNA polymerase II CTD peptides used in this study. *S*, phosphorylated serine.*

6.2 MOLECULAR CLONING

6.2.1 Preparation of chemically competent cells

Chemically competent DH5α, WK6, Rosetta (DE3) and BL21(DE3) *E. coli* cells were generated following the protocol outlined below. Initially, 500 ml of antibiotic-free 2x YT was inoculated with 3-5 ml of an overnight culture and grown until reaching an OD_{600} of 0.3–0.4. Upon reaching this optical density, the culture was promptly cooled on ice for 15 min, pelleted at 3,500 g for 10 min at 4°C, and subsequently resuspended in 100 ml of ice-cold 0.1 M $MgCl_2$. Following a 5-10 min incubation at 4°C with gentle shaking, the cells were pelleted again at 3,500 g for 10 min at 4°C and resuspended in 100 ml of ice-cold 0.1 M $CaCl_2$. Following a 5-10 min incubation at 4°C with gentle shaking, the cells were pelleted again at 3,500 g for 10 min at 4°C, resuspended in 20 ml of ice-cold 0.1 M $CaCl_2$ +15% (v/v) glycerol and shaken gently on ice until fully resuspended. The resulting cell suspension was then portioned into 50-100 µl aliquots in sterilised Eppendorf tubes, snap-frozen in liquid N_2 and stored at -80°C.

6.2.2 Transformation

E. coli cells were transformed through a heat-shock method for 45 s, at 42°C, followed by a recovery period for 1 h, at 37°C in SOC medium (Thermo Scientific), if necessary. Transformed cells were selected by growth on LB agar plates supplemented with the appropriate antibiotics or in 5 ml LB supplemented with the appropriate antibiotics.

6.2.3 Preparation of plasmid DNA

Plasmid DNA was extracted using midiprep or miniprep kits (Qiagen), in accordance with the manufacturer's instructions.

6.2.4 Site-directed mutagenesis

Site-directed mutagenesis PCR was conducted using Pfu Turbo DNA Polymerase (Agilent) in 50 µl reaction mixtures prepared in accordance with the manufacturer's instructions. The reaction mixtures were supplemented with 1 U of Phusion High-Fidelity DNA Polymerase (NEB). Following the PCR, reactions were treated with 1 µl of DpnI (NEB) at 37°C for 1 h. The resulting products were then transformed into DH5α *E. coli* cells and grown on LB agar plates supplemented with 100 µg/ml ampicillin. The successful introduction of mutations was verified through Sanger Sequencing (Source BioScience).

6.2.5 Gibson assembly

PCR was conducted utilizing Phusion High-Fidelity DNA Polymerase (NEB) in accordance with the manufacturer's instructions. The resulting PCR products were analysed using electrophoresis on a 0.8-1.5% agarose gel pre-stained with SYBR Safe (Invitrogen). Subsequently, the desired DNA fragments were excised and purified using the QIAquick Gel Extraction Kit (Qiagen), in accordance with the manufacturer's instructions. For Gibson assembly, the NEBuilder HiFi DNA Assembly mix (NEB) was employed as per the

manufacturer's instructions. The resultant assembly products were transformed into DH5α *E. coli* cells and plated on LB agar supplemented with 100 μg/ml ampicillin. Verification of successful mutations was carried out through Sanger Sequencing (Source BioScience).

6.2.6 Megabody cloning

Megabody genes within the expression vector pMESP23NO (GenBank MT338522) were generated through the subcloning of nanobodies, enabling their expression as megabodies using the Mb-cYgjKNO scaffold within the periplasm of *E. coli*, as previously outlined (Uchański et al., 2021). In brief, the cloning and expression of megabodies from nanobodies involved the PCR amplification with primers TU89 and EP230 (Table 6.2), of gene fragments encoding β-strands B-G of the parental nanobodies. These fragments were then cloned as by SapI (BioLabs) restriction enzyme digestion into the designated expression vector pMESP23NO.

6.2.7 Mutagenesis in pCAGGS plasmids

Mutagenesis in pCAGGS plasmids was executed through an approach described below: In brief, mutagenesis PCR was first conducted using two sets of primers to generate two half segments of the mutated gene, pCAGGS_1918_PB1_PF and pCAGGS-PB1a Rev or pCAGGS_1918_PB1_PR and pCAGGS-PB1a Fwd (Table 6.2) in separate 50 μl reactions. Pfu Turbo DNA Polymerase (Agilent) was employed in accordance with the manufacturer's instructions. The resulting PCR products were separated on an agarose gel and extracted using the QIAquick Gel Extraction Kit (Qiagen). To generate a full-length mutated gene segment, a

subsequent PCR reaction was performed by combining the gel extracted products with the two external primers, pCAGGS_1918_PB1_PF and pCAGGS_1918_PB1_PR. The resulting products were again separated on an agarose gel and extracted using the QIAquick Gel Extraction Kit (Qiagen). The extracted mutated gene fragment and pCAGGS plasmid underwent separate digestion by Xho1 and XbaI (BioLabs), followed by ligation using Quick Ligase (BioLabs) at a ratio of 3:1 or 10:1 at room temperature for 1 h, in accordance with the manufacturer's instructions.

The ligation product was transformed into DH5α *E. coli* and plated on LB agar supplemented with 100 µg/ml ampicillin. Successful mutations were verified through Sanger Sequencing (Source BioScience).

6.3 CELLS AND VIRUSES

6.3.1 Bacterial culture

Unless otherwise specified, competent *E. coli* cells were cultured in LB medium supplemented with ampicillin at a working concentration of 100 μg/ml. Bacterial growth was carried out at 37°C, with shaking at 180 rpm. The bacterial strains utilized are summarised in Table 6.5 below:

Name	Origin
DH5-α	In-house production by me.
Rosetta (DE3)	In-house production by Alaa Baazaoui.
BL21 (DE3)	In-house production by Dr Haitian Fan.
WK6	In-house production by me.

Table 6.5. List and origins of the bacterial strains used in this study.

6.3.2 Insect cell culture

All insect cells work was conducted using Sf9 cells derived from *Spodoptera frugiperda*. Unless otherwise specified, these cells were cultured in suspension culture, with a minimum volume of the culture at least 5 times larger than the flask, at 28°C, with shaking at 120 rpm in the dark. The culture medium was Sf900-II serum-free medium (Gibco), supplemented with 1x penicillin-streptomycin (Gibco). Cultures were maintained by periodic dilution, and cell densities were kept within the range of $0.5x10^6$–$2x10^6$ cells/ml, with manual counting

performed using a hemocytometer. New cultures were obtained, typically approximately every four months, from Zhang Lab from STRUBI at the University of Oxford.

6.3.3 Mammalian cell culture

Human Embryonic Kidney 293T (HEK 293T), Madin-Darby Bovine Kidney (MDBK) and Madin-Darby Canine Kidney (MDCK) cells were procured from the Sir William Dunn School of Pathology cell bank at the University of Oxford. HEK 293T cells were cultured in Dulbecco's Modified Eagle Medium (DMEM; Sigma Aldrich) supplemented with 10% Fetal Calf Serum (FCS; Sigma Aldrich). MDBK and MDCK cells were cultured in Minimum Essential Medium (MEM; Sigma Aldrich) supplemented with 10% FCS and 1x L-glutamine (200 mM) (Gibco). The cells were maintained at 37°C in 5% CO_2 and were trypsinized with 1x TrypLE Express Enzyme (Gibco).

6.3.4 Influenza virus reverse genetics

The 8-plasmid system developed by Hoffmann et al.,2002 (Hoffmann et al., 2002) was utilized to generate either wild type (WSN) or chimeric Influenza A viruses WSN-1918$_{RNP}$ and WSN-1918$_{polymerase}$. For the WSN virus, the following plasmids were employed: pHW2000-WSN-PB1, pHW2000-WSN-PB2, pHW2000-WSN-PA, pHW2000-WSN-NP, pHW2000-WSN-HA, pHW2000-WSN-NA, pHW2000-WSN-M, and pHW2000-WSN-NS (Table 6.1). For the WSN-1918$_{RNP}$ virus, the following plasmids were used: pHW2000-1918-PB1, pHW2000-1918-PB2, pHW2000-1918-PA, pHW2000-1918-NP, pHW2000-WSN-HA, pHW2000-WSN-NA, pHW2000-WSN-M, and pHW2000-WSN-NS (Table 6.1). For the WSN-1918$_{polymerase}$

virus, the following plasmids were used: pHW2000-1918-PB1, pHW2000-1918-PB2, pHW2000-1918-PA, pHW2000-WSN-NP, pHW2000-WSN-HA, pHW2000-WSN-NA, pHW2000-WSN-M, and pHW2000-WSN-NS (Table 6.1). In brief, approximately 0.6×10^6 HEK293T HEK 293T cells were transfected with 0.5 µg of each plasmid using Lipofectamine 2000 Transfection Reagent (Thermo Scientific) in accordance with the manufacturer's instructions and plated in 6 cm dishes. For the negative control sample, the plasmid encoding PB1 was omitted. The cells were maintained at 37°C, 5% CO2 for 6 h before replacing DMEM media supplemented with 10% (v/v) FCS with fresh DMEM media supplemented with 0.5% (v/v) FCS. Subsequently, the cells were kept at 37°C, 5% CO2 for an additional 48 h. After this period, the supernatants from HEK 293T cells were collected, and cell debris was removed by centrifugation at 2000 rpm (Hettich ROTINA 420, rotor 5624). The virus stock was stored at -80°C until further use. Viral titres were determined using a plaque assay.

Name	Descriptions
WSN	Wild type influenza A/WSN/1933 (H1N1) viruses
WSN-1918$_{RNP}$	Influenza A virus containing HA, NA, M and NS segments from A/WSN/1933 (H1N1) virus, and PB1, PB2, PA and NP segments from A/Brevig Mission/1918 (H1N1).
WSN-1918$_{polymerase}$	Influenza A virus containing HA, NA, M, NS and NP segments from A/WSN/1933 (H1N1) virus, and PB1, PB2 and PA segments from A/Brevig Mission/1918 (H1N1).

Table 6.6. List and descriptions of influenza viruses produced by reverse genetics in this study.

6.3.5 Virus propagation

Influenza viruses were propagated by infecting 70-80% confluent MDBK or MDCK cells at MOI = 0.01. The culture medium was collected 48 hpi, and cell debris was removed by centrifugation at 4000 rpm for 10 min. The clarified medium was aliquoted, and the virus titre was determined using a plaque assay.

6.3.6 Plaque assay

MDBK or MDCK cells were seeded in 6-well plates and allowed to grow to roughly 80% confluency. Serial 10-fold dilutions of a viral stock were prepared in DMEM supplemented with 0.5% (v/v) FCS, and 200 µl of each dilution was applied to the cells for 1 h at room temperature. After the incubation, the virus solution was aspirated, and 2 ml of agarose overlay (2% (w/v) agarose in PBS mixed with an equal volume of DMEM supplemented with 0.5% (v/v) FCS) was applied to the cells. Plates were incubated for 15 min at RT to allow the agarose overlay to set. Subsequently, cells were incubated at 37°C, 5% CO2 for 48-72 h before the agarose overlay was removed and the cells were stained with Coomassie Brilliant Blue stain (0.1% Coomassie R250 (w/v), 10% (v/v) acetic acid, 40% (v/v) methanol).

6.4 CELL-BASED ASSAYS

6.4.1 Ribonucleoprotein reconstitution assay

Unless stated otherwise, when RNP components were encoded in pcDNA, approximately 0.16×10^6 HEK293T cells were transfected with 0.2 µg of each plasmid encoding PB1, PB2, PA, and NP, 0.2 µg of a plasmid encoding a luciferase or an NA vRNA segment, and 0.2 µg of a nanobody or an empty pcDNA3 vector. Unless stated otherwise, when RNP components were encoded in pCAGGS, an approximately 0.16×10^6 HEK293T cells were transfected with 0.04 µg of each plasmid encoding PB1, PB2, PA, and NP, 0.2 µg of a plasmid encoding a luciferase vRNA segment, and 0.8 µg of a nanobody or an empty pcDNA3 vector. When luciferase vRNA was transfected, cells were collected 20-24 h after transfection, and total cell lysates were obtained by adding 100 µl of Luciferase Cell Culture Lysis Reagent (Promega) to the cells, followed by vigorous shaking for 10 min. The presence of luciferase enzyme was determined using a luciferase-based detection system in accordance with the manufacturer's instructions (Promega). In brief, 100 µl of Luciferase Assay Reagent (Promega) was added to 20 µl of whole cell lysate in a 96-well opaque plate. The plate was immediately read in a microplate reader SpectraMax M3 (Molecular Devices). Two technical replicates were carried out for each transfection, and mean values were calculated. These mean values from three independent transfections were used in subsequent analysis in Excel (Microsoft) and Prism 8 – Prism 10 (GraphPad). Alternatively, when NA vRNA was transfected, total cellular RNA was also extracted using TRI reagent (Sigma-Aldrich) and viral RNA levels were analysed by primer extension.

For pre-infection transfection, approximately 1×10^6 HEK293T cells were transfected with 5 μg of pcDNA plasmid encoding Nb8190, Nb8191, Nb8192, Nb8207, Nb8208, and Nb8210 or an empty pcDNA3 vector, using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Where indicated, an additional 0.25 μg of each pCAGGS plasmid encoding PB1 (or PB1a), PB2, PA, and NP of A/Brevig Mission/1/1918 was transfected. Subsequently, 20-24 h after transfection, cells were infected with a reassortant influenza A/WSN/33 virus encoding polymerase and NP genes of A/Brevig Mission/1/1918 at a MOI of 0.01. Cell culture medium was collected 12, 24, 36 and 48 h after infection, and virus titres were determined by plaque assay. At each time point, total cellular RNA was also extracted using TRI reagent (Sigma Aldrich) and viral RNA levels were analysed by primer extension.

6.4.2 Primary transcription assay

Approximately $0.2x10^6$ HEK 293T cells in DMEM + 10% FCS were transfected with 1 μg of pcDNA plasmid encoding Nb8190, Nb8191, Nb8192, Nb8207, Nb8208, Nb8210, or an empty pcDNA3A vector, using Lipofectamine 2000 according to the manufacturer′s instructions. Mutant plasmids replaced the wild type plasmids where indicated. 48 h post-transfection, medium was replaced with DMEM + 0.5% FCS containing 100μg/mL cycloheximide and influenza A virus WSN-1918_{RNP} at MOI 5. Cells were incubated for a further 4 h, then total cellular RNA was extracted and analysed by primer extension.

