INSIGHTS INTO VIRAL RNA SYNTHESIS BY THE INFLUENZA A VIRUS AND SARS-COV-2 RNA POLYMERASES

ALEXANDER P. WALKER

Sir William Dunn School of Pathology

and

Lincoln College

A thesis submitted for the degree of

Doctor of Philosophy

at the University of Oxford, Trinity Term 2021

Abstract

Insights into viral RNA synthesis by the influenza A virus and SARS-CoV-2 RNA polymerases

Alexander P. Walker — Doctor of Philosophy

Lincoln College — Trinity Term, 2021

Influenza A virus and SARS-CoV-2 are RNA viruses which cause severe respiratory disease and death in humans. Both encode RNA-dependent RNA polymerases (RdRps), which are responsible for synthesising viral RNA as part of a larger complex of viral proteins; for influenza A virus this complex is called a ribonucleoprotein (RNP), and for SARS-CoV-2 the replication-transcription complex (RTC). In this work I elucidate mechanisms used by influenza A virus and SARS-CoV-2 to synthesise viral RNA.

I first present evidence that the influenza A virus RdRp maintains the structural integrity of RNPs during transcription by binding to the 3′ end of viral genomic RNA. I then go on to demonstrate that in order to initiate viral RNA synthesis, the influenza A virus RdRp requires *trans*-activation through dimerization. Furthermore, I show that the influenza A virus RdRp co-opts the host protein ANP32A to help it form a second structurally-distinct dimer, which has the role of encapsidating nascent viral RNA into progeny RNPs. Collectively, these findings allow me to present a more detailed molecular model for how influenza viral RNA is synthesised and assembled into new RNPs.

One role of the influenza A virus RNP is to hide highly immunogenic viral RNA from cellular innate immune receptors such as RIG-I. The SARS-CoV-2 RTC instead synthesises 7-methylguanosine (m^7G)-capped viral RNA which does not activate RIG-I. In this work I also

show that one of the domains in the SARS-CoV-2 RdRp is a guanylyltransferase enzyme, which resolves a missing link in the mechanism of coronavirus m^7G cap synthesis. Together, these studies improve understanding of the diverse mechanisms used by RNA viruses to synthesise viral RNA, and identify multiple functionally important sites on the influenza A virus and SARS-CoV-2 RdRps which could be exploited for future antiviral strategies.

Declaration

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

Alexander P. Walker

Oxford, June 2021

Acknowledgements

I would like to thank my supervisor, Professor Ervin Fodor, for all of his guidance and encouragement throughout my studies. Much of this work was only possible because of his wisdom in setting me on the right path, and it has been an immense pleasure to work with him over the past few years. I would also like to thank Haitian Fan, Jeremy Keown and Loïc Carrique for all of their hard work and support on our collaborative projects. I am grateful to all members of Fodor lab who created such an enjoyable working environment, in particular Jane Sharps, without whom the lab could not function. Furthermore, I would like to thank the Medical Research Council for funding my studies, and Lincoln College and the Microbiology Society for funding several of my conference visits.

I would like to thank all of my friends from Oxford and Southampton for the many fond memories forged over the past few years, as well as my parents, brother and Nana for their support and enthusiasm during my studies. Lastly, thank you Charlie. Your love and support has truly made the past few years the best of my life.

Table of Contents

Abbreviations

UNITS

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

OTHER ABBREVIATIONS

A	adenine/alanine
ACE2	angiotensin converting enzyme 2
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
BiFC	bimolecular fluorescence complementation
C	cysteine/cytidine
CBD	calmodulin binding domain
cDNA	complementary DNA
chANP32A	chicken acidic nuclear phosphoprotein 32 family member A
chANP32B	chicken acidic nuclear phosphoprotein 32 family member B
CM	convoluted membrane
CO_2	carbon dioxide
CoV	coronavirus
COVID-19	coronavirus disease 2019
cRNA	complementary RNA
cRNP	complementary ribonucleoprotein
CTD	C-terminal domain
CTP	cytidine triphosphate
D	aspartic acid
DAPI	4′,6-diamidino-2-phenylindole
dH_2O	distilled water
DMEM	Dulbecco′s modified Eagle′s medium
DMV	double membrane vesicle
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
DPP4	dipeptidyl-peptidase 4
dsDNA	double-stranded DNA
dsRNA	double-stranded RNA
DTT	dithiothreitol
E	glutamic acid/envelope protein
EAV	equine arteritis virus
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
EM	electron microscopy
ER	endoplasmic reticulum

ERGIC	endoplasmic reticulum-golgi intermediate compartment
F	phenylalanine
FCoV	feline coronavirus
FCS	foetal calf serum
FluPol	influenza virus RNA polymerase
$FluPol_A$	influenza A virus RNA polymerase
$FluPol_B$	influenza B virus RNA polymerase
$FluPol_C$	influenza C virus RNA polymerase
$FluPol_D$	influenza D virus RNA polymerase
$FluPol^E$	encapsidating influenza virus RNA polymerase
$FluPol^R$	replicating influenza virus RNA polymerase
FRET	Förster resonance energy transfer
G	glycine/guanine
GDP	guanosine diphosphate
GTase	guanylyltransferase
GTP	guanosine triphosphate
H	histidine
HA	haemagglutinin
HCl	hydrochloric acid
HCoV	human coronavirus
HEK	human embryonic kidney
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HKU	University of Hong Kong
HPAI	highly pathogenic avian influenza
HRP	horseradish peroxidase
Hsp	heat shock protein
huANP32A	human acidic nuclear phosphoprotein 32 family member A
huANP32B	human acidic nuclear phosphoprotein 32 family member B
I	isoleucine
IgG	immunoglobulin G
IP	immunoprecipitation
K	lysine

KCl	potassium chloride
LB	Luria broth
LCAR	low-complexity acidic region
LRR	leucine-rich repeat
M	matrix protein (coronavirus)
M1	matrix protein 1 (influenza virus)
M2	matrix protein 2 (influenza virus)
m^7G	7-methylguanosine
MDBK	Madin-Darby bovine kidney
MEM	minimum essential medium
MERS	middle east respiratory syndrome
MES	2-(N-morpholino)ethanesulfonic acid
$MgCl_2$	magnesium chloride
MHV	mouse hepatitis virus
mRNA	messenger RNA
N	asparagine/nucleocapsid protein
NA	neuraminidase
Nb	nanobody
NEP	nuclear export protein
NES	nuclear export signal
NiRAN	nidovirus RdRp-associated nucleotidyltransferase
NLS	nuclear localisation signal
NP	nucleoprotein
NS1	non-structural protein 1 (influenza virus)
nsNSV	non-segmented negative-strand RNA virus
nsp	non-structural protein (coronavirus)
NT	northern territory
NTP	nucleoside triphosphate
ORF	open reading frame
P	proline/phosphoprotein
P3	polymerase 3
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
Pol II	RNA polymerase II
poly(A)	polyadenylic acid
poly(U)	polyuridylic acid
PRNTase	polyribonucleotidyltransferase
Q	glutamine
qPCR	quantitative polymerase chain reaction
R	arginine
RBD	receptor-binding domain
RdRp	RNA-dependent RNA polymerase
RIG-I	retinoic acid-inducible gene I
RNA	ribonucleic acid
RNP	ribonucleoprotein
RppH	RNA 5′ pyrophosphohydrolase
rRNA	ribosomal RNA
RT	reverse transcription
RTC	replication-transcription complex
S	serine/spike protein
s.e.m.	standard error of the mean
S-5-P	serine-5-phosphorylated
SAH	S-adenosylhomocysteine
SAM	S-adenosylmethionine
SARS	severe acute respiratory syndrome
SDS	sodium dodecyl sulfate
sgRNA	subgenomic RNA

SIM	small ubiquitin-like modifier-interacting motif
ssRNA	single-stranded RNA
svRNA	small viral RNA
swANP32A	swine acidic nuclear phosphoprotein 32 family member A
T	threonine
TAP	tandem affinity purification
TBE	tris-borate-ethylenediaminetetraacetic acid
TEMED	N,N,N′,N′-tetramethylethylenediamine
TEV	tobacco etch virus protease
TMPRSS2	transmembrane serine protease 2
Tris	tris(hydroxymethyl)aminomethane
TRS-B	transcription regulatory sequence-body
TRS-L	transcription regulatory sequence-leader
U	uracil
UTP	uridine triphosphate
UTR	untranslated region
UV	ultraviolet
V	valine
VC	venus C-terminal
VN	venus N-terminal
VP	vesicle packet
vRNA	viral RNA
vRNP	viral ribonucleoprotein
VSV	vesicular stomatitis virus
WSN	Wilson Smith neurotropic
Y	tyrosine
YFP	yellow fluorescent protein
ZBD	zinc binding domain

CHAPTER 1

An introduction to influenza virus RNA synthesis

1.1 OVERVIEW OF INFLUENZA VIRUS

1.1.1 *Background on influenza viruses*

Influenza viruses are members of the *Orthomyxoviridae* family of negative-strand RNA viruses. There are four known influenza virus species, named A, B, C and D, and influenza A viruses are further divided into subtypes based on the identity of their surface proteins, haemagglutinin (HA) and neuraminidase (NA). Influenza A and B viruses are responsible for seasonal epidemics, which occur during winter in both the northern and southern hemispheres and cause up to 650000 deaths per year. Most fatalities occur in ′high risk′ groups, which include children under 5 years, the elderly, and people with chronic medical conditions (1). In addition, influenza A virus pandemics have caused tens of millions of deaths in the past, with the 1918 H1N1 Spanish influenza pandemic alone being estimated to have caused at least 50 million deaths (2).

Vaccines are normally used to control seasonal influenza epidemics. Vaccines induce production of antibodies targeting the viral surface proteins, primarily HA, and must be updated each year as HA acquires mutations through antigenic drift (3). Viruses can also undergo antigenic shift and acquire a different HA protein type, which occurs through reassortment of the segmented viral genome. Antigenic shifts can lead to the rapid emergence of novel influenza viruses for which there is no natural immunity in the population and no effective vaccines available (4). In this event, there are several antiviral drugs available to treat influenza virus infection. Zanamivir (Relenza) and oseltamivir (Tamiflu) are the two most widely used drugs which target the NA protein, and baloxavir and favipiravir target the viral RNA-dependent RNA polymerase (FluPol). However, zanamivir and oseltamivir have limited clinical efficacy and baloxavir resistance is an

emerging problem (5–8). Given the danger posed by influenza viruses there is an urgent need for new therapeutic options, which requires a more in-depth understanding of the mechanisms underlying the viral life cycle.

1.1.2 *Influenza virion and genome structure*

Influenza virions are enveloped particles that can adopt spherical, bacilliform, or filamentous morphologies with a diameter of 100nm (**Fig. 1.1A, B**). Their surface proteins are embedded in the outer membrane and include trimeric HA, used for host cell attachment, tetrameric NA, used in budding from the host cell, and tetrameric matrix 2 (M2), a proton channel (9). The major structural component of influenza virions is matrix 1 (M1) protein, which coats the inside of the virion membrane (10). Influenza virions also contain non-structural protein 1 (NS1), nuclear export protein (NEP), and a large number of host cellular proteins (11).

Inside the virion, single-stranded negative-sense viral RNA (vRNA) genome segments are packaged into viral ribonucleoproteins (vRNPs). Influenza A and B viruses have 8 different vRNA segments, and influenza C and D viruses have 7, but virions from all species contain on average 8 vRNPs (12, 13). The influenza A virus genome totals 13.6 kilobases, with segments ranging from 0.9 to 2.3 kilobases (**Fig. 1.1C**). vRNA is 5′ triphosphorylated and non-polyadenylated, and each segment contains one or two open reading frames (ORFs) for viral proteins, flanked by 5′ and 3′ untranslated regions (UTRs) and short promoter elements. In each vRNP heterotrimeric FluPol binds specifically to the 5′ and 3′ promoters of each vRNA segment, and nucleoprotein (NP) binds non-specifically along the entire length of the vRNA (13).

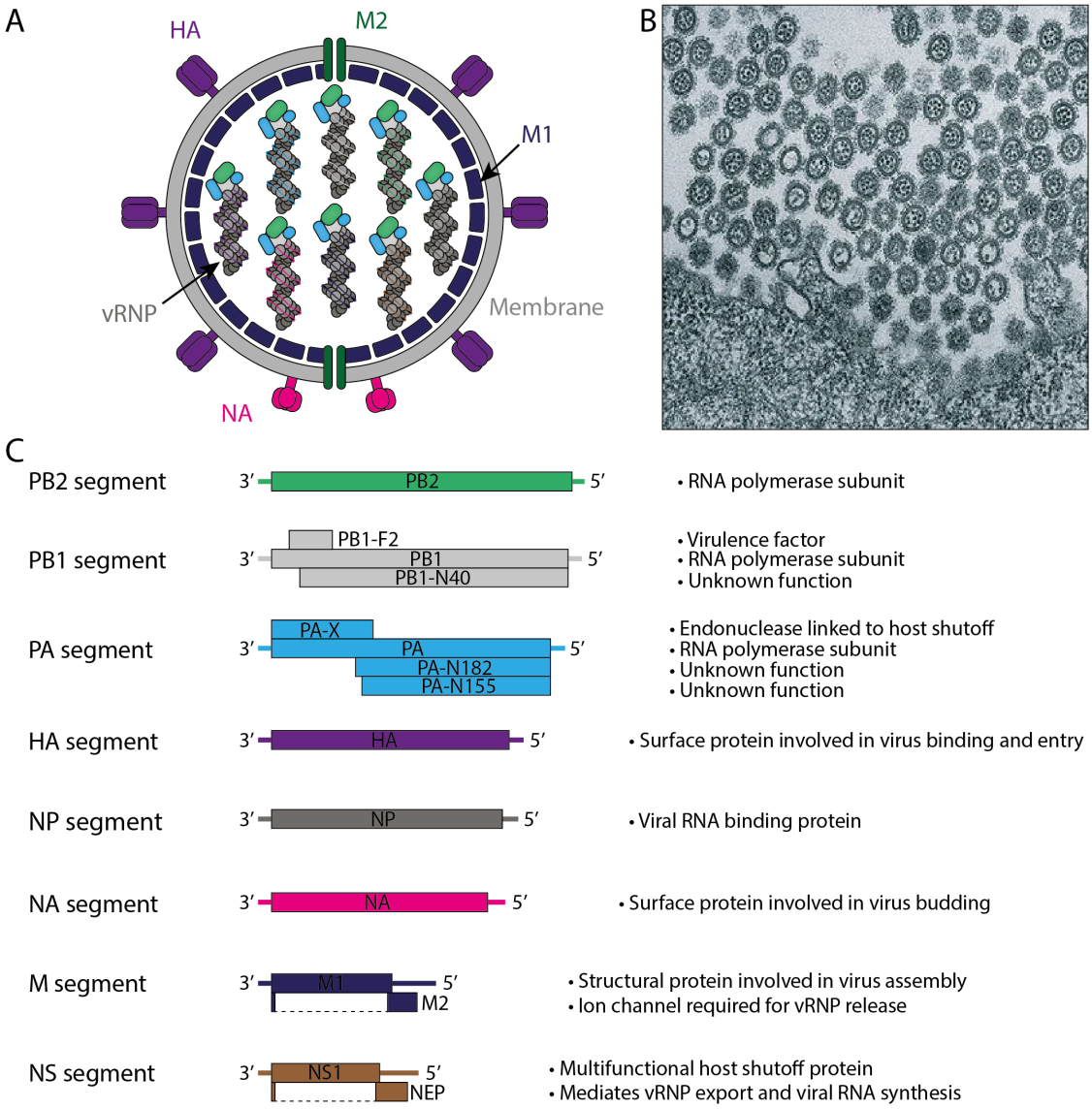

Figure 1.1 **Virion structure and viral genome segments.** A, Diagram of an influenza virion showing the major structural components. B, Electron micrograph of budding influenza A/WSN/33 (H1N1) virions. Adapted from Noda *et al*. 2006 (14). C, Influenza A virus genome segments arranged from longest to shortest (left). Negative-sense vRNA segments are shown in the 3′ to 5′ orientation, with ORFs encoded by the corresponding positive-sense viral mRNAs represented by rectangles. Dotted lines in the M and NS segments indicate spliced mRNAs. A brief description of the viral protein function is shown to the right of each ORF.

1.1.3 *Viral life cycle*

HA attaches influenza virions to target cells by binding to sialic acid residues on plasma membrane glycoproteins (**Fig. 1.2A**). HA from human-adapted influenza viruses bind to sialic acids which have an α2-6 linkage to the neighbouring galactose residue, which are located in the upper respiratory tract. Virions then enter cells through receptor-mediated endocytosis, and the resulting early endosome becomes acidified (**Fig. 1.2B**). Low pH triggers a conformational change in HA, which inserts a hydrophobic fusion peptide into the endosomal membrane. Membrane fusion is dependent on immature HA (HA0) having been cleaved into the mature proteins HA1 and HA2 by host proteases, which include TMPRSS2 (**Fig. 1.2C**) (15, 16). Simultaneously, the M2 proton channel allows acidification of the virion interior which disrupts the interaction between M1 and vRNPs (**Fig. 1.2D**). Further steps, such as M1 binding to cytoplasmic transportin-1, also promote vRNP release and debundling in the cytosol (17). After vRNP release, importin-α1 or -α5 binds to the nuclear localisation signal (NLS) of NP and mediates trafficking of the vRNPs to the cell nucleus (13).

Once in the cell nucleus, FluPol in a vRNP transcribes the negative-sense vRNA to produce viral mRNA (**Fig. 1.2E**). Viral genome transcription is dependent on cellular RNA polymerase II (Pol II) activity, as FluPol uses a cap snatching mechanism to initiate transcription. FluPol synthesises 5′ 7-methyl-guanosine (m^7G)-capped, 3′ polyadenylated viral mRNAs which are exported from the nucleus and translated using the same pathways as host mRNAs (**Fig. 1.2F**) (18). Viral proteins accumulate, stimulating replication of the viral genome at 4-6 hours post infection (**Fig. 1.2G, H**) (19). vRNPs replicate by synthesising positive-sense 5′ triphosphorylated, non-polyadenylated complementary RNA (cRNA) intermediates, which are packaged into complementary ribonucleoproteins (cRNPs). cRNPs are structurally similar to vRNPs, and synthesise vRNA which is packaged into progeny vRNPs (20).

vRNPs are exported from the nucleus by M1 and NEP, which recruits the nuclear export receptor CRM1. vRNPs are thought to traffic to the plasma membrane on recycling endosomes, which

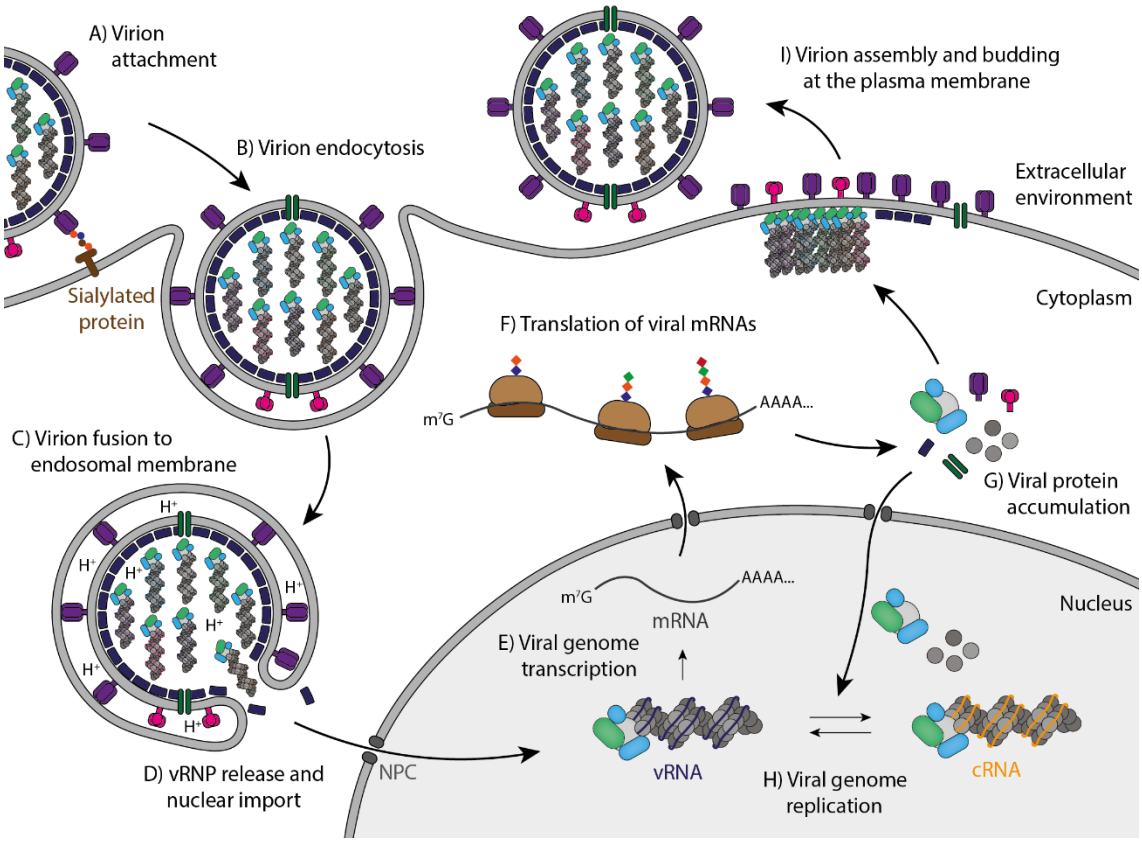

Figure 1.2 **Viral life cycle.** A, Influenza virions attach to the target cell membrane by binding to sialic acid residues on glycosylated proteins. B, Virions are endocytosed. C, Low pH causes HA proteins to induce membrane fusion. D, Low pH inside the virion, facilitated by the M2 ion channel, triggers vRNP release. vRNPs are then imported to the cell nucleus through the nuclear pore complex. E, vRNPs transcribe the viral genome segments to produce viral mRNAs. F, Viral mRNAs are exported from the nucleus and translated by host machinery. G, Viral proteins accumulate, and FluPol and NP are imported to the cell nucleus. H, The availability of newly synthesised FluPol and NP activates viral genome replication by vRNPs. cRNA is synthesised from vRNA, and progeny vRNA is synthesised from cRNA. With FluPol and NP, this process assembles new vRNPs which are exported from the cell nucleus. I, Virions assemble from newly synthesised viral proteins and vRNPs at the cell plasma membrane. Virion budding is aided by NA, which cleaves sialic acid residues from cell surface glycoproteins to prevent reattachment to the parent cell.

they associate with through an interaction between FluPol in the vRNPs and cellular Rab11 (**Fig. 1.2I**) (13). vRNPs assemble into bundles of 8 at the plasma membrane, and are thought to use RNA-RNA interactions to selectively package all of the different vRNA segments (21). Newly-synthesised HA and NA are present on lipid rafts at the plasma membrane where they associate with M1, which in turn recruits M2. M1 polymerises into a helical structure which directs formation of the budding virion (10). As the virion buds, NA cleaves the sialic acid residues of surrounding plasma membrane glycoproteins to prevent re-attachment or aggregation with other virions (9).

1.2 **INFLUENZA VIRUS RIBONUCLEOPROTEINS**

Influenza virus vRNPs consist of a vRNA segment bound at both ends by FluPol, coated with NP which binds non-specifically along the entire length of the vRNA. Several cryo-EM studies have been able to resolve the structure of the NP double helix, however, relatively little is known about the structure of the vRNP ends (**Fig. 1.3A**) (22–25).

Structurally, vRNPs are rod-like and consist of a flexible double helix of NP molecules with 5-6 NP per turn (22). NP oligomerises by the tail-loop of one molecule binding in the basic docking site of the neighbouring molecule, and this interaction drives the helical shape of the vRNP (25, 26). The putative RNA-binding groove of NP is on the outside face of the vRNP, so vRNA is thought to wrap around the NP core (24). At one end of the vRNP is a loop of NP molecules which connects the two antiparallel strands of the NP double helix, and the other end is a FluPol molecule bound to the vRNA segment. Current studies have not identified any major structural differences between cRNPs and vRNPs, however, there may be subtle distinctions that have not yet been identified (27).

vRNPs contain all the necessary viral components to transcribe their vRNA segment. FluPol is a multifunctional enzyme responsible for all the catalytic processes necessary to obtain a 5′ m^7G

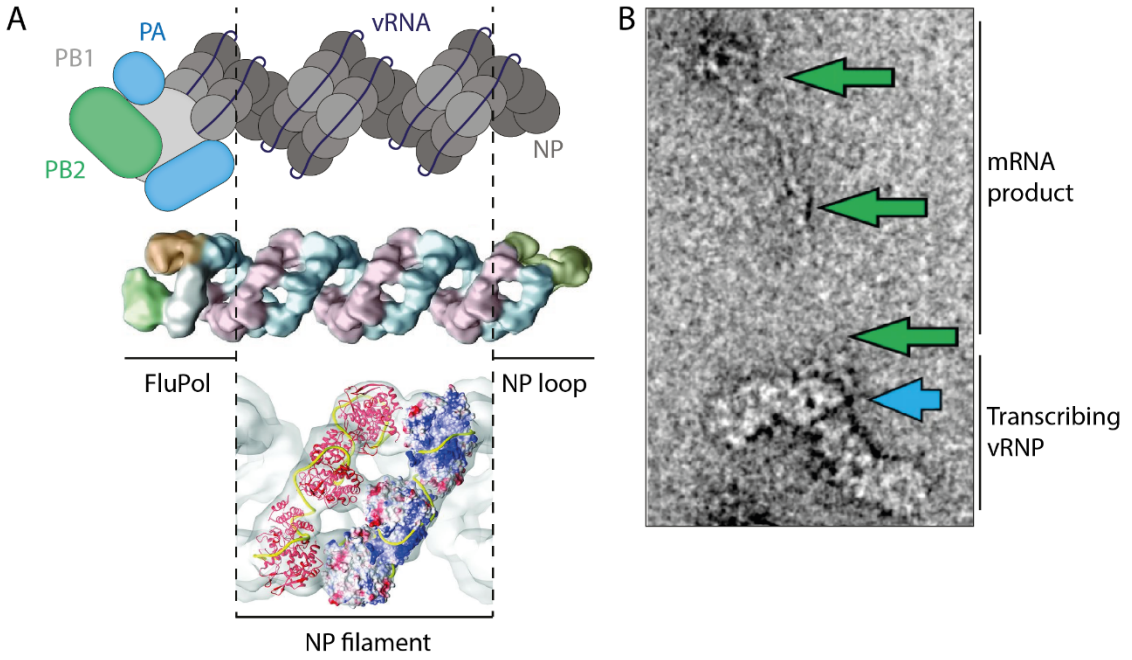

Figure 1.3 **Influenza ribonucleoprotein structure.** A, Cartoon of a vRNP showing FluPol bound to the vRNA termini, vRNA bound to a helical filament of NP, and a loop of NP at the distal end (top). This cartoon is compared with a low-resolution vRNP model derived from cryo-electron tomography of intact influenza A/WSN/33 virions (middle), adapted from (23). Cryo-EM models are sufficient to unambiguously fit NP monomers into the vRNP helical filament structure (bottom), adapted from (23). The path of the vRNA (yellow) around the vRNP can be predicted from this model. B, Cryo-EM image of a transcribing vRNP with a visible mRNA product, adapted from (24). Density suggested to correspond to FluPol is highlighted with a blue arrow, and mRNA product is indicated with green arrows.

cap, synthesise the mRNA transcript, and carry out polyadenylation (18). NP is thought to be required as an elongation factor, based on the observation that FluPol only requires NP to transcribe vRNA templates over a certain length (28). How NP performs this function is unclear, though it may involve melting secondary RNA structures in the vRNA template. vRNPs have been shown to retain the NP double helix structure during transcription, which would allow them to perform multiple rounds of RNA synthesis. These cryo-EM studies of transcribing vRNPs suggest that FluPol can move along one strand of NP without disrupting the overall structure,

which is facilitated by the flexibility of the vRNP (**Fig. 1.3B**). The ′processive helical track′ model suggests that the vRNA loops into the FluPol active site with minimal disruption to the vRNP structure, although the molecular details of this mechanism have not yet been established (24).

1.3 **THE INFLUENZA VIRUS RNA POLYMERASE**

1.3.1 *RNA polymerase structure*

FluPol is the vRNP component responsible for all catalytic activities associated with RNA synthesis. FluPol is a 255kDa heterotrimeric complex composed of the subunits polymerase basic 1 (PB1), polymerase basic 2 (PB2), and polymerase acidic (PA) (**Fig. 1.4A**). High resolution structures of FluPol from all known influenza virus species are available, solved by crystallography or cryo-EM, and these provide detailed insight into the mechanisms underlying FluPol activity (29).

PB1 is the core of the complex, and has a right-handed RNA-dependent RNA polymerase domain (RdRp) with palm, thumb and fingers subdomains (**Fig. 1.4B**). This domain includes active site motifs A-F which are required for polymerase activity; of these, residues D445 and D446 in motif C coordinate the catalytic magnesium ions. A flexible priming loop located on the PB1 thumb projects into the active site, and a long β-ribbon protrudes from the tip of the PB1 fingers which contains the bipartite nuclear localisation signal (NLS).

In addition to the canonical RdRp domain, PB1 has a C-terminal extension (C-ext) which forms a 4-helix bundle with the N-terminus of PB2. This interaction keeps PB2 associated with the rigid FluPol core composed of PB1, the PB2 N-terminus and the PA C-terminus. The PB2 lid domain forms part of the core, and is important for separating the product RNA strand from template RNA during polymerisation. The rest of PB2 is made up of a series of flexibly-linked globular domains, including the cap-binding domain (amino acid residues 319-481) which binds to m^7G

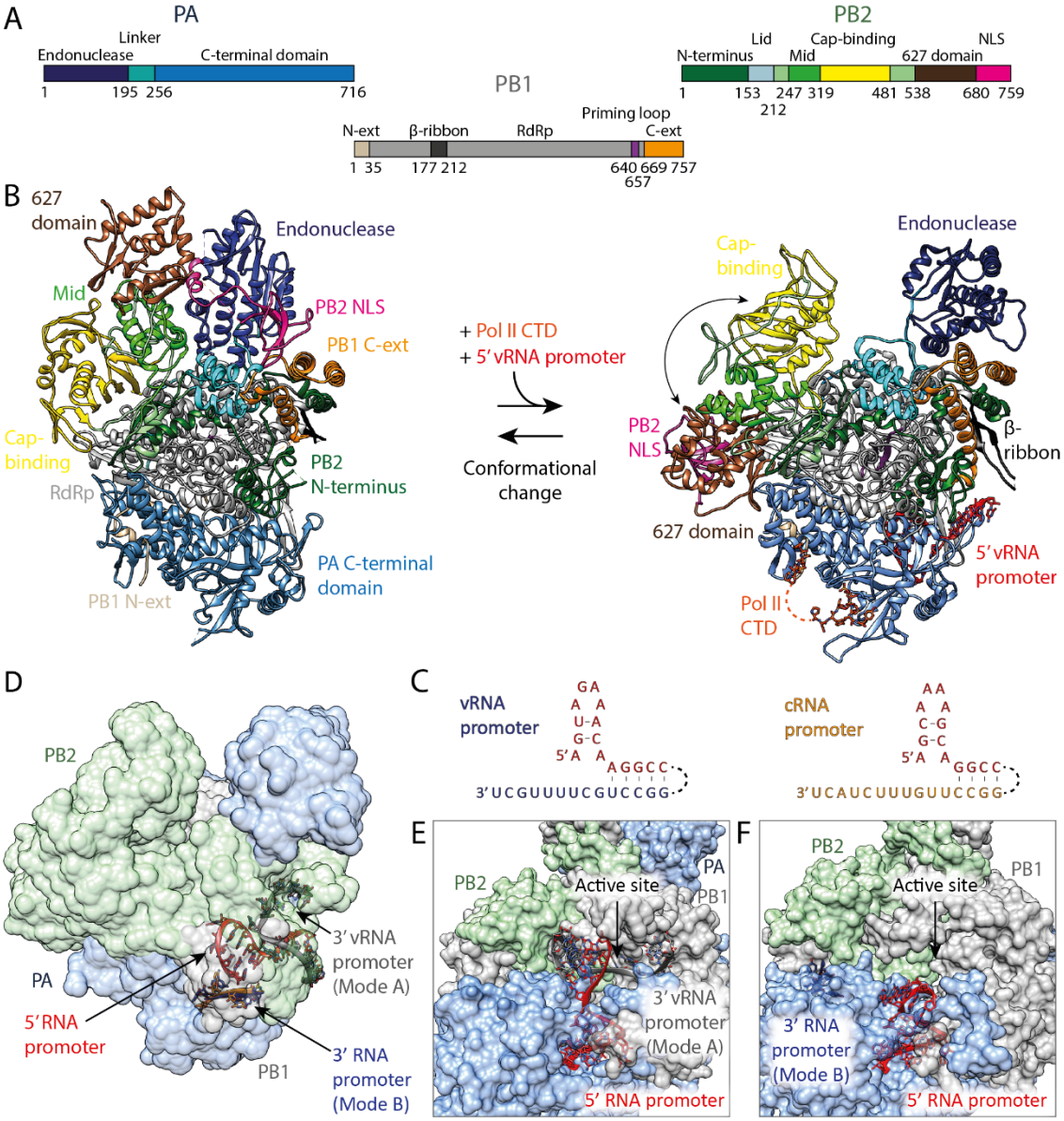
A
PA
Linker
Endonuclease
C-terminal domain
1 195 256 716
PB1
N-ext
β-ribbon
RdRp
Priming loop
C-ext
1 35 177 212 640 669 757
657
PB2
Lid
Cap-binding
NLS
N-terminus
Mid
627 domain
1 153 247 319 481 538 680 759
212
B
627 domain
Endonuclease
Mid
PB2 NLS
PB1 C-ext
Cap-binding
PB2 N-terminus
RdRp
PA C-terminal domain
PB1 N-ext
+ Pol II CTD
+ 5′ vRNA promoter
Conformational change
Cap-binding
Endonuclease
PB2 NLS
β-ribbon
627 domain
5′ vRNA promoter
Pol II CTD
C
vRNA promoter
G A
A A
U-A
G-C
5′A A
A G G C C
3′ U C G U U U U C G U C C G G
cRNA promoter
A A
A A
C-G
G-C
5′A A
G G C C
3′ U C A U C U U U G U U C C G G
D
PB2
3′ vRNA promoter (Mode A)
PA
5′ RNA promoter
3′ RNA promoter (Mode B)
PB1
E
PB2
Active site
PA
PB1
3′ vRNA promoter (Mode A)
5′ RNA promoter
F
PB2
Active site
PB1
3′ RNA promoter (Mode B)
5′ RNA promoter

Figure 1.4 **Influenza virus RNA polymerase structure and RNA binding sites.** A, Schematic of $FluPol_A$ subunit domain structures. B, Ribbon diagram of $FluPol_A$ in the transcriptionally inactive form (left; PDB: 6QNW). Binding to the 5′ vRNA promoter and S-5-P Pol II CTD promotes a conformational change into the transcriptionally active form (right; PDB: 6RR7). S-5-P Pol II CTD overlayed from PDB: 5M3H. In this conformation the PB2 627 and cap-binding domains undergo a major rearrangement so that the cap binding site is accessible. C, Schematic of the vRNA (left) and cRNA (right) promoter regions located at the 5′ and 3′ termini of viral genome segments. The 5′ RNA promoters (red) have distinct sequences but adopt similar hairpin structures and base pair with the 3′ RNA promoters. D, Semi-transparent $FluPol_A$ structure showing the binding sites for vRNA and cRNA promoters. 5′ RNA promoters (red; PDB: 4WRT) have one binding site, while the 3′ RNA promoters have two binding sites on the surface. Only the 3′ vRNA promoter has been found in the Mode A site (grey; PDB: 4WRT), but 3′ vRNA (blue; PDB: 6ABF) and 3′ cRNA (gold; PDB: 6QX3) promoters can both bind in the Mode B site. E, Position of the 3′ vRNA promoter in the Mode A site relative to the FluPol active site (PDB: 4WRT). F, Position of the 3′ RNA promoter in the Mode B site relative to the FluPol active site (PDB: 6ABF).

(20, 29). The cap-binding domain achieves specificity by using amino acid residues H357 and F404 to create an aromatic sandwich that identifies the positively-charged N7 with cation-π interactions, which is a similar mode of recognition to cellular m^7G-binding proteins (30). Downstream of the cap-binding domain is the PB2 627 domain (amino acid residues 538-680), named because it contains the PB2 627 residue implicated in influenza virus host adaption (31, 32). Finally, the C-terminus of PB2 contains the NLS. The PA subunit is divided into N- and C-terminal regions, joined by a linker which wraps around the FluPol core. The PA N-terminus is an endonuclease domain used for cap snatching, and contains a number of conserved acidic residues such as D108 which coordinate catalytic manganese ions (33). The larger PA C-terminal domain (CTD) is part of the rigid FluPol core, and is intimately associated with the PB1 palm and thumb subdomains. It has a major role in promoting transcription initiation by binding to the Pol II CTD, and may have other functions which have not yet been characterised (18).

The FluPol core is relatively static, however, PB2 and the PA endonuclease can undergo major conformational rearrangements (**Fig. 1.4B**). The 'transcriptionally active' conformation was first observed for vRNA-bound influenza A virus FluPol ($FluPol_A$) and $FluPol_B$, and is compatible with cap snatching activity since the PB2 cap-binding domain is orientated towards the PA endonuclease (34, 35). In this conformation the cap-binding domain can also be orientated towards the PB1 product exit channel, which presumably allows transcription initiation after cap snatching. Functional studies suggest that 5′ vRNA promoter and Pol II CTD binding promotes the transcriptionally active conformation, and therefore enhances transcription initiation activity (36). A ′transcriptionally inactive′ conformation has also been identified for apo $FluPol_A$, apo $FluPol_C$ and cRNA-bound $FluPol_B$, with the PB2 627 and cap-binding domains inverted and the PB2 NLS interacting with the PA endonuclease. Cap snatching is not possible in this conformation since the m^7G binding site is occluded, so it may be relevant for other FluPol functions such as viral genome replication (37–39).

FluPol in an RNP binds to the vRNA or cRNA 5′ and 3′ promoters, and this interaction is also essential for RNA polymerase activity (40–43). The 5′ and 3′ RNA promoters interact with each other through 4 or 5 Watson-Crick base pairs for cRNA or vRNA respectively, which are disrupted when RNA synthesis initiates (**Fig. 1.4C**) (44, 45). The 5′ vRNA promoter is 15 nucleotides in length and varies from the 14-nucleotide 5′ cRNA promoter at several positions. However, both 5′ promoters form an extremely similar hook-like structure by Watson-Crick base pairing, which was predicted by early functional studies and confirmed by structural studies (46–48). The 5′ vRNA and cRNA promoters both bind in a pocket formed primarily by the PB1 subunit, with the PA CTD ′arch′ wrapped around (**Fig. 1.4D**) (34, 38). The 3′ vRNA and cRNA promoters are more dynamic and bind with lower affinity than the 5′ RNA promoters (49). Structural studies have observed the 14-nucleotide 3′ vRNA promoter bound to FluPol in three distinct conformations, and current evidence suggests that the 15-nucleotide 3′ cRNA promoter can adopt similar conformations. The first conformation is in the RdRp active site, where the 3′

vRNA or cRNA promoters are templates for RNA polymerase activity. Only the 3′ vRNA promoter has been structurally resolved in the active site, bound to capped RNA or to nascent RNA product, but Förster resonance energy transfer (FRET) studies indicate that the 3′ cRNA promoter can also enter the active site (37, 44, 49, 50). The second conformation is called the ′Mode A site′, and is located on the FluPol surface near the entrance to the RdRp active site (**Fig. 1.4E**). The Mode A site has only been observed for 3′ vRNA promoter bound to $FluPol_B$, and in this structure the 3′ vRNA promoter adopts a U-shape and interacts with basic residues in the PB1 subunit (51). Several FRET-based studies suggest that the 3′ cRNA promoter can also occupy the Mode A site, although this has not been observed structurally. These studies further show that the 3′ vRNA promoter can move between the RdRp active site and the Mode A site, and therefore this site may be important for initiating RNA synthesis (49, 50). The final conformation is called the ′Mode B site′, which is a groove formed primarily by the PB1 thumb and PA CTD, with some interactions made by the PB2 N-terminus (**Fig. 1.4F**). The 3′ vRNA and cRNA promoters have both been observed in this site by cryo-EM. Further structural studies have identified the Mode B site in $FluPol_A$, $FluPol_B$, and $FluPol_D$, and a similar 3′ vRNA promoter binding site in structures of the RdRps from La Crosse orthobuynavirus and Machupo mammarenavirus (37, 52–54). This suggests a conserved function for the Mode B site, and a recent cryo-EM study supports a model where this site is important for binding to 3′ vRNA promoter after it emerges from the RdRp active site during transcription (44).

1.3.2 *Viral genome transcription*

vRNP-associated FluPol is responsible for transcribing vRNA to produce 5′ m^7G-capped, 3′ polyadenylated viral mRNA. The recent availability of high-resolution FluPol structures means that the molecular mechanisms underlying transcription are becoming increasingly well understood (18, 29).

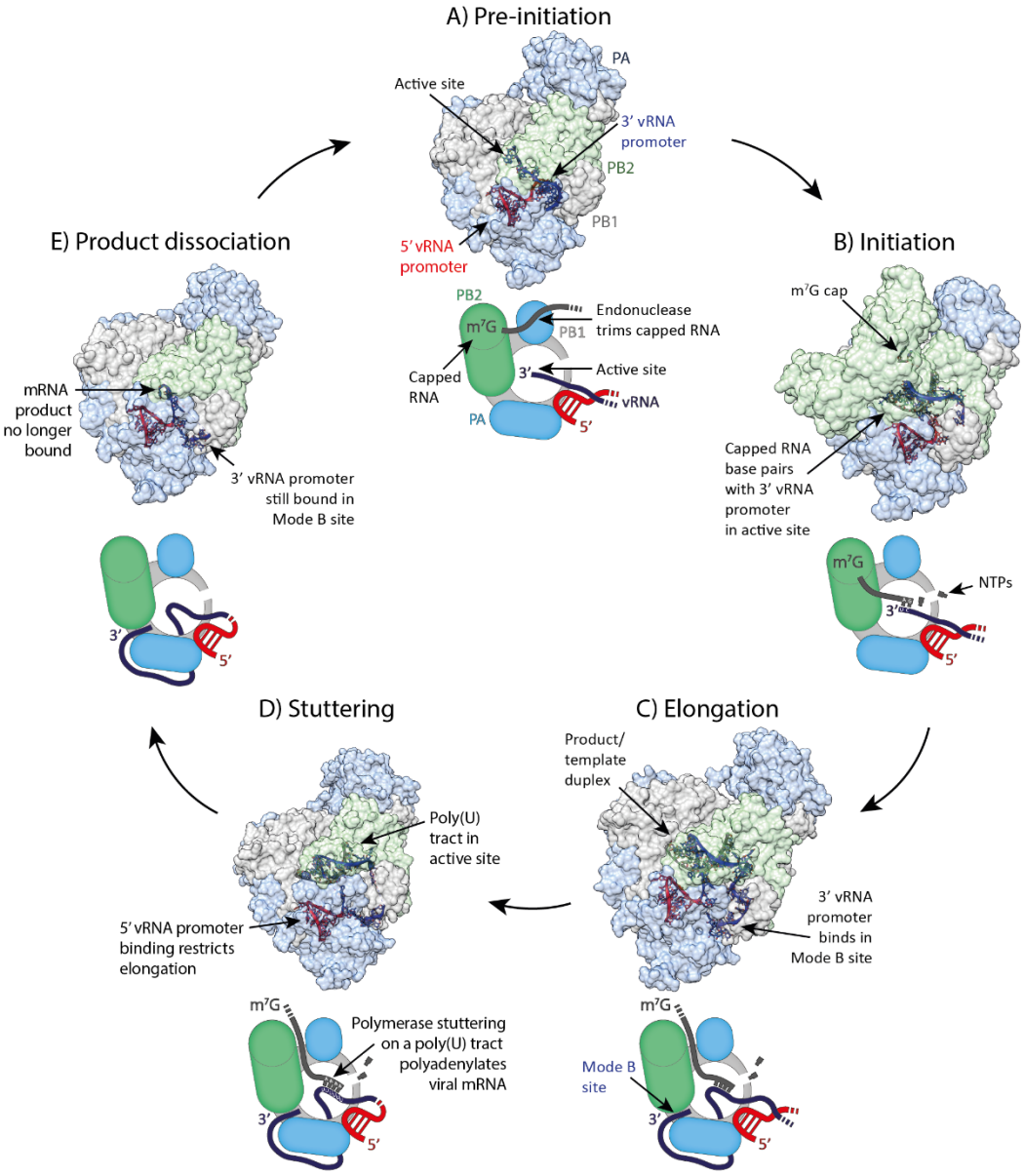
A) Pre-initiation
PA
Active site
3′ vRNA promoter
PB2
PB1
5′ vRNA promoter
B) Initiation
m7G cap
Capped RNA base pairs with 3′ vRNA promoter in active site
m7G
NTPs
3′
5′
C) Elongation
Product/ template duplex
3′ vRNA promoter binds in Mode B site
m7G
Mode B site
D) Stuttering
Poly(U) tract in active site
5′ vRNA promoter binding restricts elongation
m7G
Polymerase stuttering on a poly(U) tract polyadenylates viral mRNA
E) Product dissociation
mRNA product no longer bound
3′ vRNA promoter still bound in Mode B site
PB2
m7G
PB1
Endonuclease trims capped RNA
Capped RNA
3′
Active site
vRNA
PA
5′

Figure 1.5 **Model for viral genome transcription.** A, FluPol binds to nascent m^7G-capped RNA (grey) and cleaves it using the PA endonuclease to generate a primer for viral transcription (bottom). This reaction has not been visualised by current transcription pre-initiation structures (top; PDB: 6T0N). B, m^7G-capped RNA translocates to the FluPol active site, where it makes base pairing interactions with the 3′ vRNA promoter (blue; PDB: 6T0V). C, FluPol extends the m^7G-capped RNA along the vRNA template, generating a product-template duplex in the FluPol active site which is separated by the PB2 lid domain. After the 3′ vRNA promoter emerges from the FluPol active site it binds in the Mode B site on the FluPol surface (PDB: 6SZU). D, FluPol polyadenylates viral mRNA by stuttering on a polyU tract close to the 5′ end of the vRNA template. Stuttering is thought to occur because the strong interaction of the 5′ vRNA promoter (red) with FluPol restricts elongation (PDB: 6T0S). E, The polyadenylated viral mRNA dissociates from FluPol, leaving the single-stranded vRNA template threaded through the active site (PDB: 6T0U). The FluPol PB1 C-ext and PB2 N-terminus are thought to fold outwards to allow the vRNA template to loop out of the active site.

Before initiating transcription, FluPol binds to the serine-5-phosphorylated (S-5-P) CTD of Pol II. S-5-P CTD modification is associated with initiating Pol II, so this selectivity ensures that FluPol has access to nascent host mRNAs (55, 56). Structures of $FluPol_A$, $FluPol_B$ and $FluPol_C$ bound to Pol II CTD-mimetic peptides show that the location of the CTD binding sites on FluPol varies between influenza virus species. However, a common feature is that Pol II CTD binding stabilises FluPol in a transcriptionally active conformation, which allows FluPol to bind to nascent mRNAs as they emerge from Pol II (**Fig. 1.4B**) (36, 57).

The PB2 cap-binding domain binds to the nascent mRNA, and the PA endonuclease cleaves 10-13 nucleotides downstream of the m^7G cap (**Fig. 1.5A**) (58, 59). The capped RNA fragment is then translocated into the RdRp active site by a 70° rotation of the PB2 cap-binding domain. The length of the capped RNA fragment produced is controlled by the distance between the PB2 m^7G binding site and the PA endonuclease active site, and allows the capped RNA fragment to form 1-2 base pairs with the 3′ vRNA promoter in the RdRp active site (**Fig. 1.5B**) (35, 45, 60). Weak base-pairing in the RdRp active site can cause the capped RNA fragment to realign on the 3′

vRNA promoter multiple times during initiation (58, 61). After FluPol initiates transcription by extending the capped RNA fragment along the 3′ vRNA promoter, the nascent viral mRNA leaves the RdRp active site through the product exit channel. A recent cryo-EM study demonstrated that the 3′ vRNA promoter undergoes a sharp turn and extrudes the PB1 priming loop from the RdRp active site, which creates a template exit channel close to the template entry channel (**Fig. 1.5C**). Then, after exiting the active site, the 3′ vRNA promoter binds in the Mode B site on the FluPol surface (44). FluPol synthesises a 3′ poly(A) tail on viral mRNAs by polymerase stuttering, in which it continuously backtracks over a poly(U) stretch located close to the 5′ end of the vRNA template (**Fig. 1.5D**) (62–64). This mechanism was demonstrated by showing that replacement of poly(U) with poly(A) leads to synthesis of viral mRNAs with a poly(U) tail, and recent cryo-EM structures confirm that stuttering occurs because FluPol cannot dislodge the tightly-bound 5′ vRNA promoter (44, 65). After polyadenylation the viral mRNA dissociates from FluPol and the vRNA template exits the RdRp active site, which may require rearrangement of the PB1 C-terminus-PB2 N-terminus bundle (**Fig. 1.5E**) (44).

Following transcription, 5′ m^7G-capped and 3′ polyadenylated viral mRNAs are then exported from the nucleus by the NXF1/p15 pathway. The level of NXF1 dependence differs between transcripts, with HA, M1 and M2 mRNAs being the most reliant on NXF1 for their nuclear export (66, 67). M1 and NS1 viral mRNAs can also be spliced by the cellular spliceosome prior to export, which may alter the mechanism of NXF1 recruitment compared with intronless viral mRNA (68, 69).

1.3.3 *cRNA synthesis*

At 4-6 hours post-infection the rate of viral genome transcription decreases and FluPol begins replicating the viral genome, possibly induced by the accumulation of FluPol and NP or

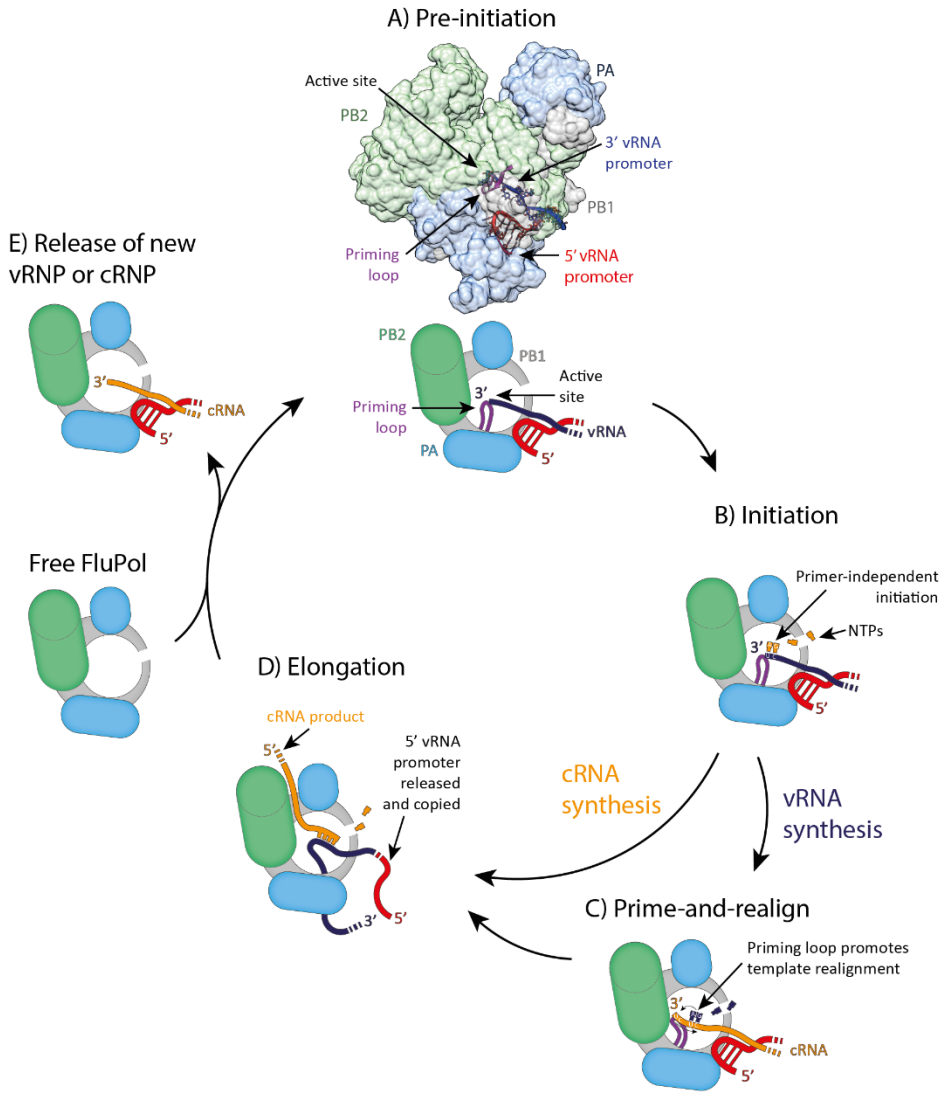

Figure 1.6 **Model for viral genome replication.** A, At the pre-initiation stage the 3′ vRNA promoter (blue) interacts with the FluPol priming loop (purple) in the active site (PDB: 6RR7). B, cRNA synthesis initiates through *de novo* synthesis of a pppApG dinucleotide (gold) from ATP and GTP, which is supported by the FluPol priming loop. C, vRNA synthesis requires an additional prime-and realign step, to move pppApG from positions 4 and 5 of the 3′ cRNA template to positions 1 and 2. D, FluPol elongates the pppApG along the vRNA template to synthesise the nascent cRNA product, and the 5′ vRNA promoter (red) is released from its binding pocket and copied. E, An additional FluPol molecule binds to the cRNA product and assembles it into a cRNP, which then performs vRNA synthesis using a similar mechanism.

production of small viral RNAs (svRNAs) (19, 70). Viral genome replication is a two-step process consisting of cRNA synthesis followed by vRNA synthesis, and both steps are initiated by FluPol *de novo* to produce 5′ triphosphorylated RNAs without a 3′ poly(A) tail. However, cRNA synthesis and vRNA synthesis have some fundamental mechanistic differences (29).

cRNA synthesis is carried out by FluPol in the context of a vRNP, using vRNA as a template (71). FluPol initiates cRNA synthesis by producing a pppApG dinucleotide *de novo* from ATP and GTP, which occurs in the FluPol active site opposite residues 1 and 2 of the 3′ vRNA promoter (**Fig. 1.6A, B**) (72). The PB1 priming loop is important for this process, in particular amino acid residue P651 is thought to be responsible for stabilising initiating ATP in the active site (73). FluPol uses pppApG as a primer to synthesise the full-length cRNA product, however, few mechanistic details are known about how elongation of the RNA product occurs in cRNA synthesis (**Fig. 1.6D**). The termination mechanism for viral genome replication is different from transcription as, instead of stuttering, FluPol releases the 5′ vRNA promoter from its binding pocket and copies the 5′ end of the vRNA template with no polyadenylation. The mechanism for 5′ vRNA promoter release is unclear, but could be related to changes in the base pairing between the 5′ and 3′ vRNA promoters (20). Nascent cRNA is thought to be encapsidated co-replicatively by newly-synthesised FluPol and NP through an unknown mechanism, generating a cRNP which is structurally similar to the parent vRNP (**Fig. 1.6E**) (27).

1.3.4 *vRNA synthesis*

In the second step of viral genome replication, FluPol in a cRNP replicates cRNA to generate progeny vRNA. vRNA synthesis has a more complex initiation process than cRNA synthesis, which may be a regulatory mechanism to control the rate of viral genome replication or prevent aberrant transcription by cRNPs (72). Since FluPol and NP in vRNPs and cRNPs is the same,

differences between the 5′ and 3′ vRNA and cRNA promoters are thought to be responsible for the different initiation mechanisms (74, 75).

To initiate vRNA synthesis, FluPol synthesises a pppApG dinucleotide *de novo* from ATP and GTP. The 3′ cRNA promoter is inserted more deeply into the RdRp active site than the 3′ vRNA promoter, so pppApG synthesis occurs opposite residues 4 and 5 of the 3′ cRNA promoter instead of residues 1 and 2 (72). This process is called internal initiation, and is not dependent on the PB1 priming loop because interactions with the 3′ cRNA promoter alone are sufficient to stabilise the initiating ATP (73). Once pppApG is synthesised it is translocated to positions 1 and 2 of the 3′ cRNA promoter, which is necessary to synthesise a full-length vRNA product (**Fig. 1.6C**). This ′realignment′ mechanism is dependent on the PB1 priming loop, which is thought to push the 3′ cRNA promoter out of the RdRp active site while protein-RNA interactions hold the pppApG in the same place (72, 76). Realignment results in pppApG being annealed to positions 1 and 2 of the 3′ cRNA promoter, and it is then extended to form a full-length vRNA product.

Termination of vRNA synthesis occurs with no polymerase stuttering, producing an exact copy of the original vRNA template. As with cRNA synthesis, the nascent vRNA product is co-replicatively encapsidated by newly-synthesised FluPol and NP through an unknown mechanism to generate a progeny vRNP (20).

1.3.5 *RNA polymerase oligomerisation*

In order to encapsidate nascent vRNA or cRNA into a vRNP or cRNP, newly-synthesised FluPol must be recruited to bind to the nascent product. For this reason, viral genome replication is hypothesised to involve interactions between newly-synthesised FluPol molecules and FluPol molecules in RNPs (20).

FluPol dimerization has been observed in cells and *in vitro*, however, the role of these interactions in viral genome replication is unknown (77, 78). One model suggests that newly-synthesised

FluPol molecules activate viral genome replication in *trans*, and do not perform RNA synthesis directly (*trans*-activation). This is supported by the observation that cRNPs only produce vRNA *in vitro* when provided with exogenous FluPol, independently of whether the exogenous FluPol is catalytically active (27). A second model suggests that newly-synthesised FluPol molecules perform RNA synthesis themselves using vRNA or cRNA from the RNP, acting in *trans* (*trans*-acting); however, there is less convincing evidence to support this model (79).

1.4 **ROLE OF THE INFLUENZA VIRUS RNA POLYMERASE IN HOST ADAPTION**

1.4.1 *Avian-human host adaption*

Influenza A viruses can infect aquatic birds and swine in addition to humans. Avian-origin influenza A viruses are not naturally well-adapted to humans, so infections are rare and often do not result in human-human transmission. When zoonotic influenza A viruses infect the human population and successfully adapt for human to human transmission, this can lead to pandemics as there is no pre-existing immunity to the novel strain. From the over 800 cases of human infection, H5N1 highly pathogenic avian influenza (HPAI) has a mortality rate of over 50% (80). The danger posed by zoonotic influenza A viruses means a more detailed understanding of the mechanisms underlying host adaptation is essential.

Transmitting efficiently in a new species requires the virus to undergo several adaptations which affect different points of its life cycle. For an avian-origin influenza A virus to efficiently infect humans it must acquire mutations in HA which alter its receptor binding preference from α-2,3-linked sialic acids (present in the avian gut) to α-2,6-linked sialic acids (present in the human upper respiratory tract). Furthermore, avian-origin $FluPol_A$ does not work efficiently in human cells, a limitation which must be overcome by acquiring host-adaptive mutations (81). A key host-adaptive amino acid residue is position PB2 627, where avian-origin viruses encode glutamic acid and human-origin viruses often encode lysine. Human-origin $PB2_{K627}$ forms part of a patch of

basic amino acid residues on the PB2 627 domain, which is interrupted in the avian-origin $PB2_{E627}$ (31, 32). Avian-origin influenza A viruses often undergo a $PB2_{E627K}$ substitution when adapting to human hosts, which enables avian-origin $FluPol_A$ to function efficiently in human cells. $PB2_{K627}$ is the most common residue found in human influenza A virus isolates, however, other non-acidic substitutions such as $PB2_{E627V}$ can occur at this position (82, 83). Alternative adaptive mutations in avian-origin $FluPol_A$ can compensate for $PB2_{E627}$, such as $PB2_{D701N}$ (84). Many of these other adaptive mutations still cluster around the PB2 627 domain, such as $PB2_{G590S}$ and $PB2_{Q591R}$, which occurred together in the 2009 pandemic H1N1 virus (85, 86).

1.4.2 *Role of the cellular protein ANP32 in host adaption*

Understanding the restriction of avian-origin $FluPol_A$ in human cells is important for preventing influenza A virus pandemic emergence. Previous studies have shown that $PB2_{E627}$ $FluPol_A$ is defective in viral genome replication but not transcription in human cells (87). This was found to be because avian-origin $FluPol_A$ is not compatible with human ANP32 (Acidic Nuclear Phosphoprotein 32kDa), a host protein which is essential for viral genome replication (88).

There are three ANP32 proteins in humans: ANP32A, ANP32B and ANP32E. ANP32 proteins interact with histones and have many endogenous functions, including transcriptional control by chromatin remodelling. They have an N-terminal Leucine-Rich Repeat (LRR) domain, the structure of which has been solved, and a disordered C-terminal Low-Complexity Acidic Region (LCAR) (89). ANP32A and ANP32B can both support FluPol activity and are functionally redundant, but ANP32E cannot support activity. FluPol from influenza A and B viruses requires ANP32, and many studies have aimed to identify a molecular mechanism for this effect (90–92). ANP32 is known to interact directly with the FluPol PB2 627 domain, and this interaction occurs in the nucleus (93). There are conflicting reports on whether ANP32 prefers to interact with apo FluPol or with RNPs, but recent data favour the latter, with inactive RNPs showing the strongest

interaction (94, 95). Multiple studies show that ANP32 is required by FluPol for viral genome replication but not transcription, and some suggest that it is specifically required for vRNA synthesis (92, 94).

Compared to chicken ANP32A (chANP32A), human ANP32A (huANP32A) lacks a 33 amino acid residue insertion in the LCAR resulting from an exon duplication event. This insertion is required for ANP32A to support avian-origin $FluPol_A$ activity, but not human-origin $FluPol_A$ (88). chANP32A interacts with $FluPol_A$ more strongly than huANP32A, however, multiple studies have found that interaction strength does not depend on whether $FluPol_A$ has $PB2_{E627}$ or $PB2_{K627}$ (95, 96). A SUMO-Interaction Motif (SIM) sequence was identified in the chANP32A LCAR insertion, and was initially suggested to be involved in supporting $PB2_{E627}$ $FluPol_A$ activity, but subsequent reports contradicted this by showing that a splice variant of chANP32A lacking the SIM motif was also able to support activity (93, 96). Unlike huANP32B, chANP32B cannot support $PB2_{E627}$ or $PB2_{K627}$ $FluPol_A$ activity because two amino acid residues in the LRR domain, I129 and N130, disrupt the interaction with $FluPol_A$ (90, 97). Swine can be infected by both human- and avian-origin influenza A viruses, so are thought to be ′mixing vessels′ for reassortment between influenza A viruses from different species. Interestingly, swine ANP32A (swANP32A) can support avian-origin $FluPol_A$ activity due to its unique V106 and S156 amino acid residues which, in combination with their expressing both α-2,3-linked and α-2,6-linked sialic acids, provides a molecular explanation for why swine can be efficiently infected by human- and avian-origin influenza A viruses (98, 99).

These studies highlight that ANP32 proteins are important not just for supporting FluPol activity, but also for determining the host range of influenza A viruses. Despite this, a molecular mechanism for ANP32 supporting FluPol activity remains elusive (81).

1.5 **OBJECTIVES**

Influenza A virus is an important human pathogen; however, its mechanism of viral genome replication remains poorly understood. The general aim of my DPhil work was to better understand how $FluPol_A$ replicates the viral genome and the involvement of key host proteins.

I began by determining how the interactions between $FluPol_A$ and its RNA promoters are important for RNA synthesis activity, taking advantage of new high-resolution structural information to guide functional studies (Chapter 2). Next, I investigated the role of $FluPol_A$ dimerization in replication of the viral genome, addressing this important question with cell-based and biochemical assays to gain detailed mechanistic insight (Chapter 3). Finally, I studied the function of the host protein ANP32A in supporting $FluPol_A$ activity in a cellular context. These studies also led to me investigating how species-specific differences in ANP32A drive particular $FluPol_A$ host adaptive mutations (Chapter 4).

CHAPTER 2

The functions of influenza virus RNA polymerase 3′ RNA promoter binding sites

Data from this chapter were published in:

***Walker AP**, Sharps J, Fodor E. 2020. Mutation of an Influenza Virus Polymerase 3′ RNA Promoter Binding Site Inhibits Transcription Elongation. J Virol 94:e00498-20.*

2.1 INTRODUCTION

Influenza virus encodes an RNA-dependent RNA polymerase (FluPol), which forms a ribonucleoprotein (RNP) complex with viral RNA and nucleoprotein (NP). FluPol in an RNP is bound to promoter sequences at the 5′ and 3′ ends of the viral RNA segment, effectively circularising viral RNA which is coated with NP (29). RNPs containing negative-sense viral RNA (vRNA) are referred to as vRNPs, while those containing positive-sense complementary RNA (cRNA) are referred to as cRNPs (27). FluPol always performs RNA synthesis in the context of RNPs, though it is unclear how this occurs without disrupting the RNP structure.

FluPol binds to the 5′ and 3′ RNA promoters using different binding pockets. 5′ vRNA and cRNA promoters form similar hook-like structures, which bind tightly in a pocket formed by the FluPol PB1 and PA subunits (**Fig. 2.1A**) (34, 38). The 3′ vRNA and cRNA promoters are more dynamic, and can bind in the FluPol active site as well as in two distinct sites on the external surface. By binding in the FluPol active site the 3′ RNA promoters can act as templates for viral RNA synthesis, however, the function of 3′ RNA promoter binding in the external sites is not as clear (37, 44). The first of these external sites is called the Mode A site, and is formed entirely by the PB1 subunit. Only the 3′ vRNA promoter has been observed in this site structurally, and initial

observations suggested that this is a ′pre-initiation′ site for vRNA transcription (35). The second external site is called the Mode B site, which binds to both 3′ RNA promoters and is formed primarily by the PB1 and PA subunits (37, 44, 53). Several related viral RNA polymerases have 3′ RNA promoter binding sites similar to the Mode B site, suggesting that it has a conserved role in viral RNA synthesis; however, when these studies were performed this function remained unknown (52, 54). Here I aimed to understand the function of 3′ RNA promoter binding in the Mode A and Mode B sites using structure-guided functional analyses.

2.2 **RESULTS**

2.2.1 *The Mode A site is involved in viral genome replication*

First, I aimed to gain insight into the function of 3′ RNA promoter binding in the $FluPol_A$ Mode A site. To achieve this, I mutated basic amino acids in the site to alanine, avoiding amino acid residues located close to the $FluPol_A$ active site or the PB1 nuclear localisation signal (NLS) on the PB1 β-ribbon (**Fig. 2.1B**) (35). I created $PB1_{R135A}$, $PB1_{R350A}$ and $PB1_{R220A}$ mutations in $FluPol_A$, and performed minigenome assays by transfecting plasmids encoding mutant $FluPol_A$, nucleoprotein (NP), and a vRNA segment into HEK 293T cells. 24 hours post-transfection, I quantified viral RNA levels by primer extension (**Fig. 2.1C**). In minigenome assays the $PB1_{R135A}$ mutation strongly inhibited $FluPol_A$ activity as measured by mRNA, cRNA and vRNA accumulation. The $PB1_{R350A}$ mutation had no significant effect on $FluPol_A$ activity, while $PB1_{R220A}$ reduced the levels of all viral RNA species to approximately 50%. To further disrupt 3′ RNA promoter binding in the Mode A site I created glutamic acid substitutions at $PB1_{R220}$ and $PB1_{R350}$ residues, which repel the negatively charged 3′ RNA promoter. The $PB1_{R220E}$ mutation reduced $FluPol_A$ activity slightly more than the $PB1_{R220A}$ mutant, and the $PB1_{R350E}$ mutant abolished $FluPol_A$ activity.

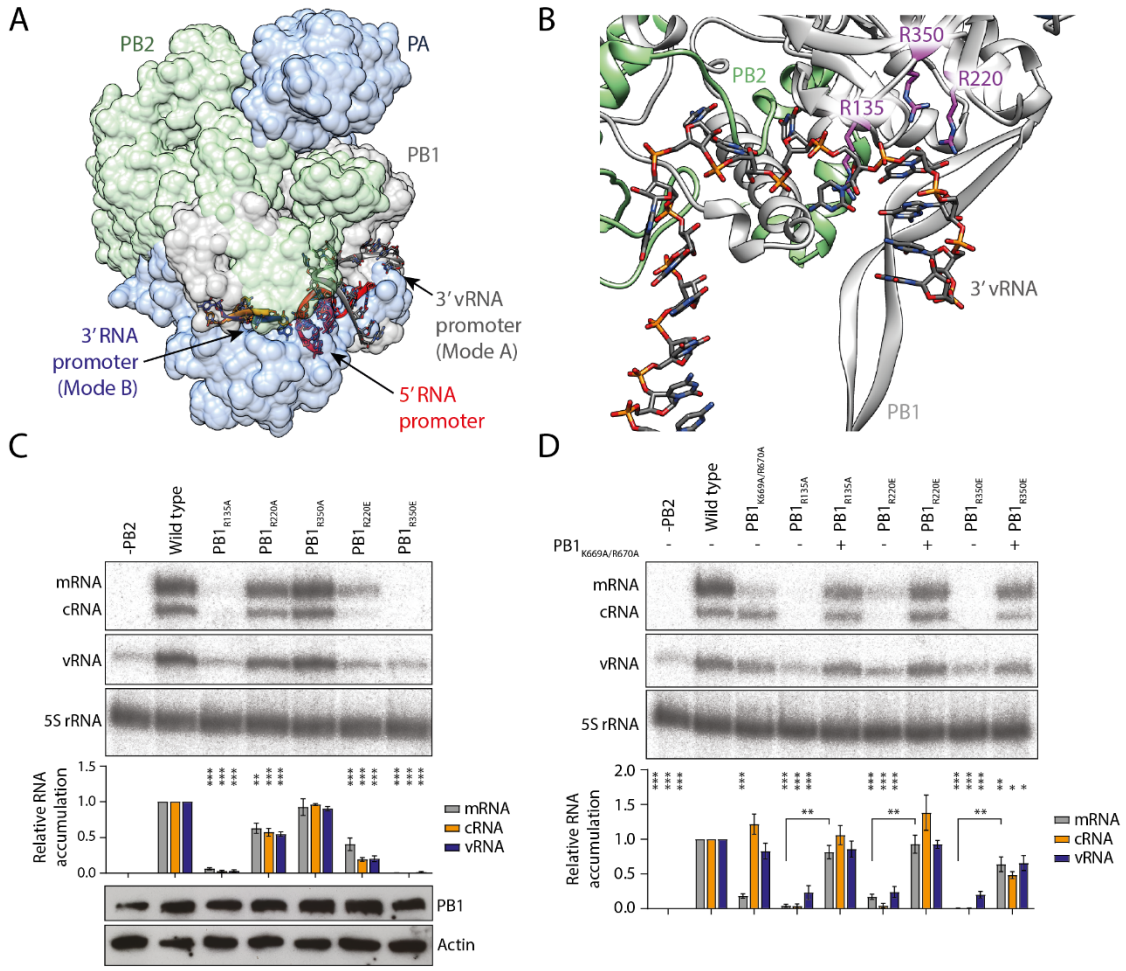

Figure 2.1 **Effect of Mode A site mutations in minigenome assays.** A, Overall structure of FluPol$_A$ with 5′ vRNA promoter (red) and 3′ vRNA promoter (grey) bound in the Mode A site (PDB: 4WRT). 3′ vRNA promoter (blue; PDB: 6ABF) and 3′ cRNA promoter (gold; PDB: 6QX3) are modelled in the Mode B site. B, Close-up view of 3′ vRNA promoter (grey) in the Mode A site (PDB: 4WRT), with several key basic amino acid residues shown (purple). C, FluPol$_A$ minigenome assay with Mode A site mutations. HEK 293T cells were transfected with plasmids encoding wild type or mutant FluPol$_A$ subunits, NP and an NA segment vRNA template. Viral RNA levels were analysed 24 hours post-transfection by primer extension (top). PB1 protein expression was analysed by western blot, with actin as a loading control (bottom). D, Complementation with transcription-deficient PB1$_{K669A/R670A}$ FluPol$_A$ was used to determine the transcription activity of Mode A site FluPol$_A$ mutants. Quantification is from n=3 independent transfections, data are mean ± s.e.m., analysed by one-way ANOVA. *P<0.05, **P<0.01, ***P<0.001. Individual mRNA signals were compared by two-tailed unpaired t-test; **P<0.01.

Next, I wanted to determine if these mutations inhibit the ability of $FluPol_A$ to carry out viral genome replication, transcription, or both. Normally, $FluPol_A$ synthesises mRNA using vRNA as a template, so if vRNA accumulation is inhibited by a mutation then mRNA synthesis is always decreased, independently of whether the mutation directly affects the transcription activity of $FluPol_A$. In other words, mRNA synthesis is always dependent on vRNA synthesis in this assay. To overcome this limitation I used a previously described complementation assay with a transcription-deficient $FluPol_A$ mutant ($PB1_{K669A/R670A}$) (87). I expressed $PB1_{K669A/R670A}$ mutant $FluPol_A$, which is able to support vRNA accumulation, then co-expressed each $FluPol_A$ mutant (**Fig. 2.1D**). $PB1_{K669A/R670A}$ $FluPol_A$ is unable to efficiently synthesise mRNA, so mRNA signal observed when other $FluPol_A$ mutants are co-expressed is the result of mRNA synthesis by those mutants using the vRNA template generated by $PB1_{K669A/R670A}$ $FluPol_A$ (100). When I co-expressed each of the Mode A site $FluPol_A$ mutants I observed a robust mRNA signal, indicating that the $PB1_{R135A}$, $PB1_{R350E}$ and $PB1_{R220E}$ $FluPol_A$ mutants are all able to efficiently produce viral mRNA by transcription. These data show that the Mode A site mutations disrupt viral genome replication by $FluPol_A$ but not transcription.

To confirm that the Mode A site mutants interfere with 3′ RNA promoter binding, I overexpressed $FluPol_A$ with $PB1_{R135A}$, $PB1_{R220E}$ or $PB1_{R350E}$ mutations in HEK 293T cells and purified the recombinant $FluPol_A$ using a protein A tag on the PB2 C-terminus (**Fig. 2.2A**). I then incubated the purified $FluPol_A$ with radiolabelled 3′ vRNA or cRNA promoter in the presence of the corresponding 5′ RNA promoter, cross-linked the RNAs to $FluPol_A$, and determined the relative quantities of radiolabelled 3′ RNA promoter bound by SDS PAGE and phosphorimaging (**Fig. 2.2B**). I observed a small decrease in 3′ vRNA promoter binding for the $PB1_{R135A}$ and $PB1_{R220E}$ mutations, while $PB1_{R350E}$ had no significant effect. Binding to the 3′ cRNA promoter appeared to be reduced by a similar level to the 3′ vRNA promoter, although this was only statistically significant for the $PB1_{R220E}$ mutant. These results suggest that the Mode A site mutations affect 3′ RNA promoter binding, however, most of the 3′ vRNA or cRNA promoter remains bound.

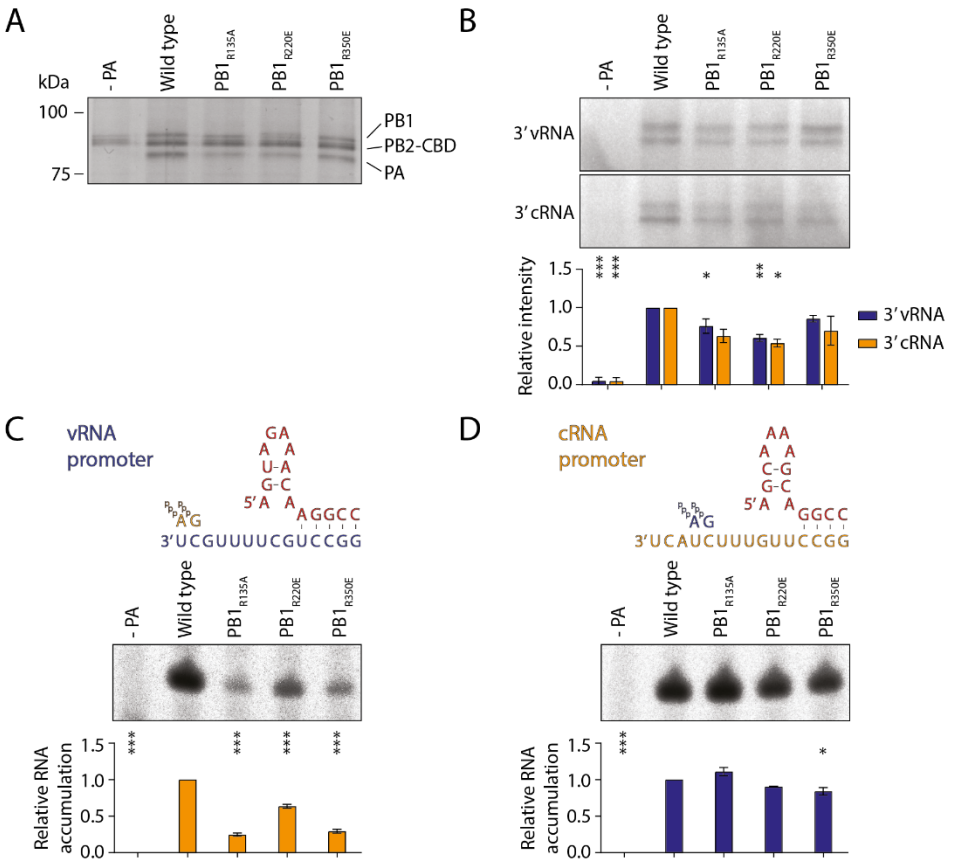

Figure 2.2 ***In vitro* assays with Mode A site FluPol$_A$ mutants.** A, Protein A-tagged Mode A site FluPol$_A$ mutants were expressed in HEK 293T cells and purified, then analysed by SDS PAGE. After protein A cleavage a Calmodulin Binding Domain (CBD) remains on the PB2 subunit. B, Binding of radiolabelled 3′ vRNA and cRNA promoters to purified FluPol$_A$ was assayed by cross-linking and SDS PAGE, followed by phosphorimaging. 3′ vRNA and cRNA promoters cross-link to multiple FluPol$_A$ subunits, resulting in multiple bands on the SDS gel. C, *De novo* cRNA synthesis initiation (pppApG synthesis) assays were performed on a model vRNA template using purified FluPol$_A$, a schematic of the initiation mechanism on this template is shown (top). Dinucleotide products were resolved by alkaline phosphatase treatment then denaturing PAGE (bottom). D, *De novo* vRNA synthesis initiation assays were performed on a model cRNA template using purified FluPol$_A$, a schematic of the initiation mechanism on this template is shown (top). Dinucleotide products were resolved by alkaline phosphatase treatment then denaturing PAGE (bottom). Quantification is from n=3 independent protein preparations, data are mean ± s.e.m., analysed by one-way ANOVA. *P<0.05, **P<0.01, ***P<0.001.

This is most likely because the 3′ RNA promoter can still bind to other sites on $FluPol_A$, such as the Mode B site, despite the Mode A site mutations. Collectively, these data show that $FluPol_A$ requires the Mode A site for viral genome replication, but mutations in this site do not affect viral genome transcription.

To examine the function of the Mode A site in more detail, I tested the ability of the Mode A site $FluPol_A$ mutants to initiate viral genome replication. $FluPol_A$ initiates viral genome replication in a primer-independent manner by *de novo* synthesis of a pppApG dinucleotide from ATP and GTP. During cRNA synthesis this occurs opposite positions 1 and 2 of the 3′ vRNA promoter, while for vRNA synthesis the pppApG is synthesised opposite positions 4 and 5 of the 3′ cRNA promoter (72). By supplying purified $FluPol_A$ with model 5′ and 3′ vRNA or cRNA promoters, ATP and radiolabelled GTP, I examined the efficiency of dinucleotide synthesis by the Mode A site $FluPol_A$ mutants *in vitro* (**Fig. 2.2C, D**). All Mode A site mutations strongly reduced cRNA dinucleotide synthesis by $FluPol_A$, with $PB1_{R135A}$ and $PB1_{R350E}$ showing the greatest reduction. Interestingly, all of the Mode A site mutants had little to no effect on vRNA dinucleotide synthesis. These data suggest that while the Mode A site is important for cRNA synthesis initiation by $FluPol_A$, it is not involved in initiating vRNA synthesis.