6.4.3 cRNA stabilisation assay

Approximately $0.2x10^6$ HEK 293T cells in DMEM + 10% FCS were transfected with 1µg of pcDNA plasmid encoding Nb8190, Nb8191, Nb8192, Nb8207, Nb8208, Nb8210, or an empty pcDNA3 vector, and 0.4 µg of pCAGGS-1918-PB1a, pCAGGS-1918-PB2, pCAGGS-1918-PA and pCAGGS-1918-NP using Lipofectamine 2000 according to the manufacturer′s instructions. 48 h post-transfection, medium was replaced with DMEM + 0.5% FCS containing 100 µg/ml cycloheximide (Sigma Aldrich) and influenza A virus WSN-1918_{RNP} at MOI 5. Cells were incubated for a further 4 h, then total cellular RNA was extracted and analysed by primer extension.

6.4.4 Replication assay

Approximately $0.2x10^6$ HEK 293T cells in DMEM + 10% FCS were transfected with with 1µg of pcDNA plasmid encoding Nb8190, Nb8191, Nb8192, Nb8207, Nb8208, Nb8210, or an empty pcDNA3 vector, and 0.4 µg of pCAGGS-1918-PB1, pCAGGS-1918-PB2, pCAGGS-1918-PA and pCAGGS-1918-NP using Lipofectamine 2000 according to the manufacturer′s instructions. 48 h post-transfection, medium was replaced with DMEM + 0.5% FCS containing 100 µg/ml cycloheximide (Sigma Aldrich) and influenza A virus WSN-1918_{RNP} at MOI 5. Cells were incubated for a further 4 h, then total cellular RNA was extracted and analysed by primer extension.

6.4.5 Viral growth assay

Approximately 1×10^6 HEK293T cells were transfected with 5 μg of pcDNA plasmid encoding Nb8190, Nb8191, Nb8192, Nb8207, Nb8208, Nb8210, or an empty pcDNA3 vector, using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Subsequently, 20 h after transfection, cells were infected with WSN-1918_{RNP} at a multiplicity of infection of 0.01. Cell culture medium was collected 12, 24, 36 and 48 hpi, and virus titres were determined by plaque assay. At each time point, total cellular RNA was also extracted and analysed by primer extension.

6.4.6 Cell viability assay

Approximately 1×10^6 HEK293T cells were transfected with 5 μg of pcDNA plasmid encoding Nb8190, Nb8191, Nb8192, Nb8207, Nb8208, Nb8210, or an empty pcDNA3 vector, using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Subsequently, 20 h after transfection, cells were infected with WSN-1918_{RNP} at a multiplicity of infection of 0.01. Cell viability was determined 20 h post-transfection, or 48 hpi using a CellTiter-Glo luminescent assay (Promega) according to the manufacturer's instructions.

6.5 PROTEIN EXPRESSION AND PURIFICATION

6.5.1 Insect cell expression system

Sf9 insect cells, which had been passaged the previous day, at a density of approximately 2 x 10^6 cells/ml, were infected with recombinant baculoviruses expressing influenza virus polymerases (Table 6.7) at an MOI 0.5-1 for 72 h at 27°C. The insect cells were harvested at the end of infection through centrifugation at 800 g for 15-20 min at 4°C.

Name	Descriptions
NT60	Baculovirus made by Dr Narin Hengrung expressing wild type influenza viral polymerase of A/Northern Territory/60/1968 (H3N2)
NT60 PB1a	Baculovirus made by Dr Haitian Fan expressing the active-site mutant influenza viral polymerase of A/Northern Territory/60/1968 (H3N2)
1918	Baculovirus made by Dr Jeremy Keown expressing wild type influenza viral polymerase of A/Brevig Mission/1/1918 (H1N1)
Fj	Baculovirus made by Dr Haitian Fan expressing wild type influenza viral polymerase of A/duck/Fujian/01/2002 (H5N1)
FluC	Baculovirus made by Dr Narin Hengrung expressing wild type influenza viral polymerase of C/Johannesburg/1/1966
WSN	Baculovirus made by Dr Haitian Fan expressing wild type influenza viral polymerase of A/WSN/1933 (H1N1)
WSN PB1a	Baculovirus made by Dr Haitian Fan expressing the active-site mutant influenza viral polymerase of A/WSN/1933 (H1N1)

Table 6.7. List and descriptions of baculoviruses used in this study for protein expression in Sf9 cells.

6.5.2 Bacterial cell expression system

1-2 L of bacterial cultures for protein expression were grown at 37°C in antibiotic-supplemented LB or 2x YT using 1. freshly transformed bacteria recovered overnight at 37°C in appropriate antibiotics, 2. a streak of freshly transformed bacterial colonies from LB plates supplemented with the appropriate antibiotics, or 3. starting culture from glycerol stock of transformed bacteria grown overnight at 37°C in appropriate antibiotics. The culture was allowed to grow to OD_{600} = 0.6–0.9 before being cooled down to 18°C and induced by 1 mM isopropyl β- d-1-thiogalactopyranoside (IPTG) overnight at 18°C, and shaking at 180 rpm. The bacterial cells were harvested at the end of induction through centrifugation at 4000 g for 15-20 min at 4°C.

6.5.3 Influenza virus polymerase purification

Viral polymerase was purified as previously described (Fan et al., 2019). In brief, insect cells from 1 L of culture expressing influenza viral polymerase were resuspended in approximately 30 ml of resuspension buffer (50 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 10% (v/v) glycerol, 0.05% (w/v) OTG, 1 mM DTT, 100 µg/ml ribonuclease A (RNaseA, Qiagen) and 1X cOmplete EDTA-free protease inhibitor (Roche)) at room temperature for 45 min. Resuspended cells were lysed by sonication on ice at 15 µm oscillation, 3x 45 s on and 30 s off and centrifuged at 35000 g for 45 min at 4°C. The lysate supernatant was incubated with 2 ml of pre-washed IgG Sepharose 6 fast flow beads (1 ml slurry, GE Healthcare), rotating at 4°C for 3 h. Following this, the beads were washed three times in 10 ml wash buffer (50 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 10% (v/v) glycerol) by pelleting at 2000 g for 3 min

after each wash. For the cleavage reaction, the beads were resuspended in 7 ml wash buffer, supplemented with 1 mM final concentration of DTT, and incubated with 0.5 mg home-made TEV protease with rotation overnight at 4°C. After cleavage, the beads were removed by centrifugation at 4500 g for 1 min at 4°C and washed with 5 ml wash buffer. The resulted protein sample was concentrated using an Amicon Ultra centrifugal filter (Merck Millipore, 100 kDa MWCO) by centrifuging at 4000 g, to reduce the volume to below 500 µl. The concentrated protein sample was centrifuged for 3 min at 13000 g at 4°C to remove the precipitation and further purified by SEC on a Superdex 200 increase 10/300 gel filtration column (Cytiva) in polymerase SEC buffer (25 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 10% (v/v) glycerol) connected to an ÄKTA fast protein liquid chromatography system (GE Healthcare). The peak fractions containing the viral polymerase were combined and concentrated to the desired concentration. Purified viral polymerases were aliquoted, snap frozen in liquid N_2 and stored at -80°C for long-term usage.

6.5.4 *Influenza virus nucleoprotein purification*

NP was purified as previously described (Wang et al., 2022). In brief, cell pellet from 2 L of bacterial culture expressing viral NP was resuspended in 20 ml resuspension buffer (50 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 10% (v/v) glycerol, 0.05% (w/v) OTG, 1 mM DTT, 100 µg/ml ribonuclease A (RNaseA, Qiagen) and 1X cOmplete EDTA-free protease inhibitor (Roche)) supplemented with 1 mg/ml lysozyme (Sigma Aldrich) at room temperature for 45 min. Resuspended cells were lysed by sonication on ice at 15 µm oscillation, 4x 30 s on and 30 s off and centrifuged at 35000 g for 45 min at 4°C. The lysate supernatant was incubated with 2 ml of pre-washed Glutathione Sepharose 4B beads (GE Healthcare) (2 ml slurry per 1

L culture) rotating at 4°C for 3 h. Following this, the beads were washed one time in 10 ml wash buffer (50 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 10% (v/v) glycerol) and three times in 10 ml high salt wash buffer (50 mM HEPES-NaOH pH 7.5, 1.5 M NaCl, 10% (v/v) glycerol) by pelleting at 2000 g for 3 min after each wash, in order to remove bacterial RNA associated to NP. For the cleavage reaction, the beads were resuspended in 7 ml wash buffer, supplemented with 1 mM final concentration of DTT, and incubated with 0.5 mg home-made HRV 3C protease with rotation overnight at 4°C. After cleavage, the beads were removed by centrifugation at 4500 g for 1 min at 4°C and washed with 5 ml wash buffer. The resulted protein sample was concentrated using an Amicon Ultra centrifugal filter (Merck Millipore, 30 kDa MWCO) by centrifuging at 4000 g, to reduce the volume to below 500 µl. The concentrated protein sample was centrifuged for 3 min at 13000 g at 4°C to remove the precipitation and further purified by SEC on a Superdex 200 increase 10/300 gel filtration column (Cytiva) in NP SEC buffer (25 mM HEPES-NaOH pH 7.5, 150 mM NaCl, 10% (v/v) glycerol) connected to an ÄKTA fast protein liquid chromatography system (GE Healthcare). The peak fractions containing NP were combined and concentrated to the desired concentration. Purified NP was aliquoted, snap frozen in liquid N_2 an stored at -80°C for long-term usage.

6.5.5 ANP32 family protein purification

ANP32 family proteins were purified as previously described (Carrique et al., 2020). In brief, bacterial cells from per 2 L of culture expressing ANP32 was resuspended in 20 ml resuspension buffer (50 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 10% (v/v) glycerol, 0.05% (w/v) OTG, 1 mM DTT, 100 µg/ml ribonuclease A (RNaseA, Qiagen) and 1X cOmplete EDTA-free protease inhibitor (Roche)) supplemented with 1 mg/ml lysozyme (Sigma Aldrich)

at room temperature for 45 min. Resuspended cells were lysed by sonication on ice at 15 µm oscillation, 4x 30 s on and 30 s off and centrifuged at 35000 g for 45 min at 4°C. The lysate supernatant was incubated with 2 ml of pre-washed Glutathione Sepharose 4B beads (GE Healthcare) (2 ml slurry per 1 L culture) rotating at 4°C for 3 h. Following this, the beads were washed three times in 10 ml wash buffer (50 mM HEPES-NaOH pH 7.5, 500 mM NaCl, 10% (v/v) glycerol) by pelleting at 2000 g for 3 min after each wash. For the cleavage reaction, the beads were resuspended in 7 ml wash buffer, supplemented with 1 mM final concentration of DTT, and incubated with 0.5 mg home-made HRV 3C protease with rotation overnight at 4°C. After cleavage, the beads were removed by centrifugation at 4500 g for 1 min at 4°C and washed with 5 ml wash buffer. The resulted protein sample was concentrated using an Amicon Ultra centrifugal filter (Merck Millipore, 10 kDa MWCO) by centrifuging at 4000 g, to reduce the volume to below 500 µl. The concentrated protein sample was centrifuged for 3 min at 13000 g at 4°C to remove the precipitation and further purified by SEC on a Superdex 200 increase 10/300 gel filtration column (Cytiva) in ANP32 SEC buffer (25 mM HEPES-NaOH pH 7.5, 150 mM NaCl, 10% (v/v) glycerol) connected to an ÄKTA fast protein liquid chromatography system (GE Healthcare). The peak fractions containing ANP32 were combined and concentrated to the desired concentration. Purified ANP32 was aliquoted, snap frozen in liquid N_2 an stored at -80°C for long-term usage.

6.5.6 Nanobody purification

Nanobodies were expressed and purified as previously described. In brief, transformed WK6 cells were grown on LB agar plates, supplemented with 100 µg/ml ampicillin and 1% (w/v) glucose. For each expression, 1 L 2x YT supplemented with 0.1% (w/v) glucose, 1 mM $MgCl_2$

and 100 µg/ml ampicillin was inoculated with 10 to 100 ml overnight culture (freshly transformed colony or glycerol stock grown in 20 ml LB, supplemented with 2% (w/v) glucose, 1 mM $MgCl_2$ and 100 µg/ml ampicillin). This was grown at 37°C, with shaking at 180 rpm, until the OD_{600} reached 0.6-0.8. Expression was then induced by adding IPTG at 1 mM final concentration, and incubating the culture overnight at 28°C with shaking at 120 rpm. Post-induction bacterial cells were harvested by centrifugation at 4000 g for 15-20 min.

Nanobodies were purified from the periplasm of the expressing cells by osmotic shock. For each 1 L culture, bacterial cells were resuspended in 15 ml TES buffer (200 mM Tris-HCl pH 8.0, 0.5 mM EDTA, 0.5 M sucrose). This mixture was left rotating at 4°C for 1 h to fully resuspend the cells. Osmotic shock was performed by adding double the resuspended-culture volume of 4x diluted TES buffer, and leaving rotating at 4°C for 45 min. After this, the mixture was centrifuged at 4000 g for 30 min at 4°C, and the supernatant was added to 1 ml settled Ni-NTA agarose (0.5 ml slurry, Qiagen) that had been equilibrated with nanobody wash buffer (100 mM Tris-HCl pH7.5, 500 mM NaCl, 30mM imidazole). This mixture was rotated for 1 h, at 4°C, then the beads were washed three times with 10 ml nanobody wash buffer, and eluted in 7 ml then 5ml of nanobody elution buffer (100 mM Tris-HCl pH7.5, 500 mM NaCl, 500mM imidazole). Nanobodies were concentrated using an Amicon Ultra centrifugal filter unit (Merck Millipore, NMWL of 10 kDa) and buffer exchanged into nanobody storage buffer (25 mM Tris-HCl pH7.5, 150 mM NaCl). Purified nanobodies were aliquoted, snap frozen in liquid N_2 an stored at -20°C for long-term usage.

6.5.7 Megabody purification

Megabody purification protocol was adapted from Uchanski et al., 2021 and Coupland et al., 2021 (Coupland et al., 2021; Uchański et al., 2021). In brief, 100 ng - 200 ng of megabody expression plasmids were freshly transformed into 20 µl to 50 µl competent WK6 cells by heat shock. Transformed WK6 cells was grown in SOC medium (Sigma Aldrich) at 37°C shaking vigorously for 45 min before being added into 20ml of 2x YT supplemented with 1% (w/v) glucose and 100 µg/ml ampicillin. Inoculation culture is grown overnight in a 50 ml falcon at 37°C, with shaking at 180 rpm. For each expression, 1 L 2x YT supplemented with 0.1% (w/v) glucose, 1 mM $MgCl_2$ and 100 µg/ml ampicillin was inoculated with 10 to 20 ml overnight culture (freshly transformed colony or glycerol stock grown in 20 ml LB, supplemented with 2% (w/v) glucose, 1 mM $MgCl_2$ and 100 µg/ml ampicillin). This was grown at 37°C, with shaking at 120 rpm, until the OD_{600} = 0.6-0.9. Expression was then induced by adding IPTG at 1 mM final concentration, and incubating the culture overnight at 28°C with shaking at 120 rpm. Post-induction bacterial cells were harvested by centrifugation at 4000 g for 15-20 min.