2.2.2 *The Mode B site is involved in viral genome transcription and replication*

Next, I aimed to determine the function of 3′ RNA promoter binding in the Mode B site. I targeted key amino acid residues highlighted in a cryo-EM structure of $FluPol_A$ bound to the 3′ cRNA promoter, mutating these to alanine (**Fig. 2.3A, B**) (37). I focused on the PB1 and PA subunits, as these interact most extensively with the 3′ RNA promoter in the Mode B site, and confirmed expression of these $FluPol_A$ mutants before testing their activity in minigenome assays. None of the single amino acid substitutions significantly reduced $FluPol_A$ activity as measured by mRNA, vRNA and cRNA accumulation; however since the Mode B site is an extended groove and there

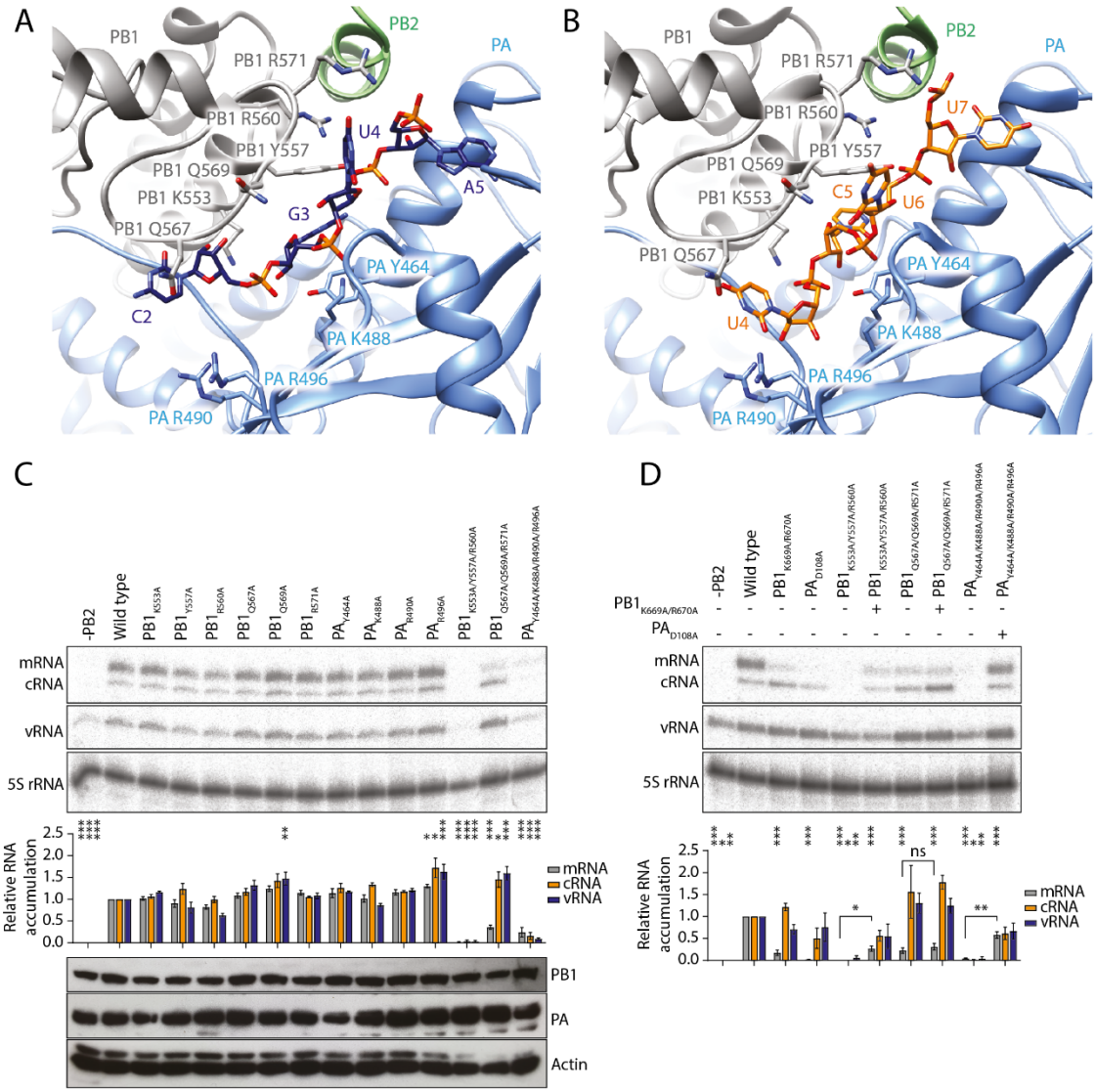

Figure 2.3 **Effect of Mode B site mutations in minigenome assays.** A, Close-up view of 3′ vRNA promoter (blue) in the Mode B site (PDB: 6ABF), with key amino acids shown. B, Close-up view of 3′ cRNA promoter (gold) in the Mode B site (PDB: 6QX3), with key amino acids shown. C, FluPol$_A$ minigenome assays were performed using plasmids encoding wild type or Mode B site mutant FluPol$_A$ subunits. Viral RNA levels were assayed by primer extension 24 hours post-transfection (top), and PB1 or PA protein expression was analysed by western blot with actin as a loading control (bottom). D, Complementation with transcription-deficient PB1$_{K669A/R670A}$ or PA$_{D108A}$ FluPol$_A$ was used to determine the transcription activity of Mode B site FluPol$_A$ mutants. Quantification is from n=3 independent transfections, data are mean ± s.e.m., analysed by one-way ANOVA. *P<0.05, **P<0.01, ***P<0.001. Individual mRNA signals were compared by two-tailed unpaired t-test; *P<0.05, **P<0.01.

are many protein-RNA interactions, I reasoned that single amino acid substitutions may not sufficiently disrupt 3′ RNA promoter binding (**Fig. 2.3C**). Therefore, I made $FluPol_A$ constructs with multiple Mode B site amino acid substitutions and tested their activity in minigenome assays. $PB1_{K553A/Y557A/R560A}$ and $PA_{Y464A/K488A/R490A/R496A}$ mutations abolished mRNA, vRNA and cRNA synthesis, while the $PB1_{Q567A/Q569A/R571A}$ mutation significantly reduced mRNA synthesis but caused an increase in vRNA and cRNA accumulation.

Next, I tested the transcription activity of $FluPol_A$ with Mode B site mutations using the same complementation assay as I used for the Mode A site mutations. I used $PB1_{K669A/R670A}$ $FluPol_A$ as the transcription-deficient mutant for complementing PB1 mutations, or PA_{D108A} $FluPol_A$ as the transcription-deficient mutant for complementing PA mutations (**Fig. 2.3D**) (100, 101). When I co-expressed $PB1_{K553A/Y557A/R560A}$ or $PB1_{Q567A/Q569A/R571A}$ $FluPol_A$ with $PB1_{K669A/R670A}$ $FluPol_A$ I did not observe a robust mRNA signal, indicating that these mutations inhibit $FluPol_A$ transcription activity. However, when I co-expressed $PA_{Y464A/K488A/R490A/R496A}$ $FluPol_A$ with PA_{D108A} $FluPol_A$ I observed a strong mRNA signal, indicating that the $PA_{Y464A/K488A/R490A/R496A}$ mutation interferes with viral genome replication by $FluPol_A$ but not transcription.

To confirm that these mutations disrupt 3′ RNA promoter binding, I purified the overexpressed $FluPol_A$ mutants from HEK 293T cells, then tested then ability of the $FluPol_A$ mutants to bind to radiolabelled 3′ vRNA and cRNA promoter by cross-linking and SDS PAGE (**Fig. 2.4A, B**). All of the Mode B site mutations significantly reduced 3′ vRNA and cRNA promoter binding, with $PA_{Y464A/K488A/R490A/R496A}$ having the strongest effect. Collectively, these data show that 3′ RNA promoter binding in the Mode B site is involved in both viral genome replication and transcription.

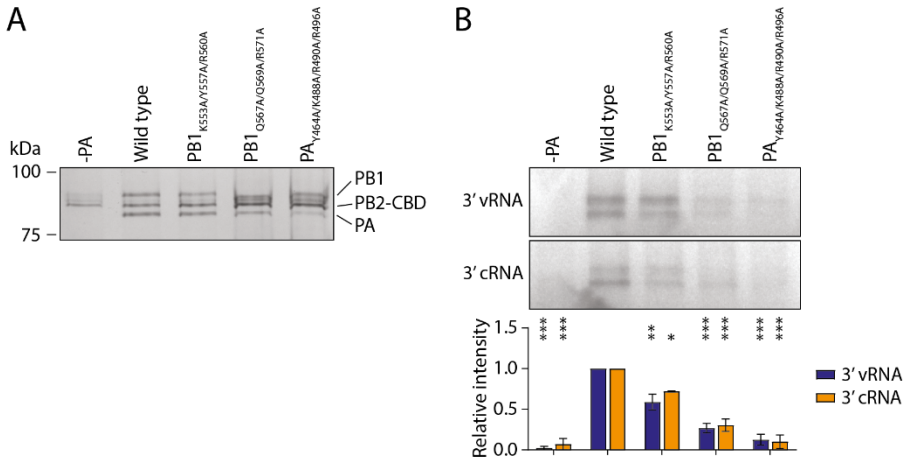

Figure 2.4 **Purification and RNA binding affinity of Mode B site FluPol$_A$ mutants.** A, Protein A-tagged Mode B site FluPol$_A$ mutants were purified from HEK 293T cells and analysed by SDS PAGE. B, Radiolabelled 3′ vRNA and cRNA promoter binding to purified FluPol$_A$ was assayed by cross-linking and SDS PAGE, followed by phosphorimaging. Quantification is from n=3 independent protein preparations, data are mean ± s.e.m., analysed by one-way ANOVA. *P<0.05, **P<0.01, ***P<0.001.

2.2.3 *The Mode B site is required for elongation during transcription*

To gain further insight into the role of the 3′ RNA promoter binding in the Mode B site, I took advantage of the transcription-specific phenotype of the PB1$_{Q567A/Q569A/R571A}$ FluPol$_A$ mutant. This mutation interferes with mRNA synthesis but not cRNA or vRNA synthesis, meaning that transcriptionally-inactive PB1$_{Q567A/Q569A/R571A}$ mutant vRNPs can still accumulate in cells and could be purified. Therefore, I overexpressed PB1$_{Q567A/Q569A/R571A}$ mutant FluPol$_A$, NP and a vRNA template in HEK 293T cells and purified the mutant vRNPs using a protein A tag on the PB2 subunit (**Fig. 2.5A**). I then performed *in vitro* transcription assays by incubating the purified vRNPs with a capped RNA primer in the presence or absence of NTPs. When I analysed the reaction products by primer extension, using a radiolabelled primer targeting the 5′ end of mRNA products (NA 160*), I observed a similar mRNA signal for wild type and PB1$_{Q567A/Q569A/R571A}$ mutant vRNPs (**Fig. 2.5B, C**). I also performed reverse-transcription using the same primer

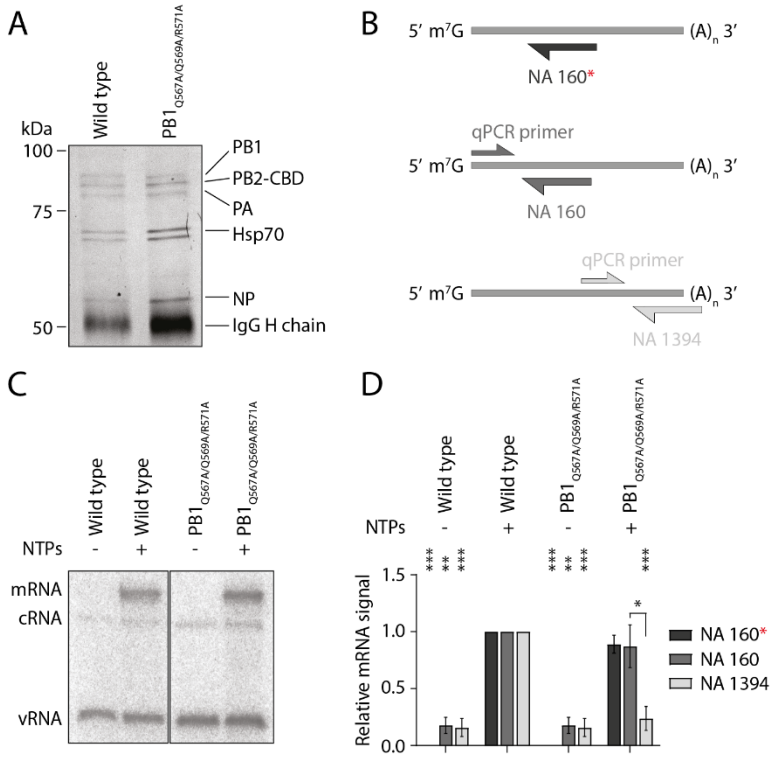

Figure 2.5 **Purification and transcription activity of Mode B site mutant vRNPs.** A, Protein A-tagged wild type or Mode B site mutant $FluPol_A$ was expressed in HEK 293T cells with NP and a vRNA template, then purified and analysed by SDS PAGE. B, Schematic of the viral mRNA product denoting annealing sites of the NA 160 (top and middle) and NA 1394 primers (bottom). A red asterisk indicates the ^{32}P radiolabel used in the primer extension assay (top), and positions of the reverse qPCR primers are also shown (middle and bottom). C, Capped RNA-primed *in vitro* transcription assays were carried out using wild type and Mode B site mutant vRNPs, and mRNA products were visualised by primer extension using the NA 160* primer. D, *In vitro* transcription products were analysed by reverse transcription and qPCR using either the NA 160 or NA 1394 primers. Quantification is from n=3-4 independent reactions using 2 different vRNP preparations; data are mean ± s.e.m., analysed by one-way ANOVA. **P<0.01, ***P<0.001. Individual mRNA signals were compared by two-tailed unpaired t-test; *P<0.05.

without a radiolabel (NA 160), followed by qPCR, and obtained similar results (**Fig. 2.5B, D**). However, when I performed reverse transcription using a primer directed against the 3′ end of mRNA products (NA 1394) followed by qPCR, I observed a significant decrease in mRNA signal from the $PB1_{Q567A/Q569A/R571A}$ mutant vRNPs compared to wild type vRNPs (**Fig. 2.5B, D**). This

indicates that while the $PB1_{Q567A/Q569A/R571A}$ mutant vRNPs can initiate transcription as well as wild type, they are less able to synthesise full length mRNA products. Collectively, these data suggest that the Mode B site is required for elongation during viral genome transcription.

2.3 **DISCUSSION**

With this study, I aimed to determine the function of 3′ RNA promoter binding in the Mode A and Mode B sites on the external surface of $FluPol_A$. I found that mutating the Mode A site interferes with viral genome replication, specifically preventing the *de novo* initiation of cRNA synthesis. On the other hand, I found that the Mode B site is important for both viral genome transcription and replication. Furthermore, using *in vitro* biochemical analyses I found that the Mode B site is dispensable for initiating viral genome transcription but is essential for the production of full-length viral mRNA.

My minigenome assay results suggest that the Mode A site is not important for transcription of the viral genome, contradicting previous hypotheses that 3′ vRNA promoter binds in the Mode A site to initiate viral genome transcription (**Fig. 2.1**) (35). My follow-up analyses found that all Mode A site mutations tested caused a defect in *de novo* initiation; specifically, the initiation of cRNA synthesis by vRNA-bound $FluPol_A$ (**Fig. 2.2C, D**). This finding is consistent with a previous biophysical study which found that the 3′ vRNA promoter can access the FluPol active site from the Mode A site, but the 3′ cRNA promoter cannot (49). Despite the striking difference in utilisation of the Mode A site by 3′ vRNA and 3′ cRNA promoters, I observed that Mode A site mutations do not differentially affect 3′ vRNA and cRNA promoter binding (**Fig. 2.2B**). These results suggest that the 3′ cRNA promoter can bind in the Mode A site under the *in vitro* conditions used, but this binding may not be functionally important.

My data also show that the Mode B site is not required for initiating viral mRNA synthesis, but is required to synthesise full-length viral mRNA transcripts. These data support a model in which

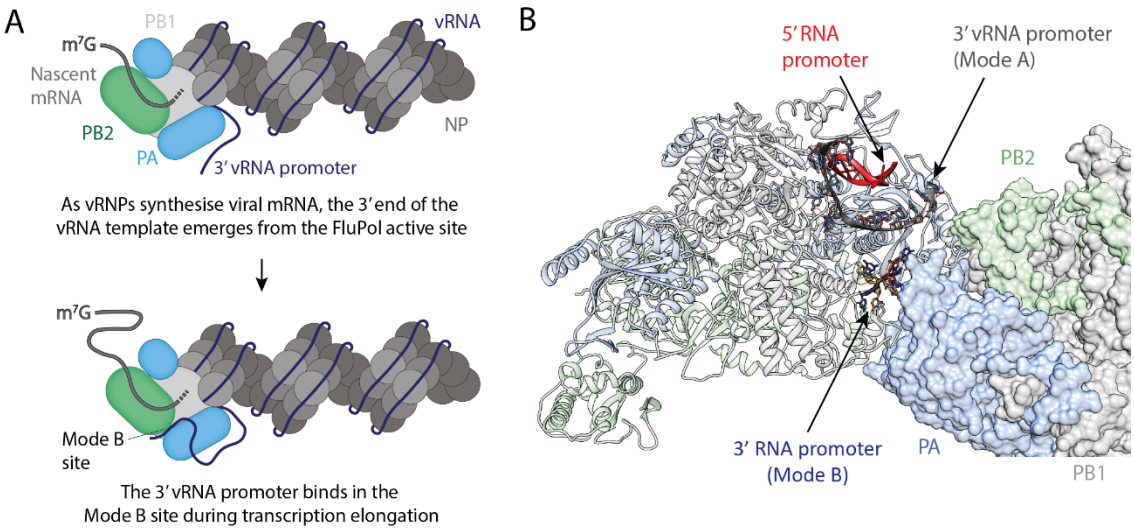

Figure 2.6 **Model for the function of the Mode B site.** A, Model for the function of the Mode B site during viral genome transcription. FluPol extends the capped mRNA primer using vRNA in the vRNP as a template, and the 3′ end of the vRNA template leaves the FluPol active site after it has been copied (top). Once it has exited the FluPol active site the 3′ vRNA promoter binds in the Mode B site, where it remains bound throughout transcription elongation (bottom). B, Structure of the $FluPol_A$ *trans*-activating dimer with one FluPol heterotrimer shown as ribbon cartoon and the other as a surface (PDB: 6QNW). 3′ vRNA promoter (blue; PDB: 6ABF) and 3′ cRNA promoter (gold; PDB: 6QX3) are modelled in the Mode B site. 5′ vRNA promoter and 3′ vRNA promoter in the Mode A site are also shown (red and grey; PDB: 4WRT).

the 3′ end of the vRNA template binds in the Mode B site after it has passed through the FluPol active site during transcription elongation, as suggested previously (**Fig. 2.6A**) (37). This binding would maintain the structure of the vRNP throughout transcription, which is necessary for two reasons: firstly, vRNA associated with a vRNP cannot be detected by innate immune sensors such as RIG-I, and major disruptions to vRNP integrity would risk releasing immunogenic vRNA into the cellular environment (102). Secondly, the interaction between FluPol and NP is important for processivity during transcription, therefore a vRNP must maintain its structure to be capable of multiple sequential rounds of transcription (28). This was clearly demonstrated by a recent cryo-EM study which found that transcribing vRNPs maintain the ordered double-helical arrangement

of NP (24). My model for the function of the Mode B site is also strongly supported by a recent independent study, which showed that when $FluPol_A$ transcribes an artificial vRNA template the 3′ vRNA promoter binds in the Mode B site after passing through the active site. This study also found that the Mode B site remains occupied until transcription is completed, and proposed a recycling step in which the 3′ vRNA promoter moves from the Mode B site back into the FluPol active site (44).

It is likely that the Mode B site has a similar function in viral genome transcription and replication, and indeed my minigenome assay data show that some Mode B site amino acid residues are important for viral genome replication. Interestingly, I observed different effects by mutating different regions of the Mode B site; for example, the $PB1_{Q567A/Q569A/R571A}$ mutation only affected transcription while the $PA_{Y464A/K488A/R490A/R496A}$ mutation only affected viral genome replication (**Fig. 2.3C, D**). These data suggest that the transcription and replication processes have differences in how they utilise the Mode B site, or require different affinities for the 3′ RNA promoter, which results in a dependence on different amino acid residues. Intriguingly, the Mode B site is located at a $FluPol_A$ *trans*-activating dimer interface which is important for vRNA synthesis (see Chapter 3), raising the possibility that interplay between 3′ RNA promoter binding and FluPol dimerization occurs during viral genome replication (**Fig. 2.6B**) (37).

These results provide the first insights into the functional relevance of FluPol 3′ RNA promoter binding sites. I find that 3′ vRNA promoter binding in the Mode A site is important for *de novo* initiation of cRNA synthesis, which challenges existing hypotheses. I also find that 3′ RNA promoter binding in the Mode B site is required for viral genome transcription and replication, and using more detailed analyses I show that this is important for the elongation of mRNA products by transcribing vRNPs. From these data I present a model for the function of the Mode B site, which is now well supported by an independent study (44). My findings not only improve understanding of influenza A virus RNA synthesis, but may also be applicable to other negative-strand RNA viruses with similar RNA polymerases.

CHAPTER 3

The role of influenza virus RNA polymerase dimerization in viral genome replication

Data from this chapter were published in:

Fan H, **Walker AP***, Carrique L*, Keown JR*, Serna Martin I, Karia D, Sharps J, Hengrung N, Pardon E, Steyaert J, Grimes JM, Fodor E. 2019. Structures of influenza A virus RNA polymerase offer insight into viral genome replication. Nature 573:287-290.*

** Joint authorship*

3.1 INTRODUCTION

The influenza virus RNA-dependent RNA polymerase (FluPol) is responsible for replicating and transcribing the viral RNA genome. Recent structural data have provided detailed insights into the mechanisms underlying viral genome transcription, however, the process of viral genome replication remains poorly understood (29).

Viral genome replication is a two-step process involving synthesis of a positive-sense complementary RNA (cRNA) intermediate, then viral RNA (vRNA) synthesis using the cRNA as a template. FluPol initiates both steps of viral genome replication without a primer, by *de novo* synthesis and extension of a pppApG dinucleotide (29). However, there are some differences between the two steps, such as a ′prime-and-realign′ mechanism which is used to initiate vRNA synthesis but not cRNA synthesis. In this mechanism the initiating pppApG dinucleotide relocates from positions 4 and 5 on the 3′ cRNA promoter, where it is synthesised, to positions 1 and 2

(72). The molecular mechanism of prime-and-realign is not fully understood, although the PB1 priming loop is thought to be involved (76).

Viral genome replication occurs later in the course of an infection, after newly synthesised viral proteins have accumulated (19). FluPol and nucleoprotein (NP) are required to assemble nascent cRNA and vRNA into ribonucleoproteins (RNPs), which has led to the suggestion that FluPol oligomerisation is a requirement for viral genome replication. Several studies support this, although they provide little insight into the exact function of FluPol oligomerisation and present some conflicting results (27, 79). Here I aimed to identify the precise role, if any, of FluPol oligomerisation in viral genome replication.

3.2 **RESULTS**

3.2.1 *Detection of influenza virus RNA polymerase dimerization in cells*

Previous work from our group showed that purified FluPol from influenza A virus ($FluPol_A$) forms dimers of heterotrimers in solution (**Fig. 3.1A**) (37). This symmetrical dimer interface is primarily formed by the PA subunit, with PB2 and PB1 also contributing some interactions. A key feature of the dimer interface is a series of charged amino acid residues from PA 352-356, which interact with the same region on the opposing PA subunit (**Fig. 3.1B**).

To examine the functional relevance of $FluPol_A$ dimerization I first determined whether the $FluPol_A$ dimer observed in solution could form in cells. Since dimerization is likely to be dependent on $FluPol_A$ concentration I used a bimolecular fluorescence complementation (BiFC) approach, which detects protein-protein interactions using a split fluorescent reporter and does not require cell lysis (103). To detect $FluPol_A$ dimerization I fused N-terminal (1-155; VN) and C-terminal (156-229; VC) fragments of Venus, a yellow fluorescent protein (YFP) variant optimised for BiFC, to the N-terminus of PB1 (**Fig. 3.1C**) (104). This design places the Venus fragments close enough to interact upon $FluPol_A$ dimerization and ensures that no signal is

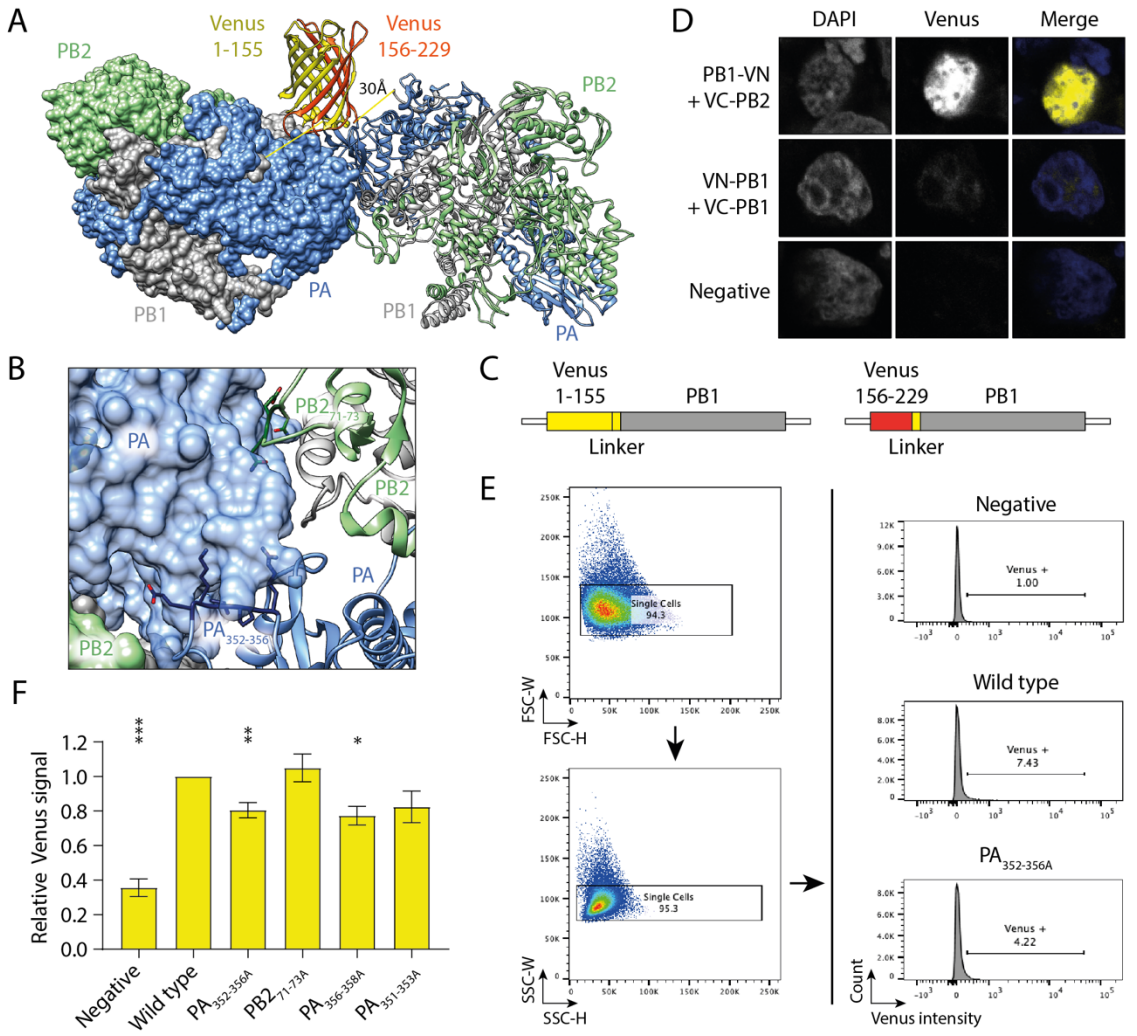

Figure 3.1 **Analysis of FluPol$_A$ dimerization using Bimolecular Fluorescence Complementation.** A, Structure of a FluPol$_A$ dimer of heterotrimers observed in solution (PDB: 6QNW). One heterotrimer is shown as a surface and the other as ribbons. Reconstituted split Venus protein is modelled across the dimer interface, with flexible linker peptides represented as solid yellow lines (PDB: 1YFP). B, Close-up view of the PA 352-356 loop and PB2 71-73 loop interactions at the dimer interface. PA 352-356 and PB2 71-73 amino acid residues for the cartoon FluPol$_A$ molecule are shown in dark blue and dark green respectively. C, Schematic of PB1 constructs with Venus N- and C- terminal fragments fused. D, Cells were transfected with plasmids encoding split Venus tagged FluPol$_A$ and stained with DAPI, then visualised using fluorescence microscopy. As a negative control the plasmids encoding PB2 and PA were omitted. E, Dimer interface mutations were introduced to split Venus tagged FluPol$_A$, and signal was quantified using flow cytometry. Gating was used to remove clumped cells (left) and quantify Venus signal (right). F, Venus signal was quantified from n=3-9 independent transfections, data are mean ± s.e.m., analysed by one-way ANOVA. *P<0.05, **P<0.01, ***P<0.001.

observed from interactions between $FluPol_A$ subunits in a single heterotrimer (**Fig. 3.1A**). I transfected plasmids encoding $FluPol_A$ fused to VN and VC into HEK 293T cells and examined Venus signal 24 hours later by confocal fluorescence microscopy (**Fig. 3.1D**). As a positive control for BiFC I used constructs with VN and VC at the PB1 C-terminus and PB2 N-terminus respectively, regions which interact closely within a single $FluPol_A$ heterotrimer. As expected, the positive control constructs produced strong Venus signal in cell nuclei. For wild type $FluPol_A$ with Venus fragments at the PB1 N-termini, faint Venus signals were also visible in cell nuclei. To confirm that this Venus signal resulted from interactions between $FluPol_A$ heterotrimers and not from intrinsic affinity between VN and VC fragments, I also overexpressed VN- and VC-tagged PB1 in the absence of PB2 or PA subunits. In this instance I did not observe any Venus signal, confirming that the BiFC system is suitable for detecting interactions between $FluPol_A$ heterotrimers.

To investigate whether the $FluPol_A$ interactions detected by BiFC correspond to the dimer observed in solution, I created a series of mutations at the dimer interface focusing on the key PA 352-356 loop interaction. Jane Sharps confirmed equal expression of the mutant $FluPol_A$ subunits and I expressed wild type or dimer interface mutant $FluPol_A$ with the VN-PB1 and VC-PB1 BiFC tags in HEK 293T cells, then used flow cytometry to quantify Venus signal for each sample relative to the wild type (**Fig. 3.1E, F**). The $PA_{352\text{-}356A}$ and $PA_{356\text{-}358A}$ mutations reduced Venus signal, while the $PB2_{71\text{-}73A}$ and $PA_{351\text{-}353A}$ mutations had no significant effect. Collectively, these data suggest that $FluPol_A$ forms dimers of heterotrimers in cells and the PA 352-356 loop is important for this interaction.

3.2.2 *Influenza virus RNA polymerase dimerization is required for initiating vRNA synthesis*

I next aimed to determine the effect of mutating the dimer interface on $FluPol_A$ function. I therefore transfected HEK 293T cells with plasmids encoding wild type or $PA_{352\text{-}356A}$ mutant $FluPol_A$, NP, and a vRNA segment, then quantified viral RNA levels 24 hours post-transfection by primer extension. In this minigenome assay the $PA_{352\text{-}356A}$ mutation caused a significant reduction in mRNA, vRNA and cRNA synthesis levels compared to wild type (**Fig. 3.2A**). $PB2_{71\text{-}73A}$, $PA_{351\text{-}535A}$, and $PA_{356\text{-}358A}$ mutations all caused similar reductions in $FluPol_A$ activity, indicating that the $FluPol_A$ dimer interface is functionally important (**Fig. 3.2B**).

One of the functions of $FluPol_A$ is to perform viral genome transcription, synthesising viral mRNA using vRNA as a template. If a mutation interferes with the ability of $FluPol_A$ to synthesise vRNA in a minigenome assay, mRNA synthesis is always reduced because less vRNA template is available. Therefore, reduced mRNA signal in a minigenome assay does not necessarily mean that a mutation interferes with $FluPol_A$ transcription activity. To overcome this limitation and determine whether the $PA_{352\text{-}356A}$ mutation directly affects $FluPol_A$ transcription activity, I used a previously described complementation approach (**Fig. 3.2A**) (87). $FluPol_A$ with a PA_{D108A} mutation cannot synthesise viral mRNA but can synthesise vRNA, which can then be used as a template for mRNA synthesis by co-expressed $FluPol_A$ (101). When I co-expressed PA_{D108A} $FluPol_A$ and $PA_{352\text{-}356A}$ $FluPol_A$ in a minigenome assay I observed a robust mRNA signal. This result indicates that when provided with sufficient vRNA template, $PA_{352\text{-}356A}$ mutant $FluPol_A$ is able to efficiently perform viral genome transcription. Therefore, these data show that the $FluPol_A$ dimer interface is important for viral genome replication, but not for transcription.

To gain a more detailed insight into the role of the $FluPol_A$ dimer interface in viral genome replication, I used *in vitro* activity assays to individually examine each reaction performed by $FluPol_A$ (**Fig. 3.2C**). First, I tested the transcription activity of purified $FluPol_A$ by incubating with a model 5′ and 3′ vRNA promoter and radiolabelled capped RNA in the presence of NTPs

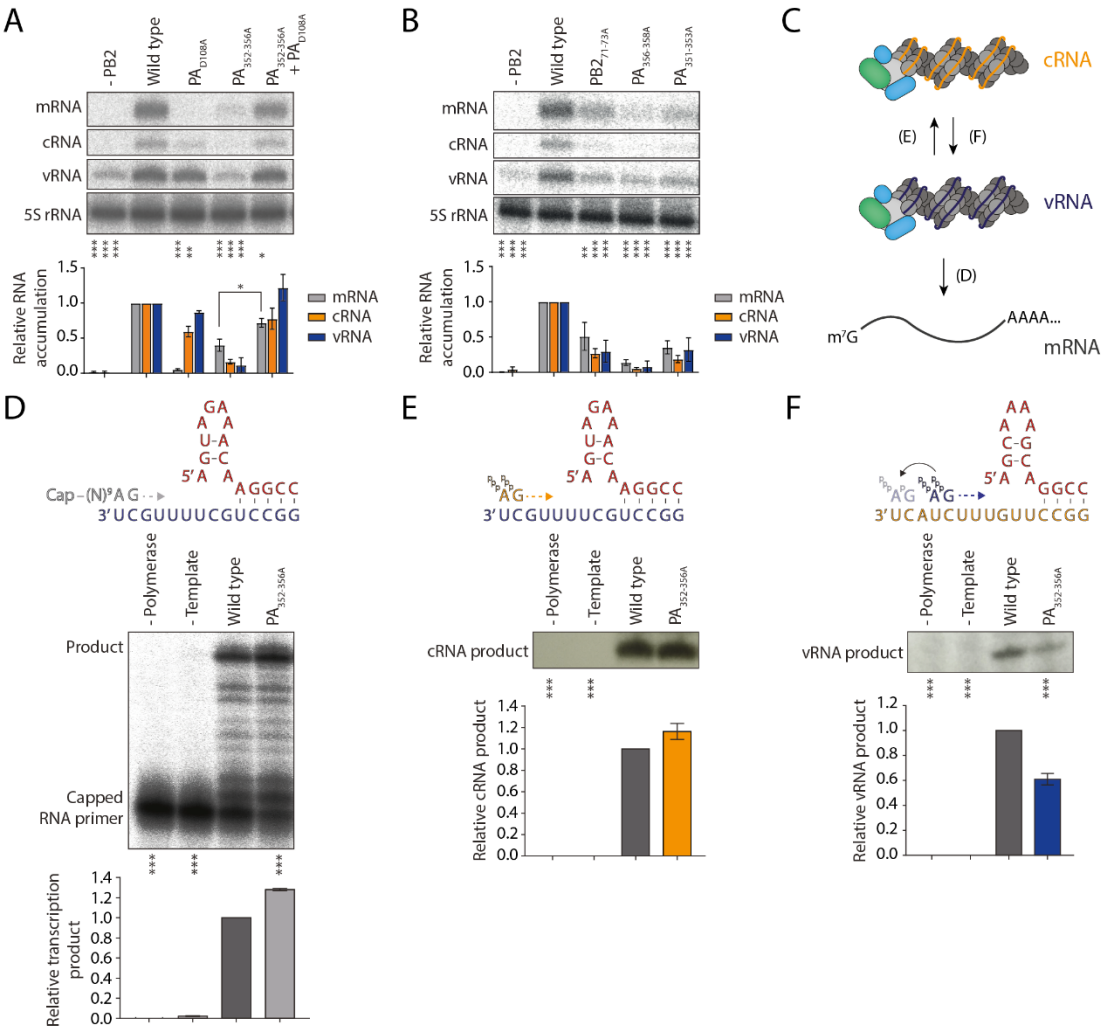
A
- PB2
Wild type
PA D108A
PA 352-356A
PA 352-356A + PA D108A
mRNA
cRNA
vRNA
5S rRNA
Relative RNA accumulation
mRNA
cRNA
vRNA
B
- PB2
Wild type
PB2 71-73A
PA 356-358A
PA 351-353A
mRNA
cRNA
vRNA
5S rRNA
Relative RNA accumulation
mRNA
cRNA
vRNA
C
cRNA
(E)
(F)
vRNA
(D)
m7G
AAAA...
mRNA
D
Cap–(N)9A G
3'UCGUUUUCGUCCGG
- Polymerase
- Template
Wild type
PA 352-356A
Product
Capped RNA primer
Relative transcription product
E
3'UCGUUUUCGUCCGG
- Polymerase
- Template
Wild type
PA 352-356A
cRNA product
Relative cRNA product
F
3'UCAUCUUUGUUCCGG
- Polymerase
- Template
Wild type
PA 352-356A
vRNA product
Relative vRNA product

Figure 3.2 **Effect of dimer interface mutations on FluPol$_A$ activity.** A, Minigenome assay with the PA$_{352-356A}$ FluPol$_A$ mutant. HEK 293T cells were transfected with plasmids encoding wild type or mutant FluPol$_A$ subunits, NP and a vRNA template. Viral RNA levels were analysed 24 hours post-transfection by primer extension. Complementation with transcription-deficient PA$_{D108A}$ FluPol$_A$ was used to determine the transcription activity of PA$_{352-356A}$ FluPol$_A$. B, Minigenome assays were performed with further FluPol$_A$ dimer interface mutants. C, Schematic of FluPol viral genome transcription and replication. D, Schematic of capped RNA primer binding to model vRNA promoter (top). *In vitro* transcription assays were performed with PA$_{352-356A}$ mutant FluPol$_A$ (bottom). E, Schematic of *de novo* initiation on a model vRNA promoter (top). *In vitro* cRNA synthesis assays were performed with PA$_{352-356A}$ mutant FluPol$_A$ (bottom). F, Schematic of *de novo* initiation and prime-and-realign on a model cRNA promoter (top). *In vitro* vRNA synthesis assays were performed with PA$_{352-356A}$ mutant FluPol$_A$ (bottom). Quantification is from n=3 independent transfections or reactions, data are mean ± s.e.m., analysed by one-way ANOVA. *P<0.05, **P<0.01, ***P<0.001. Individual mRNA signals were compared by two-tailed unpaired t-test; *P<0.05, **P<0.01, ***P<0.001.

(**Fig. 3.2D**). The PA$_{352-356A}$ mutation did not reduce FluPol$_A$ transcription activity, which is consistent with the minigenome assay result. Next, I tested cRNA synthesis activity by incubating FluPol$_A$ with the model vRNA promoters and NTPs in the absence of a primer, to allow *de novo* initiation (**Fig. 3.2E**). Under these conditions, the PA$_{352-356A}$ mutation did not reduce cRNA synthesis activity. Finally, I tested vRNA synthesis activity by incubating FluPol$_A$ with model 5′ and 3′ cRNA promoters and NTPs in the absence of a primer (**Fig. 3.2F**). In this assay, the PA$_{352-356A}$ mutation caused a significant decrease in FluPol$_A$ vRNA synthesis activity by approximately 50%. Collectively, these data indicate that the FluPol$_A$ dimer interface is important for the vRNA synthesis step of viral genome replication, but is dispensable for mRNA or cRNA synthesis.