Megabodies purification was carried out as follow: For each 1 L culture, bacterial cells were resuspended in 30 ml PE buffer (50 mM Tris-HCl pH 8.0, 1 mM EDTA, 150 mM NaCl, 20% (w/v) sucrose), supplemented freshly with 1 mg/ml lysozyme from chicken egg white (Sigma Aldrich), one SIGMAFAST™ protease inhibitor cocktail tablet (Sigma Aldrich), 1X PMSF (Sigma Aldrich) and 1X RNAse (0.2 mg/ml). To resuspend the cells this mixture was left rotating at 4°C for 30 min, and pipetted vigorously to completely homogenise the culture. After this, the mixture was centrifuged for 4000 g for 25-40 min at 4°C, and the supernatant was topped up to 500 mM NaCl and 5 mM $MgCl_2$ before being added to 0.5 ml settled Ni-NTA agarose (Qiagen) that had been equilibrated with megabody wash buffer (100 mM Tris-HCl

pH7.5, 500 mM NaCl, 30 mM imidazole). Binding was carried out at 4°C rotating for 45-75 min. Then, the beads were washed three times with 10 ml megabody wash buffer, and eluted in 7 ml then 5ml of megabody elution buffer (100 mM Tris-HCl pH7.5, 500 mM NaCl, 500 mM imidazole). Megabodies were concentrated using an Amicon Ultra centrifugal filter unit (Merck Millipore, NMWL of 10 kDa). Concentrated megabodies were then buffer exchanged twice into megabody storage buffer (50mM HEPES-Na, pH7.5, 150 mM NaCl) or further purified on a Superdex 200 increase 10/300 gel filtration column (Cytiva) in megabody storage buffer connected to an ÄKTA fast protein liquid chromatography system (GE Healthcare). The peak fractions containing megabody were combined and concentrated to the desired concentration. Purified megabodies were aliquoted, snap frozen in liquid N_2 an stored at -80°C for long-term usage.

6.6 RNP PURIFICATION

6.6.1 Purification of recombinant RNP from mammalian cells

Approximately $1.6x10^7$ HEK 293T cells in a 15 cm dish were transfected with 5 μg of each of pCAGGS-1918-PB1, pCAGGS-1918-PB2-TAP, pCAGGS-1918-PA, pCAGGS-1918-NP and the pPOLI-NA using Lipofectamine 2000 in accordance with the manufacturer′s instructions.

Whole cell lysates were harvested 48 h post-transfection, by centrifugation at 1000 rpm for 5 min and lysed in 2.5 ml ice-cold lysis buffer (50 mM HEPES-NaOH pH 8.0, 200 mM NaCl, 25% glycerol, 0.5% Igepal (NP40), 1 mM β-mercaptoethanol, 1x PMSF, 1x Protease Inhibitor Cocktail tablet (Sigma Aldrich) for 10 min at 4°C with occasional flicking. Lysate was cleared by centrifugation at 13000 g for 5 min, diluted 1/5 in salt water to maintain a final NaCl

concentration at 150 mM. The diluted whole cell lysate was incubated with pre-washed IgG sepharose beads (GE Healthcare) for 2 h, at 4°C. Thereafter, beads were washed 2x in 10 ml wash buffer (10 mM HEPES- NaOH, pH 8.0, 150 mM NaCl, 0.1% v/v Igepal (NP40), 1× PMSF) and 1x in 2 ml cleavage buffer (10 mM HEPES- NaOH, pH 8.0, 150 mM NaCl, 0.1% v/v Igepal (NP40), 1× PMSF, 1 mM DTT), then cleaved overnight at 4°C using 10–20 U AcTEV protease (Thermo Scientific) in a total volume of 2000 µl cleavage buffer. Beads were cleared by centrifugation at 13000 g for 1 min and eluates were analysed by SDS-PAGE followed by silver staining. Purified RNP was snap-frozen and stored at -80°C for short term usage.

6.6.2 Purification of vRNP from virions

Virion-derived vRNP was prepared as described in previous literature (Vreede & Brownlee, 2007). Briefly, MDBK cell culture supernatant from infected cells was loaded onto precooled 30% sucrose cushion (30% sucrose, 100 mM NaCl, 10 mM Tris-HCl pH 7.4, 1 mM EDTA) in thick-walled ultracentrifuge tubes. Virus was pelleted by centrifugation at 105000 g for 1.5 h, at 4˚C and resuspended in chilled resuspension buffer (100 mM NaCl, 10 mM Tris-HCl pH 7.4, 1 mM EDTA) at 4˚C. Resuspended virus was lysed by incubation in a disruption buffer (100 mM Tris-HCl pH 7.4, 100 mM NaCl, 5 mM $MgCl_2$, 1% Triton X-100, 5% glycerol, 0.5% Igepal (NP40), and freshly added 20 mg/ml lysophosphatidylcholine (Sigma Aldrich) and 1.5 mM DTT) at 31˚C for 30 min with vigorous shaking. Content of the viral lysate was fractionated by centrifugation on a discontinuous glycerol gradient (1 ml each 70%, 50%, 40%, and 33% glycerol in 50 mM Tris-HCl pH 7.4 and 150 mM NaCl). The gradient was centrifuged at 250000 g in a Beckman SW55Ti rotor for 4 h, at 4˚C. Fractions of approximately 250 µl

were collected dropwise from the bottom of the tube. Proteins were analysed by SDS-PAGE analysis followed by silver staining. Around three to four fractions enriched with vRNP were combined and pelleted in an Optimax-XP TLA 100.3 fixed angle rotor in thick-walled polycarbonate tubes at 550000 g for 4 h, at 4˚C. vRNP pellet was resuspended in vRNP buffer (100 mM HEPES-NaOH pH 8.0, 150 mM NaCl, 10% glycerol, 1× PMSF) to 1 mg/ml measured concentration on Nanodrop, aliquoted, snap-frozen in liquid N_2, and stored at −80˚C for mid-term usage. The molarity of purified vRNP was estimated by comparing it to a serial dilution of recombinant viral polymerase of known concentration.

6.7 *IN VITRO* ASSAYS

6.7.1 In vitro virion-derived vRNP activity assay

Approximately 100 ng of recombinant vRNP and indicated recombinant proteins diluted in vRNP buffer (100 mM HEPES-NaOH pH 8.0, 150 mM NaCl, 10% glycerol, 1x PMSF) were mixed with 40 ng rabbit globin mRNA (Sigma-Aldrich) or 0.5 mM pppApG (MedChemExpress) as primer donor (in the primed reaction), 1 mM ATP, 0.5 mM GTP, 0.5 mM CTP, 0.5 mM UTP, 5 mM $MgCl_2$, 1 mM DTT, and 2 U RNasin (Promega) in a reaction volume of 20 µl. After a 4-h incubation at 30˚C, RNA was extracted, 40% of which per reaction was used in primer extension analysis. Approximately 1000 ng of recombinant polymerase (WT and PB1a), 130 ng of recombinant ANP32B, and 800 ng of recombinant NP (WT, M, 8A, and M8A) were added when indicated. In titration experiments, 50 ng, 100 ng, 200 ng, 500 ng, 1,000 ng, and 2,000 ng of FluPol; 6.5 ng, 13 ng, 26 ng, 65 ng, 130 ng, and 260 ng ofANP32B; and 40 ng, 80 ng, 200 ng, 300 ng, 400 ng, and 800 ng of NP were used as indicated. In ANP32

family protein titration experiment, 26 ng, 65 ng, and 130 ng ofchANP32A; 23 ng, 57 ng, and 114 ng of human huANP32A/B; and 24.5 ng, 61 ng, and 122 ng of huANP32E were used. In ANP32 protein truncation experiment, 116 ng of 1–250, 108 ng of 1–235, 102 ng of 1–220, 96 ng of 1–208, 86 ng of 1–188, 72 ng of 1–158, and 68 ng of 1–149 were used (equal mol). Experiments were performed with three independently purified batches of vRNPs.

6.7.2 In vitro recombinant RNP activity assay

A total of 10 µl of recombinant RNP purified from mammalian cells was mixed with 20 ng rabbit globin mRNA (Sigma Aldrich) or 0.5 mM pppApG (MedChemExpress) as primer donor, 200 ng of purified nanobody or equivalent volume of nanobody storage buffer, 1mM ATP, 0.5 mM GTP, 0.5 mM CTP, 0.5 mM UTP, 5 mM $MgCl_2$, 1mM DTT, and 2U RNasin in a reaction volume of 20 µl. After a 4-h incubation at 30°C, RNA was extracted and 40% of extracted RNA from each reaction was analysed using primer extension. Experiments were performed with three independently purified batches of RNPs.

6.7.3 Polymerase transcription assay

In vitro polymerase transcription assays were carried out as previously described (Te Velthuis et al., 2016). In short, a cap structure was added to the 5’ terminus of a synthetic 11 nt long RNA (5 ′ -ppGAAUACUCAAG-3 ′) or a 20 nt long RNA (5 ′ -ppAAUCUAUAAUAGCAUUAUCC-3′) (ChemGenes) by mixing 1 µM of RNA with 0.25 µM α32P-GTP (3000 Ci/mmol; Perkin-Elmer), 0.8 mM S-adenosylmethionine, 0.5 U/µl

vaccinia virus capping enzyme (NEB) and 2.5 U/µl 2′-O-methyltransferase (NEB) in a 20 µl reaction at 37°C for 1 h. The product was isolated using a 16% (w/v) denaturing PAGE, excised, eluted overnight in distilled H_2O and desalted using NAP-10 columns (GE Healthcare). Transcription reactions were performed using the capped RNA primer in a 3 µl reaction mixture with 33 ng/ µl Sf9-purified influenza viral polymerase, 1 mM ATP, 0.5 mM UTP, 0.5 mM CTP, 0.5 mM GTP, 5 mM $MgCl_2$, 1mM DTT, 2U/µl RNasin, 0.5 µM 5′ and 4’ vRNA promoter (Table 6.3). Reactions were incubated at 30°C for 1 h and kept at 0°C overnight. Equal volume of 2x RNA loading dye (80% formamide, 1 mM EDTA and bromophenol blue and xylene cyanol) was added, followed by incubation at 95°C for 3 min. Products were resolved by 20% (w/v) denaturing PAGE and visualised via phosphorimaging on an FLA-5000 scanner (Fuji). Experiments were performed with three independently purified batches of polymerase expressed in insect cells.

6.7.4 Polymerase replication assay

Primer-independent replication assays were conducted in 3µl reactions, comprising 5mM $MgCl_2$, 0.5 µM 5′ and 3′ vRNA or cRNA promoter (Table 6.3), 1mM DTT, 1U RNasin, 1 mM ATP, 0.5 mM CTP, 0.5 mM UTP, 0.1 µM GTP, 0.15 µl α32P-GTP (3000Ci/mmol, Perkin-Elmer), and 33 ng/µl Sf9-purified influenza viral polymerase. In cases where specified, Nb8205 was included at a concentration of 16 ng/µl. Reactions were carried out for 4 h at 30°C and halted by adding 3µl loading dye (80% formamide, 1mM EDTA, bromophenol blue, xylene cyanol), followed by heating to 95°C for 3 min. The resulting reaction products were separated using a 20% (w/v) denaturing PAG and visualised via phosphorimaging on an FLA-

5000 scanner (Fuji). Experiments were performed with three independently purified batches of polymerase expressed in insect cells.

6.8 PROTEIN ANALYSIS

6.8.1 Protein Quantification

Concentrations of protein samples were measured by their absorbance at 280 nm with a Nanodrop 2000 spectrophotometer (Thermo Scientific). The exact concentration was calculated according to the Beer-Lambert equation, with an extinction coefficient estimated using Richard's Protein Calculator applied to the absorbance measurement.

6.8.2 SDS polyacrylamide gel electrophoresis (SDS-PAGE), Coomassie or silver staining

Protein samples were combined with an equal volume of SDS protein loading dye (500 mM Tris-HCl pH 6.8, 4% (w/v) SDS, 40% glycerol, 2 mM DTT, bromophenol blue) and heated to 95°C for 5 min. Gels were assembled with a 4% (v/v) 29:1 acrylamide:bis-acrylamide (AccuGel; Geneflow) stacking layer (125mM Tris-HCl pH 6.8, 0.1% SDS) and an 8-15% 29:1 acrylamide:bis-acrylamide separating layer (375mM Tris-HCl pH 8.8, 0.1% (w/v) SDS), solidified by the addition of 0.1% (w/v) APS and 0.01% (v/v) TEMED. The gels were run at 120 V for 10 min and then at 150-180 V in 1x SDS running buffer (25 mM Tris, 250 mM glycine, 0.1% (w/v) SDS). Protein visualization was achieved through Coomassie staining by

heating the SDS-PAGE gel in PageBlue Protein staining solution (Thermo Scientific) for 10-20 min, followed by destaining for 30 min in distilled H_2O. In instances where Coomassie staining reached its detection limit, proteins were visualized by silver staining using SilverXpress (Invitrogen) in accordance with the manufacturer's instructions.

6.8.3 Western blots

Instead of using Coomassie or silver staining, proteins were transferred from the gel to a nitrocellulose membrane via wet transfer in 1x transfer buffer (25 mM Tris, 250 mM glycine, 20% ethanol) at 100 V at 4°C. The membranes were then blocked overnight at 4°C, or for 1 h at room temperature, in blocking buffer (5% non-fat milk, 0.1% Tween-20, 1x PBS). Subsequently, the membranes were incubated with primary antibodies, at a concentration recommended by the manufacturer, in blocking buffer for 1 h at room temperature, or overnight at 4°C, and washed 3 times for 5-15 min each in PBS + 0.1% Tween-20. When necessary, secondary antibodies, conjugated to horseradish peroxidase (HRP), were applied to the membranes at a 1 in 10000 dilution in blocking buffer and incubated for 45 min-1 h at room temperature. The membranes were then washed 3 times for 5-15 min each in PBS + 0.1% Tween-20, and HRP was detected by incubation for 1 min with Amersham ECL Western Blotting Detection Reagents (GE Healthcare),) or 5 min with Chemiluminescent HRP Substrate (Millipore). The film was immediately exposed on Super RC Fuji Medical X-ray film (Kodak). Commercially available polyclonal antibodies were used to blot nanobodies (Jackson ImmunoResearch), PB1 (Genetex), PB2 (Genetex), and actin (Sigma Aldrich), while a custom-made antibody was employed for blotting PA and NP (a gift from the Digard Lab)

(Carrasco et al., 2004; Engelhardt et al., 2005). In all cases, a goat anti-rabbit antibody conjugated to HRP served as the secondary antibody.

6.8.4 Size exclusion chromatography (SEC)

SEC was carried out in individually specified SEC buffers on a Superdex 200 Increase 10/300 GL column using ÄKTA fast protein liquid chromatography system (GE Healthcare) at 4°C, with a flow rate at 0.5 ml/min. 280 nm UV absorbance was plotted against elution volume (ml). And the fractions of interest were analysed and concentrated using Amicon Ultra centrifugal filter (Merck Millipore).

6.8.5 CTD-binding assay

The determination of purified viral polymerase binding to a synthetic Pol II CTD mimic peptide, phosphorylated on serine-5 of the heptad repeats, in the presence of nanobodies was conducted as previously described (Martínez-Alonso et al., 2016). In brief, 20 µg of the Pol II CTD mimic peptide was immobilized onto 10 µl streptavidin agarose resin (Thermo Scientific) for 2 h, at 4°C. After two washes with wash buffer (10 mM HEPES-NaOH pH 8.0, 150 mM NaCl, 0.1 % (v/v) Igepal, 1 mM PMSF), the resin was blocked with 1% (w/v) bovine serum albumin (BSA) in wash buffer. Subsequently, 4 µg of the purified 1918 polymerase, with or without a two-fold excess of nanobodies over polymerase, was added to the peptide-coated beads, followed by incubation for 1 h, at 4°C. The formed complexes underwent five washes with wash buffer, were then heated to 95°C for 5 min in SDS protein sample buffer, and were then analysed by

SDS-PAGE. Silver staining was performed according to the manufacturer's instructions (Invitrogen).

6.8.6 Gradient cross-linking assay (Grafix)

Grafix was modified based on the procedures outlined by Kastner et al., 2008 (Kastner et al., 2008). In brief, a sucrose gradient was prepared by combining light sucrose stock (25 mM HEPES-NaOH pH 8.0, 150 mM NaCl, 5% (w/v) sucrose, 2 mM DTT) and heavy sucrose stock (25 mM HEPES-NaOH pH 8.0, 150 mM NaCl, 20% (w/v) sucrose, 2 mM DTT, 0.06% Glutaraldehyde (Alfa Aesar)) on a BioCompTM gradient making station. The sample mixture was loaded on top of each pre-prepared gradient and centrifuged at 30000 rpm, at 4°C, for 16 h, with maximum acceleration and deceleration on a SW55Ti rotor (Beckman). The cross-linking reaction was halted by adding Tris-HCl pH 7.5 at a final concentration of 100 mM. 500 µl fractions were taken from the top. Protein content from each fraction was analysed by SDS-PAGE, and fractions containing protein complexes of interest were combined, concentrated, and subjected to buffer exchange for cryo-EM analysis.

6.8.7 Mass spectrometry (MS)

Tandem MS (MS-MS) and intact MS analyses were performed at the Mass Spectrometry facility at the Structural Genomics Consortium, University of Oxford. Samples for MS-MS were submitted in the form of bands excised from fixed SDS-PAGE gels stored in distilled water. For intact mass spectrometry, samples were submitted in their original storage buffers at a concentration of 1 mg/ml.

6.9 RNA ANALYSIS

6.9.1 RNA purification

From 24-well or 6-well plates: The cell monolayer was resuspended in 250–500 µl TRI Reagent (Sigma Aldrich), followed by adding 75-150 µl of chloroform, and the mixture was centrifuged at 13000 g for 15 min at 4°C. The clear supernatant was carefully removed and combined with same volume of isopropanol to precipitate the RNA. After centrifugation at 13000 g for 30 min at 4°C, the supernatant was discarded, and the RNA pellet was washed in 500-1000 µl of 70% ethanol, followed by centrifugation at 13000 g for 10 min at 4°C. The supernatant was discarded again, and the pellet was briefly air-dried. Finally, the RNA was dissolved in 10 µl nuclease-free water (Ambion).

From *in vitro* activity assays: 500 µl TRI Reagent (Sigma Aldrich), followed by adding 150 µl of chloroform, and the mixture was centrifuged at 13000 g for 15 min at 4°C. The clear supernatant was carefully removed and combined with 250 µl isopropanol to precipitate the RNA. 1 µl of glycogen is added to help with RNA precipitation. After centrifugation at 13000 g for 30 min at 4°C, the supernatant was discarded, and the RNA pellet was washed in 1000 µl of 70% ethanol, followed by centrifugation at 13000 g for 10 min at 4°C. The supernatant was discarded again, and the pellet was briefly air-dried. Finally, the RNA was dissolved in 10 µl nuclease-free water (Ambion).

6.9.2 Radiolabelling oligonucleotides

A 1µM oligonucleotide was incubated at 37°C for 1 h with 10 U of T4 Polynucleotide Kinase (NEB), 1x kinase buffer A, and 1µl of γ^{32}P-ATP (3000 Ci/mmol; Perkin Elmer) in a total reaction volume of 10µl. Excess nucleotides were removed using the QIAquick Nucleotide Removal Kit (Qiagen) following the manufacturer's instructions, and radiolabelled primers were eluted in 30-50µl dH_2O.

6.9.3 Radiolabelling capped oligonucleotides

In a total reaction volume of 20 µl, 1 µl of 20 nt or 11 nt oligonucleotides at 10 µM was combined with 1.5 µl of γ^{32}P-GTP (3000 Ci/mmol, Perkin Elmer), 0.5 µl of SAM methyl donor, 1 µl of methyltransferase, 1 µl of capping enzyme (Epicentre Biotechnologies), and 2 µl of 10x buffer. The reaction was incubated at 37°C for 1 h and then heated to 95°C for 4 min. Subsequently, 20 µl of formamide loading dye (80% formamide, 1mM EDTA, bromophenol blue, xylene cyanol) was added, followed by an additional incubation at 95°C for 4 min. The samples were then cooled on ice and separated by RNA PAGE. The products were visualized by autoradiography on films (Biomax MR film, Carestream), and the radiolabelled 20 nt or 11nt bands were excised from the gel, eluted in 1 ml nuclease-free water (Ambion). The eluate was further desalted using NAP-10 columns, dried, and reconstituted in 100 µl nuclease-free water (Ambion).

6.9.4 Primer extension assay

In the first step of the primer extension reaction, 4 µl of each RNA sample was mixed with 0.25 µl of each radiolabelled primer in a total volume of 5 µl. When RNA sample was extracted from cells, as a control, an additional 0.1 µl of the radiolabelled 5S rRNA primer and 0.4 µl of the unlabelled 5S rRNA primer were added in a total volume of 5 µl. The primers were then annealed to the sample RNA by incubating these mixtures at 95°C for 5 min, cooling on ice, then warming to 50°C. Primer extension was initiated by adding 5 µl of reverse transcription mix (2x RT buffer, 20 mM DTT, 1 mM dNTP mix (Fermentas), 0.25 µl/reaction Superscript III reverse transcriptase (Invitrogen)) to each reaction. These were incubated at 50°C for 1 h. The products were denatured by boiling at 98°C for 5 min in10 µl formamide loading dye (80% formamide, 1mM EDTA, bromophenol blue, xylene cyanol) before separation by PAGE on a denaturing 6-20% denaturing PAGE.

6.9.5 Denaturing polyacrylamide gel electrophoresis (PAGE)

Radiolabelled RNA or DNA was separated by 6-20% polyacrylamide gel electrophoresis (PAGE) using AccuGel 19:1 acrylamide:bis-acrylamide (Geneflow) with 7 M urea in 1x Tris-Borate EDTA (TBE; 90 mM Tris, 90 mM boric acid, 2.5 mM EDTA). Gels were polymerised with 0.1% ammonium persulfate (APS) and 0.1% N, N, N′, N′-tetramethylethane-1,2-diamine (TEMED) and run at 1250V or 35W in 1x TBE using a vertical sequencing gel apparatus. Radiolabelled products were visualized by exposure to phosphorimaging screens, which were then scanned on FLA-5000 scanner (Fuji). Quantification of products was performed using ImageJ, and data analysis was carried out using Prism 8 - Prism 10 (GraphPad). For primer

extension analysis of cellular RNA samples, the signal was normalized to the 5S rRNA loading control, and the vRNA signal in the negative control lane was subtracted from all subsequent lanes.

6.10 ELECTRON MICROSCOPY

6.10.1 Negative staining electron microscopy

Nanogold labelling of the megabodies was conducted as follows: Virion-isolated vRNPs, concentrated from combined glycerol gradient fractions, were incubated with 5 nm Ni-NTA-Nanogold (Nanoprobes) at a 1:2 molar ratio for 30 min at 4°C before preparing grids for negative staining electron microscopy. Negative staining electron microscopic grids preparation and data collection were performed by the Dunn School Electron Microscopy Facility. In brief, protein samples were diluted to a final concentration of 0.01-0.1 mg/ml and applied to glow-discharged 200 mesh carbon film-coated copper grids (Agar Scientific). The protein was incubated on the grid for 1 min before blotting off the excess. The grid was washed in ddH_2O, followed by a 30-s incubation in 0.75% uranium formate and allowed to dry. Stained grids were imaged on an FEI Tecnai T12 transmission electron microscope (120 keV).

6.10.2 Cryo-electron microscopy (cryo-EM)

Cryo-EM grid preparations and data collection were carried out on a Vitrobot Mark IV (FEI). In brief, protein was diluted to 0.25-0.4 mg/ml before applying 3.3 µl onto C-flat holey carbon

copper grids (2.0 µm holes, 1.2 µm spacing, 200 mesh, Electron Microscopy Sciences) and incubating for ~30 s before plunging into liquid ethane. The Vitrobot was set to plunge at 20°C, humidity 100%, blot-force -25, and blotting time 3.5 s. Cryo-EM grids screening and data collection were carried out by Dr Jeremy Keown, Dr Loïc Carrique, or Dr Haitian Fan at the Oxford Particle Imaging Centre, Division of Structural Biology, University of Oxford. Data were collected on a 300 keV Titan Krios G3i microscope (Thermo Scientific) using a K2 summit direct electron detector (Gatan) fitted with a GIF Quantum energy filter or on a 200 keV Glacios (Thermo Scientific) microscope with a Falcon 3 detector. Automated data collection was performed using either SERIALEM or EPU versions 2.3-2.5 software. Movies were collected in counting mode with a total dose of around 40-70 e/Å^2 spread over the total number of frames. Data processing was carried out using cryoSPARC versions V2.12-3.0. Briefly, the movies were aligned using Motion Correction, keeping all the frames and dose weighting up to the total exposure. The contrast transfer function (CTF) of full dose non-weighted micrographs was estimated using Gctf-v1.06 (Zhang et al., 2016). Non-weighted micrographs were imported into cryoSPARC, where particles were automatically picked using the blob picker and 2D classified. Selected classes were used as a template for automatic picking with the template picker and 2D classified. Three *ab initio* models were generated using the selected 2D classes and refined using heterogenous refinement. Each class was refined independently using non-uniform refinement. Cryo-EM maps generated this way were imported into ChimeraX (Pettersen et al., 2021) to be fitted with structural models.

References

Almond, J. (1977). *A single gene determines the host range of influenza virus*. *270*(December), 617–618.

Arragain, B., Krischuns, T., Pelosse, M., Drncova, P., Blackledge, M., Naffakh, N., & Cusack, S. (2024). *Structures of influenza A and B replication complexes explain avian to human host adaption and reveal a role of ANP32 as an electrostatic chaperone for the apo-polymerase. April*, 4–6.

Arranz, R., Coloma, R., Chichón, F. J., Conesa, J. J., Carrascosa, J. L., Valpuesta, J. M., Ortín, J., & Martín-Benito, J. (2012). The structure of native influenza virion ribonucleoproteins. *Science*, *338*(6114), 1634–1637. https://doi.org/10.1126/science.1228172

Bang, F. B., & Andrewes, C. H. (1955). *A Short Description of the Myxovirus and Related Viruses) Group (Influenza At the International Microbiological Conference held in Rome in September 1953 , the Virus Subcommittee of the International Nomen- clature Committee made certain proposals conce*. *184*, 176–184.

Bannas, P., Hambach, J., & Koch-Nolte, F. (2017). Nanobodies and nanobody-based human heavy chain antibodies as antitumor therapeutics. *Frontiers in Immunology*, *8*(NOV), 1–13. https://doi.org/10.3389/fimmu.2017.01603

Barr, J. N., & Fearns, R. (2010). How RNA viruses maintain their genome integrity. *Journal of General Virology*, *91*(6), 1373–1387. https://doi.org/10.1099/vir.0.020818-0

Baudin, F., Bach, C., Cusack, S., & Ruigrok, R. W. H. (1994). Structure of influenza virus RNP.