The mechanism for initiating vRNA synthesis is distinct from cRNA synthesis as it requires a prime-and-realign step (72). I hypothesised that since the FluPol$_A$ dimer interface is only important for vRNA synthesis, it could be involved in prime-and-realign. To test this hypothesis, I performed *in vitro* assays with an ApG dinucleotide primer which imitates the pppApG

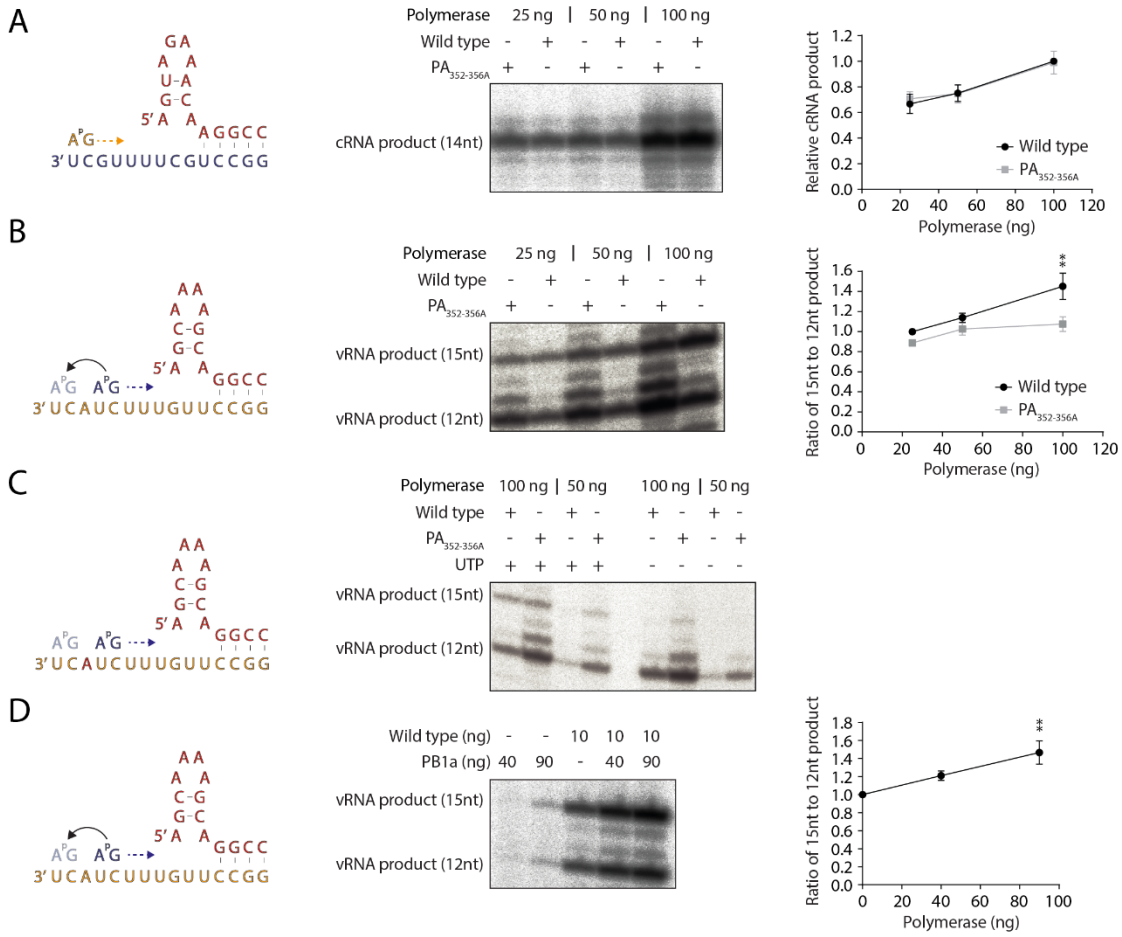

Figure 3.3 **Prime-and-realign activity with dimer interface mutant FluPol$_A$.** A, Schematic of ApG-primed initiation on a model vRNA template (left). ApG-primed cRNA synthesis assays were performed with PA$_{352-356A}$ mutant FluPol$_A$ (right). B, Schematic of terminal and internal ApG-primed initiation on a model cRNA template (left). ApG-primed vRNA synthesis assays were performed with PA$_{352-356A}$ mutant FluPol$_A$ (right). C, Schematic of terminal and internal ApG-primed initiation on a model cRNA template with no UTP (left). ApG-primed vRNA synthesis assays were performed in the presence or absence of UTP (right). D, Schematic of terminal and internal ApG-primed initiation on a model cRNA template (left). ApG-primed vRNA synthesis assays were performed with added catalytically inactive FluPol$_A$ (right). Quantification is from n=3 independent reactions, data are mean ± s.e.m., analysed by one-way or two-way ANOVA. **P<0.01

dinucleotide formed during *de novo* initiation. The $PA_{352\text{-}356A}$ mutation had no effect on ApG-primed cRNA synthesis, which consistent with the primer-independent assay (**Fig. 3.3A**). In ApG-primed vRNA synthesis assays I observed two major products, which I assigned as full-length (15nt), or resulting from the failure of $FluPol_A$ to realign the initiating ApG from template positions 4 and 5 to positions 1 and 2 (12nt) (**Fig. 3.3B**). Additional products at 13nt and 14nt could either result from non-templated extension of the 12nt product, or from aberrant realignment of ApG to other positions on the template. In these assays the $PA_{352\text{-}356A}$ mutation did not reduce the overall level of product synthesised; however, on comparison of the 12nt and 15nt products, I found that the $PA_{352\text{-}356A}$ mutation significantly reduced the amount of 15nt product synthesised compared to the 12nt product. This effect was dependent on the concentration of $FluPol_A$, which is consistent with involvement of an intermolecular interaction. To confirm that the 12nt product results from failed prime-and-realign I performed assays in the absence of UTP, which prevented full-length product synthesis but allowed 12nt product synthesis since the only adenosine in the 3′ cRNA promoter is at position 3 (**Fig. 3.3C**). These data show that the 12nt product results from $FluPol_A$ extending the ApG from positions 4 and 5 upon failure to realign to positions 1 and 2.

Finally, to confirm the conclusion that an intermolecular interaction between $FluPol_A$ molecules promotes prime-and-realign, I added increasing concentrations of catalytically inactive (PB1a) $FluPol_A$ to a constant concentration of wild type $FluPol_A$ (**Fig. 3.3D**). As expected, the quantity of 15nt vRNA product synthesised compared to the 12nt product increased significantly in a concentration-dependent manner. Collectively, these data show that $FluPol_A$ dimerization through the PA 352-356 loop is important for vRNA synthesis by promoting prime-and-realign.

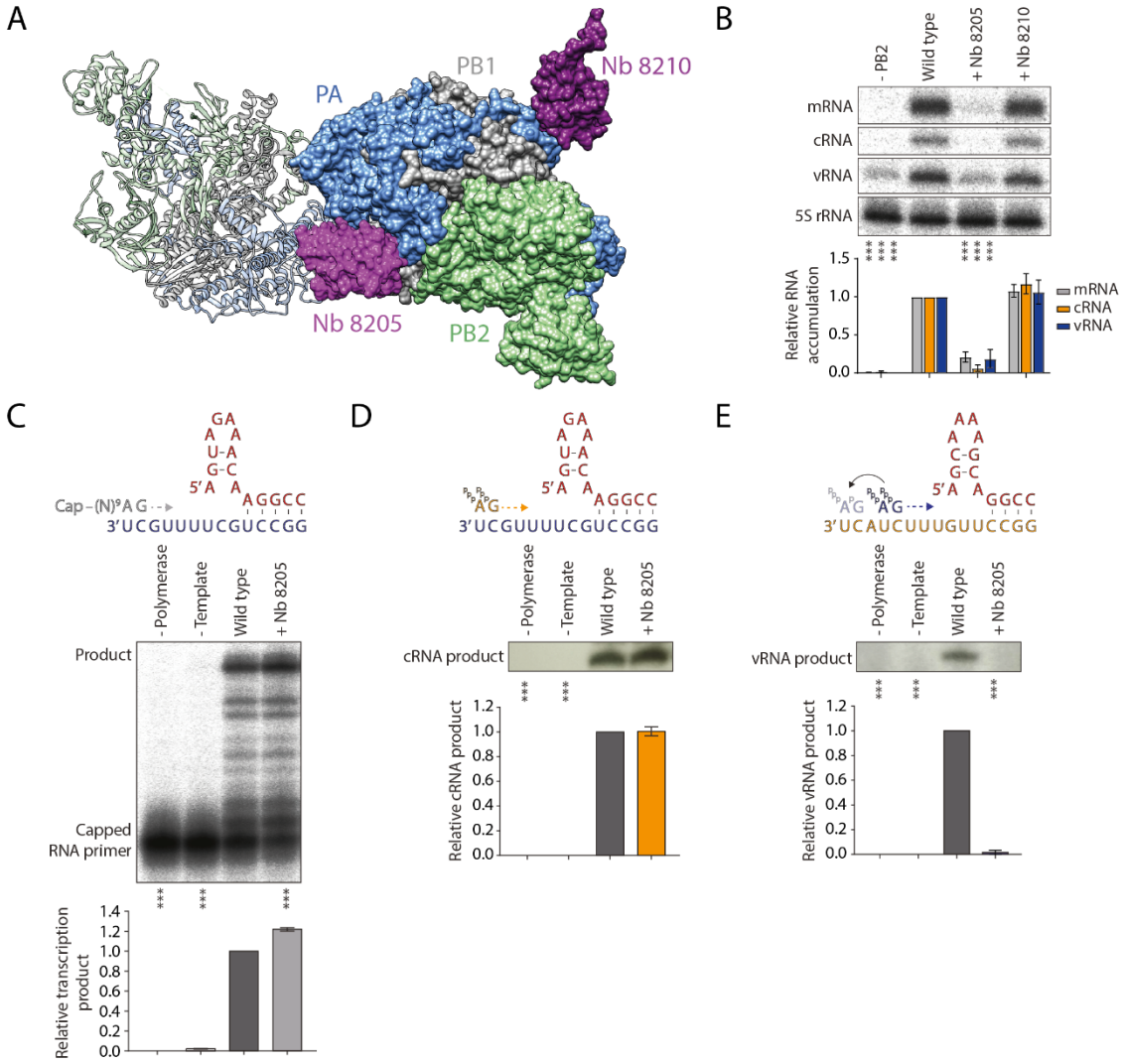

Figure 3.4 **The effect of Nb 8205 on FluPol$_A$ activity.** A, Model of the FluPol$_A$ dimer with Nb 8205 (magenta; PDB: 6QX3) and Nb 8210 (dark magenta; unpublished data courtesy of Jeremy Keown) binding sites shown. B, FluPol$_A$ minigenome assays were performed in the presence or absence of Nb 8205 and Nb 8210. C, Schematic of capped RNA primer binding to model vRNA promoter (top). *In vitro* transcription assays were performed in the presence or absence of Nb 8205 (bottom). D, Schematic of *de novo* initiation on a model vRNA promoter (top). *In vitro* cRNA synthesis assays were performed in the presence or absence of Nb 8205 (bottom). E, Schematic of *de novo* initiation and prime-and-realign on a model cRNA promoter (top). *In vitro* vRNA synthesis assays were performed in the presence or absence of Nb 8205 (bottom). Quantification is from n=3 independent transfections or reactions, data are mean ± s.e.m., analysed by one-way ANOVA. ***P<0.001.

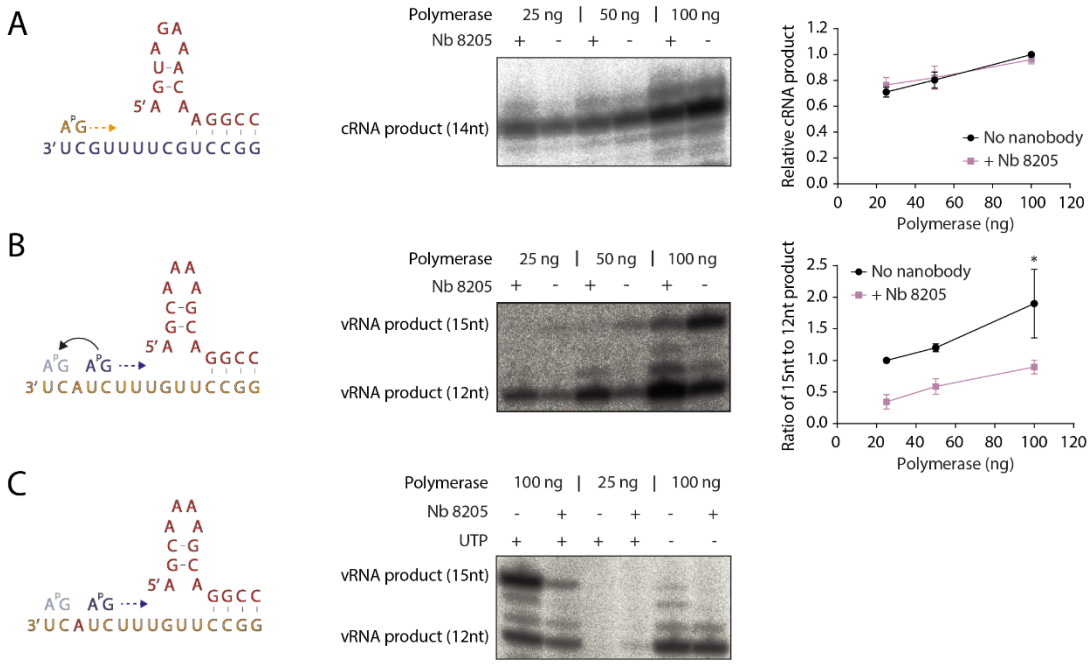

Figure 3.5 **FluPol$_A$ prime-and-realign activity in the presence of Nb 8205.** A, Schematic of ApG-primed initiation on a model vRNA template (left). ApG-primed cRNA synthesis assays were performed in the presence or absence of Nb 8205 (right). B, Schematic of terminal and internal ApG-primed initiation on a model cRNA template (left). ApG-primed vRNA synthesis assays were performed in the presence or absence of Nb 8205 (right). Quantification is from n=3 independent reactions, data are mean ± s.e.m., analysed by two-way ANOVA. *P<0.05. C, Schematic of terminal and internal ApG-primed initiation on a model cRNA template with no UTP (left). ApG-primed vRNA synthesis assays were performed in the presence or absence of UTP and Nb 8205 (right).

3.2.3 *Disruption of influenza virus RNA polymerase dimerization using a nanobody*

Next, I aimed to study the effect on FluPol$_A$ function of disrupting dimerization using a nanobody. Nanobodies are single-domain antibodies derived from camelids, and their epitope recognition loops are all located on a single 12-15kDa immunoglobulin domain (105). Nb 8205 binds to the FluPol$_A$ PA subunit close to the dimer interface, and has been shown to disrupt FluPol$_A$ dimerization in solution and in cells (**Fig. 3.4A**) (37). To study the functional consequences of this, I first performed FluPol$_A$ minigenome assays in the presence of co-expressed Nb 8205 (**Fig.**

3.4B). Nb 8205 abolished $FluPol_A$ activity as measured by mRNA, vRNA and cRNA accumulation, whereas a control nanobody (Nb 8210) which binds to the PA linker had no effect on $FluPol_A$ activity (**Fig. 3.4A**). To provide further evidence for the function of $FluPol_A$ dimerization, I performed *in vitro* $FluPol_A$ activity assays in the presence of Nb 8205. Similar to the $PA_{352\text{-}356A}$ mutation, Nb 8205 did not reduce mRNA synthesis or primer-independent cRNA synthesis (**Fig. 3.4C, D**). However, addition of Nb 8205 abolished the vRNA synthesis activity of purified $FluPol_A$ (**Fig. 3.4E**).

Next, I determined the effect of Nb 8205 on prime-and-realign during vRNA synthesis. Nb 8205 had no effect on ApG-primed cRNA synthesis, which is consistent with the primer-independent cRNA synthesis assay (**Fig. 3.5A**). However, in ApG-primed vRNA synthesis assays, Nb 8205 strongly reduced synthesis of the full length 15nt product when compared to the truncated 12nt product, with a more potent effect than the $PA_{352\text{-}356A}$ mutation (**Fig. 3.5B**). To once again confirm that the 12nt product observed resulted from failed prime-and-realign, I performed assays in the absence of UTP, which prevented ApG extension from positions 1 and 2 on the template. With no UTP present very little 15nt product was synthesised, confirming that the 12nt product resulted from failed prime-and-realign (**Fig. 3.5C**). Collectively, these data show that disrupting $FluPol_A$ dimerization using a nanobody causes a defect in prime-and-realign during vRNA synthesis initiation.

Finally, to determine the effect of blocking $FluPol_A$ dimerization on influenza A virus growth, I transfected HEK 293T cells with plasmids encoding Nb 8205 or Nb 8210, or an empty plasmid. 24 hours post-transfection, I infected the cells with influenza A/WSN/33 virus and assayed viral growth for 48 hours post-infection (**Fig. 3.6A**). Nb 8205 significantly reduced virus titre at 24 and 32 hours post-infection, while Nb 8210 had no significant effect at any time point. To confirm that the reduced virus titre in the presence of Nb 8205 resulted from decreased viral RNA synthesis, I extracted total cellular RNA at the time points indicated and examined viral RNA accumulation by primer extension (**Fig. 3.6B**). mRNA, cRNA and vRNA accumulation was

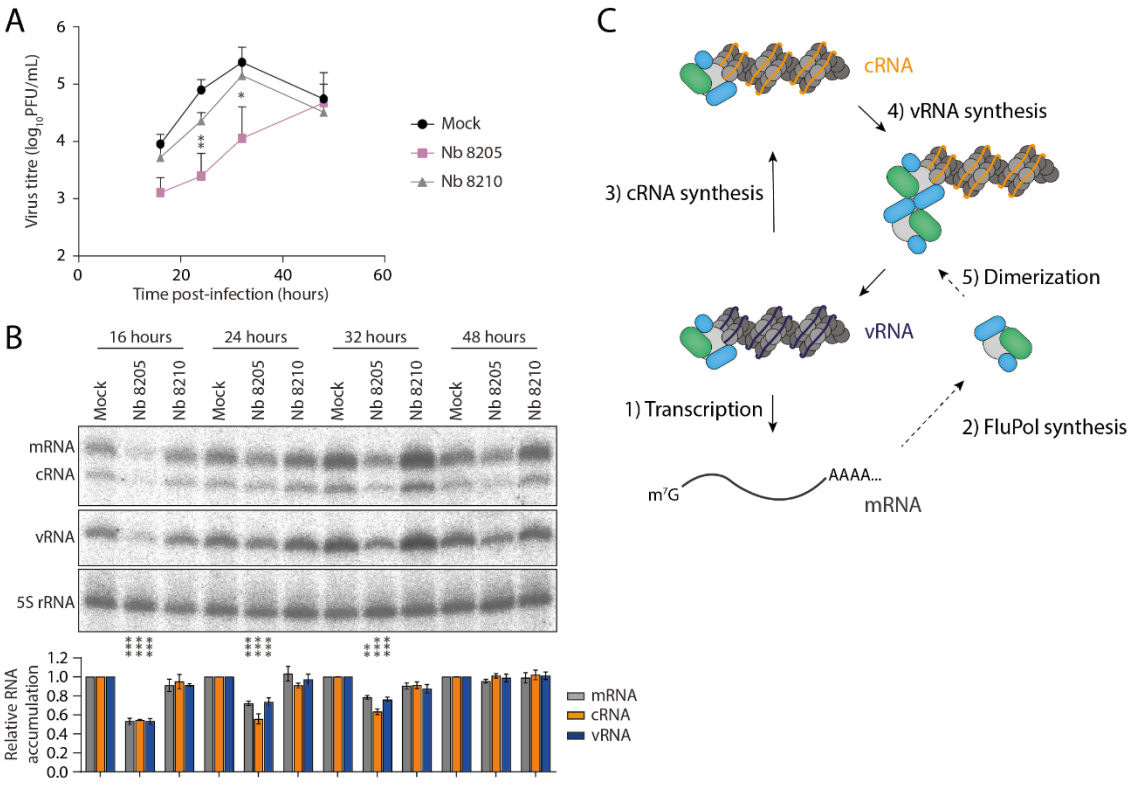

Figure 3.6 **Influenza A virus growth in the presence of Nb 8205 or Nb 8210, and a model for the role of FluPol$_A$ dimerization in viral genome replication.** A, HEK 293T cells were transfected with plasmids encoding Nb 8205 or Nb 8210, then infected with influenza A virus. Virus titres in cell culture media were determined by plaque assay up to 48 hours post-infection. B, Viral RNA accumulation in total cellular RNA was determined by primer extension. Quantification is from n=3 independent transfections, data are mean ± s.e.m., analysed by one-way or two-way ANOVA. *P<0.05, **P<0.01, ***P<0.001. C, Schematic of viral genome transcription and replication, including the role of FluPol$_A$ dimerization in modulating vRNA synthesis.

significantly reduced by Nb 8205 at 16, 24 and 32 hours post-infection, while Nb 8210 had only a minor effect on viral RNA accumulation at 32 hours post-transfection. As with virus titres, at 48 hours post-infection viral RNA accumulation in the presence of Nb 8205 was the same as the mock transfected sample. Collectively, these data demonstrate that blocking the FluPol$_A$ dimer interface reduces viral RNA accumulation and therefore influenza A virus growth.

3.3 **DISCUSSION**

With this study, I aimed to identify the role of $FluPol_A$ oligomerisation in viral genome replication. I first demonstrated that $FluPol_A$ forms dimers of heterotrimers in cells, and identified a critical region of the PA subunit for this interaction. Using both mutagenesis and a nanobody, I then showed that $FluPol_A$ dimerization is required for viral genome replication, specifically for initiating vRNA synthesis. Finally, I found that blocking the $FluPol_A$ dimer interface with a nanobody inhibits vRNA accumulation and influenza A virus growth.

Collectively, these data support a model in which newly synthesised FluPol activates vRNA synthesis by forming a dimer with cRNA-associated FluPol (**Fig. 3.6C**). This model is further supported by a more recent study, which demonstrated that recombinant influenza A viruses with aberrant FluPol dimerization have a defect in vRNA synthesis (106). The proposed model could allow FluPol dimerization to function as regulatory mechanism to modulate the level of vRNA synthesis based on the availability of free FluPol in the cell. In other words, FluPol dimerization would act as a molecular clock, ensuring that vRNA synthesis only takes place later in the course of an infection once sufficient FluPol is available to encapsidate nascent vRNA into vRNPs. This would help to avoid the production of naked vRNAs that would otherwise trigger an antiviral response due to their 5′ triphosphate activating innate immune receptors (102).

I find that $FluPol_A$ dimerization promotes vRNA synthesis by facilitating prime-and-realign, a key step during the initiation of vRNA synthesis (72). Cryo-EM structures of $FluPol_A$ from our group suggest a possible molecular mechanism for how dimerization is linked to prime-and-realign; briefly, intermolecular interactions which occur upon dimerization cause rearrangement of a helical bundle consisting of the PB1 thumb and the N-terminus of PB2. The PB1 priming loop, previously implicated in prime-and-realign, is attached to this helical bundle and pulls away from the $FluPol_A$ active site by approximately 7Å (76). The priming loop interacts with the 3′ cRNA promoter, so this movement could translocate the 3′ cRNA promoter such that the initiating pppApG dinucleotide relocates from positions 4 and 5 to positions 1 and 2. Intriguingly, several

other polymerases from negative-strand RNA viruses also utilise prime-and-realign mechanisms, such as Tacaribe arenavirus and Hantaan bunyavirus, raising the possibility that this is a shared mechanism for regulating vRNA synthesis (107–109).

Previous studies have presented several possible models for the function of FluPol oligomers (described in detail in section 1.3.5) (27, 79). My findings are consistent with $FluPol_A$ dimerization playing a *trans*-activating role in vRNA synthesis, since the prime-and-realign mechanism promoted by dimerization would not require the newly-synthesised FluPol molecule to be catalytically active. However, it is unlikely that this interaction is responsible for assembling new RNPs during viral genome replication, as dimerization is only required for vRNA synthesis and RNP assembly is a fundamental requirement for both cRNA and vRNA synthesis (29). Therefore, the primary function of $FluPol_A$ dimerization appears to be regulatory, and RNP assembly may occur by a different mechanism. Such a mechanism could involve further interactions between FluPol molecules, either indirectly or through the formation of other structurally distinct dimers of heterotrimers.

In summary, I have identified that the role of $FluPol_A$ dimerization in viral genome replication is promoting prime-and-realign during vRNA synthesis, which could be a key regulatory mechanism to control the rate of viral genome replication and avoid aberrant vRNA synthesis. I also showed that blocking the $FluPol_A$ dimer interface is detrimental to influenza A virus growth, implicating $FluPol_A$ dimerization as a possible target for antiviral drugs against influenza A virus.

CHAPTER 4

The role of host protein ANP32A in supporting influenza viral genome replication

Data from this chapter were published in:

Carrique L, Fan H*, **Walker AP***, Keown JR*, Sharps J, Staller E, Barclay WS, Fodor E, Grimes JM. 2020. Host ANP32A mediates the assembly of the influenza virus replicase. Nature 587:638–643.*

** Joint authorship*

4.1 INTRODUCTION

Influenza A viruses have the ability to infect many different hosts, including humans and aquatic birds. Adaption of avian influenza A viruses to human hosts can result in novel viruses such as H5N1 highly pathogenic avian influenza virus. These viruses are particularly dangerous because the human population has no pre-existing immunity, so they have pandemic potential (110).

In order for avian influenza A viruses to replicate efficiently in human cells, their RNA-dependent RNA polymerases ($FluPol_A$) must undergo one of several point mutations, with the most-studied example being $PB2_{E627K}$ (81). Previous studies found that the purpose of this mutation is to make avian-origin $FluPol_A$ functionally compatible with ANP32A (Acidic Nuclear Phosphoprotein 32kDa A) and ANP32B in human cells. Most avian ANP32A proteins, including chicken ANP32A (chANP32A), have a 33 amino acid residue insertion compared to human ANP32A (huANP32A) and huANP32B, which enables them to support avian-origin ($PB2_{E627}$) $FluPol_A$ activity. Since huANP32A and huANP32B lack this insertion they cannot normally support

avian-origin $FluPol_A$ activity. On the other hand, human-origin ($PB2_{K627}$) $FluPol_A$ activity can be supported by huANP32A, huANP32B or chANP32A (88).

The huANP32A and huANP32B host cellular proteins are functionally redundant but essential for $FluPol_A$ activity, to the point where huANP32A/huANP32B double knockout cells are completely unable to support influenza A virus growth (92). Past studies have implicated ANP32 proteins in supporting viral genome replication, however, the precise molecular role of this interaction remains elusive (90–92). Recently, our group were able to solve cryo-EM structures of the $FluPol_C$-huANP32A and $FluPol_C$-chANP32A complexes (111). Here I study the function of these complexes to gain insights into the role of ANP32 proteins in host adaption and viral genome replication.

4.2 **RESULTS**

4.2.1 *Characterisation of the influenza virus RNA polymerase-ANP32A interface*

I first aimed to gain insight into the functional relevance of the FluPol-ANP32A interface identified in the $FluPol_C$-huANP32A complex cryo-EM structure (**Fig. 4.1A**) (111). The huANP32A N-terminal leucine-rich repeat (LRR) domain is well resolved and binds in a cavity formed by an asymmetric dimer of two $FluPol_C$ heterotrimers, henceforth referred to individually as the replicating FluPol ($FluPol^R$) and encapsidating FluPol ($FluPol^E$). I identified several key interactions between the huANP32A LRR domain and $FluPol_C$, and made substitutions at equivalent amino acid residues in $FluPol_A$ as identified by sequence and structural alignments. Amino acid residues $P3_{E513}$, $P3_{K391}$ and $PB2_{H631}$ in $FluPol_C$ were found to be equivalent to PA_{D529}, PA_{K413} and $PB2_{T609}$ respectively in $FluPol_A$, which I used for functional assays as it is a better-characterised polymerase and a more relevant background for studies on avian-human host adaption by influenza A viruses (**Fig. 4.1B, C, D, E**). An early version of the cryo-EM structure suggested that amino acid residues $P3_{N515}$ and $P3_{E317}$, equivalent to $FluPol_A$ residues PA_{R531} and

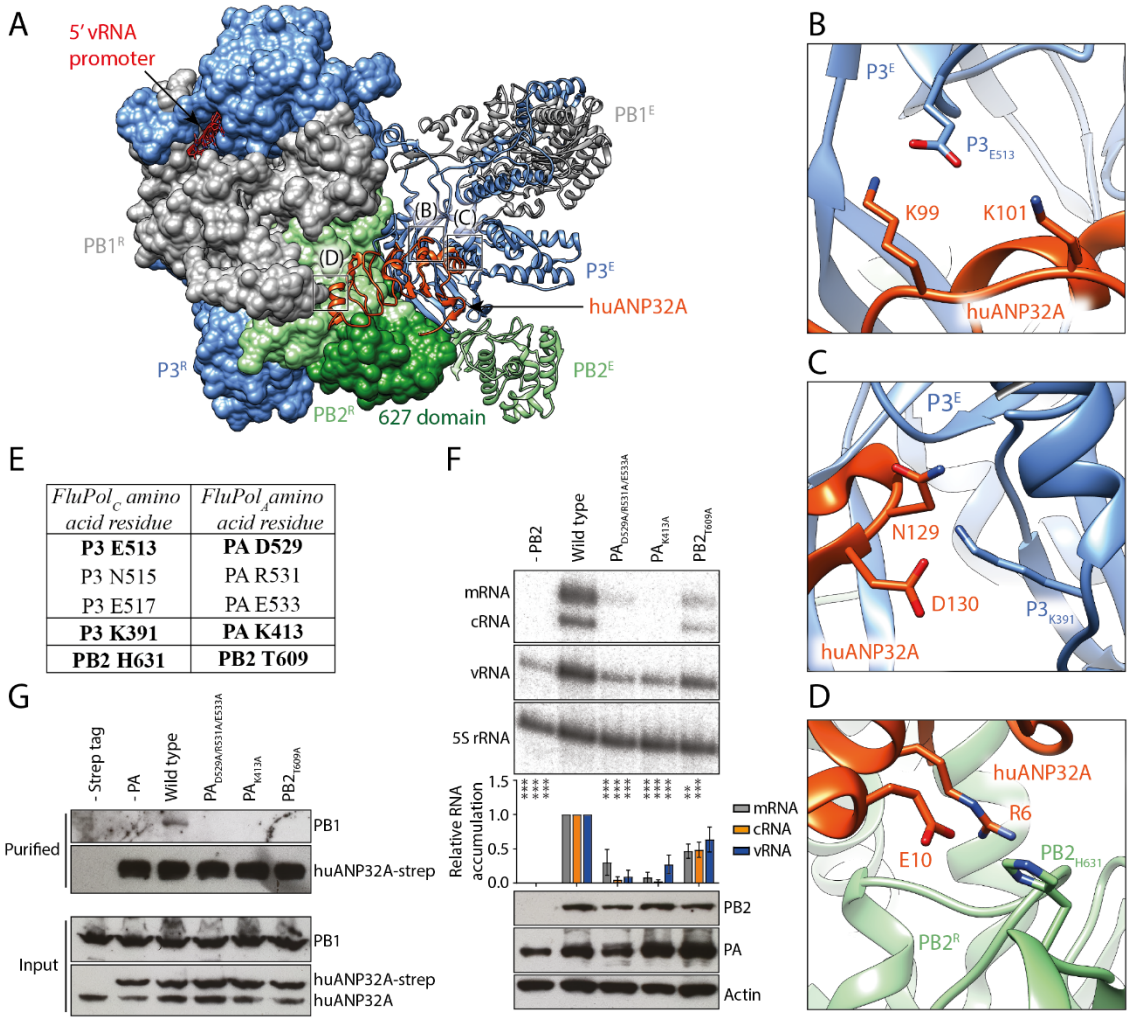

Figure 4.1 **Functional analysis of FluPol-huANP32A interface mutations.** A, Structure of the FluPol$_C$-huANP32A complex (PDB: 6XZQ) bound to vRNA (red). FluPol^R (replicating FluPol) and FluPol^E (encapsidating FluPol) are shown as ribbon and surface representations respectively. Boxes indicate key interactions between FluPol$_C$ and huANP32A. B, C, D, Close-up views of key interactions between huANP32A (orange) and FluPol$_C$ PB2 (green) or PA (blue) subunits. E, Equivalent amino acid residues in FluPol$_C$ and FluPol$_A$, with key interacting residues in bold. F, HEK 293T cells were transfected with plasmids encoding wild type or mutant FluPol$_A$ subunits, NP and a vRNA template. Viral RNA levels were analysed 24 hours post-transfection by primer extension (top), and PB2 or PA protein expression was analysed by western blot with actin as a loading control (bottom). Quantification is from n=3 independent transfections, data are mean ± s.e.m., analysed by one-way ANOVA. **P<0.01, ***P<0.001. G, HEK 293T cells were transfected with plasmids encoding wild type or mutant FluPol$_A$ subunits and huANP32A-strep. huANP32A-strep was purified 48 hours post-transfection, and co-purified FluPol$_A$ was detected by western blot for PB1 protein.

PA_{E533}, could also be important for huANP32 LRR binding. These amino acid residues were therefore mutated along with PA_{D529}, although the final cryo-EM structure indicates that they do not interact directly with the LRR.

I began by studying the effect of FluPol-huANP32A interface mutations on $FluPol_A$ activity. I transfected HEK 293T cells with plasmids encoding wild type or mutant $FluPol_A$, nucleoprotein (NP) and a viral RNA (vRNA) template, and analysed viral RNA levels by primer extension (**Fig. 4.1F**). In these minigenome assays all huANP32A interface mutations significantly disrupted $FluPol_A$ activity, as measured by mRNA, cRNA and vRNA accumulation. To confirm that the mutations successfully disrupted the $FluPol_A$-huANP32A interaction, I transfected HEK 293T cells with plasmids encoding strep-tagged huANP32A and wild type or mutant $FluPol_A$, then analysed co-purification of $FluPol_A$ with strep-tagged huANP32A (**Fig. 4.1G**). None of the mutant $FluPol_A$ proteins co-purified with strep-tagged huANP32A, confirming that the FluPol-huANP32A interface identified in the $FluPol_C$-huANP32A cryo-EM structure is functionally relevant for $FluPol_A$ activity.

chANP32A can support $PB2_{E627}$ $FluPol_A$ activity because it contains a 33 amino acid residue insertion in the low-complexity acidic region (LCAR) domain compared to huANP32A (88). Analysis of the $FluPol_C$-chANP32A cryo-EM structure suggested that chANP32A residues 176-183, located in the LCAR insertion, are positioned close to $FluPol^R$ residue $PB2_{K649}$ (equivalent to $PB2_{K627}$ in $FluPol_A$) (**Fig. 4.2A**). Therefore, I hypothesised that chANP32A residues 176-183 are necessary to form a productive interaction with $PB2_{E627}$ $FluPol_A$. To test this hypothesis I replaced amino acid residues 176-183 of huANP32A with the corresponding residues of chANP32A, resulting in the $huANP32A_{176-183}$ chimeric construct (**Fig. 4.2B**). When I co-transfected the $huANP32A_{176-183}$ construct with $PB2_{E627}$ $FluPol_A$ in minigenome assays, I found that it was fully able to support $FluPol_A$ activity with comparable efficiency to wild type chANP32A (**Fig. 4.2C**). Collectively, these data show that huANP32A residues 176-183 are responsible for the low activity of $PB2_{E627}$ $FluPol_A$ in human cells.

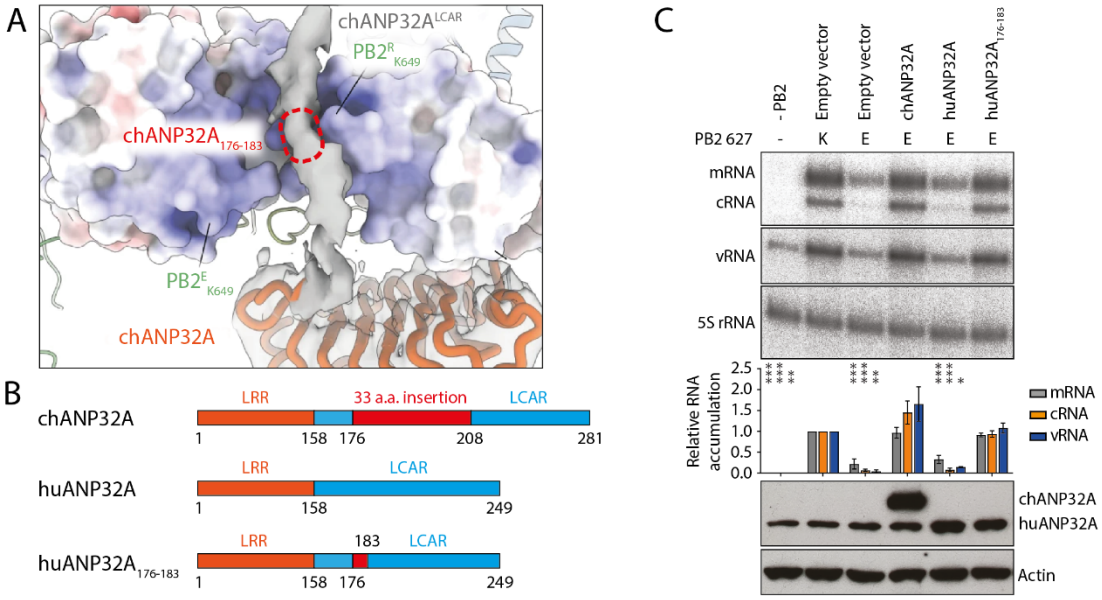

Figure 4.2 **Effect of ANP32A LCAR interactions on FluPol$_A$ activity.** A, Cryo-EM density of the chANP32A LCAR (grey) interacting with FluPol^R and FluPol^E PB2 627 domains. PB2 627 domains are coloured according to surface charge, with positive charges in blue and negative charges in red. The approximate location of chANP32A amino acid residues 176-183 are highlighted (red box) next to PB2$_{K649}$ residues, which are equivalent to PB2$_{K627}$ in FluPol$_A$. Adapted from Carrique *et al*. 2020 (111). B, Protein domain schematics of chANP32A, huANP32A and the huANP32A$_{176-183}$ construct, not to scale. The avian-ANP32 specific 33a.a. insertion (red) is highlighted in the LCAR (blue). C, PB2$_{K627}$ or PB2$_{E627}$ FluPol$_A$ minigenome assays were performed in HEK 293T cells with or without co-expressed ANP32A. Viral RNA levels were assayed by primer extension 24 hours post-transfection (top), and ANP32A protein expression was analysed by western blot with actin as a loading control (bottom). Quantification is from n=3 independent transfections, data are mean ± s.e.m., analysed by one-way ANOVA. *P<0.05, **P<0.01, ***P<0.001.

4.2.2 *The asymmetric influenza virus RNA polymerase dimer bridged by ANP32A encapsidates nascent cRNA*

The $FluPol_C$-huANP32A cryo-EM structure contains an asymmetric dimer of $FluPol_C$ molecules, which is structurally distinct from the symmetrical *trans*-activating $FluPol_A$ dimer discussed in Chapter 3 (111). Based on this observation I hypothesised that $FluPol_A$ can undergo multiple functionally important dimerization interactions during viral genome replication.

To examine this possibility I first performed a cRNA encapsidation assay, which has been described previously (19, 71). Briefly, I transfected HEK 293T cells with plasmids encoding catalytically inactive $FluPol_A$ and NP, then infected the cells with influenza A/WSN/33 virus in the presence of actinomycin D, which prevents cellular mRNA synthesis. In this assay the infecting virions release vRNPs into the cell which begin synthesising viral mRNA and cRNA. If pre-expressed $FluPol_A$ is present, it encapsidates the nascent cRNA and causes it to accumulate; otherwise, the cRNA is degraded (**Fig. 4.3A**). When I performed this assay with pre-expressed wild type $FluPol_A$, cRNA accumulated as expected. Inclusion of the $PA_{352\text{-}356A}$ *trans*-activating dimer interface mutation did not prevent cRNA accumulation, however, deletion of the PB2 627 domain ($PB2_{535\text{-}667\Delta}$) in the pre-expressed $FluPol_A$ abolished cRNA accumulation which is consistent with previous observations (112). Next, I examined the effect of pre-expressed $FluPol_A$ on primary transcription, which is the synthesis of viral mRNA by incoming vRNPs, using a similar experimental setup (**Fig. 4.3B**). Interestingly, I found that that pre-expression of wild type or $PA_{352\text{-}356A}$ $FluPol_A$ significantly reduced primary transcription, while inclusion of the $PB2_{535\text{-}667\Delta}$ mutation in pre-expressed $FluPol_A$ alleviated this effect. Taken together, these data raise the possibility that an interaction between pre-expressed $FluPol_A$ and vRNPs, distinct from the *trans*-activating $FluPol_A$ dimer, promotes cRNA encapsidation and reduces transcription.