I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *EMBO Journal*, *13*(13), 3158–3165. https://doi.org/10.1002/j.1460-2075.1994.tb06614.x

Beaton, A. R., & Krug, R. M. (1986). Transcription antitermination during influenza viral template RNA synthesis requires the nucleocapsid protein and the absence of a 5' capped end. *Proceedings of the National Academy of Sciences of the United States of America*, *83*(17), 6282–6286. https://doi.org/10.1073/pnas.83.17.6282

Biswas, S K, & Nayak, D. P. (1994). Mutational analysis of the conserved motifs of influenza A virus polymerase basic protein 1. *Journal of Virology*, *68*(3), 1819–1826. https://doi.org/10.1128/jvi.68.3.1819-1826.1994

Biswas, Siddhartha K., Boutz, P. L., & Nayak, D. P. (1998). Influenza Virus Nucleoprotein Interacts with Influenza Virus Polymerase Proteins. *Journal of Virology*, *72*(7), 5493–5501. https://doi.org/10.1128/jvi.72.7.5493-5501.1998

Bloch, J. S., Mukherjee, S., Kowal, J., Filippova, E. V., Niederer, M., Pardon, E., Steyaert, J., Kossiakoff, A. A., & Locher, K. P. (2021). Development of a universal nanobody-binding Fab module for fiducial-assisted cryo-EM studies of membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(47). https://doi.org/10.1073/pnas.2115435118

Blümel, J., Burger, R., Drosten, C., Gröner, A., Gürtler, L., Heiden, M., Hildebrandt, M., Jansen, B., Klamm, H., Montag-Lessing, T., Offergeld, R., Pauli, G., Seitz, R., Schlenkrich, U., Schottstedt, V., Willkommen, H., Von König, C. H. W., & Schweiger, B. (2009). Influenza virus. *Transfusion Medicine and Hemotherapy*, *36*(1), 32–39. https://doi.org/10.1159/000197314

Bouloy, M., Plotch, S. J., & Krug, R. M. (1978). Globin mRNAs are primers for the

transcription of influenza viral RNA in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, *75*(10), 4886–4890. https://doi.org/10.1073/pnas.75.10.4886

Bouvier, N. M., & Palese, P. (2008). The biology of influenza viruses. *Vaccine*, *26*(SUPPL. 4), 49–53. https://doi.org/10.1016/j.vaccine.2008.07.039

Calder, L. J., Wasilewski, S., Berriman, J. A., & Rosenthal, P. B. (2010). Structural organization of a filamentous influenza A virus. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(23), 10685–10690. https://doi.org/10.1073/pnas.1002123107

Carrasco, M., Amorim, M. J., & Digard, P. (2004). Lipid raft-dependent targeting of the influenza A virus nucleoprotein to the apical plasma membrane. *Traffic*, *5*(12), 979–992. https://doi.org/10.1111/j.1600-0854.2004.00237.x

Carrat, F., Vergu, E., Ferguson, N. M., Lemaitre, M., Cauchemez, S., Leach, S., & Valleron, A. J. (2008). Time lines of infection and disease in human influenza: A review of volunteer challenge studies. *American Journal of Epidemiology*, *167*(7), 775–785. https://doi.org/10.1093/aje/kwm375

Carrique, L., Fan, H., Walker, A. P., Keown, J. R., Sharps, J., Staller, E., Barclay, W. S., Fodor, E., & Grimes, J. M. (2020). Host ANP32A mediates the assembly of the influenza virus replicase. *Nature*, *587*(B), 638–643. https://doi.org/10.1038/s41586-020-2927-z

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., & 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), 7337–7345. https://doi.org/10.1128/jvi.02474-09

Chang, S., Sun, D., Liang, H., Wang, J., Li, J., Guo, L., Wang, X., Guan, C., Boruah, B. M., Yuan, L., Feng, F., Yang, M., Wang, L., Wang, Y., Wojdyla, J., Li, L., Wang, J., Wang, M., Cheng, G., … Liu, Y. (2015). Cryo-EM Structure of Influenza Virus RNA Polymerase Complex at 4.3Å Resolution. *Molecular Cell*, *57*(5), 925–935. https://doi.org/10.1016/j.molcel.2014.12.031

Chase, G. P., Rameix-Welti, M. A., Zvirbliene, A., Zvirblis, G., Götz, V., Wolff, T., Naffakh, N., & Schwemmle, M. (2011). Influenza virus ribonucleoprotein complexes gain preferential access to cellular export machinery through chromatin targeting. *PLoS Pathogens*, *7*(9). https://doi.org/10.1371/journal.ppat.1002187

Chaturvedi, U. C., & Shrivastava, R. (2005). Interaction of viral proteins with metal ions: Role in maintaining the structure and functions of viruses. *FEMS Immunology and Medical Microbiology*, *43*(2), 105–114. https://doi.org/10.1016/j.femsim.2004.11.004

Chenavas, S., Estrozi, L. F., Slama-Schwok, A., Delmas, B., Di Primo, C., Baudin, F., Li, X., Crépin, T., & Ruigrok, R. W. H. (2013). Monomeric Nucleoprotein of Influenza A Virus. *PLoS Pathogens*, *9*(3). https://doi.org/10.1371/journal.ppat.1003275

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., & 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), eadj9974. https://doi.org/10.1126/sciadv.adj9974

Chu, C., Fan, S., Li, C., Macken, C., Kim, J. H., Hatta, M., Neumann, G., & Kawaoka, Y. (2012). Functional analysis of conserved motifs in influenza virus PB1 protein. *PloS One*, *7*(5), 3–10. https://doi.org/10.1371/journal.pone.0036113

Cianci, C., Tiley, L., & Krystal, M. (1995). Differential activation of the influenza virus

polymerase via template RNA binding. *Journal of Virology*, *69*(7), 3995–3999. https://doi.org/10.1128/jvi.69.7.3995-3999.1995

Coloma, R., Arranz, R., de la Rosa-Trevín, J. M., Sorzano, C. O. S., Munier, S., Carlero, D., Naffakh, N., Ortín, J., & Martín-Benito, J. (2020). Structural insights into influenza A virus ribonucleoproteins reveal a processive helical track as transcription mechanism. *Nature Microbiology*, *5*(5), 727–734. https://doi.org/10.1038/s41564-020-0675-3

Coloma, R., Valpuesta, J. M., Arranz, R., Carrascosa, J. L., Ortín, J., & Martín-Benito, J. (2009). The structure of a biologically active influenza virus ribonucleoprotein complex. *PLoS Pathogens*, *5*(6), 1–10. https://doi.org/10.1371/journal.ppat.1000491

Compans, R. W., Content, J., & Duesberg, P. H. (1972). Structure of the Ribonucleoprotein of Influenza Virus. *Journal of Virology*, *10*(4), 795–800. https://doi.org/10.1128/jvi.10.4.795-800.1972

Compans, R. W., & Oldstone, M. B. A. (2014). Influenza Pathogenesis and Control - Volume I. In *Current Topics in Microbiology and Immunology* (Vol. 385). /pmc/articles/PMC7122299/?report=abstract%0Ahttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC7122299/%0Ahttp://link.springer.com/10.1007/978-3-319-11155-1

Coupland, C. E., Andrei, S. A., Ansell, T. B., Carrique, L., Kumar, P., Sefer, L., Schwab, R. A., Byrne, E. F. X., Pardon, E., Steyaert, J., Magee, A. I., Lanyon-Hogg, T., Sansom, M. S. P., Tate, E. W., & Siebold, C. (2021). Structure, mechanism, and inhibition of Hedgehog acyltransferase. *Molecular Cell*, *81*(24), 5025-5038.e10. https://doi.org/10.1016/j.molcel.2021.11.018

Curran, J., Marq, J. B., & Kolakofsky, D. (1995). An N-terminal domain of the Sendai paramyxovirus P protein acts as a chaperone for the NP protein during the nascent chain assembly step of genome replication. *Journal of Virology*, *69*(2), 849–855.

https://doi.org/10.1128/jvi.69.2.849-855.1995

Dadonaite, B., Gilbertson, B., Knight, M. L., Trifkovic, S., Rockman, S., Laederach, A., Brown, L. E., Fodor, E., & Bauer, D. L. V. (2019). The structure of the influenza A virus genome. *Nature Microbiology*, *4*(11), 1781–1789. https://doi.org/10.1038/s41564-019-0513-7

Dawson, A. R., Wilson, G. M., Coon, J. J., & Mehle, A. (2020). Post-Translation Regulation of Influenza Virus Replication. *Annual Review of Virology*, *7*, 167–187. https://doi.org/10.1146/annurev-virology-010320-070410

Dawson, A. R., Wilson, G. M., Freiberger, E. C., Mondal, A., Coon, J. J., & Mehle, A. (2020). Phosphorylation controls RNA binding and transcription by the influenza virus polymerase. *PLoS Pathogens*, *16*(9), 1–21. https://doi.org/10.1371/journal.ppat.1008841

de Beer, M. A., & Giepmans, B. N. G. (2020). Nanobody-Based Probes for Subcellular Protein Identification and Visualization. *Frontiers in Cellular Neuroscience*, *14*(November), 1–16. https://doi.org/10.3389/fncel.2020.573278

De Castro Martin, I. F., Fournier, G., Sachse, M., Pizarro-Cerda, J., Risco, C., & Naffakh, N. (2017). Influenza virus genome reaches the plasma membrane via a modified endoplasmic reticulum and Rab11-dependent vesicles. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-01557-6

De Vlugt, C., Sikora, D., Rocheleau, L., & Pelchat, M. (2019). Priming and realignment by the influenza a virus RdRp is dependent on the length of the host primers and the extent of base pairing to viral RNA. *Virology*, *536*(June), 91–100. https://doi.org/10.1016/j.virol.2019.08.002

Deng, T., Engelhardt, O. G., Thomas, B., Akoulitchev, A. V., Brownlee, G. G., & Fodor, E. (2006). Role of Ran Binding Protein 5 in Nuclear Import and Assembly of the Influenza

Virus RNA Polymerase Complex. *Journal of Virology*, *80*(24), 11911–11919. https://doi.org/10.1128/jvi.01565-06

Deng, T., Vreede, F. T., & Brownlee, G. G. (2006). Different De Novo Initiation Strategies Are Used by Influenza Virus RNA Polymerase on Its cRNA and Viral RNA Promoters during Viral RNA Replication. *Journal of Virology*, *80*(5), 2337–2348. https://doi.org/10.1128/jvi.80.5.2337-2348.2006

Dias, A., Bouvier, D., Crépin, T., McCarthy, A. A., Hart, D. J., Baudin, F., Cusack, S., & Ruigrok, R. W. H. (2009). The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature*, *458*(7240), 914–918. https://doi.org/10.1038/nature07745

Diekmann, Y., & Pereira-Leal, J. B. (2013). Evolution of intracellular compartmentalization. *The Biochemical Journal*, *449*(2), 319–331. https://doi.org/https://doi.org/10.1042/BJ20120957

Domingues, P., & Hale, B. G. (2017). Functional Insights into ANP32A-Dependent Influenza A Virus Polymerase Host Restriction. *Cell Reports*, *20*(11), 2538–2546. https://doi.org/10.1016/j.celrep.2017.08.061

Donchet, A., Oliva, J., Labaronne, A., Tengo, L., Miloudi, M., C.A. Gerard, F., Mas, C., Schoehn, G., W.H. Ruigrok, R., Ducatez, M., & 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), 1–14. https://doi.org/10.1038/s41598-018-37306-y

Eisfeld, A. J., Neumann, G., & Kawaoka, Y. (2015). At the centre: Influenza A virus ribonucleoproteins. *Nature Reviews Microbiology*, *13*(1), 28–41. https://doi.org/10.1038/nrmicro3367

Elton, D., Medcalf, E., Bishop, K., & Digard, P. (1999). Oligomerization of the influenza virus

nucleoprotein: Identification of positive and negative sequence elements. *Virology*, *260*(1), 190–200. https://doi.org/10.1006/viro.1999.9818

Engelhardt, O. G., Smith, M., & Fodor, E. (2005). Association of the Influenza A Virus RNA-Dependent RNA Polymerase with Cellular RNA Polymerase II. *Journal of Virology*, *79*(9), 5812–5818. https://doi.org/10.1128/jvi.79.9.5812-5818.2005

Etkind, P. R., & Krug, R. M. (1974). *Influenza Viral Messenger RNA*. *45*, 38–45.

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., & Fodor, E. (2019). Structures of influenza A virus RNA polymerase offer insight into viral genome replication. *Nature*, *573*(7773), 287–290. https://doi.org/10.1038/s41586-019-1530-7

Fereidouni, S., Starick, E., Karamendin, K., Genova, C. Di, Scott, S. D., Khan, Y., Harder, T., & Kydyrmanov, A. (2023). Genetic characterization of a new candidate hemagglutinin subtype of influenza A viruses. *Emerging Microbes and Infections*, *12*(2), 1–9. https://doi.org/10.1080/22221751.2023.2225645

Fodor, E, Pritlove, D. C., & Brownlee, G. G. (1994). The influenza virus panhandle is involved in the initiation of transcription. *Journal of Virology*, *68*(6), 4092–4096. https://doi.org/10.1128/jvi.68.6.4092-4096.1994

Fodor, Ervin, Devenish, L., Engelhardt, O. G., Palese, P., Brownlee, G. G., & García-Sastre, A. (1999). Rescue of Influenza A Virus from Recombinant DNA. *Journal of Virology*, *73*(11), 9679–9682. https://doi.org/10.1128/jvi.73.11.9679-9682.1999

Fodor, Ervin, Mingay, L. J., Crow, M., Deng, T., & Brownlee, G. G. (2002). A Single Amino Acid Mutation in the PA Subunit of the Influenza Virus RNA Polymerase Promotes the Generation of Defective Interfering RNAs. *Journal of Virology*, *77*(8), 5017–5020. https://doi.org/10.1128/jvi.77.8.5017-5020.2003

Fournier, E., Moules, V., Essere, B., Paillart, J. C., Sirbat, J. D., Isel, C., Cavalier, A., Rolland, J. P., Thomas, D., Lina, B., & Marquet, R. (2012). A supramolecular assembly formed by influenza A virus genomic RNA segments. *Nucleic Acids Research*, *40*(5), 2197–2209. https://doi.org/10.1093/nar/gkr985

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., & Klenk, H. D. (2011). Differential use of importin-α isoforms governs cell tropism and host adaptation of influenza virus. *Nature Communications*, *2*(1). https://doi.org/10.1038/ncomms1158

Gastaminza, P., Perales, B., Falcón, A. M., & Ortín, J. (2003). Mutations in the N-Terminal Region of Influenza Virus PB2 Protein Affect Virus RNA Replication but Not Transcription. *Journal of Virology*, *77*(9), 5098–5108. https://doi.org/10.1128/jvi.77.9.5098-5108.2003

Gavazzi, C., Yver, M., Isel, C., Smyth, R. P., Rosa-Calatrava, M., Lina, B., Moulès, V., & Marquet, R. (2013). A functional sequence-specific interaction between influenza A virus genomic RNA segments. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(41), 16604–16609. https://doi.org/10.1073/pnas.1314419110

Giese, S., Ciminski, K., Bolte, H., Moreira, É. A., Lakdawala, S., Hu, Z., David, Q., Kolesnikova, L., Götz, V., Zhao, Y., Dengjel, J., Chin, Y. E., Xu, K., & Schwemmle, M. (2017). Role of influenza A virus NP acetylation on viral growth and replication. *Nature Communications*, *8*(1). https://doi.org/10.1038/s41467-017-01112-3

Goto, H., & Kawaoka, Y. (1998). A novel mechanism for the acquisition of virulence by a human influenza A virus. *Proceedings of the National Academy of Sciences of the*

United States of America, *95*(17), 10224–10228. https://doi.org/10.1073/pnas.95.17.10224

Gui, X., Li, R., Zhang, X., Shen, C., Yu, H., Guo, X., Kang, Y., Chen, J., Chen, H., Chen, Y., & Xia, N. (2014). An important amino acid in nucleoprotein contributes to influenza A virus replication by interacting with polymerase PB2. *Virology*, *464–465*(1), 11–20. https://doi.org/10.1016/j.virol.2014.06.033

Günl, F., Krischuns, T., Schreiber, J. A., Henschel, L., Wahrenburg, M., Drexler, H. C. A., Leidel, S. A., Cojocaru, V., Seebohm, G., Mellmann, A., Schwemmle, M., Ludwig, S., & Brunotte, L. (2023). The ubiquitination landscape of the influenza A virus polymerase. *Nature Communications*, *14*(1). https://doi.org/10.1038/s41467-023-36389-0

Hanke, L., Knockenhauer, K. E., Brewer, R. C., van Diest, E., Schmidt, F. I., Schwartz, T. U., & Ploegh, H. L. (2016). The antiviral mechanism of an influenza a virus nucleoprotein-specific single-domain antibody fragment. *MBio*, *7*(6), 1–11. https://doi.org/10.1128/mBio.01569-16