I hypothesised that the $FluPol^R$-$FluPol^E$ asymmetric dimer interface observed in the $FluPol_C$-huANP32A structure could be responsible for cRNA encapsidation, as the product exit channel of $FluPol^R$ is located close to the 5′ RNA promoter binding site of $FluPol^E$, meaning that the

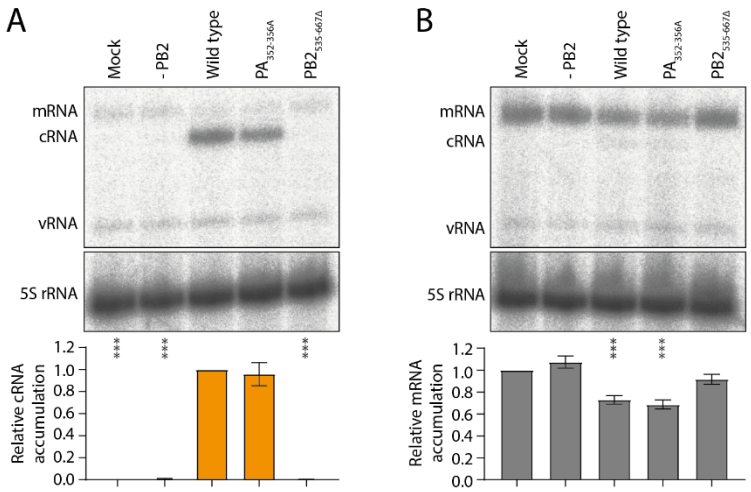

Figure 4.3 **The effect of FluPol$_A$ expression on vRNP RNA synthesis activity.** A, cRNA encapsidation assays were performed by transfecting HEK 293T cells with plasmids encoding wild type or mutant catalytically inactive FluPol$_A$ and NP, then infecting 48 hours later with influenza A/WSN/33 virus in the presence of actinomycin D. Mock samples were transfected with empty pcDNA3A vector. Viral RNA levels were analysed 6 hours post-transfection by primer extension. B, Primary transcription assays were performed by transfecting HEK 293T cells with plasmids encoding wild type or mutant catalytically inactive FluPol$_A$, then infecting 48 hours later with influenza A/WSN/33 virus in the presence of cycloheximide. Viral RNA levels were analysed 4 hours post-transfection by primer extension. Quantification is from n=4-5 independent transfections, data are mean ± s.e.m., analysed by one-way ANOVA. ***P<0.001.

complex is ideally arranged to pass nascent RNA products from vRNP-associated FluPol^R to a free polymerase, FluPol^E (**Fig. 4.4A**). In addition, this hypothesis is consistent with my observation that the PB2 627 domain is important for cRNA encapsidation, as the FluPol^R-FluPol^E dimer interface involves interactions between the PB2 627 domains of both FluPol molecules. To directly implicate the FluPol^R-FluPol^E dimer in cRNA encapsidation, I created a series of alanine mutations in FluPol$_A$ at key sites across the dimer interface (**Fig. 4.4B, C, D, E**). These included FluPol$_C$ amino acid residues P3$_{R229}$, P3$_{N318}$/P3$_{Q319}$ and PB2$_{E139}$, equivalent to PA$_{K324}$, PA$_{K339}$/PA$_{Q340}$ and PB2$_{P132}$ respectively in FluPol$_A$. Early cryo-EM structures indicated that P3$_{K231}$ and P3$_{D232}$, equivalent to PA$_{H326}$ and PA$_{E327}$ in FluPol$_A$, could also be important amino acid

residues for the FluPol^{R}-FluPol^{E} dimer interface. Substitutions were therefore made at these residues along with PA$_{K324}$, although the final structures indicate that they do not form any key interactions.

I began by testing the activity of these FluPol$_{A}$ mutants in minigenome assays, and found that all FluPol^{R}-FluPol^{E} dimer interface mutations caused a significant loss of activity as measured by mRNA, cRNA and vRNA accumulation (**Fig. 4.4F**). Next, I carried out cRNA encapsidation assays as described above. Both FluPol^{R}-FluPol^{E} dimer interface mutations tested caused a significant decrease in the level of cRNA encapsidated, suggesting that the FluPol^{R}-FluPol^{E} interface is important for nascent cRNA encapsidation (**Fig. 4.4G**). Note that the PB2$_{P132A}$ mutation could not be tested in this assay, as this mutation is located on the FluPol^{R} side of the dimer interface and pre-expressed FluPol$_{A}$ in the assay is equivalent to FluPol^{E}. Next, I wanted to confirm that these mutations disrupt the FluPol^{R}-FluPol^{E} interaction. Attempts to develop an assay to directly measure the FluPol^{R}-FluPol^{E} interaction were unsuccessful; therefore, I took advantage of the fact that the ANP32A LRR domain binding site is formed by the FluPol^{R}-FluPol^{E} interaction, so ANP32A binding could be used as an indirect readout for the FluPol^{R}-FluPol^{E} interaction. I performed co-purification assays using strep-tagged chANP32A and wild type or mutant FluPol$_{A}$, and found that all FluPol^{R}-FluPol^{E} dimer interface mutations reduced chANP32A binding, with the PA$_{K339A/Q340A}$ mutation having the strongest effect (**Fig. 4.4H**).

To provide further evidence for the involvement of the FluPol^{R}-FluPol^{E} dimer in cRNA encapsidation I used a nanobody, Nb 8192, which blocks the FluPol^{R}-FluPol^{E} dimer interface (**Fig. 4.4A**). I performed cRNA encapsidation assays using pre-expressed FluPol$_{A}$ and Nb 8192, or a control nanobody (Nb 8210) which does not block the FluPol^{R}-FluPol^{E} dimer interface, and found that co-expression of Nb 8192 strongly reduced cRNA encapsidation while Nb 8210 had no significant effect (**Fig. 4.4I**). Collectively, these data strongly suggest that the FluPol^{R}-FluPol^{E} asymmetric dimer is required for cRNA encapsidation.

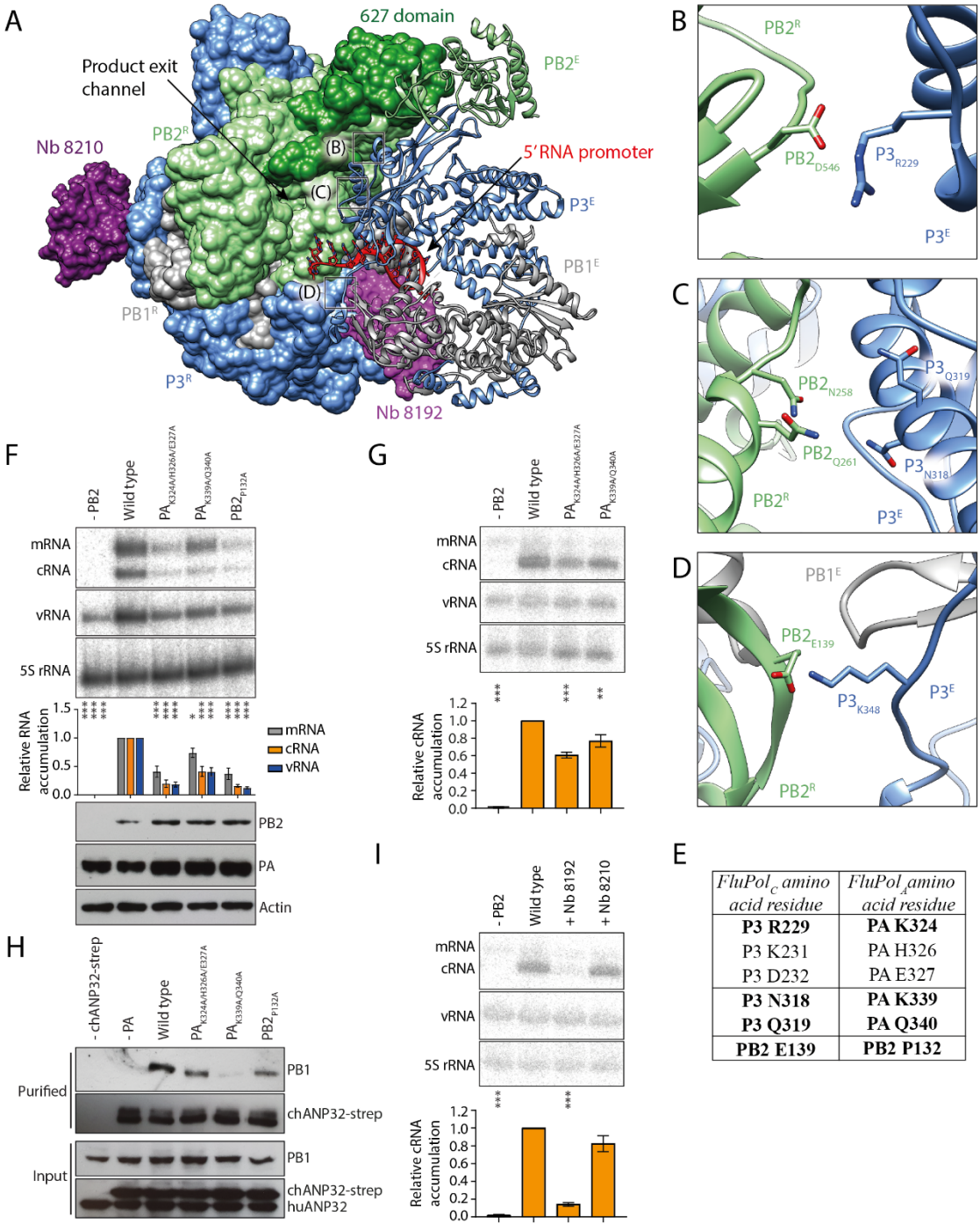
A
627 domain
Product exit channel
PB2E
PB2R
Nb 8210
(B)
5'RNA promoter
(C)
P3E
PB1E
(D)
PB1R
P3R
Nb 8192
B
PB2R
PB2D546
P3R229
P3E
C
P3Q319
PB2N258
PB2Q261
P3N318
PB2R
P3E
D
PB1E
PB2E139
P3K348
P3E
PB2R
E
FluPol_C amino acid residue
FluPol_A amino acid residue
P3 R229
PA K324
P3 K231
PA H326
P3 D232
PA E327
P3 N318
PA K339
P3 Q319
PA Q340
PB2 E139
PB2 P132
F
- PB2
Wild type
PAK324A/H326A/E327A
PAK339A/Q340A
PB2P132A
mRNA
cRNA
vRNA
5S rRNA
Relative RNA accumulation
mRNA
cRNA
vRNA
PB2
PA
Actin
G
- PB2
Wild type
PAK324A/H326A/E327A
PAK339A/Q340A
mRNA
cRNA
vRNA
5S rRNA
Relative cRNA accumulation
H
- chANP32-strep
- PA
Wild type
PAK324A/H326A/E327A
PAK339A/Q340A
PB2P132A
Purified
Input
PB1
chANP32-strep
PB1
chANP32-strep
huANP32
I
- PB2
Wild type
+ Nb 8192
+ Nb 8210
mRNA
cRNA
vRNA
5S rRNA
Relative cRNA accumulation

Figure 4.4 **The FluPol^{R}-FluPol^{E} dimer is responsible for cRNA encapsidation.** A, Structure of the FluPol$_{C}$-huANP32A complex (PDB: 6XZQ) with Nb 8192 and Nb 8210 (magenta and dark magenta respectively; unpublished data courtesy of Jeremy Keown) binding sites shown on FluPol^{R}. FluPol^{R} and FluPol^{E} are shown as surface and ribbon representations respectively. 5′ vRNA promoter (PDB: 4WSB) is modelled into the binding site on FluPol^{E}, and the nearby product exit channel of FluPol^{R} is indicated. Boxes indicate key interactions between FluPol^{R} and FluPol^{E}. B, C, D, Close-up views of key interactions between FluPol^{R} PB2 (green) and FluPol^{E} PA (blue) in the FluPol$_{C}$-huANP32A cryo-EM structure. E, Equivalent amino acid residues in FluPol$_{C}$ and FluPol$_{A}$, with key interacting residues in bold. F, Minigenome assays were performed in HEK 293T cells using plasmids encoding wild type or mutant FluPol$_{A}$, NP and vRNA. Viral RNA levels were assayed by primer extension 24 hours post-transfection (top), and PB2 or PA protein expression was analysed by western blot with actin as a loading control (bottom). G, cRNA encapsidation assays were performed using wild type or mutant catalytically inactive FluPol$_{A}$. Viral RNA levels were analysed 6 hours post-transfection by primer extension. H, Wild type or mutant FluPol$_{A}$ and chANP32A-strep were co-expressed in HEK 293T cells. chANP32A-strep was purified 48 hours post-transfection, and co-purified FluPol$_{A}$ was assayed by western blotting for PB1 protein. I, cRNA encapsidation assays were performed using wild type catalytically inactive FluPol$_{A}$ with co-expressed Nb 8192 or Nb 8210. Quantification is from n=2-3 independent transfections, data are mean ± s.e.m., analysed by one-way ANOVA. *P<0.05, **P<0.01, ***P<0.001.

4.2.3 *Interactions between influenza virus RNA polymerase molecules promote viral genome replication*

These data demonstrate that the FluPol^{R}-FluPol^{E} dimer interface is important for nascent cRNA encapsidation, however, I also wanted to determine whether this interaction directly promotes the cRNA synthesis activity of vRNPs. Therefore, I purified vRNPs from influenza A/WSN/33 virions and performed *in vitro* activity assays. vRNPs supplied with β-globin mRNA readily synthesised both mRNA and cRNA *in vitro* in the presence of NTPs, however, I found that addition of exogenously purified wild type FluPol$_{A}$ promoted cRNA accumulation in a concentration-dependent manner (**Fig. 4.5A**). Addition of FluPol$_{A}$ with the *trans*-activating dimer

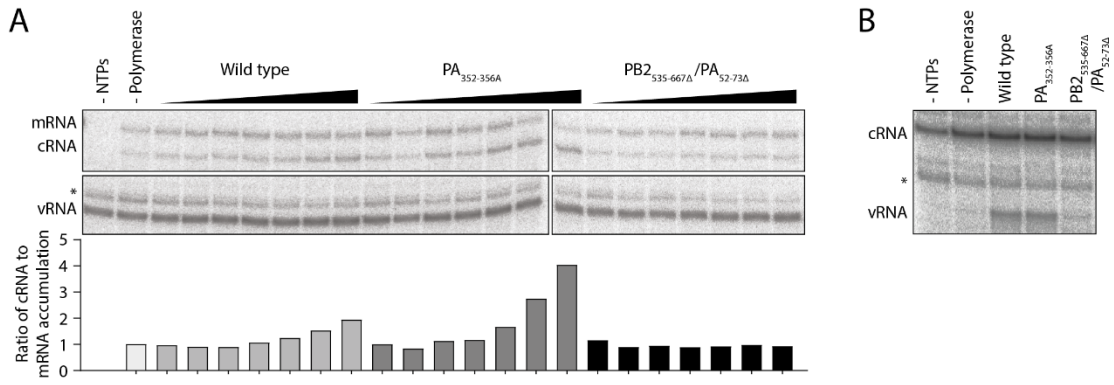

Figure 4.5 **Exogenous FluPol$_A$ enhances RNP viral genome replication activity.** A, *In vitro* activity assays were performed using vRNPs purified from influenza A virions in the presence of β-globin mRNA, with 5-100ng/µl wild type or mutant FluPol$_A$ added where indicated. Reaction products were analysed by primer extension, the asterisk denotes an artefactual product band arising from primers cross-reacting with *E. coli* tRNA. B, *In vitro* cRNP activity assays were performed in the presence of ApG dinucleotide, with 100ng/µl wild type or mutant FluPol$_A$ added where indicted. Reaction products were analysed by primer extension, the asterisk denotes an unknown artefactual band.

interface mutation PA$_{352\text{-}356A}$ caused a stronger enhancement in cRNA accumulation than the wild type, while addition of FluPol$_A$ with a PB2$_{535\text{-}667\Delta}$/PA$_{52\text{-}73\Delta}$ double mutation, which would disrupt the FluPol^R-FluPol^E dimer interface, had no effect.

To investigate whether the same interaction could be important for promoting vRNA synthesis by cRNPs, I performed *in vitro* activity assays using cRNPs purified with a PP7 bacteriophage RNA affinity tag. Purified cRNPs were only able to synthesise vRNA upon the addition of exogenous FluPol$_A$, which is consistent with previous reports (**Fig. 4.5B**) (27). PA$_{352\text{-}356A}$ mutant FluPol$_A$ was also able to activate vRNA synthesis by cRNPs, whereas PB2$_{535\text{-}667\Delta}$/PA$_{52\text{-}73\Delta}$ mutant FluPol$_A$ was not. Collectively, these data suggest that an interaction between FluPol$_A$ and RNPs promotes viral genome replication, and are consistent with the FluPol^R-FluPol^E dimer interface being responsible for this effect.

4.3 **DISCUSSION**

With this study, I aimed to use new $FluPol_C$-ANP32A cryo-EM structures to investigate the role of ANP32 proteins in $FluPol_A$ host adaption and viral genome replication. I found that amino acid residues 176-183 of huANP32A are responsible for its incompatibility with $PB2_{E627}$ $FluPol_A$, and therefore drive the $PB2_{E627K}$ host adaptive mutation. Furthermore, I found that the asymmetric $FluPol^R$-$FluPol^E$ dimer bridged by ANP32A is important for $FluPol_A$ viral genome replication, specifically for the encapsidation of nascent cRNA products.

By targeting specific interactions I demonstrate that the FluPol-huANP32A interface identified by cryo-EM is important for $FluPol_A$ activity. I mutated PA_{K413}, which is equivalent to $P3_{K391}$ and interacts with huANP32A residues N129 and D130 in the LRR domain (**Fig. 4.1C**). Previous studies have shown that the corresponding residues in chANP32A are essential for supporting $FluPol_A$ activity, which is consistent with my data (90, 97). In addition, I found that huANP32A residues 176-183 in the LCAR play an important role; specifically, these residues in huANP32A are responsible for the incompatibility with $PB2_{E627}$ $FluPol_A$ (**Fig. 4.2C**). This is a highly acidic sequence of residues in huANP32A and huANP32B, 176-EEEYDEDA-183 and 176-DEEDEDDE-183 respectively, while the corresponding sequence in chANP32A contains a more diverse mixture of residues (176-VLSLVKDR-183). Therefore, these data could be explained by the negatively charged $PB2_{E627}$ residue failing to form a productive interaction with the huANP32 LCAR due to like-charge repulsion, whereas the more neutral chANP32A LCAR interacts with either a $PB2_{E627}$ or $PB2_{K627}$ residue. This hypothesis is consistent with a previous study indicating that avian-origin $FluPol_A$ can function in human cells by acquiring one of many non-acidic residues at position $PB2_{627}$, such as $PB2_{E627V}$ (83). As avian influenza A viruses adapt to human hosts, compensatory mutations can occur at other positions; for example, the 2009 pandemic H1N1 virus underwent a $PB2_{Q591K}$ host adaptive mutation, which is equivalent to $FluPol_C$ residue $PB2_{V614}$ (85). This is located close to ANP32A residues 176-183 so could function through a similar mechanism as $PB2_{E627}$ mutations. However, the $PB2_{D701N}$ host adaptive mutation

(equivalent to $FluPol_C$ residue $PB2_{P719}$) does not interact with any structurally resolved region of ANP32A, suggesting that there are also indirect mechanisms for overcoming a non-productive interaction between $PB2_{E627}$ and the huANP32 LCAR (113).

As well as being important for host tropism, ANP32 proteins are fundamentally required for $FluPol_A$ viral genome replication. My data strongly suggest that the $FluPol^R$-$FluPol^E$ dimer bridged by ANP32A functions as an ′encapsidating′ dimer to assemble nascent cRNA products into RNPs. Most previous studies have implicated ANP32 proteins in vRNA synthesis, however, there is also evidence for a role in cRNA synthesis; specifically, avian-origin $FluPol_A$ has been shown to generate replication-incompetent cRNA in human cells, which is fully consistent with my finding that the $FluPol^R$-$FluPol^E$ dimer is important for cRNA encapsidation (87, 92, 94). While my assays specifically address cRNA encapsidation, it is likely that a similar mechanism applies for vRNA encapsidation due to intrinsic similarities between cRNA and vRNA synthesis. In particular, both processes require nascent RNA encapsidation and result in structurally similar RNPs (27, 29). My data also suggest that encapsidating dimer interaction could have a secondary role in directly promoting viral RNA synthesis, which would mean that nascent RNA products are more likely to be made when a $FluPol^E$ molecule is ready to encapsidate them (**Fig. 4.5A, B**). The current data are insufficient to speculate on a molecular mechanism for how the $FluPol^R$-$FluPol^E$ interaction could promote viral RNA synthesis, however, it is clear that formation of the encapsidating dimer requires $FluPol^R$ to be in a transcription-incompetent conformation, which potentially explains why exogenous $FluPol_A$ reduces transcription activity by vRNPs in cells and *in vitro*.

My data provide insights into $FluPol_A$ host adaption by offering a molecular explanation for the $PB2_{E627K}$ host adaptive mutation, which has implications for understanding the emergence of pandemic influenza A viruses. Furthermore, the finding that ANP32A bridges an encapsidating FluPol dimer reveals a key step in the molecular mechanism of influenza A virus RNA synthesis, and could be exploited by novel antiviral strategies targeting host or viral proteins.

4.4 **A MODEL FOR INFLUENZA VIRAL GENOME REPLICATION**

Taken together, the studies described in Chapters 2, 3 and 4 allow me to present a more complete model for the molecular mechanism of viral genome replication by influenza A virus.

vRNPs from incoming virions carry out transcription to produce viral mRNA early in the course of infection, and the accumulation of new viral proteins in the cell, including $FluPol_A$ and NP, allow viral genome replication to take place (**Fig. 4.6A, B**) (19). Viral genome replication begins when heterotrimeric $FluPol_A$ forms an asymmetric encapsidating dimer with the RNP-associated $FluPol_A$, mediated by huANP32 proteins (**Fig. 4.6C**). The RNP-associated $FluPol^R$ then initiates RNA synthesis, which may be promoted by formation of the encapsidating dimer (Chapter 4) (**Fig. 4.6D**). If vRNA synthesis by a cRNP is taking place, then another free $FluPol_A$ molecule associates to form a *trans*-activating dimer interaction with $FluPol^R$ (**Fig. 4.6E**). This interaction is structurally compatible with the encapsidating dimer and promotes prime-and-realign, an essential step in vRNA synthesis (Chapter 3) (72). Following prime-and-realign (if relevant), the RNA product is elongated and emerges from the $FluPol^R$ active site (**Fig. 4.6F**). The $FluPol^R$ product exit channel is linked to the $FluPol^E$ 5′ RNA promoter binding site by a 50Å cavity lined with basic amino acids, meaning that the nascent RNA product can be encapsidated by $FluPol^E$ with minimal exposure to the cellular environment (Chapter 4). Simultaneously, the 3′ end of the RNA template emerges from the $FluPol^R$ template exit channel, where the 3′ RNA promoter binds in the Mode B site located at the *trans*-activating dimer interface (Chapter 2) (**Fig. 4.6G**). It is unclear how or when the *trans*-activating dimer interaction is broken, although 3′ RNA promoter binding in the Mode B site could be responsible. The 3′ RNA promoter most likely remains bound in the Mode B site until RNA synthesis is completed. As elongation of the nascent RNA continues, NP is recruited to form a new RNP containing the nascent RNA and $FluPol^E$, which eventually dissociates from the parent RNP (**Fig. 4.6H, I**). Once RNA synthesis is completed, the new RNP can perform RNA synthesis using the same mechanism.

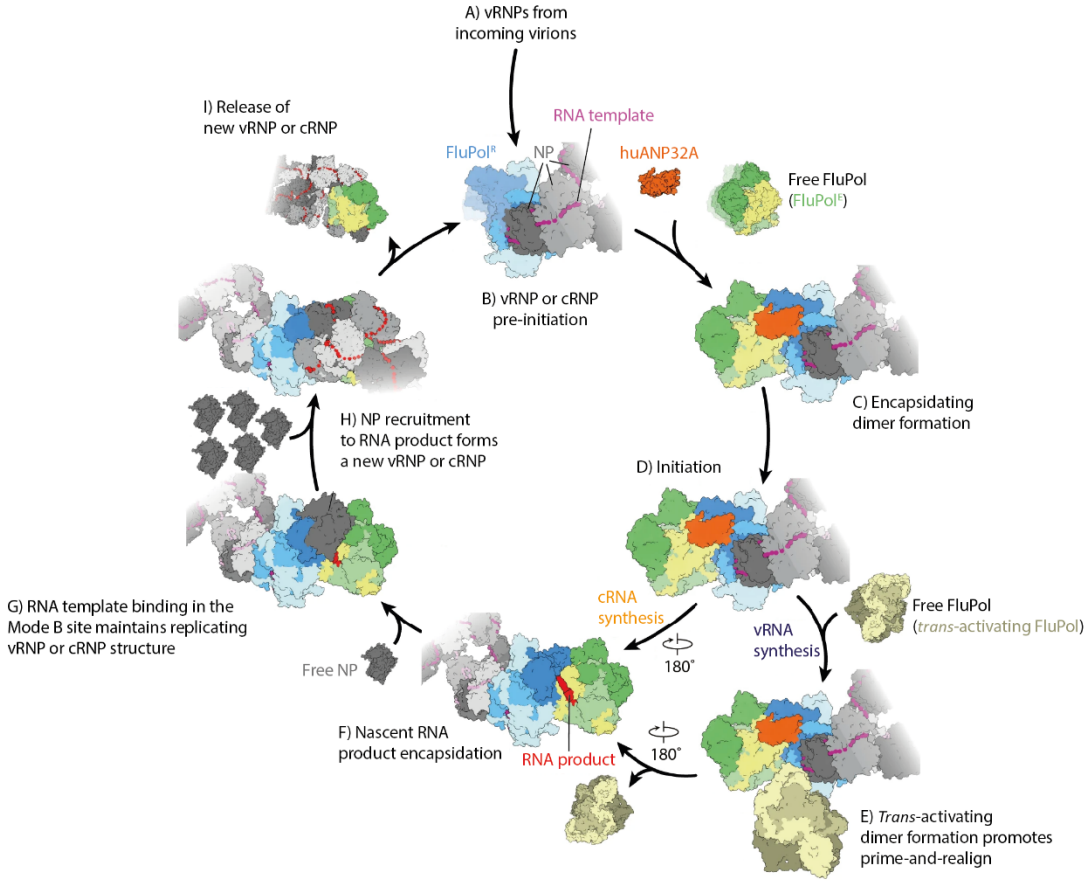

Figure 4.6 **An updated model for influenza viral genome replication.** A, vRNPs from virions synthesise mRNA early in an infection, then begin viral genome replication. B, vRNPs or cRNPs contain the replicating polymerase, FluPol^{R} (blue), along with NP (grey) and RNA (purple). C, Free FluPol molecules translated from viral mRNA act as encapsidating polymerases, FluPol^{E} (green), forming an encapsidating dimer with FluPol^{R} mediated by cellular huANP32A (orange). D, FluPol^{R} in the RNP initiates RNA synthesis. E, Free FluPol (beige) forms a *trans*-activating dimer with FluPol^{R} during vRNA synthesis, which promotes cRNA template realignment. This step is not required for cRNA synthesis. F, As the nascent RNA product (red) is elongated, it becomes encapsidated by FluPol^{E}. G, Simultaneously, the RNA template exits the FluPol^{R} active site and binds in the Mode B site. H, NP is recruited to the elongating RNA product by an unknown mechanism. I, Once RNA synthesis is completed, the newly assembled vRNP or cRNP is released. Adapted from Carrique *et al*. 2020 (111).

Central to this model is the formation of two structurally distinct $FluPol_A$ dimers. The *trans*-activating $FluPol_A$ dimer allows influenza A virus to control the rate of viral genome replication, by ensuring that vRNA synthesis cannot initiate without free $FluPol_A$ available. In addition, the encapsidating FluPol dimer prevents nascent RNA products being released into the cellular environment during viral genome replication, where they could otherwise be detected by the innate immune sensor RIG-I (102). Exploiting $FluPol_A$ dimerization allows influenza A virus to have these control measures without encoding any additional viral proteins, which would increase the size of the viral genome.

There are several key questions left unanswered by this model which could be addressed by future work. First, how NP is recruited to nascent RNA products? It is tempting to speculate that the encapsidating dimer itself forms a platform for NP assembly. The acidic LCAR of huANP32A could also aid NP recruitment by interacting with the basic RNA binding groove and acting as a molecular whip, in a mechanism analogous to the P phosphoproteins of non-segmented negative-strand RNA viruses (29). However, current data are insufficient to support or reject this hypothesis. Second, how are the encapsidating and *trans*-activating dimers regulated? Previous studies showed that excessive stabilisation of the *trans*-activating $FluPol_A$ dimer is detrimental for influenza A virus growth, suggesting that the *trans*-activating dimer should not remain associated throughout viral RNA synthesis (106). The fact that the Mode B 3′ RNA promoter binding site overlaps with the *trans*-activating dimer interface hints that RNA binding could regulate its formation, and elucidating the details of such mechanisms could provide further insights into influenza A virus RNA synthesis.

CHAPTER 5

An introduction to coronavirus RNA synthesis

5.1 OVERVIEW OF CORONAVIRUSES

5.1.1 *Background on coronaviruses*

Coronaviruses are a family of positive-strand RNA viruses in the order Nidovirales. Within the subfamily *Coronavirinae* are four genera: *Alpha*-, *Beta*-, *Gamma*- and *Deltacoronavirus*. HCoV-229E is a species of *Alphacoronavirus* which causes mild respiratory illness associated with the common cold. The genus *Betacoronavirus* includes low-pathogenic viruses such as HCoV-HKU1, as well as several notable highly pathogenic species. The highly pathogenic *Betacoronaviruses* are severe acute respiratory syndrome coronavirus (SARS-CoV), Middle East respiratory syndrome coronavirus (MERS-CoV) and SARS-CoV-2 (114, 115). SARS-CoV emerged in the human population in 2003, and had an average mortality rate of 10% across a total of 8000 infections (116). MERS-CoV emerged in 2012 and small numbers of cases are still reported as of December 2020; in this time, 2500 cases have been reported with an average mortality rate of 35% (117). SARS-CoV-2 emerged in 2019 and went on to cause the coronavirus disease 2019 (COVID-19) pandemic. The mortality rate of SARS-CoV-2 is substantially lower than SARS-CoV and MERS-CoV, however, millions have been infected partly due to the ability of the virus to transmit asymptomatically (115).

Like influenza viruses, coronaviruses have high zoonotic potential. *Alpha*- and *Betacoronaviruses* infect various mammalian species, and *Gamma*- and *Deltacoronaviruses* have a wide host range that includes avian species. SARS-CoV, MERS-CoV and SARS-CoV-2 are all derived from zoonotic coronaviruses which adapted to infect humans, with transmission to humans occurring from civet cats and camels for SARS-CoV and MERS-CoV respectively (114). The emergence

of novel coronaviruses poses an ongoing threat to public health, and therapeutic options for treating coronavirus infections are limited. This means that a more in-depth understanding of coronavirus replication is required to identify new targets for antiviral drugs.

5.1.2 *Coronavirus virion and genome structure*

Coronavirus virions are enveloped, spherical particles with an average diameter of around 90nm (85nm for SARS-CoV, 90nm for SARS-CoV-2) (118, 119). 50 to 100 spike (S) protein homotrimers, used for target cell attachment and entry, are embedded in the outer membrane of each virion along with envelope (E) and membrane (M) proteins, which are primarily required for virion assembly (**Fig. 5.1A, B**). The viral RNA genome is located inside the virion, assembled into ribonucleoproteins (RNPs) with many copies of nucleoprotein (N). SARS-CoV-2 virions contain an average of 38 RNPs, which all have a parallel pillar-like structure and are thought to be formed using one copy of the viral RNA genome, arranged along it like beads on a string (120). Viral RNA is modified with a 5′ 7-methylguanosine (m^7G) cap and 3′ poly(A) tail, which mimics host mRNA and therefore allows viral RNA to be translated by host machinery upon infection.

Coronaviruses have unusually large genomes for RNA viruses; 29.7 kilobases for SARS-CoV, 30.1 kilobases for MERS-CoV, and 29.8 kilobases for SARS-CoV-2. Two-thirds of the genome consists of two large open reading frames (ORFs) encoding two polyproteins, pp1a and pp1b, which autolyse to release the viral non-structural proteins (nsps) 1-16 (**Fig. 5.1C**). Further downstream in the viral genome are ORFs encoding the structural proteins S, E, M and N, and interspersed with these are a series of small ORFs encoding accessory proteins; in SARS-CoV these are 3a, 3b, 6, 7a, 7b, 8a, 8b and 9b (121, 122). The number of accessory proteins expressed can vary not only between coronavirus species but also between different strains of the same virus (123, 124).

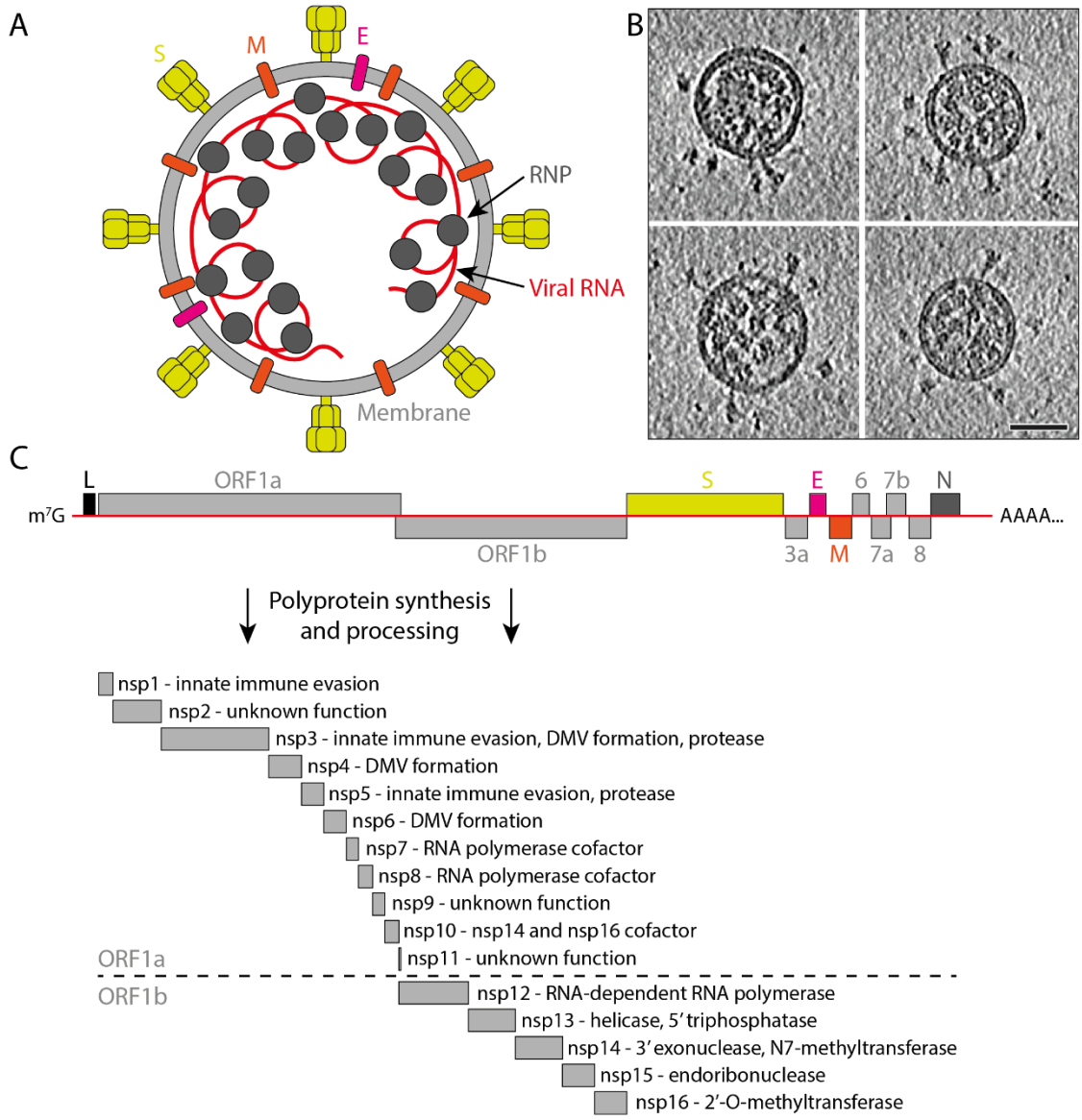

Figure 5.1 **Virion structure and viral genome architecture.** A, Diagram of a coronavirus virion showing the major structural components. B, Electron micrograph of purified SARS-CoV-2, showing heterogeneity in the numbers and conformations of S proteins. Adapted from Ze *et al*. 2020 (118). C, Schematic of the full length m^7G-capped and polyadenylated SARS-CoV-2 genome (top). The 5′ end consists of a leader sequence followed by two large ORFs encoding polyproteins. The products of polyprotein cleavage are shown with a brief description of their functions (bottom). The 3′ end of the SARS-CoV-2 genome includes ORFs encoding structural and accessory proteins, which are expressed from sgRNAs.

5.1.3 *Viral life cycle*

Attachment of a coronavirus virion to the target cell is mediated by the homotrimeric S protein, which has two functionally distinct parts, S1 and S2 (**Fig. 5.2A**). S1 contains the receptor-binding domain (RBD), which binds to the target receptor on the host cell plasma membrane. SARS-CoV and SARS-CoV-2 use angiotensin-converting enzyme 2 (ACE2) as an entry receptor, while MERS-CoV uses dipeptidyl-peptidase 4 (DPP4) (125–127). Following attachment the S protein is cleaved into S1 and S2, then further cleavage of S2 exposes the fusion peptide. This allows the membrane-proximal S2 to undergo conformational rearrangements and cause membrane fusion in a mechanism similar to HA2 of influenza virus (**Fig. 5.2B**). SARS-CoV and MERS-CoV S proteins can either be cleaved at the cell surface by the serine protease TMPRSS2, or post-endocytosis by the endosomal cysteine proteases cathepsin B or cathepsin L (128–131). SARS-CoV-2 is distinct in that it is primarily dependent on TMPRSS2, suggesting it does not use endocytosis as a major entry pathway (126, 132). SARS-CoV-2 also has a polybasic cleavage site at the S1-S2 boundary, which allows efficient cleavage by additional proteases and expands the zoonotic potential of the virus. The acquisition of a polybasic cleavage site also allows SARS-CoV-2 to use neuropilin-1 as a cofactor for attachment, which has been suggested to contribute to increased infectivity over SARS-CoV (133, 134).