Hara, K., Schmidt, F. I., Crow, M., & Brownlee, G. G. (2006). Amino Acid Residues in the N-Terminal Region of the PA Subunit of Influenza A Virus RNA Polymerase Play a Critical Role in Protein Stability, Endonuclease Activity, Cap Binding, and Virion RNA Promoter Binding. *Journal of Virology*, *80*(16), 7789–7798. https://doi.org/10.1128/jvi.00600-06

Harris, A., Cardone, G., Winkler, D. C., Heymann, J. B., Brecher, M., White, J. M., & Steven, A. C. (2006). Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(50), 19123–19127. https://doi.org/10.1073/pnas.0607614103

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., & Fodor, E. (2015). Crystal structure of the

RNA-dependent RNA polymerase from influenza C virus. *Nature*, *527*(7576), 114–117. https://doi.org/10.1038/nature15525

Hoffmann, E., Mahmood, K., Yang, C. F., Webster, R. G., Greenberg, H. B., & Kemble, G. (2002). Rescue of influenza B virus from eight plasmids. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(17), 11411–11416. https://doi.org/10.1073/pnas.172393399

Hoffmann, E., Neumann, G., Kawaoka, Y., Hobom, G., & Webster, R. G. (2000). A DNA transfection system for generation of influenza A virus from eight plasmids. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(11), 6108–6113. https://doi.org/10.1073/pnas.100133697

Huarte, M., Falcón, A., Nakaya, Y., Ortín, J., García-Sastre, A., & Nieto, A. (2003). Threonine 157 of Influenza Virus PA Polymerase Subunit Modulates RNA Replication in Infectious Viruses. *Journal of Virology*, *77*(10), 6007–6013. https://doi.org/10.1128/jvi.77.10.6007-6013.2003

Hutchinson, E. C., Charles, P. D., Hester, S. S., Thomas, B., Trudgian, D., Martínez-Alonso, M., & Fodor, E. (2014). Conserved and host-specific features of influenza virion architecture. *Nature Communications*, *5*(May), 4816. https://doi.org/10.1038/ncomms5816

Hutchinson, E. C., Denham, E. M., Thomas, B., Trudgian, D. C., Hester, S. S., Ridlova, G., York, A., Turrell, L., & Fodor, E. (2012). Mapping the Phosphoproteome of Influenza A and B Viruses by Mass Spectrometry. *PLoS Pathogens*, *8*(11). https://doi.org/10.1371/journal.ppat.1002993

Hutchinson, E. C., Orr, O. E., Liu, S. M., Engelhardt, O. G., & Fodor, E. (2011). Characterization of the interaction between the influenza A virus polymerase subunit

PB1 and the host nuclear import factor ran-binding protein 5. *Journal of General Virology*, *92*(8), 1859–1869. https://doi.org/10.1099/vir.0.032813-0

Hutchinson, E. C., von Kirchbach, J. C., Gog, J. R., & Digard, P. (2010). Genome packaging in influenza A virus. *Journal of General Virology*, *91*(2), 313–328. https://doi.org/10.1099/vir.0.017608-0

Iglesias, N. G., Filomatori, C. V., & Gamarnik, A. V. (2011). The F1 Motif of Dengue Virus Polymerase NS5 Is Involved in Promoter-Dependent RNA Synthesis. *Journal of Virology*, *85*(12), 5745–5756. https://doi.org/10.1128/jvi.02343-10

Jabour, A. M., Varghese, J., Damad, A. H., Ghailan, K. Y., & Mehmood, A. M. (2021). Examining the correlation of google influenza trend with hospital data: Retrospective study. *Journal of Multidisciplinary Healthcare*, *14*, 3073–3081. https://doi.org/10.2147/JMDH.S322185

Jensen, S., & Thomsen, A. R. (2012). Sensing of RNA Viruses: a Review of Innate Immune Receptors Involved in Recognizing RNA Virus Invasion. *Journal of Virology*, *86*(6), 2900–2910. https://doi.org/10.1128/jvi.05738-11

Jeronimo, C., Collin, P., & Robert, F. (2016). The RNA Polymerase II CTD: The Increasing Complexity of a Low-Complexity Protein Domain. *Journal of Molecular Biology*, *428*(12), 2607–2622. https://doi.org/10.1016/j.jmb.2016.02.006

Jorba, N., Coloma, R., & Ortín, J. (2009). Genetic trans-complementation establishes a new model for influenza virus RNA transcription and replication. *PLoS Pathogens*, *5*(5), e1000462. https://doi.org/10.1371/journal.ppat.1000462

Kashiwagi, T., Leung, B. W., Deng, T., Chen, H., & Brownlee, G. G. (2009). The N-terminal region of the PA subunit of the RNA polymerase of influenza A/HongKong/156/97 (H5N1) influences promoter binding. *PLoS ONE*, *4*(5).

https://doi.org/10.1371/journal.pone.0005473

Kastner, B., Fischer, N., Golas, M. M., Sander, B., Dube, P., Boehringer, D., Hartmuth, K., Deckert, J., Hauer, F., Wolf, E., Uchtenhagen, H., Urlaub, H., Herzog, F., Peters, J. M., Poerschke, D., Lührmann, R., & Stark, H. (2008). GraFix: Sample preparation for single-particle electron cryomicroscopy. *Nature Methods*, *5*(1), 53–55. https://doi.org/10.1038/nmeth1139

Kawaguchi, A., Momose, F., & Nagata, K. (2011). Replication-Coupled and Host Factor-Mediated Encapsidation of the Influenza Virus Genome by Viral Nucleoprotein. *Journal of Virology*, *85*(13), 6197–6204. https://doi.org/10.1128/jvi.00277-11

Kawakami, E., Watanabe, T., Fujii, K., Goto, H., Watanabe, S., Noda, T., & Kawaoka, Y. (2011). Strand-specific real-time RT-PCR for distinguishing influenza vRNA, cRNA, and mRNA. *Journal of Virological Methods*, *173*(1), 1–6. https://doi.org/10.1016/j.jviromet.2010.12.014

Keown, J. R., Zhu, Z., Carrique, L., Fan, H., Walker, A. P., Serna Martin, I., Pardon, E., Steyaert, J., Fodor, E., & Grimes, J. M. (2022). Mapping inhibitory sites on the RNA polymerase of the 1918 pandemic influenza virus using nanobodies. *Nature Communications*, *13*(1), 251. https://doi.org/10.1038/s41467-021-27950-w

Klumpp, K., Ruigrok, R. W. H., & Baudin, F. (1997). Roles of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. *EMBO Journal*, *16*(6), 1248–1257. https://doi.org/10.1093/emboj/16.6.1248

Knight, M. L., Fan, H., Bauer, D. L. V., Grimes, J. M., Fodor, E., & Keown, J. R. (2021). Structure of an H3N2 influenza virus nucleoprotein. *Acta Crystallographica Section F: Structural Biology Communications*, *77*, 208–214. https://doi.org/10.1107/S2053230X2100635X

Koromyslova, A. D., & Hansman, G. S. (2015). Nanobody Binding to a Conserved Epitope Promotes Norovirus Particle Disassembly. *Journal of Virology*, *89*(5), 2718–2730. https://doi.org/10.1128/jvi.03176-14

Kouba, T., Drncová, P., & Cusack, S. (2019). Structural snapshots of actively transcribing influenza polymerase. *Nature Structural and Molecular Biology*, *26*(6), 460–470. https://doi.org/10.1038/s41594-019-0232-z

Koutsakos, M., Wheatley, A. K., Laurie, K., Kent, S. J., & Rockman, S. (2021). Influenza lineage extinction during the COVID-19 pandemic? *Nature Reviews Microbiology*, *19*(12), 741–742. https://doi.org/10.1038/s41579-021-00642-4

Krischuns, T., Arragain, B., Isel, C., Paisant, S., Budt, M., Wolff, T., Cusack, S., & Naffakh, N. (2024). The host RNA polymerase II C-terminal domain is the anchor for replication of the influenza virus genome. *Nature Communications*, *15*(1), 1–19. https://doi.org/10.1038/s41467-024-45205-2

Krischuns, T., Isel, C., Drncova, P., Lukarska, M., Pflug, A., Paisant, S., Navratil, V., Cusack, S., & Naffakh, N. (2022). Type B and type A influenza polymerases have evolved distinct binding interfaces to recruit the RNA polymerase II CTD. *PLoS Pathogens*, *18*(5), 1–24. https://doi.org/10.1371/journal.ppat.1010328

Krug, R. M., Morgan, M. A., & Shatkin, A. J. (1976). Influenza viral mRNA contains internal N6-methyladenosine and 5'-terminal 7-methylguanosine in cap structures. *Journal of Virology*, *20*(1), 45–53. https://doi.org/10.1128/jvi.20.1.45-53.1976

Kumlin, U., Olofsson, S., Dimock, K., & Arnberg, N. (2008). Sialic acid tissue distribution and influenza virus tropism. *Influenza and Other Respiratory Viruses*, *2*(5), 147–154. https://doi.org/10.1111/j.1750-2659.2008.00051.x

Lamb, R. A., & Choppin, P. W. (1977). Synthesis of influenza virus polypeptides in cells

resistant to alpha-amanitin: evidence for the involvement of cellular RNA polymerase II in virus replication. *Journal of Virology*, *23*(3), 816–819. https://doi.org/10.1128/jvi.23.3.816-819.1977

Leyrat, C., Jensen, M. R., Ribeiro, E. A., Gérard, F. C. A., Ruigrok, R. W. H., Blackledge, M., & Jamin, M. (2011). The N0-binding region of the vesicular stomatitis virus phosphoprotein is globally disordered but contains transient α-helices. *Protein Science*, *20*(3), 542–556. https://doi.org/10.1002/pro.587

Li, Y., Sun, L., Zheng, W., Madina·Mahesutihan, Li, J., Bi, Y., Wang, H., Liu, W., & Luo, T. R. (2018). Phosphorylation and dephosphorylation of threonine 188 in nucleoprotein is crucial for the replication of influenza A virus. *Virology*, *520*(March), 30–38. https://doi.org/10.1016/j.virol.2018.05.002

Lidschreiber, M., Leike, K., & Cramer, P. (2013). Cap Completion and C-Terminal Repeat Domain Kinase Recruitment Underlie the Initiation-Elongation Transition of RNA Polymerase II. *Molecular and Cellular Biology*, *33*(19), 3805–3816. https://doi.org/10.1128/mcb.00361-13

Liu, J. W., Lin, S. H., Wang, L. C., Chiu, H. Y., & Lee, J. A. (2021). Comparison of Antiviral Agents for Seasonal Influenza Outcomes in Healthy Adults and Children: A Systematic Review and Network Meta-analysis. *JAMA Network Open*, *4*(8), 1–14. https://doi.org/10.1001/jamanetworkopen.2021.19151

Long, J. S., Giotis, E. S., Moncorgé, O., Frise, R., Mistry, B., James, J., Morisson, M., Iqbal, M., Vignal, A., Skinner, M. A., & Barclay, W. S. (2016). Species difference in ANP32A underlies influenza A virus polymerase host restriction. *Nature*, *529*(7584), 101–104. https://doi.org/10.1038/nature16474

Long, J. S., Idoko-Akoh, A., Mistry, B., Goldhill, D., Staller, E., Schreyer, J., Ross, C.,

Goodbourn, S., Shelton, H., Skinner, M. A., Sang, H., McGrew, M. J., & Barclay, W. (2019). Species specific differences in use of ANP32 proteins by influenza A virus. *ELife*, *8*, 1–22. https://doi.org/10.7554/eLife.45066.001

Lukarska, M., Fournier, G., Pflug, A., Resa-Infante, P., Reich, S., Naffakh, N., & Cusack, S. (2017). Structural basis of an essential interaction between influenza polymerase and Pol II CTD. *Nature*, *541*(7635), 117–121. https://doi.org/10.1038/nature20594

Mänz, B., de Graaf, M., Mögling, R., Richard, M., Bestebroer, T. M., Rimmelzwaan, G. F., & Fouchier, R. A. M. (2016). Multiple Natural Substitutions in Avian Influenza A Virus PB2 Facilitate Efficient Replication in Human Cells. *Journal of Virology*, *90*(13), 5928–5938. https://doi.org/10.1128/jvi.00130-16

Marklund, J. K., Ye, Q., Dong, J., Tao, Y. J., & Krug, R. M. (2012). Sequence in the Influenza A Virus Nucleoprotein Required for Viral Polymerase Binding and RNA Synthesis. *Journal of Virology*, *86*(13), 7292–7297. https://doi.org/10.1128/jvi.00014-12

Martínez-Alonso, M., Hengrung, N., & Fodor, E. (2016). RNA-Free and Ribonucleoprotein-Associated Influenza Virus Polymerases Directly Bind the Serine-5-Phosphorylated Carboxyl-Terminal Domain of Host RNA Polymerase II. *Journal of Virology*, *90*(13), 6014–6021. https://doi.org/10.1128/jvi.00494-16

Masters, P. S., & Banerjee, A. K. (1988). Resolution of multiple complexes of phosphoprotein NS with nucleocapsid protein N of vesicular stomatitis virus. *Journal of Virology*, *62*(8), 2651–2657. https://doi.org/10.1128/jvi.62.8.2651-2657.1988

Mavrakis, M., Méhouas, S., Réal, E., Iseni, F., Blondel, D., Tordo, N., & Ruigrok, R. W. H. (2006). Rabies virus chaperone: Identification of the phosphoprotein peptide that keeps nucleoprotein soluble and free from non-specific RNA. *Virology*, *349*(2), 422–429. https://doi.org/10.1016/j.virol.2006.01.030

McMahon, C., Baier, A. S., Pascolutti, R., Wegrecki, M., Zheng, S., Ong, J. X., Erlandson, S. C., Hilger, D., Rasmussen, S. G. F., Ring, A. M., Manglik, A., & Kruse, A. C. (2018). Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nature Structural and Molecular Biology*, *25*(3), 289–296. https://doi.org/10.1038/s41594-018-0028-6

Medalia, O., Weber, I., Frangakis, A. S., Nicastro, D., Gerisch, G., & Baumeister, W. (2002). Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography. *Science*, *298*(5596), 1209–1213. https://doi.org/10.1126/science.1076184

Moeller, A., Kirchdoerfer, R. N., Potter, C. S., Carragher, B., & Wilson, I. A. (2012). Organization of the influenza virus replication machinery. *Science*, *338*(6114), 1631–1634. https://doi.org/10.1126/science.1227270

Mondal, A., Potts, G. K., Dawson, A. R., Coon, J. J., & Mehle, A. (2015). Phosphorylation at the Homotypic Interface Regulates Nucleoprotein Oligomerization and Assembly of the Influenza Virus Replication Machinery. *PLoS Pathogens*, *11*(4), 1–24. https://doi.org/10.1371/journal.ppat.1004826

Mosley, R. T., Edwards, T. E., Murakami, E., Lam, A. M., Grice, R. L., Du, J., Sofia, M. J., Furman, P. A., & Otto, M. J. (2012). Structure of Hepatitis C Virus Polymerase in Complex with Primer-Template RNA. *Journal of Virology*, *86*(12), 6503–6511. https://doi.org/10.1128/jvi.00386-12

Muyldermans, S. (2013). Nanobodies: Natural single-domain antibodies. *Annual Review of Biochemistry*, *82*(March), 775–797. https://doi.org/10.1146/annurev-biochem-063011-092449

Narkpuk, J., Jaru-Ampornpan, P., Subali, T., Bertulfo, F. C. T., Wongthida, P., & Jongkaewwattana, A. (2015). Mechanistic study of intertypic nucleoprotein complex

formation and its inhibitory effect toward influenza A virus. *Virology*, *485*, 104–115. https://doi.org/10.1016/j.virol.2015.06.032

Ng, A. K.-L., Lam, M. K.-H., Zhang, H., Liu, J., Au, S. W.-N., Chan, P. K.-S., Wang, J., & Shaw, P.-C. (2012). Structural Basis for RNA Binding and Homo-Oligomer Formation by Influenza B Virus Nucleoprotein. *Journal of Virology*, *86*(12), 6758–6767. https://doi.org/10.1128/jvi.00073-12

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., & 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). https://doi.org/10.1371/journal.pone.0036415

Ng, A. K. L., Wang, J. H., & Shaw, P. C. (2009). Structure and sequence analysis of influenza A virus nucleoprotein. *Science in China, Series C: Life Sciences*, *52*(5), 439–449. https://doi.org/10.1007/s11427-009-0064-x

Ng, A. K., Zhang, H., Tan, K., Li, Z., Liu, J., Chan, P. K., Li, S., Chan, W., Au, S. W., Joachimiak, A., Walz, T., Wang, J., & Shaw, P. (2008). Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *The FASEB Journal*, *22*(10), 3638–3647. https://doi.org/10.1096/fj.08-112110

Nilsson-Payant, B. E., Sharps, J., Hengrung, N., & Fodor, E. (2018). The Surface-Exposed PA 51-72 -Loop of the Influenza A Virus Polymerase Is Required for Viral Genome Replication . *Journal of Virology*, *92*(16), 1–11. https://doi.org/10.1128/jvi.00687-18

Noda, T., Murakami, S., Nakatsu, S., Imai, H., Muramoto, Y., Shindo, K., Sagara, H., & Kawaoka, Y. (2018). Importance of the 1+7 configuration of ribonucleoprotein complexes for influenza A virus genome packaging. *Nature Communications*, *9*(1), 1–10. https://doi.org/10.1038/s41467-017-02517-w

Ortega, J., Martín-Benito, J., Zürcher, T., Valpuesta, J. M., Carrascosa, J. L., & Ortín, J. (2000). Ultrastructural and Functional Analyses of Recombinant Influenza Virus Ribonucleoproteins Suggest Dimerization of Nucleoprotein during Virus Amplification. *Journal of Virology*, *74*(1), 156–163. https://doi.org/10.1128/jvi.74.1.156-163.2000

Osterholm, M. T., Kelley, N. S., Sommer, A., & Belongia, E. A. (2012). Efficacy and effectiveness of influenza vaccines: A systematic review and meta-analysis. *The Lancet Infectious Diseases*, *12*(1), 36–44. https://doi.org/10.1016/S1473-3099(11)70295-X

Oymans, J., & te Velthuis, A. J. W. (2017). A Mechanism for Priming and Realignment during Influenza A Virus Replication. *Journal of Virology*, *92*(3), 1–14. https://doi.org/10.1128/jvi.01773-17

Paget, J., Caini, S., Del Riccio, M., van Waarden, W., & Meijer, A. (2022). Has influenza B/Yamagata become extinct and what implications might this have for quadrivalent influenza vaccines? *Eurosurveillance*, *27*(39), 1–6. https://doi.org/10.2807/1560-7917.ES.2022.27.39.2200753

Pardon, E., Laeremans, T., Triest, S., Rasmussen, S. G. F., Wohlkönig, A., Ruf, A., Muyldermans, S., Hol, W. G. J., Kobilka, B. K., & Steyaert, J. (2014). *A general protocol for the generation of Nanobodies for structural biology*. https://doi.org/10.1038/nprot.2014.039

Peacock, T. P., Sheppard, C. M., Staller, E., & Barclay, W. S. (2019). Host Determinants of Influenza RNA Synthesis. *Annual Review of Virology*, *6*, 215–233. https://doi.org/10.1146/annurev-virology-092917-043339

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., & Shi, Y. (2019). Structural insight into RNA synthesis by influenza D polymerase. *Nature Microbiology*, *4*(10), 1750–1759.

https://doi.org/10.1038/s41564-019-0487-5

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*, *30*(1), 70–82. https://doi.org/10.1002/pro.3943

Peukes, J., Xiong, X., Erlendsson, S., Qu, K., Wan, W., Calder, L. J., Schraidt, O., Kummer, S., Freund, S. M. V., Kräusslich, H. G., & Briggs, J. A. G. (2020). The native structure of the assembled matrix protein 1 of influenza A virus. *Nature*, *587*(7834), 495–498. https://doi.org/10.1038/s41586-020-2696-8

Pflug, A., Guilligay, D., Reich, S., & Cusack, S. (2014). Structure of influenza A polymerase bound to the viral RNA promoter. *Nature*, *516*(7531), 355–360. https://doi.org/10.1038/nature14008

Phan, T., Fay, E. J., Lee, Z., Aron, S., Hu, W.-S., & Langlois, R. A. (2021a). Segment-Specific Kinetics of mRNA, cRNA, and vRNA Accumulation during Influenza Virus Infection. *Journal of Virology*, *95*(10), 1–31. https://doi.org/10.1128/jvi.02102-20

Phan, T., Fay, E. J., Lee, Z., Aron, S., Hu, W., & Langlois, R. A. (2021b). Accumulation during In fl uenza Virus Infection. *Journal of Virology*, *95*(10), 1–14.

Pichlmair, A., Schulz, O., Tan, C. P., Naslund, T. I., Liljestrom, P., Weber, F., & Sousa, C. R. e. (2006). RIG-I–Mediated Antiviral Responses to Single-Stranded RNA Bearing 5’-Phosphates. *Science*, *314*(15018), 997–1001.

Plotch, S. J., Bouloy, M., Ulmanen, I., & 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), 847–858. https://doi.org/10.1016/0092-8674(81)90449-9

Poon, L. L. M., Pritlove, D. C., Fodor, E., & 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), 3473–3476. https://doi.org/10.1128/jvi.73.4.3473-3476.1999

Pritlove, D. C., Fodor, E., Seong, B. L., & Brownlee, G. G. (1995). In vitro transcription and polymerase binding studies of the termini of influenza A virus cRNA: Evidence for a cRNA panhandle. *Journal of General Virology*, *76*(9), 2205–2213. https://doi.org/10.1099/0022-1317-76-9-2205

Pritlove, David C., Poon, L. L. M., Fodor, E., Sharps, J., & Brownlee, G. G. (1998). Polyadenylation of Influenza Virus mRNA Transcribed In Vitro from Model Virion RNA Templates: Requirement for 5′ Conserved Sequences. *Journal of Virology*, *72*(2), 1280–1286. https://doi.org/10.1128/jvi.72.2.1280-1286.1998

Pumroy, R. A., Ke, S., Hart, D. J., Zachariae, U., & Cingolani, G. (2015). Molecular determinants for nuclear import of influenza A PB2 by importin α isoforms 3 and 7. *Structure*, *23*(2), 374–384. https://doi.org/10.1016/j.str.2014.11.015

Pymm, P., Adair, A., Chan, L. J., Cooney, J. P., Mordant, F. L., Allison, C. C., Lopez, E., Haycroft, E. R., O'Neill, M. T., Tan, L. L., Dietrich, M. H., Drew, D., Doerflinger, M., Dengler, M. A., Scott, N. E., Wheatley, A. K., Gherardin, N. A., Venugopal, H., Cromer, D., … Tham, W. H. (2021). Nanobody cocktails potently neutralize SARS-CoV-2 D614G N501Y variant and protect mice. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(19), 1–12. https://doi.org/10.1073/pnas.2101918118

Reich, S., Guilligay, D., & Cusack, S. (2017). An in vitro fluorescence based study of initiation of RNA synthesis by influenza B polymerase. *Nucleic Acids Research*, *45*(6), 3353–3368. https://doi.org/10.1093/nar/gkx043

Reich, S., Guilligay, D., Pflug, A., Malet, H., Berger, I., Crepin, T., Hart, D., Lunardi, T., Nanao, M., Ruigrok, R. W. H., & Cusack, S. (2014). Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature*, *516*(7531), 361–366. https://doi.org/10.1038/nature14009

Robb, N. C., Smith, M., Vreede, F. T., & Fodor, E. (2009). NS2/NEP protein regulates transcription and replication of the influenza virus RNA genome. *Journal of General Virology*, *90*(6), 1398–1407. https://doi.org/10.1099/vir.0.009639-0

Robb, N. C., Te Velthuis, A. J. W., Fodor, E., & Kapanidis, A. N. (2019). Real-time analysis of single influenza virus replication complexes reveals large promoter-dependent differences in initiation dynamics. *Nucleic Acids Research*, *47*(12), 6466–6477. https://doi.org/10.1093/nar/gkz313

Roberts, M. J., Bentley, M. D., & Harris, J. M. (2012). Chemistry for peptide and protein PEGylation. *Advanced Drug Delivery Reviews*, *64*(SUPPL.), 116–127. https://doi.org/10.1016/j.addr.2012.09.025

Robertson, J. S., Schubert, M., & Lazzarini, R. A. (1981). Polyadenylation sites for influenza virus mRNA. *Journal of Virology*, *38*(1), 157–163. https://doi.org/10.1128/jvi.38.1.157-163.1981

Schmidt, F. I., Hanke, L., Morin, B., Brewer, R., Brusic, V., Whelan, S. P. J., & Ploegh, H. L. (2016). Phenotypic lentivirus screens to identify functional single domain antibodies. *Nature Microbiology*, *1*(8). https://doi.org/10.1038/nmicrobiol.2016.80

Serna Martin, I., Hengrung, N., Renner, M., Sharps, J., Martínez-Alonso, M., Masiulis, S., Grimes, J. M., & Fodor, E. (2018). A Mechanism for the Activation of the Influenza Virus Transcriptase. *Molecular Cell*, *70*(6), 1101-1110.e4. https://doi.org/10.1016/j.molcel.2018.05.011

Shao, W., Li, X., Goraya, M. U., Wang, S., & Chen, J. L. (2017). Evolution of influenza a virus by mutation and re-assortment. *International Journal of Molecular Sciences*, *18*(8). https://doi.org/10.3390/ijms18081650

Shapiro, G. I., & Krug, R. M. (1988). Influenza virus RNA replication in vitro: synthesis of viral template RNAs and virion RNAs in the absence of an added primer. *Journal of Virology*, *62*(7), 2285–2290. https://doi.org/10.1128/jvi.62.7.2285-2290.1988

Sheppard, C. M., Goldhill, D. H., Swann, O. C., Penn, R., Staller, E., Baillon, L., Frise, R., Peacock, T. P., Fodor, E., & Barclay, W. S. (2023). An Influenza A virus can evolve to use human ANP32E through altering dimerization of the polymerase. *Under Review*, 1–16. https://doi.org/10.1038/s41467-023-41308-4

Short, K. R., Richard, M., Verhagen, J. H., van Riel, D., Schrauwen, E. J. A., van den Brand, J. M. A., Mänz, B., Bodewes, R., & Herfst, S. (2015). One health, multiple challenges: The inter-species transmission of influenza A virus. *One Health*, *1*, 1–13. https://doi.org/10.1016/j.onehlt.2015.03.001

Skorko, R., Summers, D. F., & Galarza, J. M. (1991). Influenza A virus in vitro transcription: Roles of NS1 and NP proteins in regulating RNA synthesis. *Virology*, *180*(2), 668–677. https://doi.org/10.1016/0042-6822(91)90080-U

Staller, E., Carrique, L., Swann, O. C., Fan, H., Keown, J. R., Sheppard, C. M., Barclay, W. S., Grimes, J. M., & Fodor, E. (2024). Structures of H5N1 influenza polymerase with ANP32B reveal mechanisms of genome replication and host adaptation. *Nature Communications*, *15*(1), 1–10. https://doi.org/10.1038/s41467-024-48470-3

Subbarao, E. K., London, W., & Murphy, B. R. (1993). A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. *Journal of Virology*, *67*(4), 1761–1764. https://doi.org/10.1128/jvi.67.4.1761-1764.1993

Subbarao, K., Chen, H., Swayne, D., Mingay, L., Fodor, E., Brownlee, G., Xu, X., Lu, X., Katz, J., Cox, N., & Matsuoka, Y. (2003). Evaluation of a genetically modified reassortant H5N1 influenza a virus vaccine candidate generated by plasmid-based reverse genetics. *Virology*, *305*(1), 192–200. https://doi.org/10.1006/viro.2002.1742

Sugiyama, K., Kawaguchi, A., Okuwaki, M., & Nagata, K. (2015). PP32 and APRIL are host cell-derived regulators of influenza virus RNA synthesis from cRNA. *ELife*, *4*(OCTOBER2015), 1–19. https://doi.org/10.7554/eLife.08939

Takizawa, N., Watanabe, K., Nouno, K., Kobayashi, N., & Nagata, K. (2006). Association of functional influenza viral proteins and RNAs with nuclear chromatin and sub-chromatin structure. *Microbes and Infection*, *8*(3), 823–833. https://doi.org/10.1016/j.micinf.2005.10.005

Tang, Y. S., Xu, S., Chen, Y. W., Wang, J. H., & 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), 4144–4154. https://doi.org/10.1093/nar/gkab203

Tarendeau, F., Boudet, J., Guilligay, D., Mas, P. J., Bougault, C. M., Boulo, S., Baudin, F., Ruigrok, R. W. H., Daigle, N., Ellenberg, J., Cusack, S., Simorre, J. P., & Hart, D. J. (2007). Structure and nuclear import function of the C-terminal domain of influenza virus polymerase PB2 subunit. *Nature Structural and Molecular Biology*, *14*(3), 229–233. https://doi.org/10.1038/nsmb1212

Taubenberger, J. K., & Morens, D. M. (2006). 1918 Influenza: the Mother of All Pandemics. *Emerging Infectious Diseases*, *12*(1), 15–22. https://doi.org/10.3201/eid1209.050979