S2-mediated membrane fusion releases the positive-sense viral RNA genome into the cell cytoplasm. Following release the m^7G-capped viral RNA is translated, generating pp1a and pp1b polyproteins from ORF1a and ORF1b respectively (**Fig. 5.2C**). Translation of ORF1b requires a -1 ribosomal frameshift at the end of ORF1a, which is stimulated by an RNA pseudoknot structure (135, 136). pp1a is cleaved into nsps1-11, and pp1b is cleaved into nsps1-10 and 12-16 by proteases encoded in nsp3 (PL^{pro}) and nsp5 (M^{pro}). Nsps1-16 have three major functions: firstly, nsp1 and nsp3 are involved in innate immune evasion by interfering with host functions such as mRNA translation (121, 137). Secondly, nsp3, nsp4 and nsp6 are involved in establishing cytoplasmic double-membrane vesicle (DMV) compartments which are used as sites of viral

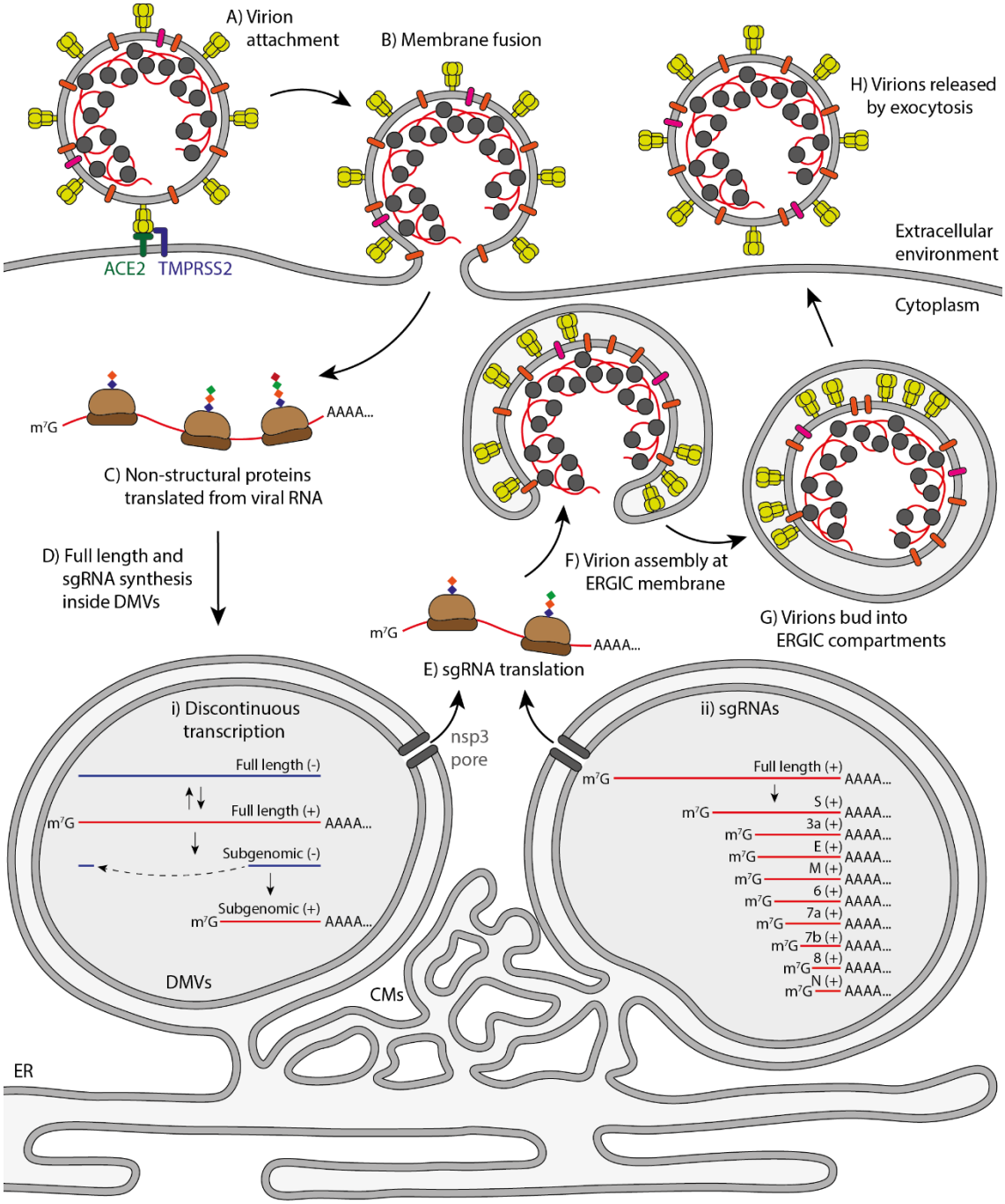
A) Virion attachment
B) Membrane fusion
H) Virions released by exocytosis
ACE2
TMPRSS2
Extracellular environment
Cytoplasm
m7G
AAAA...
C) Non-structural proteins translated from viral RNA
D) Full length and sgRNA synthesis inside DMVs
F) Virion assembly at ERGIC membrane
G) Virions bud into ERGIC compartments
E) sgRNA translation
i) Discontinuous transcription
Full length (-)
Full length (+)
Subgenomic (-)
Subgenomic (+)
DMVs
nsp3 pore
CMs
ii) sgRNAs
Full length (+)
S (+)
3a (+)
E (+)
M (+)
6 (+)
7a (+)
7b (+)
8 (+)
N (+)
ER

Figure 5.2 **Viral life cycle.** A, Coronavirus virions attach to their cell surface receptor and the S protein is cleaved by a host protease. In SARS-CoV-2 ACE2 and TMPRSS2 are the receptor and protease, respectively. B, S protein cleavage triggers membrane fusion, releasing the viral genome into the cytoplasm. C, The full-length viral genome is translated by host ribosomes to generate two polyproteins, which are autolysed to generate the viral non-structural proteins. D, The nsps trigger endoplasmic reticulum remodelling and assemble into the viral RTC, which performs viral RNA synthesis inside DMVs. The RTC synthesises new full-length viral genomes, or performs discontinuous transcription to generate the viral sgRNAs (i). SARS-CoV-2 encodes 9 sgRNAs, which contain the ORFs for structural and accessory proteins (ii). E, Viral RNAs are thought to exit DMVs through pores formed by nsp3, and are then translated. F, Virions assemble at the ERGIC membrane through the interactions of structural proteins with N protein-coated viral genomic RNA. G, Virions bud into the ERGIC, and structural proteins redistribute evenly around the virion membrane following budding. H, Exocytosis releases virions into the extracellular environment.

RNA synthesis (138, 139). Finally, most nsps are components of the replication-transcription complex (RTC) which performs viral RNA synthesis (140).

The RTC assembles in DMVs, where it uses positive-sense viral RNA genomes as templates to synthesise negative-sense RNA antigenomes (**Fig. 5.2D**). These are then used as templates for synthesising new positive-sense viral genomes, which are translated or assembled into new virions. In addition, the coronavirus RTC synthesises a number of subgenomic RNAs (sgRNAs) (122). Negative-sense sgRNAs are the templates for sgRNA synthesis, and these are made by discontinuous transcription of positive-sense viral RNA genomes. During transcription the RTC encounters transcription regulatory sequences in the viral RNA (TRS-B), which are located upstream of each structural and accessory protein ORF and have the consensus sequence 5′-ACGAAC-3′ in SARS-CoV and SARS-CoV-2. At the TRS-B the nascent RNA product dissociates from the template, and the RTC translocates to a transcription regulatory sequence 70 nucleotides away from the 5′ end of the template (TRS-L). The RTC re-initiates at the TRS-L, producing a negative-sense sgRNA with a large deletion and a shared ′leader′ sequence at the 5′

end. Transcription of the various negative-sense sgRNAs produces positive-sense sgRNAs encoding structural and accessory proteins (**Fig. 5.2E**) (141–143).

The roles of coronavirus accessory proteins are poorly understood, and expression of these proteins varies widely between virus species. In general they are thought to be involved in immune evasion functions, as shown for SARS-CoV proteins 3b, 6, and 9b (144–146). Structural proteins are more conserved between virus species, and they all have roles in virus assembly. Coronavirus virions assemble inside the cell at the cytosolic side of endoplasmic reticulum-golgi intermediate compartments (ERGICs), which contain newly expressed S, E, and M proteins (**Fig. 5.2F**). M protein recruits RNP-associated N protein to these membranes and virions bud into the ERGIC lumen, a process which is dependent on the presence of E protein for SARS-CoV (**Fig. 5.2G**) (120, 147, 148). After budding the virions are released into the extracellular environment by exocytosis (**Fig. 5.2H**).

5.2 **CORONAVIRUS REPLICATION COMPARTMENTS**

The replication complexes of positive-strand RNA viruses are generally associated with modified intracellular membranes, and coronaviruses are no exception. DMVs were the first type of membranous structure associated with SARS-CoV infection, and further studies identified convoluted membranes (CMs) as a second distinct type of structure (**Fig. 5.3A, B**) (138, 149, 150). DMVs and CMs are highly dynamic and become increasingly abundant through the course of an infection. At late time points structures similar to DMVs called vesicle packets (VPs) can be observed, which have a single outer membrane but contain multiple inner vesicles. The outer membranes of DMVs, CMs and VPs are always connected with the rough endoplasmic reticulum (ER), forming an elaborate network within the cytoplasm of an infected cell (138). It is unclear how these structures are formed, only that the transmembrane proteins nsp3, nsp4 and nsp6 are involved (151).

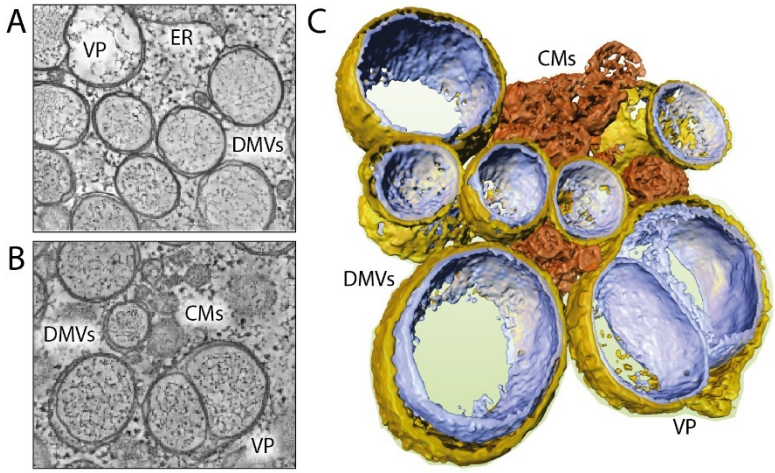

Figure 5.3 **Structure of coronavirus replication compartments.** A, B, Electron micrographs of CMs, DMVs and VPs originating from the ER membrane during a SARS-CoV infection. C, 3-dimensional reconstruction of CMs, DMVs and VPs from micrographs in (B). Adapted from Knoops *et al*. 2008 (138).

Double-stranded RNA (dsRNA) can be detected inside DMVs, so their formation is thought to be induced by coronaviruses to help avoid detection of viral RNA synthesis intermediates (**Fig. 5.3A**) (120, 138). However, the exact site of viral RNA synthesis is still unknown, since further studies found that the RTC component nsp8 was distributed mostly in CMs rather than DMVs (138). Until recently it was unclear how newly-synthesised viral RNA could be released from DMVs to reach the sites of translation and virus budding. This issue was resolved by a cryo-electron tomography study on murine hepatitis virus (MHV) infected cells which identified a pore complex composed of nsp3, presumably used for viral RNA export, spanning the double membrane of DMVs (152).

5.3 **THE CORONAVIRUS REPLICATION-TRANSCRIPTION COMPLEX**

5.3.1 *RNA polymerase structure and function*

Coronaviruses require RTCs with a multitude of distinct catalytic activities to synthesise their unusually large RNA genomes. The core of the coronavirus RTC is the RNA-dependent RNA polymerase (RdRp), located in nsp12 (**Fig. 5.4A**) (140). The nsp12 RdRp domain is a canonical right-handed fold with palm, thumb and fingers subdomains. Motifs A-F line the active site, with motif C amino acid residues D760 and D761 coordinating catalytic magnesium ions in the active site. At the nsp12 N-terminus, amino acid residues 1-250 encode the nidovirus RdRp-associated nucleotidyltransferase (NiRAN) domain which packs against the palm subdomain (153). The NiRAN domain is linked to the RdRp through a 115-amino acid residue interface domain (154, 155).

Recombinant nsp12 has limited RdRp activity *in vitro*, and this is strongly enhanced by nsp7 and nsp8 which promote RNA binding (156, 157). Initial cryo-EM structures of the SARS-CoV and SARS-CoV-2 nsp7/8/12 complexes revealed that two copies of nsp8 interact with nsp12. One nsp8 molecule interacts directly with the nsp12 fingers and interface, while the interaction of the other nsp8 is mediated entirely by nsp7, which interacts with the fingers. These structures were able to show that nsp8 uses its C-terminal head domain to interact with nsp12, but the N-termini of both nsp8 molecules (amino acid residues 1-77) could not be resolved (154, 158). Later structures of the nsp7/8/12 complex in the presence of RNA were able to resolve these regions, and showed that the N-termini of nsp8 are elongated helices (**Fig. 5.4B**). These helices form a scaffold around the RNA template-product duplex emerging from the nsp12 active site, and make a number of non-specific interactions with the negatively-charged RNA backbone using basic amino acids (159–161).

Current structures of the nsp7/8/12 complex are thought to mimic the elongation phase of RNA synthesis rather than initiation. Early biochemical studies on nsp12 concluded that RNA

polymerase activity is exclusively primer-dependent, and RNA primers are synthesised by another enzyme (157). This hypothesis would be consistent with several studies proposing that nsp8 is a secondary RNA polymerase capable of *de novo*-initiated RNA synthesis, however, the validity of this finding is disputed by structural studies which show that nsp8 does not have any hallmarks of known RNA polymerases (162–165). Furthermore, later biochemical characterisation of the nsp7/8/12 complex found that nsp12 is in fact capable of *de novo* initiation when in complex with nsp7 and nsp8 (156). Unfortunately, structures of the nsp7/8/12 complex do not corroborate this model either since nsp12 does not have a priming loop, which is generally required for *de novo* initiation to act as a platform for initiating NTPs (154, 155). Therefore, the mechanism of coronavirus RNA synthesis initiation is still an open question.

5.3.2 *RNA helicase structure and function*

One of the enzymes required to support coronavirus RNA synthesis is an RNA helicase, located in nsp13 (166). Nsp13 has a nidovirus-specific zinc binding domain (ZBD) from amino acid residues 1-100, linked to two tandem RecA domains through a stalk domain (**Fig. 5.4C**). Tandem RecA domains are a hallmark of helicases belonging to the SF1-SF2 helicase superfamilies, and further sequence and structural analysis showed that nsp13 is closely related to Upf1 helicases (167). Structures of nsp13 from SARS-CoV, MERS-CoV, and SARS-CoV-2 show that the NTP binding pocket is located between the two RecA domains, as is typical in the SF1-SF2 superfamilies (168–170). Helicases require NTP hydrolysis to function, and biochemical analyses demonstrate that SARS-CoV nsp13 helicase activity is relatively non-specific, as it readily hydrolyses all natural NTPs and dNTPs to unwind dsRNA or dsDNA in a 5′ to 3′ direction (171).

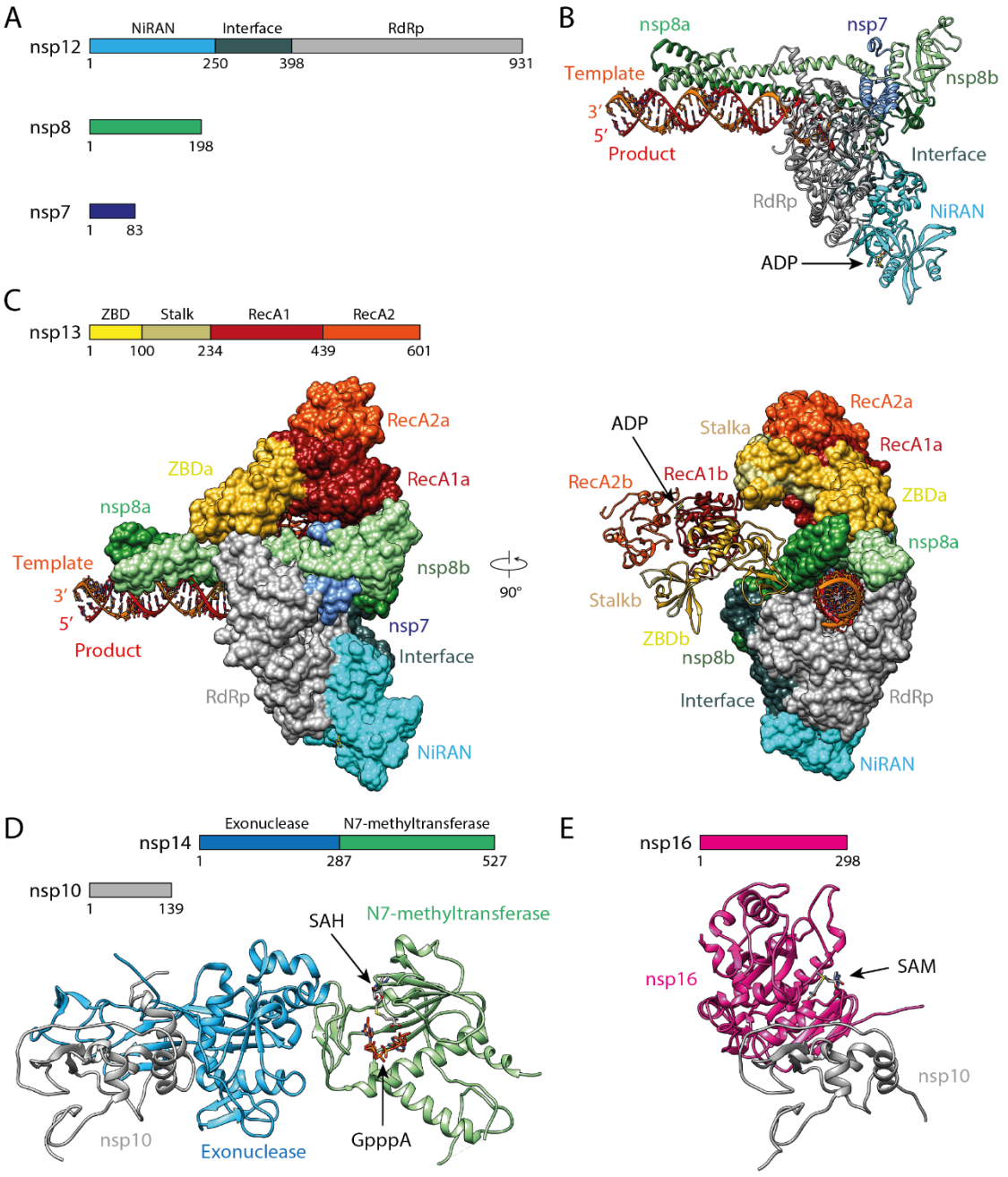
A
NiRAN
Interface
RdRp
nsp12
1
250
398
931
nsp8
1
198
nsp7
1
83
B
nsp8a
nsp7
Template
nsp8b
3′
5′
Product
Interface
RdRp
NiRAN
ADP
C
ZBD
Stalk
RecA1
RecA2
nsp13
1
100
234
439
601
RecA2a
ZBDa
RecA1a
nsp8a
Template
nsp8b
90°
3′
5′
Product
nsp7
Interface
RdRp
NiRAN
ADP
Stalka
RecA2a
RecA1a
RecA2b
RecA1b
ZBDa
nsp8a
Stalkb
ZBDb
nsp8b
Interface
RdRp
NiRAN
D
Exonuclease
N7-methyltransferase
nsp14
1
287
527
nsp10
1
139
SAH
N7-methyltransferase
nsp10
Exonuclease
GpppA
E
nsp16
1
298
nsp16
SAM
nsp10

Figure 5.4 **Coronavirus RNA polymerase structure and cofactors.** A, Schematic of SARS-CoV-2 nsp12, nsp8 and nsp7 domain structures. B, Ribbon diagram of the nsp7/8/12 RNA polymerase complex from SARS-CoV-2 bound to an RNA template-product duplex (PDB: 6XEZ). In this structure ADP was found to bind to the nsp12 NiRAN domain. C, Schematic of the SARS-CoV-2 nsp13 domain structure (top). Structure of two nsp13 molecules (each bound to ADP) in complex with SARS-CoV-2 nsp7/8/12 (bottom; PDB: 6XEZ). D, Schematic of nsp10 and nsp14 domain structures (top), with the structure of the nsp10/14 complex from SARS-CoV bound to a GpppA substrate shown below (PDB: 5C8S). E, Schematic of nsp16 (top) with the structure of the nsp10/16 complex from SARS-CoV-2 shown below (PDB: 7C2I).

Biochemical experiments show that SARS-CoV nsp13 interacts directly with nsp12 and this interaction promotes helicase activity (168, 172). A cryo-EM structure of SARS-CoV-2 nsp13 bound to the nsp7/8/12 complex offers some insight into this interaction, showing that the nsp13 ZBD interacts with the nsp12 thumb subdomain, and this interaction is further stabilised by contacts between the nsp13 ZBD and the N-terminal helix of one nsp8 molecule (**Fig. 5.4C**). The ZBD of a second nsp13 molecule simultaneously interacts with the nsp7/8/12 complex, exclusively contacting the N-terminal helix of the other nsp8 molecule. The orientation of the nsp13 molecules suggests that their function in this complex is to denature the RNA template before it enters the active site, either separating dsRNA or melting structures in single-stranded RNA (ssRNA) (169). Unlike the influenza virus RNA polymerase, the coronavirus nsp7/8/12 complex does not intrinsically separate the product and template RNA strands as they emerge from the active site. The current structure of SARS-CoV-2 nsp13 bound to the nsp7/8/12 complex does not suggest that it is involved in separating the product and template RNA strands, so the mechanism behind this process remains unknown.

5.3.3 *3′-5′ exonuclease structure and function*

SARS-CoV nsp14 was hypothesised to be an exonuclease based on sequence homology with DEDD superfamily nucleases, and 3′-5′ exonuclease activity was later confirmed experimentally (173, 174). Addition of nsp10 was found to promote exonuclease activity, and a crystal structure of the SARS-CoV nsp10/14 complex demonstrated that this occurs through a direct interaction with the N-terminal exonuclease domain of nsp14 (**Fig. 5.4D**) (175, 176). Several studies support the suggestion that the nsp10/14 complex is a proofreading exonuclease, as mutant viruses lacking nsp14 exonuclease activity demonstrate lower replication fidelity or increased susceptibility to lethal mutagenesis (177, 178). Furthermore, the nsp10/14 complex can remove mismatched bases from dsRNA *in vitro*; although paradoxically, its 3′-5′ exonuclease function is more potent on perfectly base-paired dsRNA (175).

Coronaviruses have particularly large genomes compared to other RNA viruses, ranging from 29-30 kilobases (140). Small genomes are normally necessary for RNA viruses, as RdRp enzymes are naturally error-prone so can cause lethal mutagenesis on large genomes. Encoding a proofreading exonuclease helps coronaviruses to reduce their mutation rate, making large genomes possible. In addition to promoting genome stability, the proofreading function of nsp14 can confer resistance certain nucleotide analogue drugs. For example, the nucleoside analogue drug remdesivir was found to be significantly less potent in wild type MHV compared to a mutant virus lacking exonuclease activity (179).

5.3.4 *7-methylguanosine cap synthesis*

The coronavirus RTC synthesises positive-sense viral RNA with a 5′ m^7G cap, which requires many distinct catalytic functions. There are several possible m^7G cap synthesis pathways, and identification of the enzymes involved in coronavirus m^7G cap synthesis has led to the conclusion that they utilise a canonical pathway, similar to eukaryotes (180).

The substrate for this capping pathway is 5′ triphosphorylated RNA, and in the first step this is dephosphorylated by a 5′ triphosphatase. SARS-CoV nsp13 was found to have 5′ triphosphatase activity, and this reaction utilises the same NTPase active site used for helicase activity (171). The resulting 5′ diphosphorylated RNA is then linked to GTP by a guanylyltransferase enzyme, forming an unmethylated cap-like structure. Nsp12 is hypothesised to be a guanylyltransferase based on the observation that it can bind to GDP, and a recent study proposed that nsp9 acts as a regulator of guanylyltransferase activity (181). The cap-like structure then undergoes two methylation reactions: first, the guanine N7 position is methylated by nsp14, using S-adenosylmethionine (SAM) as a methyl donor and producing S-adenosylhomocysteine. The N7-methyltransferase activity of nsp14 was first identified for SARS-CoV, and is mediated by a methyltransferase domain located downstream of the exonuclease domain (**Fig. 5.4D**) (182). Nsp10 does not interact with the nsp14 methyltransferase domain, so has no effect on methyltransferase activity (176). The second methylation reaction is performed by the 2′-O-methyltransferase nsp16, which methylates the 2′ hydroxyl of the first nucleotide in the RNA. This activity was predicted from sequence analysis, and Feline coronavirus (FCoV) nsp16 was the first to have 2′-O-methyltransferase activity demonstrated experimentally (173, 183). Further structural and biochemical studies on SARS-CoV nsp16 found that nsp10 strongly stimulates activity by stabilising RNA and SAM substrate binding (**Fig. 5.4E**) (184–186). Some negative-strand RNA viruses perform m^7G capping through a non-conventional pathway by carrying out guanine-N7-methylation after 2′-O-methylation, however, SARS-CoV nsp16 is highly selective for guanine-N7-methylated substrates so coronaviruses are thought to follow a conventional pathway (186). Following 2′-O-methylation, the resulting m^7G-capped RNA is referred to as cap-1 RNA. Higher eukaryotes produce mRNA with a cap-1 structure, and have evolved innate immune sensing pathways to detect m^7G-capped RNA without 2′-O-methylation, called cap-0 RNA (187). Therefore, the 2′-O-methyltransferase function of coronavirus nsp16 is essential primarily as an innate immune evasion strategy.

5.4 **OBJECTIVES**

The novel coronavirus SARS-CoV-2 emerged during my DPhil, leading to significant morbidity, mortality, and economic strain. I therefore used my background working on influenza A virus to investigate the poorly-understood mechanism of coronavirus RNA synthesis, with the aim of potentially exposing new targets for antiviral drugs. In particular I studied the function of the SARS-CoV-2 nsp12 NiRAN domain, using biochemical approaches to identify novel catalytic activities which could be important for viral RNA synthesis (Chapter 6).

CHAPTER 6

The function of the SARS-CoV-2 RNA polymerase NiRAN domain

Data from this chapter were published in:

***Walker AP**, Fan H, Keown JR, Margitich V, Grimes JM, Fodor E, te Velthuis AJW. 2020. Enisamium is a small molecule inhibitor of the influenza A virus and SARS-CoV-2 RNA polymerases. bioRxiv 2020.04.21.053017.*

***Walker AP**, Fan H, Keown JR, Grimes JM, Fodor E. 2020. Identification of guanylyltransferase activity in the SARS-CoV-2 RNA polymerase. bioRxiv 2021.03.17.435913.*

6.1 **INTRODUCTION**

SARS-CoV-2 is a positive-sense RNA virus responsible for the coronavirus disease 2019 (COVID-19) pandemic (115). Two-thirds of the 30-kilobase viral genome encode non-structural proteins (nsps) 1-16, which include all components of the replication-transcription complex (RTC) responsible for 7-methylguanosine (m^7G)-capped viral RNA synthesis (140).

In order to synthesise an m^7G cap the RTC requires an RNA 5′ triphosphatase, N7-methyltransferase and 2′-O-methyltransferase, functions performed by nsp13, nsp14 and nsp16 respectively (171, 173, 182, 183). m^7G cap synthesis also requires a guanylyltransferase enzyme (GTase), which had not yet been identified in the coronavirus RTC when the experiments described in this chapter were performed.

Nsp12 is the viral RNA-dependent RNA polymerase (RdRp), and it must form a complex with nsp7 and nsp8 to function as a processive RNA polymerase (156). The nsp12 RdRp domain shares

structural similarity with other viral RNA polymerases, making it a promising target for repurposed antiviral drugs against SARS-CoV-2 (155). In addition to the RdRp domain, nsp12 has an N-terminal nidovirus RdRp-associated nucleotidyltransferase (NiRAN) domain which covalently binds to nucleotides (153). With this study I aimed to functionally characterise the SARS-CoV-2 RNA polymerase, and identify the role of the NiRAN domain in viral RNA synthesis.

6.2 **RESULTS**

6.2.1 *Establishment of a SARS-CoV-2 RNA polymerase activity assay*

To gain insight into the function of the SARS-CoV-2 RNA polymerase I first established an *in vitro* RNA synthesis assay, based on extension of a radiolabelled 20 nucleotide (nt) primer (LS2) along a 40nt template (LS1) (**Fig. 6.1A**). A reconstituted complex of wild type SARS-CoV-2 nsp7, nsp8 and nsp12 proteins was able to extend the LS2 primer in a time-dependent manner in the presence of RNA and nucleotides, producing a 40nt major product (**Fig. 6.1B, C**). Nsp12 amino acid residues D760 and D761 coordinate catalytic magnesium ions in the RdRp active site, and when these residues were mutated to alanine the resulting nsp7/8/12 complex was unable to extend the LS2 primer (155).

I then used this *in vitro* assay to test the efficacy of several drugs suggested to inhibit the SARS-CoV-2 RNA polymerase, which included nucleotide analogue and non-nucleotide analogue drugs (**Fig. 6.1D**). The nucleotide analogues acyclovir triphosphate and remdesivir triphosphate, active metabolites of acyclovir and remdesivir respectively, both inhibited RNA polymerase activity with IC_{50} values of 0.95mM and 1.68mM. Enisamium is an influenza virus RNA polymerase inhibitor which is thought to be metabolised *in vivo* to a more active compound, VR17-04 (188). While enisamium only inhibited SARS-CoV-2 RNA polymerase activity in the high millimolar range, the inhibitory activity of VR17-04 was comparable to the nucleotide analogue drugs tested.

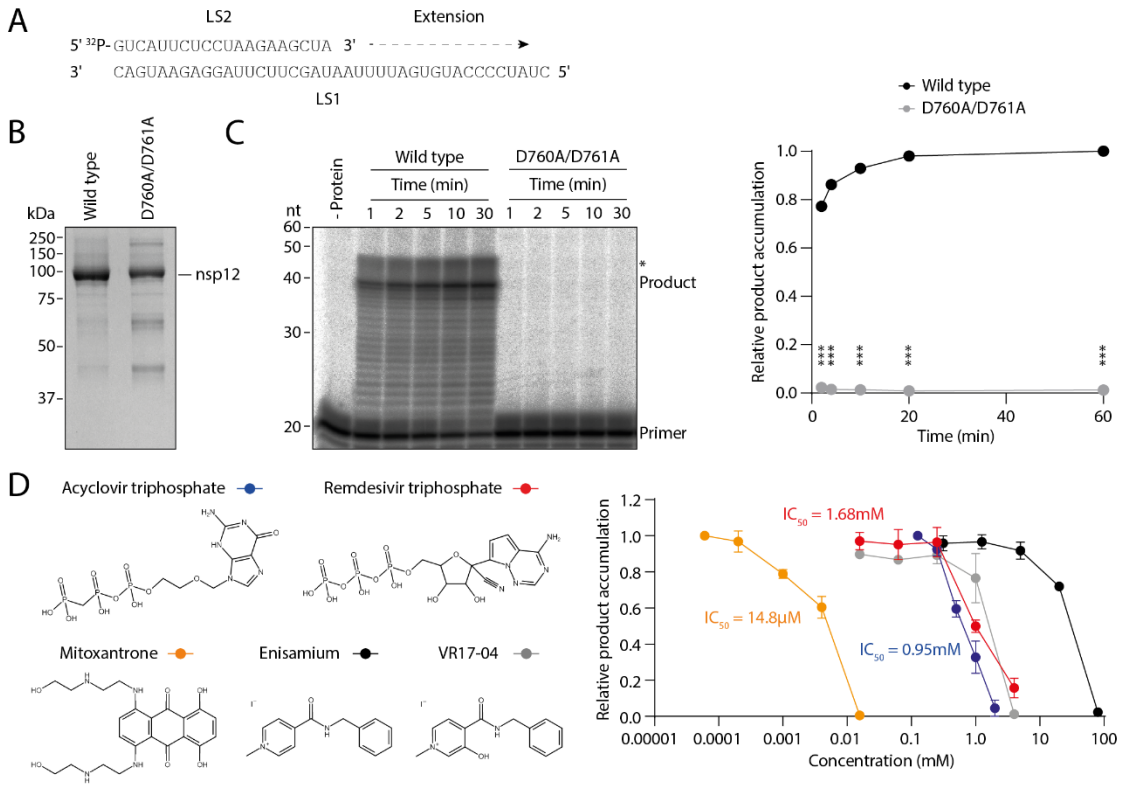

Figure 6.1 **Establishment of a SARS-CoV-2 RNA polymerase activity assay.** A, Schematic of the 40nt LS1 template and radiolabelled 20nt LS2 primer. B, SDS gel of purified wild type or D760A/D761A SARS-CoV-2 nsp12. C, SARS-CoV-2 nsp7, nsp8 and wild type or D760A/D761A nsp12 were incubated with LS1/LS2 RNA and NTPs. Reaction products were resolved by denaturing PAGE (left) and quantified (right). Size markers are indicated on the left of the gel, and the asterisk indicates a smear on the gel resulting from incomplete denaturation of radiolabelled products. Quantification is from n=3 independent reactions, error bars are too narrow to be displayed. D, Several candidate RNA polymerase inhibitor drugs (left) were tested in RNA polymerase activity assays. Inhibition curves are shown (right), and IC_{50} values are derived from fitting data to dose-response inhibition curves by nonlinear regression in GraphPad Prism 9. Note that accurate IC_{50} values could not be calculated for enisamium and VR17-04. Quantification is from n=2-3 independent reactions. Data are mean ± s.e.m., analysed by two-way ANOVA. ***P<0.001.

The most potent drug tested was mitoxantrone, a DNA intercalating agent used in cancer chemotherapy, which inhibited activity with an IC_{50} of 14.8μM. These data demonstrate that the SARS-CoV-2 RNA polymerase activity assay I have established can be used to compare the efficacy of RNA polymerase inhibitors.

6.2.2 *SARS-CoV-2 viral RNAs have a 7-methylguanosine cap*

The SARS-CoV-2 RNA polymerase is thought to synthesise m^7G-capped full-length viral RNA, but there is very little direct evidence to support this suggestion (189–191). I therefore established a reverse transcription-qPCR (RT-qPCR) protocol to characterise SARS-CoV-2 viral RNA from infected Vero CCL-81 cells, using a primer specific to the full-length or each subgenomic RNA (sgRNA) paired with a primer directed against the shared leader (L) sequence (**Fig. 6.2A**). Using this approach I was able to detect all viral RNAs except 7b sgRNA (data not shown), as well as negative-sense viral antigenomes (**Fig. 6.2B**).

As expected, all SARS-CoV-2 viral RNAs examined accumulated rapidly between 3 hours and 12 hours post-infection (**Fig. 6.2C**). To determine which viral RNAs are m^7G-capped I performed immunoprecipitation (IP) on RNA extracted from infected cells, using either an anti-m^7G antibody or an anti-his antibody as a negative control (**Fig. 6.2D, E**). Full-length viral RNA was significantly enriched by the anti-m^7G IP compared to 5S rRNA, which is consistent with the presence of an m^7G cap. All sgRNAs followed a similar pattern to the full-length viral RNA, although enrichment was only statistically significant for 3a, 6 and 8. Interestingly, negative-sense antigenomes were also significantly enriched by the anti-m^7G IP. These data suggest that all SARS-CoV-2 viral RNAs, including sgRNAs and antigenomes, are m^7G-capped.

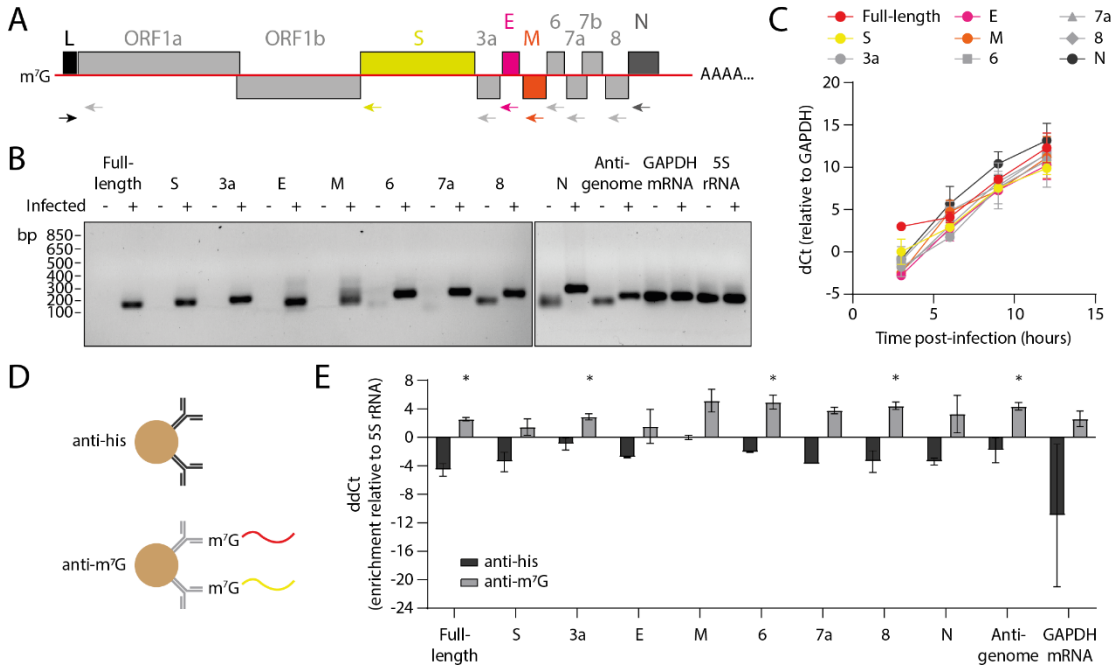

Figure 6.2 **Identification of m^{7}G-capped SARS-CoV-2 viral RNAs.** A, Schematic of the full-length SARS-CoV-2 genome with primer binding sites shown as arrows. B, RT-qPCR was performed on total cellular RNA from SARS-CoV-2 infected or mock infected Vero CCL-81 cells. qPCR reaction products were resolved by agarose gel electrophoresis, expected products range from 100-200bp. C, SARS-CoV-2 full-length and sgRNA accumulation was measured at different time points post-infection by RT-qPCR. RNA levels were quantified for n=2 independent infections, data are mean ± s.e.m. D, Schematic of RNA IP using anti-m^{7}G or anti-his antibody bound to magnetic beads. E, m^{7}G-capped RNA was immunoprecipitated from total cellular RNA of SARS-CoV-2 infected cells. RNA levels presented are relative to 5S rRNA, with the input sample set to 0. RNA levels were quantified for n=1-2 independent infections, data are mean ± s.e.m., analysed by two-tailed unpaired t-test. *P<0.05. In all cases, n=2 qPCR reactions were performed for each sample and averaged to generate the values shown.

6.2.3 *The NiRAN domain is an enzyme involved in cap synthesis*

SARS-CoV-2 is thought to utilise a canonical pathway for m^{7}G cap synthesis, which requires an unidentified GTase enzyme to covalently link GTP to the 5′ end of diphosphorylated RNA. This reaction involves a nucleotidylated enzyme (Gp-E) intermediate and produces a cap-like structure (GpppN-RNA), which is subsequently methylated (**Fig. 6.3A**) (140, 180). The NiRAN domain of

the equine arteritis virus (EAV) RNA polymerase has been shown to nucleotidylate with GTP and UTP, so I hypothesised that the NiRAN domain in the SARS-CoV-2 RNA polymerase may be responsible for performing the GTase reaction (153).

Nsp13 is thought to function directly upstream of the putative GTase, generating a diphosphorylated RNA substrate using its 5′ triphosphatase activity (171). I therefore confirmed the 5′ triphosphatase activity of purified SARS-CoV-2 nsp13 on a γ-^{32}P-ATP substrate *in vitro*, then used it to generate a diphosphorylated 20nt RNA substrate for GTase reactions (**Fig. 6.3B, C**). I incubated the diphosphorylated RNA substrate with α-^{32}P-GTP, nsp7, nsp8 and nsp12, separated the reaction products by denaturing PAGE, and observed a radiolabelled product in the presence of nsp12 which ran slightly slower than the 20nt marker and accumulated over time (**Fig. 6.3D, E**). I also performed reactions under the same conditions with vaccinia capping enzyme, a known GTase, which synthesised a radiolabelled product with the same mobility as the product made by nsp12 (**Fig. 6.3E**) (192). Since the EAV RNA polymerase NiRAN domain can nucleotidylate with GTP or UTP, I also performed reactions in the presence of α-^{32}P-UTP (**Fig. 6.3F**) (153). Under these conditions nsp12 did not make any radiolabelled product using α-^{32}P-UTP.

To confirm that the radiolabelled product made by nsp12 resulted from formation of a GpppN-RNA structure on the 20nt RNA, I performed a series of enzymatic digestions. First, I treated the diphosphorylated RNA substrate with alkaline phosphatase (AP) or RNA 5′ pyrophosphohydrolase (RppH) to produce dephosphorylated or monophosphorylated RNA respectively (**Fig. 6.3G**). Nsp12 and vaccinia capping enzyme were both unable to efficiently synthesise a radiolabelled product using either of these substrates, indicating that a diphosphorylated substrate is essential. Next, I treated the reaction products with AP or RppH (**Fig. 6.3H**). AP had no effect on the products made by nsp12 and vaccinia capping enzyme, indicating no exposed 5′ phosphates were present. However, RppH was able to degrade the products, indicating the presence of 5′ triphosphates (193). This enzymatic profile is characteristic

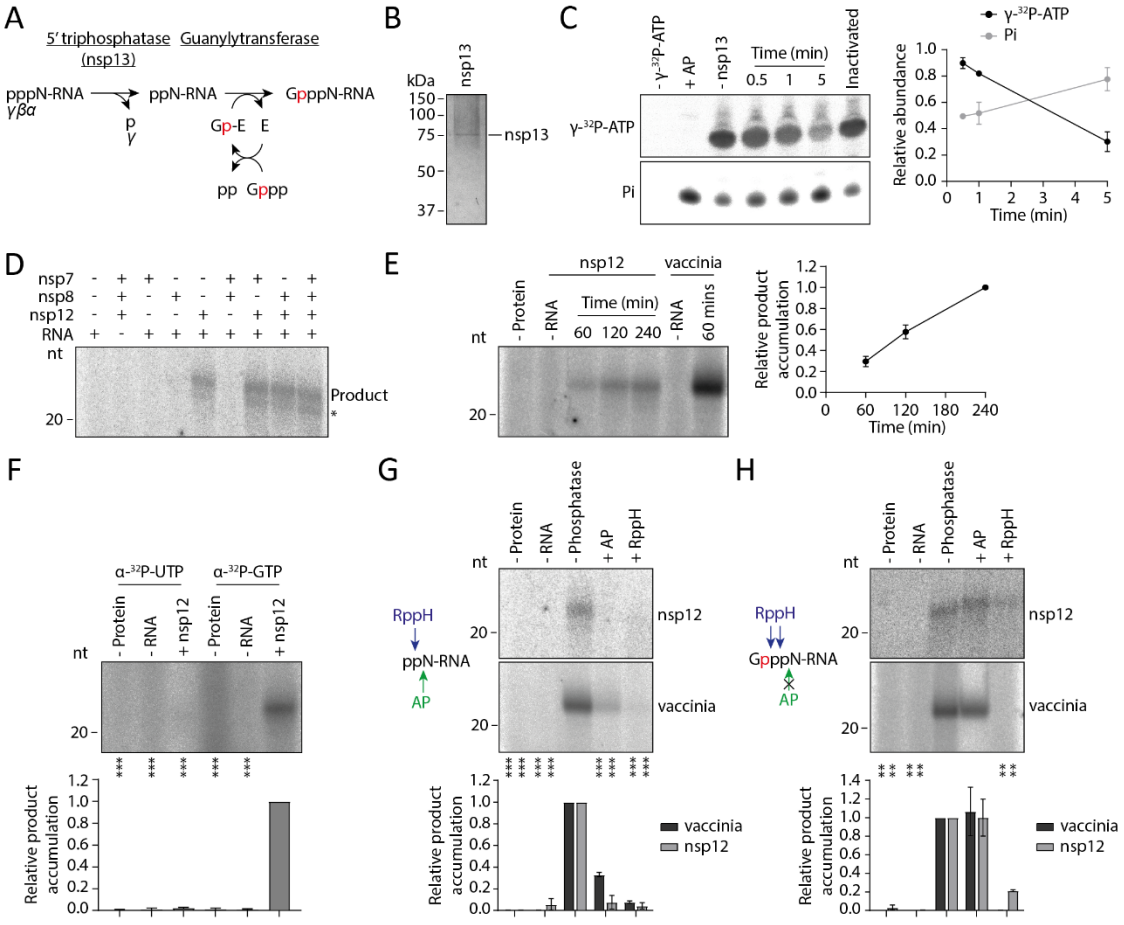
A
5′ triphosphatase (nsp13)
Guanylytransferase
pppN-RNA
γβα
ppN-RNA
GpppN-RNA
P
γ
Gp-E
E
pp Gppp
B
kDa
150
100
75
50
37
nsp13
nsp13
C
- γ-32P-ATP
+ AP
- nsp13
Time (min)
0.5 1 5
Inactivated
γ-32P-ATP
Pi
γ-32P-ATP
Pi
Relative abundance
Time (min)
D
nsp7
nsp8
nsp12
RNA
nt
20
Product
*
E
nsp12
vaccinia
- Protein
- RNA
Time (min)
60 120 240
- RNA
60 mins
nt
20
Relative product accumulation
Time (min)
F
α-32P-UTP
α-32P-GTP
- Protein
- RNA
+ nsp12
- Protein
- RNA
+ nsp12
nt
20
***
Relative product accumulation
G
- Protein
- RNA
- Phosphatase
+ AP
+ RppH
nt
RppH
ppN-RNA
AP
20
nsp12
vaccinia
***
Relative product accumulation
vaccinia
nsp12
H
- Protein
- RNA
- Phosphatase
+ AP
+ RppH
nt
RppH
GpppN-RNA
AP
20
nsp12
vaccinia
**
Relative product accumulation
vaccinia
nsp12

Figure 6.3 **SARS-CoV-2 nsp12 has guanylyltransferase activity.** A, Schematic of nsp13 5′ triphosphatase activity and the putative guanylyltransferase (GTase) activity. The GTase enzyme (E) becomes nucleotidylated and transfers GMP to the substrate. The location of radiolabelled phosphate in the α-^{32}P-GTP substrate is highlighted in red. B, SDS gel of purified His6-Zbasic-tagged SARS-CoV-2 nsp13. C, Purified nsp13 was incubated with γ-^{32}P-ATP, then the γ-^{32}P-ATP substrate and inorganic phosphate (Pi) product was resolved by denaturing PAGE (left) and quantified (right). Inactivated nsp13 was heated to 70°C for 5 minutes prior to the reaction, and AP was used as a positive control. D, Nsp7, nsp8 and nsp12 were incubated with diphosphorylated 20nt RNA and α-^{32}P-GTP, then reaction products were resolved by denaturing PAGE. Size markers are indicated on the left of the gel, the asterisk denotes a faster-moving product which could result from kinase contamination or decapping of the product of interest. E, Nsp12 or vaccinia capping enzyme was incubated with diphosphorylated 20nt RNA and α-^{32}P-GTP, then reaction products were resolved by denaturing PAGE (left) and quantified (right). F, Nsp12 was incubated with diphosphorylated 20nt RNA and α-^{32}P-UTP or α-^{32}P-GTP, then reaction products were resolved by denaturing PAGE (top) and quantified (bottom). G, Diphosphorylated RNA substrates were treated with alkaline phosphatase (AP) or RNA 5′ pyrophosphohydrolase (RppH) (left), then used in GTase reactions with nsp12 or vaccinia capping enzyme (right). H, Nsp12 or vaccinia capping enzyme GTase reactions were run and then products were treated with AP or RppH (left) before resolving by denaturing PAGE (right). Quantification is from n=2-3 independent reactions, data are mean ± s.e.m., analysed by one-way ANOVA. **P<0.01, ***P<0.001.

of a GpppN-RNA product, and collectively these results show that SARS-CoV-2 nsp12 has GTase activity *in vitro*.

To determine whether the NiRAN domain in nsp12 is responsible for the GTase activity, I designed point mutations at amino acid residues R116, D126 and D218 in the nucleotide binding pocket (**Fig. 6.4A, B, C**) (153, 181). None of these mutations reduced RNA polymerase activity in the presence of nsp7 and nsp8, however, all caused a significant loss of GTase activity (**Fig. 6.4D, E**). The D218A mutation had the most potent effect on GTase activity, which is consistent with structural data showing that D218 plays a key role in nucleotide binding by coordinating a

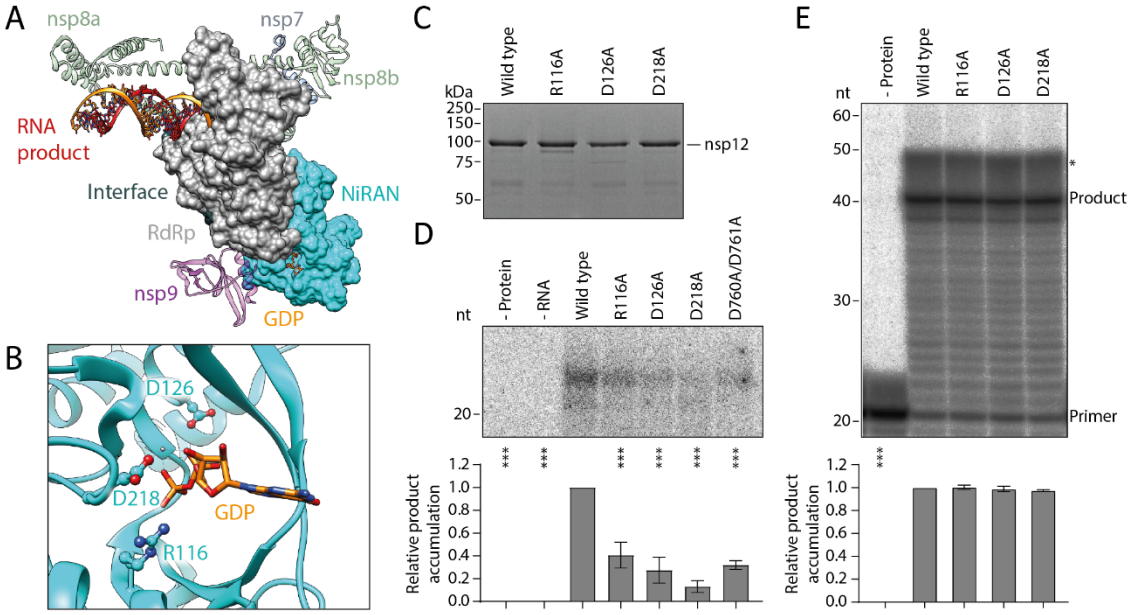

Figure 6.4 **Nsp12 NiRAN domain mutations disrupt guanylyltransferase activity.** A, Structure of the SARS-CoV-2 nsp7/8/12 complex bound to nsp9 (magenta) and RNA product (red), with GDP (orange) bound in the nsp12 NiRAN domain (cyan) (PDB: 7CYQ). B, Close-up view of the nsp12 NiRAN domain nucleotide binding pocket with GDP bound (PDB: 7CYQ). C, SDS gel of purified nsp12 mutants. D, GTase activity assays were performed using mutant nsp12 proteins. E, RNA polymerase activity assays were performed using mutant nsp12 proteins in the presence of nsp7 and nsp8, the asterisk indicates a smear on the gel resulting from incomplete denaturation of radiolabelled products. Quantification is from n=3 independent reactions, data are mean ± s.e.m., analysed by one-way ANOVA. ***P<0.001.

magnesium ion (**Fig. 6.4B**) (169, 181). Interestingly, the D760A/D761A RdRp active site mutation also significantly reduced GTase activity to approximately 40% of wild type nsp12.

To further confirm that the NiRAN domain has GTase activity, I performed reactions using either full-length nsp12 or the NiRAN domain alone. The isolated NiRAN domain was able to synthesise a radiolabelled product with the same mobility as that of full-length nsp12, but with approximately 30% of the efficiency (**Fig. 6.5A**). To confirm that this was the product of GTase activity I then performed enzymatic digestions of the substrate and product RNAs. The NiRAN domain did not make any radiolabelled product when provided with dephosphorylated or

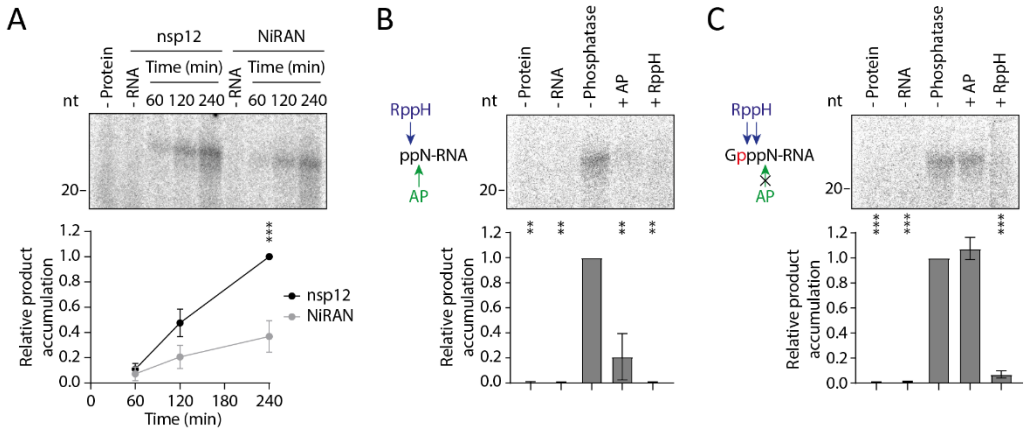

Figure 6.5 **The nsp12 NiRAN domain is an independent guanylyltransferase enzyme**. A, GTase activity assays were performed using full-length nsp12 or NiRAN domain (top) and products were quantified (bottom). Quantification is from n=3 independent reactions, data are mean ± s.e.m., analysed by two-way ANOVA. ***P<0.001. B, Diphosphorylated RNA substrates were treated with AP or RppH, then used in GTase reactions with the isolated NiRAN domain. C, NiRAN domain GTase reactions were run and then products were treated with AP or RppH. Quantification is from n=2 independent reactions, data are mean ± s.e.m., analysed by one-way ANOVA. **P<0.01, ***P<0.001.

monophosphorylated substrate RNA, produced by pre-treatment with AP or RppH respectively (**Fig. 6.5B**). In addition, the radiolabelled product was sensitive to degradation by RppH but not AP, which matches the results obtained using full-length nsp12 and vaccinia capping enzyme (**Fig. 6.5C**). Together, these data show that the nsp12 NiRAN domain can function as a GTase enzyme *in vitro*.

6.2.4 *The NiRAN domain has ATP-dependent RNA ligase activity*

While establishing the RNA polymerase activity assay I also tested nsp12 activity in the absence of nsp7 and nsp8, as previous reports suggested that nsp12 displays limited RNA polymerase activity under these conditions (157).

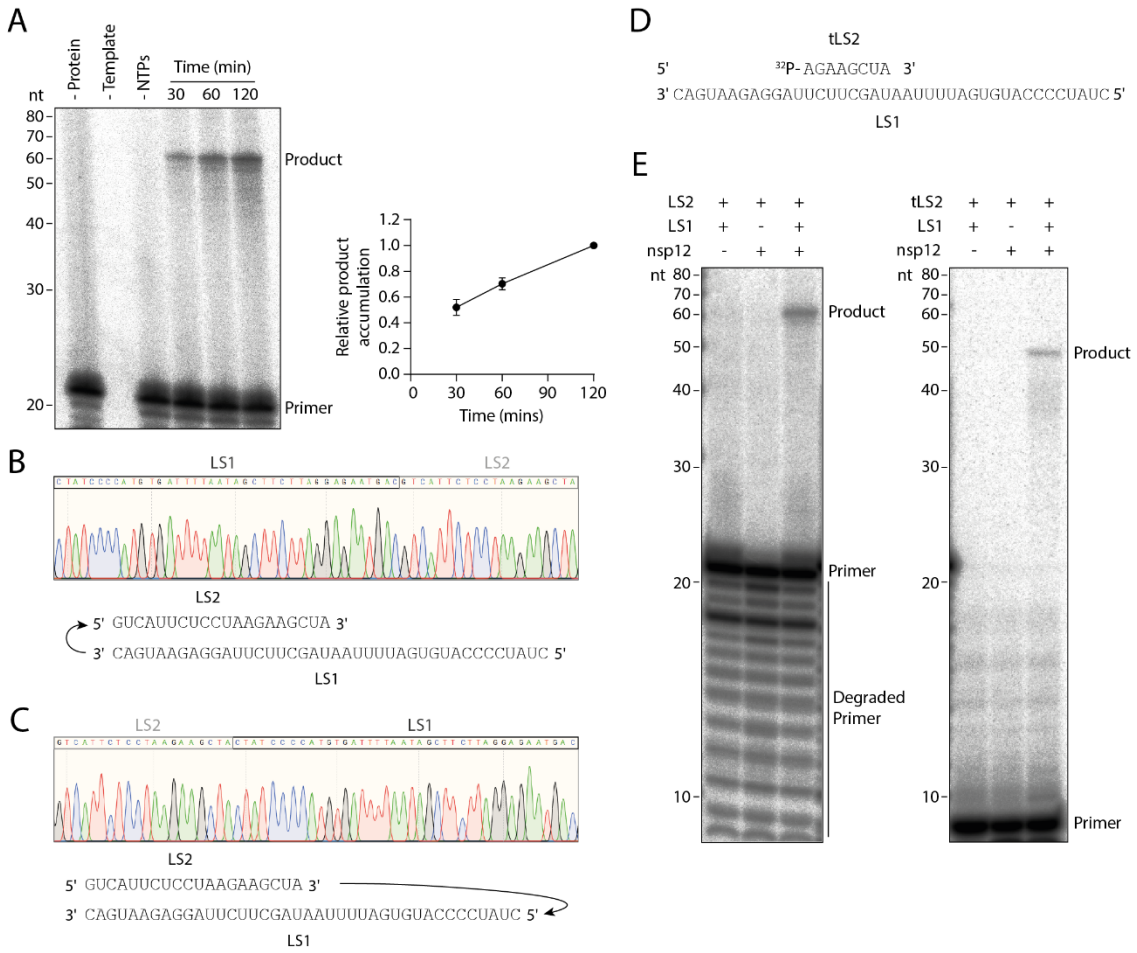

Figure 6.6 **Identification of RNA ligase activity in SARS-CoV-2 nsp12.** A, Nsp12 was incubated with LS1/LS2 template and NTPs, then reaction products were resolved by denaturing PAGE (left) and quantified (right). Quantification is from n=3 independent reactions, data are mean ± s.e.m. B, C, Nsp12 reaction products were gel purified and amplified by RT-PCR, then cloned into pcDNA3A and sequenced. Sanger sequencing traces are shown for the only two 60nt products identified (top), LS1-LS2 (B) and LS2-LS1 (C), with models depicting their synthesis from the LS1/LS2 template RNAs by nsp12 (bottom). Representative data are shown for n=2 (B) or n=4 (C) Sanger sequencing traces. D, Schematic of the radiolabelled 8nt tLS2 primer. E, Nsp12 RNA ligase activity was tested on the combination of LS2 and LS1 RNAs (left) or tLS2 and LS1 RNAs (right). Note that some degraded primer is visible below the LS2 primer.

Intriguingly, I found that nsp12 alone produces a product of 60nt in RNA polymerase activity assays, which exceeds the length of the 40nt LS1 template (**Fig. 6.6A**). I therefore gel purified the nsp12 reaction product and amplified it by RT-PCR, then cloned it into a pcDNA3A plasmid for Sanger sequencing. Sequencing identified two different products of 60nt, both comprised of the conjoined sequences of the 40nt LS1 template and the 20nt LS2 primer (**Fig. 6.6B, C**). One 60nt product appeared to be formed by joining the 3′ end of LS1 to the 5′ end of LS2, while the other product was formed by joining the 3′ end of LS2 to the 5′ end of LS1. These data suggest that nsp12 is able to covalently link the LS1/LS2 template RNAs in the absence of nsp7 and nsp8. To confirm this observation, I incubated nsp12 with LS1 and either full length LS2 or a truncated 8nt version of LS2 (tLS2) (**Fig. 6.6D**). As expected, the reaction with LS1 and LS2 produced a 60nt product; however, the reaction with LS1 and tLS2 produced a product of approximately 48nt (**Fig. 6.6E**). These data demonstrate that nsp12 can covalently link the RNA primer to the RNA template, effectively functioning as an RNA ligase *in vitro*.

Next, I aimed to characterise this activity further. I first examined the effect of nsp7 and nsp8 by mixing them with nsp12 and incubating the sub-complexes with LS1/LS2 RNAs and NTPs (**Fig. 6.7A**). In these assays nsp7 and nsp8 had opposite effects on RNA ligase activity; nsp7 slightly promoted activity, while nsp8 significantly reduced RNA ligase activity and appeared to promote RNA polymerase activity, causing limited extension of the LS2 primer. The combination of nsp7 and nsp8 abolished RNA ligase activity and strongly enhanced RNA polymerase activity.

RNA ligases are ATP-dependent, with the exception of GTP-dependent RNA ligase RtcB; therefore, to examine whether the nsp12 RNA ligase is also ATP-dependent, I performed reactions in the presence of each nucleotide individually (**Fig. 6.7B**) (194). No activity was detected for any nucleotide except ATP, which facilitated RNA ligase activity to the same level as a mixture of all nucleotides.

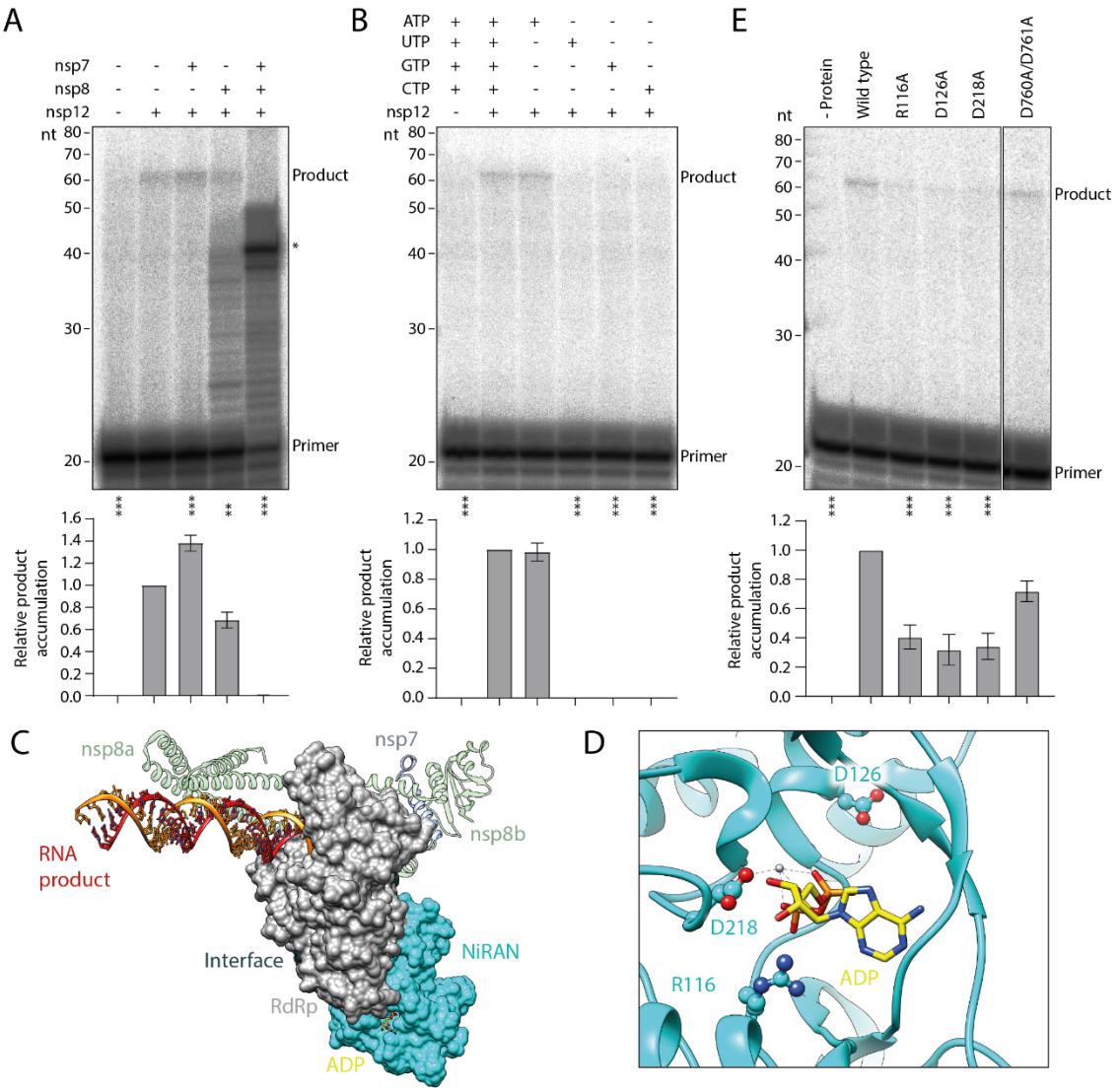

Figure 6.7 **Characterisation of the nsp12 RNA ligase activity.** A, Nsp12 RNA ligase activity was tested in the presence of nsp7 and nsp8, the asterisk denotes the 40nt major product of RNA polymerase activity. B, Nsp12 RNA ligase activity was tested in the presence of each nucleotide. C, Structure of the SARS-CoV-2 nsp7/8/12 complex bound to RNA product (red), with ADP (yellow) bound in the nsp12 NiRAN domain (cyan) (PDB: 6XEZ). D, Close-up view of the nsp12 NiRAN domain nucleotide binding pocket with ADP bound (PDB: 6XEZ). E, RNA ligase activity assays were performed using mutant nsp12 proteins. Quantification is from n=3 independent reactions, data are mean ± s.e.m., analysed by one-way ANOVA. **P<0.01, ***P<0.001.

The nsp12 NiRAN domain has been shown to bind to ADP, and this interaction involves similar amino acid residues as the interaction with GDP (**Fig. 6.7C, D**) (169). To assess the involvement of the NiRAN domain in nsp12 RNA ligase activity, I performed RNA ligase reactions using the mutant nsp12 proteins described above (**Fig. 6.7E**). All NiRAN domain mutations significantly reduced activity to approximately 30% of wild type nsp12, while the D760A/D761A RdRp active site mutation had no significant effect. Collectively, these data suggest that in addition to having GTase activity, the nsp12 NiRAN domain is involved in performing an ATP-dependent RNA ligase reaction *in vitro*.

RNA ligase and GTase reactions both require nucleotide binding, meaning that they could be inhibited by nucleotide analogue drugs binding to the NiRAN domain. I therefore titrated remdesivir triphosphate into nsp12 RNA ligase reactions, where it inhibited activity with an IC_{50} of 0.49mM (**Fig. 6.8A, D**). I also titrated the drug into nsp12 GTase activity assays and found that it inhibited activity with a strikingly similar IC_{50} of 0.43mM, which is consistent with both reactions using the same active site (**Fig. 6.8B, D**). To provide evidence that remdesivir triphosphate inhibits GTase activity by binding to the NiRAN domain, I then titrated the drug into GTase activity assays performed using only the NiRAN domain (**Fig. 6.8C, D**). Under these conditions the drug also inhibited activity, but with a much lower IC_{50} of 4.06μM. These data show that RNA ligase and GTase reactions performed by the SARS-CoV-2 nsp12 NiRAN domain can be inhibited by remdesivir triphosphate *in vitro*.

6.3 **DISCUSSION**

Coronaviruses are thought to synthesise m^7G-capped viral RNA, but the identity of the viral GTase enzyme has remained elusive (140). In this study I provide direct evidence that SARS-CoV-2 viral RNAs are m^7G-capped, and demonstrate that the nsp12 NiRAN domain can function

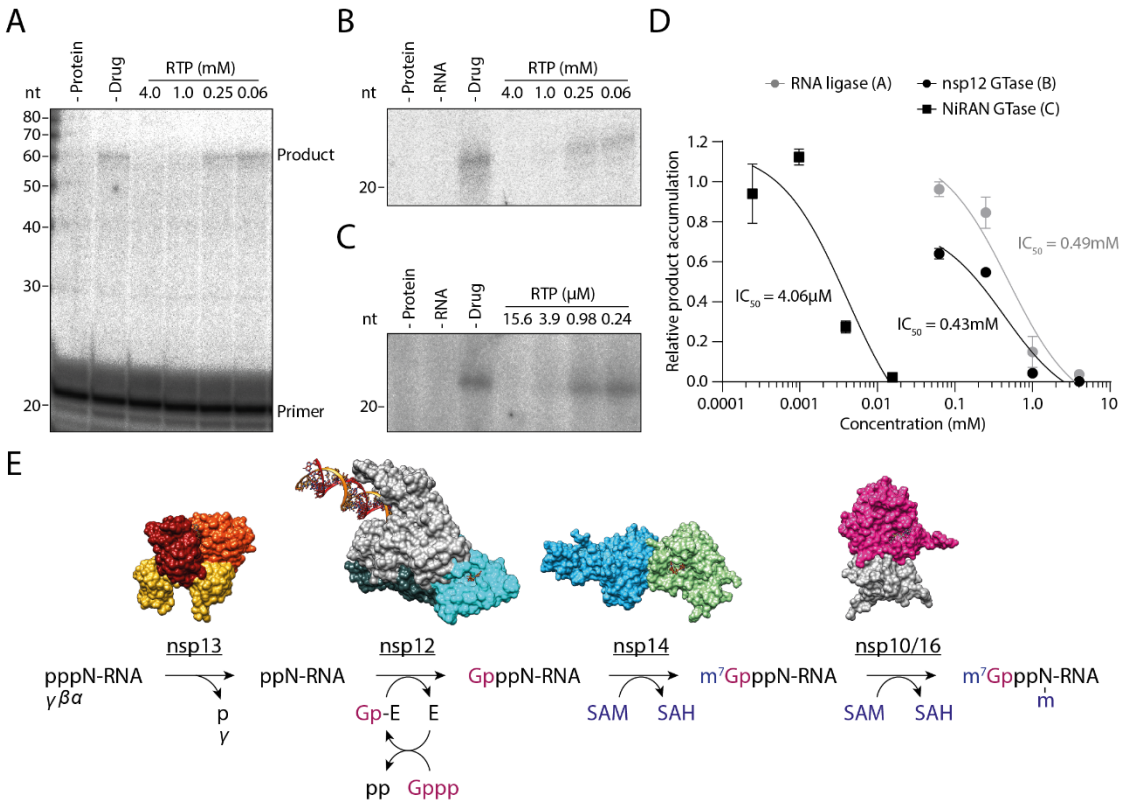

Figure 6.8 **Inhibition of RNA ligase and guanylyltransferase activities by remdesivir triphosphate, and a model for coronavirus m^{7}G cap synthesis.** A, Remdesivir triphosphate (RTP) was titrated into nsp12 RNA ligase activity assays. B, Remdesivir triphosphate was titrated into full-length nsp12 GTase activity assays. C, Remdesivir triphosphate was titrated into NiRAN domain GTase activity assays. D, Inhibition curves for remdesivir triphosphate in RNA ligase and GTase activity assays. IC_{50} values are derived from fitting data to dose-response inhibition curves by nonlinear regression in GraphPad Prism 9. Quantification is from n=3 independent reactions, data are mean ± s.e.m. E, Model for coronavirus m^{7}G cap synthesis. Nsp13 (PDB: 6XEZ) removes the γ-phosphate from nascent viral RNA using its 5′ triphosphatase activity. The resulting diphosphorylated RNA is linked to GTP (magenta) by the nsp12 (PDB: 7CYQ) GTase, forming a cap-like (GpppN-RNA) structure. Nsp14 (PDB: 5C8S) methylates the guanine cap at N7, and a complex of nsp16 and nsp10 (PDB: 7C2I) methylates the 2′ hydroxyl of the nucleotide at position +1. Both reactions consume S-adenosylmethionine (SAM; blue) and produce S-adenosylhomocysteine (SAH). The final product is viral RNA with a 5′ cap-1 structure.

as a GTase. Together, these findings allow me to propose a complete model for coronavirus m^7G cap synthesis (**Fig. 6.8E**) (181).

Nascent viral RNA is the substrate for m^7G cap synthesis, and reports of *de novo* initiation for coronavirus RNA polymerases suggest that it could be 5′ triphosphorylated (156, 162, 163, 195). The γ-phosphate of nascent viral RNA is removed by the 5′ triphosphatase activity of nsp13, resulting in a diphosphorylated substrate for the nsp12 GTase (171). GTase activity involves a GMP-enzyme covalent intermediate, which I was unable to identify for SARS-CoV-2 nsp12 but has been demonstrated for the EAV RNA polymerase (153). After the GTase reaction nsp14 and nsp16 methylate GpppN-RNA to produce the cap-1 structure, both enzymes utilising S-adenosylmethionine (SAM) as a methyl donor and producing S-adenosylhomocysteine (SAH) (173, 182, 183).

My identification of SARS-CoV-2 nsp12 as a GTase is corroborated by another independent study, and I further demonstrate that the nsp12 NiRAN domain alone is a functional GTase enzyme. It is worth noting that the NiRAN domain performs less efficiently in the GTase reaction than full-length nsp12, which is also consistent with my finding that RdRp active site mutations reduce GTase activity (**Fig. 6.5A**). These observations raise the possibility that the nsp12 RdRp domain is indirectly involved in the GTase reaction, for example it could be required for binding to substrate RNA. The presence of a GTase domain in nsp12 is reminiscent of non-segmented negative-strand RNA viruses (nsNSVs) such as VSV, which have a PRNTase domain in their RNA polymerases to perform an equivalent function to the GTase (29). This association between RNA synthesis and m^7G cap synthesis machinery is thought to promote co-transcriptional viral RNA capping (196). Interestingly, other proteins involved in coronavirus m^7G cap synthesis have RNA synthesis-related functions, such as nsp13 and nsp14 which are an RNA helicase and a proofreading exonuclease respectively (171, 174, 177). Therefore, coronaviral m^7G cap synthesis may also occur co-transcriptionally, which would minimise the risk of uncapped viral RNA detection by innate immune sensors including RIG-I (187).

My data also demonstrate that nsp12 can also perform an ATP-dependent RNA ligase reaction *in vitro*, which is dependent on the same amino acid residues as the GTase reaction and is inhibited by remdesivir triphosphate with a very similar IC_{50} (**Fig. 6.8D**). Canonical RNA ligase and GTase mechanisms both utilise a nucleotidylated enzyme intermediate, raising the possibility that the same active site could be involved in both reactions. A similar strategy is employed by nsp13, which has a shared active site for its RNA helicase and 5′ triphosphatase activities (171). Current models of coronavirus RNA synthesis do not require the RTC to have RNA ligase activity, but it has been suggested that an RNA ligase could function alongside a viral endoribonuclease, such as nsp15, in a yet-unknown pathway (153). It is also possible that an RNA ligase could be required to seal nicked RNA generated during viral RNA proofreading. This suggestion is supported by a recent study which indicates that the SARS-CoV-2 proofreading exonuclease, nsp14, can also function as an endonuclease (197).

My *in vitro* assays determined that remdesivir triphosphate inhibits SARS-CoV-2 RNA polymerase activity with an IC_{50} value of 1.68mM, which is substantially higher than the 0.77μM IC_{50} reported for SARS-CoV-2 virus growth in cell culture (**Fig. 6.1D**) (198). This discrepancy could arise from the *in vitro* assay not being fully representative of RNA polymerase function *in vivo*; for example, this assay detects extension of an RNA primer by 20nt, whereas the RTC must processively synthesise 29-kilobase SARS-CoV-2 genomes in *vivo* (199). Alternatively, this discrepancy could be explained by a dual mechanism of action, as here I show that remdesivir triphosphate inhibits multiple SARS-CoV-2 catalytic activities. Interestingly, I find that remdesivir triphosphate inhibits the GTase activity of the isolated NiRAN domain with a much lower IC_{50} than full-length nsp12, which could be explained by the drug binding non-productively to the nsp12 RdRp domain (**Fig. 6.8D**).

In summary, I have shown that the SARS-CoV-2 RNA polymerase NiRAN domain can perform GTase and RNA ligase reactions, the former allowing me to present a complete model of viral m^7G cap synthesis. I also demonstrated that these reactions are inhibited by remdesivir

triphosphate, highlighting the potential of the NiRAN domain as target for novel or repurposed antiviral drugs. The assays described here could prove useful for assessing the ability of such compounds to inhibit NiRAN domain function.

CHAPTER 7

Materials and Methods

7.1 MOLECULAR CLONING

7.1.1 *Oligonucleotide sequences*

DNA oligonucleotides were purchased from Sigma-Aldrich:

Table 7.1 Sequences of DNA oligonucleotides used for molecular cloning

Oligonucleotide	*Sequence (5' to 3')*
pBiFC-PB1-VN Fwd	CGGCAAAAAATCTCTCGAGGTACCGGAGG
pBiFC-PB1-VN Rev	ACATCCATCTCGGTCGACCGAATTCG
pcDNA-PB1-VN Fwd	CGACCGAGATGGATGTCAATCCGACTTTACTT TTCT
pcDNA-PB1-VN Rev	TCGAGAGATTTTTTGCCGTCTGAGCTCTTCAAT G
pBiFC-VC-PB2 Fwd	ATCAATTAGTAAGCGGCCGCGG
pBiFC-VC-PB2 Rev	ACCACCAGATTGAAAATACAAATTTTCACCCT TGTACAGCTCGTCCATGCC
pcDNA-VC-PB2 Fwd	GGTGAAAATTTGTATTTTCAATCTGGTGGTATG GAAAGAATAAAAGAACTAAGGAATCTA
pcDNA-VC-PB2 Rev	GGCCGCTTACTAATTGATGGCCATCC
pBiFC-VN-PB1 Fwd	AAAAATAGTGAGCGGCCGCGGG
pBiFC-VN-PB1 Rev	ACCACCAGATTGAAAATACAAATTTTCACCGG CGGTGAGATAGACGTTGT
pcDNA-VN-PB1 Fwd	GGTGAAAATTTGTATTTTCAATCTGGTGGTATG GATGTCAATCCGACTTT
pcDNA-VN-PB1 Rev	GGCCGCTCACTATTTTTGCCGTCTGAG
pBiFC-VC-PB1 Fwd	CAAAAATAGTTAAGCGGCCGCGGG
pBiFC-VC-PB1 Rev	ACCACCAGATTGAAAATACAAATTTTCACCCT TGTACAGCTCGTCCATGCC
pcDNA-VC-PB1 Fwd	GGTGAAAATTTGTATTTTCAATCTGGTGGTATG GATGTCAATCCGACTTTAC
pcDNA-VC-PB1 Rev	GCGGCCGCTTAACTATTTTTGCCGTCTGAGC
pcDNA-huANP32A 176-183 Fwd	GTGCTGTCCCTGGTGAAGGACCGGCAGGTGGT GGAAGATGAAGAGGATGAAGAT
pcDNA-huANP32A 176-183 Rev	CCGGTCCTTCACCAGGGACAGCACGTCCTCGT CCTCTTCCTCGTCGTCC
pcDNA-PA 352-356A Fwd	GCAGCAGCCGCAGCTACTAAAAATATGAAGA AAACGAGTCAGT
pcDNA-PA 352-356A Rev	AGCTGCGGCTGCTGCCTCATTCTCAATGTCCTG C

pcDNA-PB1 R135A Fwd	CAGCAGGCTGGTTCGCATTTAGAGTCCAGTCA TAGGTCTGT
pcDNA-PB1 R135A Rev	ACAGACCTATGACTGGACTCTAAATGCGAACC AGCCTGCTG
pcDNA-PB1 R220A Fwd	GTTCAGGGTTAATGCCGCAATTAGATAACTCC TTTTGTTCAATCTCTGC
pcDNA-PB1 R220A Rev	GCAGAGATTGAACAAAAGGAGTTATCTAATTG CGGCATTAACCCTGAAC
pcDNA-PB1 R220E Fwd	GTTCAGGGTTAATGCCTCAATTAGATAACTCC TTTTGTTCAATCTCTGC
pcDNA-PB1 R220E Rev	GCAGAGATTGAACAAAAGGAGTTATCTAATTG AGGCATTAACCCTGAAC
pcDNA-PB1 R350A Fwd	CATGTACCCCTTTCCCAGTGCCGCCATTTTGTT TGAGAAC
pcDNA-PB1 R350A Rev	GTTCTCAAACAAAATGGCGGCACTGGGAAAG GGGTACATG
pcDNA-PB1 R350E Fwd	CTCAAACATGTACCCCTTTCCCAGCTCCGCCAT TTTGTTTGAGAACATTAT
pcDNA-PB1 R350E Rev	ATAATGTTCTCAAACAAAATGGCGGAGCTGGG AAAGGGGTACATGTTTGAG
pcDNA-PB1 K553A Fwd	GTACGTGTACCTGTAATCTGCGATGAACAGCT GAAGGGCC
pcDNA-PB1 K553A Rev	GGCCCTTCAGCTGTTCATCGCAGATTACAGGT ACACGTAC
pcDNA-PB1 Y557A Fwd	TCTATGGCACCGGTACGTGGCCCTGTAATCTTT GATGAAC
pcDNA-PB1 Y557A Rev	GTTCATCAAAGATTACAGGGCCACGTACCGGT GCCATAGA
pcDNA-PB1 R560A Fwd	GTGTCACCTCTATGGCACGCGTACGTGTACCT GTAATC
pcDNA-PB1 R560A Rev	GATTACAGGTACACGTACGCGTGCCATAGAGG TGACAC
pcDNA-PB1 Q567A Fwd	TGATCTTCGGGTTTGTATTGCTGTGTCACCTCT ATGGCAC
pcDNA-PB1 Q567A Rev	GTGCCATAGAGGTGACACAGCAATACAAACCC GAAGATCA
pcDNA-PB1 Q569A Fwd	CAAATGATCTTCGGGTTGCTATTTGTGTGTCAC CTCTATGGC
pcDNA-PB1 Q569A Rev	GCCATAGAGGTGACACACAAATAGCAACCCG AAGATCATTTG
pcDNA-PB1 R571A Fwd	CAGTTTCTTTATTTCAAATGATCTTGCGGTTTG TATTTGTGTGTCACCTCT
pcDNA-PB1 R571A Rev	AGAGGTGACACACAAATACAAACCGCAAGAT CATTTGAAATAAAGAAACTG
pcDNA-PA Y464A Fwd	CATTAAGCAAGGCAGTATTGATGGCCACCCCC TTCATTATGTATTCTG
pcDNA-PA Y464A Rev	CAGAATACATAATGAAGGGGGTGGCCATCAAT ACTGCCTTGCTTAATG
pcDNA-PA K488A Fwd	CCTTCCCTCCTTAGTTCTACACGCGCTTATCAT TGGAATCAATTGG
pcDNA-PA K488A Rev	CCAATTGATTCCAATGATAAGCGCGTGTAGAA CTAAGGAGGGAAGG

pcDNA-PA R490A Fwd	TCGCCTTCCCTCCTTAGTTGCACACTTGCTTATCATTGGA
pcDNA-PA R490A Rev	TCCAATGATAAGCAAGTGTGCAACTAAGGAGGGAAGGCGA
pcDNA-PA R496A Fwd	ACCGTACAAATTGGTCTTTGCCCTTCCCTCCTTAGTTCTA
pcDNA-PA R496A Rev	TAGAACTAAGGAGGGAAGGGCAAAGACCAATTTGTACGGT
pcDNA-PB1 K553A/Y557A/R560A Fwd	TGTCACCTCTATGGCACGCGTACGTGGCCCTGTAATCTGCGATGAACAGCTGAAGGG
pcDNA-PB1 K553A/Y557A/R560A Rev	CCCTTCAGCTGTTCATCGCAGATTACAGGGCCACGTACGCGTGCCATAGAGGTGACA
pcDNA-PB1 Q567A/Q569A/R571A Fwd	CACAGTTTCTTTATTTCAAATGATCTTGCGGTTGCTATTGCTGTGTCACCTCTATGGCACCGGTACG
pcDNA-PB1 Q567A/Q569A/R571A Rev	CGTACCGGTGCCATAGAGGTGACACAGCAATAGCAACCGCAAGATCATTTGAAATAAAGAAACTGTG
pcDNA-PA D529A/R531A/E533A Fwd	CTTCTCCCATTTGTGTGGTGCAAGTGCTGGGGCAGTGAGGGAAAACTCCAT
pcDNA-PA D529A/R531A/E533A Rev	ATGGAGTTTTCCCTCACTGCCCCAGCACTTGCACCACACAAATGGGAGAAG
pcDNA-PA K413A Fwd	AATCGGTCAGTTCACATGCCGCGTTGAACTCATTCTGAATCC
pcDNA-PA K413A Rev	GGATTCAGAATGAGTTCAACGCGGCATGTGAACTGACCGATT
pcDNA-PB2 T609A Fwd	TGAGCGGTATCAAATGCCCCAAGCACATCCCTC
pcDNA-PB2 T609A Rev	GAGGGATGTGCTTGGGGCATTTGATACCGCTCA
pcDNA-PA K324A/H326A/E327A Fwd	CAGAAGATAATTTGGATTTATTCCCTTTGCGGCTGGTGCAACAACATTGGGTTCCTTCCATCCAAAG
pcDNA-PA K324A/H326A/E327A Rev	CTTTGGATGGAAGGAACCCAATGTTGTTGCACCAGCCGCAAAGGGAATAAATCCAAATTATCTTCTG
pcDNA-PA K339A/Q340A Fwd	TGTCCTGCAGTTCTGCCAGTACTGCCGCCCATGACAGAAGATAATTTGGAT
pcDNA-PA K339A/Q340A Rev	ATCCAAATTATCTTCTGTCATGGGCGGCAGTACTGGCAGAACTGCAGGACA
pcDNA-PB2 P132A Fwd	GGTTTCTAAAATGGACAGCGCCAAAGGTTCCATGTTT
pcDNA-PB2 P132A Rev	AAACATGGAACCTTTGGCGCTGTCCATTTTAGAAACC

7.1.2 *Plasmids*

Table 7.2 Descriptions of the origins of plasmids used in this study.