Taubenberger, J. K., & Morens, D. M. (2008). The pathology of influenza virus infections. *Annual Review of Pathology: Mechanisms of Disease*, *3*, 499–522.

https://doi.org/10.1146/annurev.pathmechdis.3.121806.154316

Te Velthuis, A. J. W., & Fodor, E. (2016). Influenza virus RNA polymerase: Insights into the mechanisms of viral RNA synthesis. *Nature Reviews Microbiology*, *14*(8), 479–493. https://doi.org/10.1038/nrmicro.2016.87

te Velthuis, A. J. W., Long, J. C., Bauer, D. L. V., Fan, R. L. Y., Yen, H. L., Sharps, J., Siegers, J. Y., Killip, M. J., French, H., Oliva-Martín, M. J., Randall, R. E., de Wit, E., van Riel, D., Poon, L. L. M., & Fodor, E. (2018). Mini viral RNAs act as innate immune agonists during influenza virus infection. *Nature Microbiology*, *3*(11), 1234–1242. https://doi.org/10.1038/s41564-018-0240-5

Te Velthuis, A. J. W., Robb, N. C., Kapanidis, A. N., & Fodor, E. (2016). The role of the priming loop in influenza A virus RNA synthesis. *Nature Microbiology*, *1*(5), 1–16. https://doi.org/10.1038/nmicrobiol.2016.29

Thierry, E., Guilligay, D., Kosinski, J., Bock, T., Gaudon, S., Round, A., Pflug, A., Hengrung, N., El Omari, K., Baudin, F., Hart, D. J., Beck, M., & Cusack, S. (2016). Influenza Polymerase Can Adopt an Alternative Configuration Involving a Radical Repacking of PB2 Domains. *Molecular Cell*, *61*(1), 125–137. https://doi.org/10.1016/j.molcel.2015.11.016

Tsurumura, T., Qiu, H., Yoshida, T., Tsumori, Y., Hatakeyama, D., Kuzuhara, T., & Tsuge, H. (2013). Conformational polymorphism of m7GTP in crystal structure of the PB2 middle domain from human influenza a virus. *PLoS ONE*, *8*(11), 1–9. https://doi.org/10.1371/journal.pone.0082020

Tumpey, T. M., Basler, C. F., Aguilar, P. V., Zeng, H., Solorzano, A., Swayne, D. E., Cox, N. J., Katz, J. M., Taubenberger, J. K., Palese, P., & Garcia-Satre, A. (2005). Characterization of the Reconstructed 1918 Spanish Influenza Pandemic Virus. *Science*, *October*, 77–80.

Turrell, L., Hutchinson, E. C., Vreede, F. T., & Fodor, E. (2015). Regulation of Influenza A Virus

Nucleoprotein Oligomerization by Phosphorylation. *Journal of Virology*, *89*(2), 1452–1455. https://doi.org/10.1128/jvi.02332-14

Turrell, L., Lyall, J. W., Tiley, L. S., Fodor, E., & Vreede, F. T. (2013). The role and assembly mechanism of nucleoprotein in influenza A virus ribonucleoprotein complexes. *Nature Communications*, *4*, 1–11. https://doi.org/10.1038/ncomms2589

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., & Steyaert, J. (2021). Megabodies expand the nanobody toolkit for protein structure determination by single-particle cryo-EM. *Nature Methods*, *18*(1), 60–68. https://doi.org/10.1038/s41592-020-01001-6

Van Riel, D., Munster, V. J., De Wit, E., Rimmelzwaan, G. F., Fouchier, R. A. M., Osterhaus, A. D. M. E., & Kuiken, T. (2007). Human and avian influenza viruses target different cells in the lower respiratory tract of humans and other mammals. *American Journal of Pathology*, *171*(4), 1215–1223. https://doi.org/10.2353/ajpath.2007.070248

Vasin, A. V., Temkina, O. A., Egorov, V. V., Klotchenko, S. A., Plotnikova, M. A., & Kiselev, O. I. (2014). Molecular mechanisms enhancing the proteome of influenza A viruses: An overview of recently discovered proteins. *Virus Research*, *185*, 53–63. https://doi.org/10.1016/j.virusres.2014.03.015

Venkataraman, S., Prasad, B. V. L. S., & Selvarajan, R. (2018). RNA dependent RNA polymerases: Insights from structure, function and evolution. *Viruses*, *10*(2), 1–23. https://doi.org/10.3390/v10020076

Vidic, J., Noiray, M., Bagchi, A., & Slama-Schwok, A. (2016). Identification of a Novel Complex between the Nucleoprotein and PA(1-27) of Influenza A Virus Polymerase. *Biochemistry*, *55*(31), 4259–4262. https://doi.org/10.1021/acs.biochem.6b00514

Vreede, F. T., & Brownlee, G. G. (2007). Influenza Virion-Derived Viral Ribonucleoproteins Synthesize both mRNA and cRNA In Vitro. *Journal of Virology*, *81*(5), 2196–2204. https://doi.org/10.1128/jvi.02187-06

Vreede, F. T., Jung, T. E., & Brownlee, G. G. (2004). Model Suggesting that Replication of Influenza Virus Is Regulated by Stabilization of Replicative Intermediates. *Journal of Virology*, *78*(17), 9568–9572. https://doi.org/10.1128/jvi.78.17.9568-9572.2004

Vreede, F. T., Ng, A. K.-L., Shaw, P.-C., & Fodor, E. (2011). Stabilization of Influenza Virus Replication Intermediates Is Dependent on the RNA-Binding but Not the Homo-Oligomerization Activity of the Viral Nucleoprotein. *Journal of Virology*, *85*(22), 12073–12078. https://doi.org/10.1128/jvi.00695-11

Vreede, Frank T., Chan, A. Y., Sharps, J., & Fodor, E. (2010). Mechanisms and functional implications of the degradation of host RNA polymerase II in influenza virus infected cells. *Virology*, *396*(1), 125–134. https://doi.org/10.1016/j.virol.2009.10.003

Walker, A. P., & Fodor, E. (2019). Interplay between Influenza Virus and the Host RNA Polymerase II Transcriptional Machinery. *Trends in Microbiology*, *27*(5), 398–407. https://doi.org/10.1016/j.tim.2018.12.013

Walker, A. P., Sharps, J., & Fodor, E. (2020). *Mutation of an Influenza Virus Polymerase 3= RNA Promoter Binding Site Inhibits Transcription Elongation*. *94*(13), 1–9.

Wandzik, J. M., Kouba, T., Karuppasamy, M., Pflug, A., Drncova, P., Provaznik, J., Azevedo, N., & Cusack, S. (2020). A Structure-Based Model for the Complete Transcription Cycle of Influenza Polymerase. *Cell*, *181*(4), 877-893.e21. https://doi.org/10.1016/j.cell.2020.03.061

Wang, F., Sheppard, C. M., Mistry, B., Staller, E., Barclay, W. S., Grimes, J. M., Fodor, E., & 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), 5713–5725. https://doi.org/10.1093/nar/gkac410

Wanitchang, A., Jengarn, J., & Jongkaewwattana, A. (2011). The N terminus of PA polymerase of swine-origin influenza virus H1N1 determines its compatibility with PB2 and PB1 subunits through a strain-specific amino acid serine 186. *Virus Research*, *155*(1), 325–333. https://doi.org/10.1016/j.virusres.2010.10.032

Waterson, A. P. (1962). Two kinds of myxovirus. *Nature*, *193*(4821), 1163–1164. https://doi.org/10.1038/1931163a0

Wheeler, R. J., & Hyman, A. A. (2018). Controlling compartmentalization by non-membrane-bound organelles. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1747). https://doi.org/10.1098/rstb.2017.0193

Xu, C., Chan, K. H., Tsang, T. K., Fang, V. J., Fung, R. O. P., Ip, D. K. M., Cauchemez, S., Leung, G. M., Peiris, J. S. M., & Cowling, B. J. (2014). Comparative Epidemiology of Influenza B Yamagata-and Victoria-Lineage Viruses in Households. *American Journal of Epidemiology*, *182*(8), 705–713. https://doi.org/10.1093/aje/kwv110

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., & Delft, F. von. (2022). Gluebodies improve crystal reliability and diversity through transferable nanobody mutations that introduce constitutive crystal contacts. *BioRxiv*, 2022.07.26.501559. https://www.biorxiv.org/content/10.1101/2022.07.26.501559v1%0Ahttps://www.biorxiv.org/content/10.1101/2022.07.26.501559v1.abstract

Ye, Q., Krug, R. M., & Tao, Y. J. (2006). The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature*, *444*(7122), 1078–1082.

https://doi.org/10.1038/nature05379

York, A., Hengrung, N., Vreede, F. T., Huiskonen, J. T., & 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 of the United States of America*, *110*(45), 1–8. https://doi.org/10.1073/pnas.1315068110

Yuan, P., Bartlam, M., Lou, Z., Chen, S., Zhou, J., He, X., Lv, Z., Ge, R., Li, X., Deng, T., Fodor, E., Rao, Z., & Liu, Y. (2009). Crystal structure of an avian influenza polymerase PA N reveals an endonuclease active site. *Nature*, *458*(7240), 909–913. https://doi.org/10.1038/nature07720

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. *Journal of Structural Biology*, *193*(1), 1–12. https://doi.org/10.1016/j.jsb.2015.11.003

Zhang, S., Wang, J., Wang, Q., & Toyoda, T. (2010). Internal initiation of influenza virus replication of viral RNA and complementary RNA in vitro. *Journal of Biological Chemistry*, *285*(52), 41194–41201. https://doi.org/10.1074/jbc.M110.130062

Zhang, Z., Zhang, H., Xu, L., Guo, X., Wang, W., Ji, Y., Lin, C., Wang, Y., & Wang, X. (2020). Selective usage of ANP32 proteins by influenza B virus polymerase: Implications in determination of host range. *PLoS Pathogens*, *16*(10). https://doi.org/10.1371/journal.ppat.1008989

Zheng, H., Lee, H. A., Palese, P., & García-Sastre, A. (1999). Influenza A Virus RNA Polymerase Has the Ability To Stutter at the Polyadenylation Site of a Viral RNA Template during RNA Replication. *Journal of Virology*, *73*(6), 5240–5243. https://doi.org/10.1128/jvi.73.6.5240-5243.1999

Zhu, Z., Fodor, E., & Keown, J. R. (2023). A structural understanding of influenza virus genome replication. *Trends in Microbiology*, *31*(3), 308–319.

https://doi.org/10.1016/j.tim.2022.09.015

Acknowledgements

The past four years have marked an extraordinary and transformative journey – a journey that, if permitted, could easily constitute the seventh chapter of my thesis. Regrettably, I am compelled to distil my profound gratitude to the remarkable individuals into a few pages. To those mentioned below, and those inadvertently omitted, my appreciation and love for you transcend the limitations of language.

First, I extend my deepest thanks to Jamie Capel, my anchor and other half, for navigating the challenges of our long-distance relationship during my DPhil and the COVID-19 pandemic. Your unwavering positivity and unconditional support for me have been my sanctuary. None of my achievement would have happened without you.

My heartfelt thanks go to the past and current Fodorians and Grimesters, who collectively created the best lab one could dream for. I want to specially thank Haitian Fan, the MVP of my DPhil, for providing invaluable support in both scientific endeavours and emotional fortitude. I also want to specially thank Jeremy Keown, whose trust and encouragement were pivotal to my growth, especially during the initiation of the *in vitro* vRNP replication project – a turning point in my DPhil and the most beautiful 'productive procrastination'. Many thanks to Jonathan Grimes, my co-supervisor, for granting me the freedom to explore and help when needed; special thanks to Alex Walker, my fellow DPhil student, for



passing down crucial lab techniques, and for being someone I can look up to; warm thanks to Jane Sharps, for maintaining order in the lab and for always looking after all of us. I owe a debt of gratitude to my lab mates who have become my 'framily': Fangzheng Wang, Alaa Baazaoui, Hana Veler, Kuang-yu Chen, Stephanie Williams, Jack Whitehead and Alison Rep – your roles as sisters and brothers have been indispensable. Special thanks to a unique lab member, our lab grandad George Brownlee, for encouraging stimulating discussion and offering your priceless advice.

Gratitude towards the senior academics extends beyond my labs. I am especially thankful to Chris Tang, Jordan Raff, William James, Georgia Isom, and Alex Borodavka, for generous mentorship and friendship on personal levels. Special gratitude to Geoffrey Smith, a mentor who empowered me since undergrad. Thank you for igniting my love for virology, and for being a cherished friend. Massive thanks also go to Nicole Robb and Ed Hutchinson, who kindly accepted my examiner invitations. Thank you for thoroughly reviewing this thesis and making my viva full of laughter and interesting discussions.

The friendly Asian community in the Dunn school, particularly Shengpan Zhang, Li Wang, Lulu Liu, Qian Lin Lee, Yiqi Zhao and Yongxu Lu, will always have a special place in my heart, and I am grateful to Loïc Carrique, Dawn Gibbons, Lucinda Risius, Errin Johnson, and Rod Chalk for technical and administrative support.

I would also like to thank WellcomeTrust, Clarendon Fund, Microbiology Society, and the GRC community for their funding and opportunities; and Zhangji, the restaurant that witnessed my failures and successes.



I deeply appreciate all my friends who has made my DPhil journey one of the best four-year of my life, and allowed me to stay positive and motivated during the COVID-19 pandemic. Special thanks to Linnea Drexhage, Helene Borrmann, Shu Ishida, Jenny Huang, Jim Zhao, Shuxiang Cao, Jordan Penn, Dong Feng and Claire Huang, for having inspired me in different ways and motivated me with your passion. Also, special thanks to Alicia Carrington and Jack Whitehead who kindly proofread this thesis.

Despite being physically separated from my family in China for five years due to the pandemic, I never failed to feel their love, trust and compassion for me, especially from my parents and all four grandparents. Additionally, I am also grateful to the Capel family for caring for and supporting me wholeheartedly during this time. Here, I offer my deepest thanks to you all for being my strongest backers.

Special thanks also go to Yuumi, my COVID-19 adopted cat, and her four unexpected wardrobe-born kittens, especially my lovely Bluey, for offering their love and companionship during and beyond my DPhil.

Finally, I would like to express my deepest gratitude to my beloved supervisor, Ervin Fodor. Thank you for everything you have done to make these four years of my life immensely enjoyable. The lessons you have imparted, the encouragement and confidence you have instilled in me, the smile that greets me whenever I knock on your door, your irrepressible excitement about science—these are all wholesome memories I will cherish during both the joyful and challenging times that lie ahead in my life. Thank you, Ervin, for being the best and loveliest mentor I could ever ask for. It has truly been a privilege to be your student.



心臓を捧げよ。