Plasmid	*Original source*
pcDNA3A	Fodor *et al*. 2002 (200)
pcDNA-PB1	Fodor *et al*. 2002 (200)
pcDNA-PB1a	Vreede *et al*.2004 (19)
pcDNA-PB2	Fodor *et al*. 2002 (200)
pcDNA-PB2 K627E	Ng *et al*. 2012 (201)
pcDNA-PB2-TAP	Deng *et al*. 2005 (202)
pcDNA-PA	Fodor *et al*. 2002 (200)
pcDNA-NP	Fodor *et al*. 2002 (200)
pPOLI-NA	Fodor *et al*. 1999 (203)
pBiFC-VN155	Kodama *et al*. 2010 (204)
pBiFC-VC155	Kodama *et al*. 2010 (204)
pBiFC-PB1-VN	Produced from pBiFC-VN155 and pcDNA-PB1 by Gibson assembly
pBiFC-VC-PB2	Produced from pBiFC-VC155 and pcDNA-PB2 by Gibson assembly
pBiFC-VN-PB1	Produced from pBiFC-VN155 and pcDNA-PB1 by Gibson assembly
pBiFC-VC-PB1	Produced from pBiFC-VC155 and pcDNA-PB1 by Gibson assembly
pcDNA-NB8205	Fan *et al*. 2019 (37)
pcDNA-NB8192	Kind gift from Itziar Serna Martín
pcDNA-NB8210	Fan *et al*. 2019 (37)
pCAGGS-chANP32A	Long *et al*. 2016 (88)
pCAGGS-chANP32A-Strep	Kind gift from Benjamin Nilsson
pCAGGS-huANP32A	Long *et al*. 2016 (88)
pCAGGS-huANP32A-Strep	Kind gift from Benjamin Nilsson
pcDNA-huANP32A	Produced from pCAGGS-huANP32A by restriction-ligation cloning
pcDNA-huANP32A 176-183	Site-directed mutagenesis from pcDNA-huANP32A
pcDNA-PA 352-356A	Site-directed mutagenesis from pcDNA-PA
pcDNA-PB2 71-73A	Kind gift from Jane Sharps
pcDNA-PA 351-353A	Kind gift from Jane Sharps
pcDNA-PA 356-358A	Kind gift from Jane Sharps
pcDNA-PB1 K669A/R670A	Kerry *et al*. 2008 (100)
pcDNA-PA D108A	Hara *et al*. 2006 (101)
pcDNA-PB1 R135A	Site-directed mutagenesis from pcDNA-PB1
pcDNA-PB1 R220A	Site-directed mutagenesis from pcDNA-PB1
pcDNA-PB1 R220E	Site-directed mutagenesis from pcDNA-PB1
pcDNA-PB1 R350A	Site-directed mutagenesis from pcDNA-PB1
pcDNA-PB1 R350E	Site-directed mutagenesis from pcDNA-PB1
pcDNA-PB1 K553A	Site-directed mutagenesis from pcDNA-PB1
pcDNA-PB1 Y557A	Site-directed mutagenesis from pcDNA-PB1
pcDNA-PB1 R560A	Site-directed mutagenesis from pcDNA-PB1
pcDNA-PB1 Q567A	Site-directed mutagenesis from pcDNA-PB1
pcDNA-PB1 Q569A	Site-directed mutagenesis from pcDNA-PB1

pcDNA-PB1 R571A	Site-directed mutagenesis from pcDNA-PB1
pcDNA-PA Y464A	Site-directed mutagenesis from pcDNA-PA
pcDNA-PA K488A	Site-directed mutagenesis from pcDNA-PA
pcDNA-PA R490A	Site-directed mutagenesis from pcDNA-PA
pcDNA-PA R496A	Site-directed mutagenesis from pcDNA-PA
pcDNA-PB1 K553A/Y557A/R560A	Site-directed mutagenesis from pcDNA-PB1
pcDNA-PB1 Q567A/Q569A/R571A	Site-directed mutagenesis from pcDNA-PB1
pcDNA-PA Y464A/K488A/R490A/R496A	Site-directed mutagenesis from pcDNA-PA
pcDNA-PA D529A/R531A/E533A	Site-directed mutagenesis from pcDNA-PA
pcDNA-PA K413A	Site-directed mutagenesis from pcDNA-PA
pcDNA-PB2 T609A	Site-directed mutagenesis from pcDNA-PB2
pcDNA-PB2 535-667Δ	Nilsson *et al*. 2017 (112)
pcDNA-PA K324A/H326A/E327A	Site-directed mutagenesis from pcDNA-PA
pcDNA-PA K339A/Q340A	Site-directed mutagenesis from pcDNA-PA
pcDNA-PB2 P132A	Site-directed mutagenesis from pcDNA-PB2

7.1.3 *Site-directed mutagenesis*

Site-directed mutagenesis PCR was carried out using Pfu Turbo DNA Polymerase (Agilent) in 50µl reaction mixtures prepared according to the manufacturer′s instructions, supplemented with 1U Phusion High-Fidelity DNA Polymerase (NEB). Thermocycling conditions were as follows: 95°C 10 mins; 50°C 30s, 68°C 18 mins, 19 cycles; 68°C 5 mins. Reactions were treated with 1µl DpnI (NEB) at 37°C for 1 hour, then transformed into DH5α *E. coli* and spread on LB agar plates supplemented with 50-100µg/mL ampicillin.

7.1.4 *Gibson assembly*

PCR was carried out with Phusion High-Fidelity DNA Polymerase (NEB) according to the manufacturer′s instructions. Reaction products were separated by 1% agarose gel electrophoresis pre-stained with SYBR Safe (Invitrogen), and target DNA fragments were excised and purified using QIAquick Gel Extraction Kit (Qiagen) according to the manufacturer′s instructions. Gibson assembly was performed using NEBuilder HiFi DNA Assembly mix (NEB) according to the

manufacturer′s instructions, and reaction products were transformed into DH5α *E. coli* then spread on LB agar plates supplemented with 50-100µg/mL ampicillin.

7.1.5 *Transformation and plasmid preparation*

20µl competent E. coli cells were incubated with 2µl plasmid DNA on ice for 1 hour. Cells were then heat shocked at 42°C for 1 min and placed back on ice for 5 mins. 150µl SOC Medium (Sigma-Aldrich) was added and cells were allowed to recover at 37°C for at least 30 mins, then spread on LB agar plates and incubated at 37°C overnight. *E. coli* colonies were inoculated into 2xYT medium supplemented with 50-100µg/mL ampicillin, then incubated at 37°C shaking at 180rpm overnight. Plasmid DNA was prepared using either QIAprep Spin Miniprep Kit (Qiagen) (3mL cultures) or Plasmid Midi Kit (Qiagen) (25mL cultures) according to the manufacturer′s instructions. Plasmids were validated by Sanger sequencing (Source Bioscience).

7.2 **CELL AND INFLUENZA VIRUS CULTURE**

7.2.1 *Mammalian cell culture*

Human Embryonic Kidney 293T (HEK 293T) and Madin-Darby Bovine Kidney (MDBK) cells were obtained from the Sir William Dunn School of Pathology cell bank at the University of Oxford. HEK 293T cells were maintained in Dulbecco′s Modified Eagle Medium (DMEM; Sigma-Aldrich) + 10% Foetal Calf Serum (FCS; Sigma-Aldrich). MDBK cells were maintained in Minimum Essential Medium (MEM; Sigma-Aldrich) + 10% FCS and 1x L-glutamine (200mM) (Gibco). Cells were grown at 37°C in 5% CO_2 and trypsinised using 1x TrypLE Express Enzyme (Gibco).

7.2.2 *Virus propagation in mammalian cells*

Influenza A/WSN/33 (H1N1) virus was propagated by infecting 70-80% confluent MDBK cells at MOI 0.01. Culture medium was collected 48 hours post-infection and cell debris was removed by centrifugation at 4000rpm for 10 mins. Medium was aliquoted and virus titre was determined by plaque assay.

7.2.3 *Plaque assay*

Approximately $0.4x10^6$ MDBK cells in MEM + 10% FCS + 1x L-glutamine (200mM) were seeded in each well of a 6-well dish. 24 hours later, duplicate 10-fold serial dilutions of virus stocks were made in MEM + 0.5% FCS + L-glutamine (200mM). Medium was removed from the seeded cells, cells were washed in PBS, then 200µl of diluted virus was added to each well. Cells were left at room temperature for 1 hour, then virus solution was removed and an overlay of 1% UltraPure Low Melting Point Agarose (ThermoFisher) in MEM + 0.5% FCS + L-glutamine (200mM) was added and allowed to set. Cells were incubated for a further 48-72 hours, then the overlay was removed and cells were stained with Coomassie Brilliant Blue (0.1% Coomassie R250, 10% acetic acid, 40% methanol).

7.3 **CELL-BASED INFLUENZA VIRAL RNA ANALYSIS**

7.3.1 *Oligonucleotide sequences*

Table 7.3 Sequences of DNA oligonucleotides used for cell-based influenza viral RNA detection

Oligonucleotide	*Sequence (5' to 3')*
NA 160	TCCAGTATGGTTTTGATTTCCG
NA 1280	TGGACTAGTGGGAGCATCAT
5S 100	TCCCAGGCGGTCTCCCATCC

7.3.2 *Ribonucleoprotein reconstitution assay*

Approximately $0.2x10^6$ HEK 293T cells in DMEM + 10% FCS in 24-well dishes were transfected with 200ng of pcDNA-PB1, pcDNA-PB2, pcDNA-PA, pPOLI-NA, and 400ng of pcDNA-NP using Lipofectamine 2000 (Invitrogen) according to the manufacturer′s instructions. Mutant plasmids replaced the wild type plasmids where indicated. In cases where additional plasmids were added to some samples, empty pcDNA3A was added to ensure an equal mass of plasmid for all samples. Cells were incubated for 24 hours post-transfection, then total cellular RNA was extracted and analysed by primer extension.

Alternatively, cells were harvested by centrifugation at 1000rpm for 5 mins and lysed in 100µl lysis buffer (50mM Tris-HCl pH 8.0, 200mM NaCl, 25% glycerol, 0.5% Igepal CA-630, 1mM DTT, 1mM phenylmethylsulfonyl fluoride (PMSF), 1x Protease Inhibitor Cocktail tablet (Sigma)) for 30 mins at 4°C, then lysate was cleared by centrifugation at 17000 x g for 5 min and proteins were analysed by sodium dodecyl sulphate (SDS) PAGE and western blotting.

7.3.3 *cRNA stabilisation assay*

Approximately $1x10^6$ HEK 293T cells in DMEM + 10% FCS in 6-well dishes were transfected with 1µg of pcDNA-PB1a, pcDNA-PB2, pcDNA-PA and pcDNA-NP using Lipofectamine 2000 according to the manufacturer′s instructions. Mutant plasmids replaced the wild type plasmids where indicated. 48 hours post-transfection, medium was replaced with DMEM + 0.5% FCS containing 5µg/mL actinomycin D and influenza A/WSN/33 virus at MOI 5. Cells were incubated for a further 6 hours, then total cellular RNA was extracted and analysed by primer extension.

7.3.4 *Primary transcription assay*

Approximately 1x10^6 HEK 293T cells in DMEM + 10% FCS in 6-well dishes were transfected with 1µg of pcDNA-PB1a, pcDNA-PB2 and pcDNA-PA using Lipofectamine 2000 according to the manufacturer′s instructions. Mutant plasmids replaced the wild type plasmids where indicated. 48 hours post-transfection, medium was replaced with DMEM + 0.5% FCS containing 100µg/mL cycloheximide and influenza A/WSN/33 virus at MOI 10. Cells were incubated for a further 4 hours, then total cellular RNA was extracted and analysed by primer extension.

7.3.5 *Viral growth assay*

Approximately 1x10^6 HEK 293T cells in DMEM + 10% FCS in 6-well dishes were transfected with 3µg of pcDNA-NB8205, pcDNA-NB8210 or pcDNA3A using Lipofectamine 2000 according to the manufacturer′s instructions. 24 hours post-transfection, medium was replaced with DMEM + 0.5% FCS containing influenza A/WSN/33 virus at MOI 0.1. Medium and cells were collected at 16, 24, 32 and 48 hours post-infection. Virus titre in the medium was determined by plaque assay, and viral RNA accumulation was determined by total cellular RNA extraction and primer extension.

7.3.6 *Radiolabelling of oligonucleotides*

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

7.3.7 *Total cellular RNA extraction and primer extension analysis*

Cells in a monolayer were lysed for 5 mins in 0.5-1mL TRI Reagent (Sigma-Aldrich), and total cellular RNA was extracted according to the manufacturer′s instructions. 1-4µl total cellular RNA was annealed to 0.25µl radiolabelled NA 160, NA 1280 and 5S 100 primers (**Table 7.3**) in 5µl total volume, by heating to 95°C for 2 mins and cooling on ice. RNA was reverse transcribed in a total reaction volume of 10µl using 50U SuperScript III Reverse Transcriptase (Invitrogen), 1x first-strand buffer and 10mM dithiothreitol (DTT), by incubating at 50°C for 1 hour followed by heating to 70°C for 10 mins. Reactions were mixed with 10µl loading dye (80% formamide, 1mM EDTA, bromophenol blue, xylene cyanol) and heated to 95°C for 3 mins before resolving on 6% denaturing PAGE.

7.3.8 *Denaturing polyacrylamide gel electrophoresis*

Radiolabelled RNA or DNA was resolved by 6-20% polyacrylamide gel electrophoresis (PAGE) (AccuGel 19:1 acrylamide:bis-acrylamide; Geneflow) with 7M urea in 1x Tris-Borate EDTA (TBE; 90mM Tris, 90mM boric acid, 2.5mM EDTA). Gels were polymerised using 0.1% ammonium persulfate (APS) and 0.1% N,N,N′,N′-tetramethylethane-1,2-diamine (TEMED), and run at 1250-2000V in 1x TBE using vertical sequencing gel apparatus. Radiolabelled products were visualised by exposure to phosphorimaging screens, which were scanned on a Fujifilm FLA-5000 scanner. Products were quantified using ImageJ and data were analysed using Prism 8 or Prism 9 (GraphPad). For primer extension analysis of cellular RNA samples, signal was normalised to the 5S rRNA loading control and vRNA signal in the negative control lane was subtracted from all subsequent lanes.

7.4 **INFLUENZA VIRUS PROTEIN INTERACTION ANALYSIS**

7.4.1 *Bimolecular fluorescence complementation*

Approximately $1x10^6$ HEK 293T cells in 6-well dishes were transfected with 1µg of pcDNA-VN-PB1, pcDNA-VC-PB1, pcDNA-PB2 and pcDNA-PA using Lipofectamine 2000 according to the manufacturer′s instructions. Mutant plasmids were substituted for wild type plasmids where indicated. 24 hours post-transfection cells were resuspended and washed in PBS, then fixed for 10 mins in 4% paraformaldehyde and washed again. Flow cytometry was carried out on an LSRFortessa X-20 cell analyser (BD Biosciences) using a 488nm laser and 530/30nm emission filter, and data were analysed using FlowJo software. Cells in each sample were classified into a Venus+ group based on 530/30nm channel intensity. The intensity threshold was placed such that 99% of cells in the negative control sample were in the Venus- group, and this threshold value was used for every subsequent sample in the same experiment. The ′Relative Venus signal′ value recorded was calculated as the percentage of cells in the Venus+ group for each sample, normalised to the wild type sample.

7.4.2 *Confocal fluorescence microscopy*

Approximately $1x10^6$ HEK 293T cells in DMEM + 10% FCS in 6-well dishes were transfected with 1µg of pcDNA-VN-PB1, pcDNA-VC-PB1, pcDNA-PB2 and pcDNA-PA using Lipofectamine 2000 according to the manufacturer′s instructions. Cells were grown on glass cover slips for 48 hours post-transfection, then washed with PBS and fixed in 4% paraformaldehyde for 10 mins. Slides were mounted in Mowiol 4-88 (Calbiochem) with 1µg/mL 4′,6-diamidino-2-phenylindole (DAPI; Sigma-Aldrich), and imaged using a 488nm laser on a Fluoview FV1200 microscope (Olympus). Images were processed using ImageJ.

7.4.3 *Affinity purification assay*

Approximately 5x10^6 HEK 293T cells DMEM + 10% FCS in 10cm dishes were transfected with 3µg pcDNA-PB1, pcDNA-PB2, pcDNA-PA and pCAGGS-huANP32A-Strep or pCAGGS-chANP32A-Strep using Lipofectamine 2000 according to the manufacturer′s instructions. Mutant plasmids replaced the wild type plasmids where indicated. 48 hours post-transfection, cells were harvested by centrifugation at 1000rpm for 5 mins and lysed in 500µl lysis buffer (50mM Tris-HCl pH 8.0, 200mM NaCl, 25% glycerol, 0.5% Igepal CA-630, 1mM DTT, 1mM PMSF, 1x Protease Inhibitor Cocktail tablet) for 30 mins at 4°C. Lysate was cleared by centrifugation at 17000 x g for 5 mins, then the supernatant was diluted in 2mL 150mM NaCl and incubated with washed Strep-Tactin beads (IBA Lifesciences) for 2 hours at 4°C. Beads were washed 3x in wash buffer (10mM Tris-HCl pH 8.0, 150mM NaCl, 10% glycerol, 0.1% Igepal CA-630, 1mM PMSF) and eluted overnight at 4°C in 200µl 1x Buffer E (IBA Lifesciences). Eluted complexes were immediately analysed by SDS PAGE and western blotting.

7.5 **INFLUENZA VIRUS PROTEIN PURIFICATION AND ANALYSIS**

7.5.1 *Protein purification from mammalian cells*

Approximately 5x10^6 HEK 293T cells were transfected with 3µg of pcDNA-PB1, pcDNA-PB2-TAP and pcDNA-PA using Lipofectamine 2000 according to the manufacturer′s instructions. For vRNP preparations, pcDNA-NP and pPOLI-NA were also transfected. Mutant plasmids replaced the wild type plasmids where indicated. 48 hours post-transfection, cells were harvested by centrifugation at 1000rpm for 5 mins and lysed in 500µl lysis buffer (50mM Tris-HCl pH 8.0, 200mM NaCl, 25% glycerol, 0.5% Igepal CA-630, 1mM DTT, 1mM PMSF, 1x Protease Inhibitor Cocktail tablet) for 30 mins at 4°C. Lysate was cleared by centrifugation at 17000 x g for 5 min and diluted in 2mL 150mM NaCl, then incubated with washed IgG sepharose beads (GE Healthcare) for 2 hours at 4°C. Beads were washed 2x in wash buffer (10mM Tris-HCl pH 8.0,

150mM NaCl, 10% glycerol, 0.1% Igepal CA-630, 1mM PMSF) and 1x in cleavage buffer (10mM Tris-HCl pH 8.0, 150mM NaCl, 40% glycerol, 0.1% Igepal CA-630, 1mM DTT, 1mM PMSF), then cleaved overnight at 4°C using AcTEV protease in a total volume of 200µl. Beads were cleared by centrifugation at 17000 x g for 1 min and elutants were analysed by 8% SDS PAGE and silver staining, then stored at -20°C.

7.5.2 *Ribonucleoprotein purification from virions*

Approximately $16.3x10^6$ MDBK cells in MEM + 0.5% FCS + L-glutamine (200mM) in T175 flasks were infected with influenza A/WSN/33 virus at MOI 0.01, and medium was harvested after 48hrs. Medium was clarified by centrifugation at 4000rpm for 10 mins, then 10000rpm for 15 mins at 4°C using a SW32 Ti rotor (Beckman Coulter). The resulting virus-containing medium was layered onto 0.22µm-filtered 30% sucrose in resuspension buffer (10mM Tris-HCl pH7.4, 100mM NaCl, 1mM EDTA) and virions were pelleted by centrifugation at 25000rpm using a SW32 Ti rotor (Beckman Coulter) for 90 mins at 4°C. Virions were resuspended overnight at 4°C in 80µl resuspension buffer, then were mixed with an equal volume of 2x disruption buffer (1x: 100mM MES, 100mM KCl, 5mM $MgCl_2$, 5% w/v glycerol, 50mM octylglucoside, 10mg/mL lysolecithin, 1.5mM DTT) and shaken for 20 mins at 31°C in a thermomixer. Disrupted virions were layered on top of a discontinuous glycerol gradient (1 ml of each fraction containing 70%, 50%, 40%, and 33% glycerol in 50 mM Tris-HCl pH 7.5, 150 mM NaCl) and the gradient was centrifuged at 45000rpm using an SW55 Ti rotor (Beckman Coulter) for 4 hours at 4°C. 300µl fractions with decreasing glycerol concentrations were collected by puncturing a hole in the bottom of the tube, and fractions were analysed by RNA extraction followed by primer extension. During the RNA extraction process 10µg of *E. coli* MRE 600 tRNA (Roche) was added immediately before isopropanol to aid precipitation. Fractions containing vRNA were pooled and stored at -20°C.

7.5.3 *SDS polyacrylamide gel electrophoresis and silver staining*

Protein samples were mixed with an equal volume of SDS loading dye (125mM Tris-HCl pH 6.8, 4% SDS, 20% glycerol, 10% β-mercaptoethanol, bromophenol blue) and heated to 95°C for 5 mins. Gels were prepared with a 3.2% 29:1 acrylamide:bis-acrylamide (AccuGel; Geneflow) stacking layer (125mM Tris-HCl pH6.8, 0.1% SDS) and 8-12% 29:1 acrylamide:bis-acrylamide separating layer (375mM Tris-HCl pH 8.8, 0.1% SDS), and set by the addition of 0.1% APS and 0.1% TEMED. Gels were run at 140V in 1x SDS running buffer (25mM Tris, 250mM glycine, 0.1% SDS) and proteins were visualised by silver staining using SilverXpress (Invitrogen) according to the manufacturer′s instructions.

7.5.4 *Western blotting*

Instead of silver staining, proteins were transferred from the gel onto a nitrocellulose membrane by wet transfer in 1x transfer buffer (25mM Tris, 190mM glycine, 20% ethanol) at 100V in 4°C. Membranes were blocked in blocking buffer (5% non-fat milk, 0.1% Tween-20, 1x PBS) overnight at 4°C or at room temperature for 1 hour. Membranes were then incubated with primary antibodies at 1 in 500 dilution in blocking buffer for 1 hour at room temperature, and washed 3x for 5 mins in PBS + 0.1% Tween-20. Secondary antibodies conjugated to horseradish peroxidase (HRP) were added to membranes at 1 in 10000 dilution in blocking buffer, and incubated for 1 hour at room temperature. Membranes were washed 3x in PBS + 0.1% Tween-20, then HRP was detected by incubation for 1 min with Amersham ECL Western Blotting Detection Reagents (GE Healthcare), followed by exposure on Super RC Fuji Medical X-ray film (Kodak).

Specific polyclonal antibodies produced in rabbits were used for western blotting. Commercially available antibodies were used to blot PB1 (Genetex), PB2 (Genetex), ANP32A (Sigma-Aldrich) and actin (Sigma-Aldrich), and a custom-made antibody was used to blot PA (55). In all cases, goat anti-rabbit antibody conjugated to HRP was used as a secondary antibody.

7.6 *IN VITRO* INFLUENZA VIRAL RNA ANALYSIS

7.6.1 *Oligonucleotide sequences*

RNA oligonucleotides were purchased from Dharmacon or Chemgenes:

Table 7.4 Sequences of RNA oligonucleotides used for influenza virus *in vitro* activity assays

Oligonucleotide	*Sequence (5′ to 3′)*
5′ vRNA promoter	AGUAGAAACAAGGCC
3′ vRNA promoter	GGCCUGCUUUUGCU
5′ cRNA promoter	AGCAAAAGCAGGCC
3′ cRNA promoter	GGCCUUGUUUCUACU
11-mer RNA	pp/pppGAAUACUCAAG
12-mer capped RNA	m^7GpppAAUCUAUAAUAG

DNA oligonucleotides were purchased from Sigma-Aldrich:

Table 7.5 Sequences of DNA oligonucleotides used for *in vitro* influenza viral RNA detection

Oligonucleotide	*Sequence (5′ to 3′)*
NA 160	TCCAGTATGGTTTTGATTTCCG
NA 160 Fwd	AGCAAAAGCAGGAGTTTAAATGAATCCAAAC C
NA 160 Rev	CCGGTTTGAATTGAATGGCTAATCCATAT
NA 1394	CCAGATCGTTCGAGTCCGTTTTTTTTTTTTTTT TGAACAAACTAC
NA 1394 Fwd	TGAATAGTGATACTGTAGATTGGTCT
NA 1394 Rev	CCAGATCGTTCGAGTCGT

7.6.2 *Cap-dependent transcription assay*

A cap-1 structure was added to 11-mer RNA (**Table 7.4**) in a 20µl reaction containing 0.5µM 11-mer RNA, 0.8mM S-adenosylmethionine, 10U Vaccinia virus capping enzyme (NEB), 50U 2′-O-methyltransferase (NEB) and 1µl α^{32}P-GTP (3000Ci/mmol; Perkin Elmer), incubated at 37°C for 1 hour. Reaction products were resolved by 20% denaturing PAGE and 11-mer capped RNA was purified by excising the band and eluting overnight in 1mL dH_2O. NAP-10 desalting columns (GE Healthcare) were washed with 15mL dH_2O, then the 1mL elutant was added, columns were washed with a further 200µl dH_2O, and 1.3mL dH_2O was added to elute the 11-mer capped RNA. Elutant was aliquoted and dried, then resuspended in 100µl dH_2O.

Transcription assays were performed in 3µl reactions containing 1µl 11-mer capped RNA, 5mM $MgCl_2$, 0.5µM 5′ and 3′ vRNA promoters (**Table 7.4**), 1mM DTT, 1U RNasin (Promega), 1mM ATP, 0.5mM CTP, 0.5mM UTP, 0.5mM GTP and 33ng/µl Sf9-purified influenza A/NT/60/68 (H3N2) $FluPol_A$ (a kind gift from Haitian Fan). Where indicated, Nb 8205 (a kind gift from Haitian Fan) was added at 16ng/µl. Reactions were run for 1 hour at 30°C, and stopped by the addition of 3µl loading dye (80% formamide, 1mM EDTA, bromophenol blue, xylene cyanol) and heating to 95°C for 3 mins. Reaction products were resolved by 20% denaturing PAGE.

7.6.3 *Primer-independent replication assay*

Primer-independent replication assays were performed in 3µl reactions containing 5mM $MgCl_2$, 0.5µM 5′ and 3′ vRNA or cRNA promoters (**Table 7.4**), 1mM DTT, 1U RNasin, 1mM ATP, 0.5mM CTP, 0.5mM UTP, 0.1µM GTP, 0.15µl α^{32}P-GTP (3000Ci/mmol), and 33ng/µl Sf9-purified influenza A/NT/60/68 (H3N2) $FluPol_A$. Where indicated, Nb 8205 was added at 16ng/µl. Reactions were run for 4 hours at 30°C, and stopped by the addition of 3µl loading dye (80% formamide, 1mM EDTA, bromophenol blue, xylene cyanol) and heating to 95°C for 3 mins. Reaction products were resolved by 20% PAGE.

7.6.4 *Dinucleotide synthesis assay*

Dinucleotide synthesis assays were performed in 3µl reactions containing 5mM $MgCl_2$, 0.5µM 5′ and 3′ vRNA or cRNA promoters (**Table 7.4**), 1mM DTT, 1U RNasin, 1mM ATP, 0.1µM GTP, 0.15µl α^{32}P-GTP (3000Ci/mmol), and 5ng/µl HEK 293T-purified influenza A/WSN/33 (H1N1) $FluPol_A$. Reactions were run for 4 hours at 30°C, then treated with 1U FastAP Thermosensitive Alkaline Phosphatase (ThermoFisher) for 1 hour at 37°C. 3µl loading dye (80% formamide, 1mM EDTA, bromophenol blue, xylene cyanol) was added and samples were heated to 95°C for 3 mins. Reaction products were resolved by 20% denaturing PAGE.

7.6.5 *Dinucleotide extension assay*

Dinucleotide extension assays were performed in 3µl reactions containing 0.5mM ApG dinucleotide (Fischer IBA), 5mM $MgCl_2$, 0.5µM 5′ and 3′ vRNA or cRNA promoters (**Table 7.4**), 1mM DTT, 1U RNasin, 1mM ATP, 0.5mM CTP, 0.5mM UTP, 0.1µM GTP, 0.15µl α^{32}P-GTP (3000Ci/mmol), and 3.3-33ng/µl Sf9-purified influenza A/NT/60/68 (H3N2) $FluPol_A$. Where indicated, Nb 8205 was added at 1.6-16ng/µl. Reactions were run for 4 hours at 30°C, and stopped by adding 3µl loading dye (80% formamide, 1mM EDTA, bromophenol blue, xylene cyanol) followed by heating to 95°C for 3 mins. Reaction products were resolved by 20% denaturing PAGE.

7.6.6 *Ribonucleoprotein activity assay*

5µl purified influenza A/WSN/33 (H1N1) vRNPs (from mammalian cells or virions) were incubated with 1mM ATP, 0.5mM UTP, 0.5mM CTP, 0.5mM GTP, 5mM $MgCl_2$, 1mM DTT, 20U RNasin and either 25mM 12-mer capped RNA (**Table 7.4**) or 4.69ng/µl rabbit β-globin mRNA (Sigma) in 10µl reactions for 4 hours at 30°C. Where indicated, 5-100ng/µl Sf9-purified

influenza A/NT/60/68 (H3N2) $FluPol_A$ was added. Activity assays using purified PP7 tag-purified cRNPs (a kind gift from Loïc Carrique) were performed using the same reaction conditions, except that 0.5mM ApG dinucleotide was added instead of 12-mer capped RNA or rabbit β-globin mRNA. After incubation, RNA was extracted and reaction products were analysed by primer extension or reverse-transcription qPCR. During the RNA extraction process 10µg of *E. coli* MRE 600 tRNA was added immediately before isopropanol to aid precipitation.

7.6.7 *Reverse-transcription qPCR*

2µl RNA was annealed to 1µM NA 160 or NA 1394 primers (**Table 7.5**) in 5µl mixtures by heating to 95°C for 2 mins and cooling on ice. Reverse transcription was performed by addition of 50U SuperScript III Reverse Transcriptase, 1x first-strand buffer, 10U RNasin and 10mM DTT to a total reaction volume of 10µl. Reactions were incubated at 50°C for 1 hour then heated to 70°C for 10 mins. 2µl cDNA was mixed with 1x qRT-PCR Brilliant III SYBR Master Mix (Aglient) and either 0.75µM NA 160 Fwd and NA 160 Rev primers, or 0.75µM NA 1394 Fwd and NA 1394 Rev primers (**Table 7.5**). qPCR was performed using a Quantstudio 5 RT-PCR machine (Applied Biosystems) with the following cycling conditions: 95°C 10 min; 95°C 15s, 60°C 1 min, 40 cycles. Data presented are $2^{(-\Delta Ct)}$, normalised to the wild type sample + NTPs, and were analysed using Prism 8.

7.6.8 *RNA cross-linking assay*

5µl HEK 293T-purified influenza A/WSN/33 (H1N1) $FluPol_A$ was incubated for 10 mins at 30°C with 1mM DTT, 20U RNasin, 0.5µM vRNA or cRNA 5′ promoters (**Table 7.5**), and 2µl radiolabelled 3′ vRNA or cRNA promoters (**Table 7.4**) in a total volume of 10µl. Mixtures in 96-well dishes were then irradiated with 254nm UV light in a UV Stratalinker (Stratagene) for 10 mins, and were mixed with 10µl SDS loading dye (125mM Tris-HCl pH 6.8, 4% SDS, 20%

glycerol, 10% β-mercaptoethanol, bromophenol blue) then heated to 95°C for 5 mins. Cross-linked complexes were separated by 8% SDS PAGE and visualised by phosphorimaging. Signals were quantified using ImageJ and analysed using Prism 8.

7.7 **CELL-BASED SARS-COV-2 VIRAL RNA ANALYSIS**

7.7.1 *Oligonucleotide sequences*

DNA oligonucleotides were purchased from Sigma-Aldrich:

Table 7.6 Sequences of DNA oligonucleotides used for SARS-CoV-2 viral RNA reverse-transcription qPCR

Oligonucleotide	*Sequence (5' to 3')*
Leader Fwd	ATTAAAGGTTTATACCTTCCCAG
FL Rev	AATTAGTTATTAATTATACTGCGTG
S Rev	GCAGGGGGTAATTGAGTT
3a Rev	GCGCGAACAAAATCTGAA
E Rev	CGCTATTAACTATTAACGTACCT
M Rev	GCTCTTCAACGGTAATAGTAC
6 Rev	CCAATCCTGTAGCGAC
7a Rev	AAAAGTACTGTTGTACCTCT
8 Rev	TGAGTACATGACTGTAAACTAC
N Rev	GTTCTCCATTCTGGTTACTG
All-anti Fwd	GTCATTCTCCTAAGAAGCT
All-anti Rev	GATCGAGTGTACAGTGAAC
GAPDH Fwd	CCATGGAGAAGGCTGGGG
GAPDH Rev	CAAAGTTGTCATGGATGACC
5S Fwd	GTCTACGGCCATACCACCCTGAACG
5S 100	TCCCAGGCGGTCTCCCATCC

7.7.2 *Reverse-transcription qPCR*

Total RNA from Vero CCL-81 cells infected with SARS-CoV-2 at MOI 1.0 was a kind gift from Michael Knight. 1µl RNA was annealed to 1µM of each reverse (Rev) primer or 5S 100 (**Table 7.6**) in separate 5µl mixtures by heating to 95°C for 2 mins and cooling on ice. Reverse transcription was performed by addition of 50U SuperScript III Reverse Transcriptase, 1x first-strand buffer, 10U RNasin and 10mM DTT to a total reaction volume of 10µl. Reactions were

incubated at 50°C for 1 hour then heated to 70°C for 10 mins. 1µl cDNA was mixed with 1x qRT-PCR Brilliant III SYBR Master Mix (Aglient) and a further 0.75µM of the reverse-transcription primer, as well as 0.75µM of the corresponding forward (Fwd) primer. For full-length (FL) and all subgenomic RNAs 0.75µM Leader Fwd primer was used (**Table 7.6**). qPCR was performed using a Quantstudio 5 RT-PCR machine (Applied Biosystems) with the following cycling conditions: 95°C 10 min; 95°C 15s, 60°C 30s, 60 cycles. Data presented are normalised according to GAPDH mRNA signal.

7.7.3 *RNA immunoprecipitation*

2µg total RNA from Vero CCL-81 cells 9 hours post-infection with SARS-CoV-2 was mixed with 2µg anti-m^7G antibody (MBL) or anti-his antibody in a total volume of 200µl wash buffer (1x PBS, 0.01% bovine serum albumin, 2mM EDTA, 1mM DTT, 0.1U/µl RNasin). The mixture was then incubated with 50µl washed Sheep Anti-Mouse IgG Dynabeads (Invitrogen) at room temperature for 30 mins on a rotating wheel. Beads were then washed 3x in 1mL wash buffer and RNA was eluted by addition of 1mL TRI reagent followed by vortexing. RNA extraction was then performed according to the manufacturer's instructions, except that 20µg glycogen (Roche) was added immediately before isopropanol to aid RNA precipitation. Extracted RNA was analysed by reverse-transcription qPCR as described in section 7.7.2. Data presented are normalised according to 5S rRNA and were analysed using Prism 9.

7.8 *IN VITRO* SARS-COV-2 VIRAL RNA ANALYSIS

7.8.1 *Oligonucleotide sequences*

RNA oligonucleotides were purchased from Dharmacon or Chemgenes:

Table 7.7 Sequences of RNA oligonucleotides used for SARS-CoV-2 *in vitro* activity assays

Oligonucleotide	*Sequence (5′ to 3′)*
LS1 template	CUAUCCCCAUGUGAUUUUAAUAGCUUCUUAG GAGAAUGAC
LS2 primer	GUCAUUCUCCUAAGAAGCUA
tLS2 primer	AGAAGCUA
20-mer RNA	pp/pppAAUCUAUAAUAGCAUUAUCC

DNA oligonucleotides were purchased from Sigma-Aldrich:

Table 7.8 Sequences of DNA oligonucleotides used for SARS-CoV-2 RNA product amplification and sequencing

Oligonucleotide	*Sequence (5′ to 3′)*
LS1-3′ RT	CAGGAAACAGCTATGACGTCATTCTCCTAAGA AGCTATTAAAATCA
LS2-3′ RT	CAGGAAACAGCTATGACTAGCTTCTTAGGAGA ATGACGTC
LS1-3′ PCR	TAATACGACTCACTATAGGGGTCATTCTCCTA AGAAGCTACTATCC
LS2-3′ PCR	TAATACGACTCACTATAGGGCTATCCCCATGT GATTTTAATAGCT
SP6 promoter	ATTTAGGTGACACTATAG

7.8.2 *RNA polymerase activity assay*

Purified SARS-CoV-2 nsp7, nsp8 and nsp12 proteins (a kind gift from Haitian Fan) were pre-mixed at a molar ratio of 5:5:1 in order to form the nsp7/8/12 complex. LS1 template and radiolabelled LS2 primer RNAs (**Table 7.7**) were mixed at a concentration of 150nM, then annealed by heating to 70°C for 5 mins and cooling to room temperature. 50nM annealed RNA was incubated with 500nM nsp7/8/12 complex, 5mM $MgCl_2$, 0.5mM of each ATP, UTP, GTP

and CTP, 10mM KCl, 1U RNasin and 1mM DTT in 3µl reaction volume for 1-30 mins at 30°C. Where drugs were included in reactions, the same volume of the appropriate solvent was added to all reactions (**Table 7.9**). Unless stated otherwise, reactions were run for 2 mins. Reactions were stopped by adding 3µl loading dye (80% formamide, 1mM EDTA, bromophenol blue, xylene cyanol), followed by heating to 95°C for 3 mins, and products were resolved by 20% denaturing PAGE.

Table 7.9 Stocks and sources of RNA polymerase inhibitor drugs tested

Drug name	*Solvent*	*Stock*	*Source*
Acyclovir triphosphate	DMSO	20mM	Kind gift from Jonathan Grimes
Remdesivir triphosphate	dH2O	20mM	MedChemExpress
Mitoxantrone dihydrochloride	dH2O	20mM	MedChemExpress
Enisamium iodide	dH2O	20mM	Kind gift from Aartjan te Velthuis
VR17-04	dH2O	20mM	Kind gift from Aartjan te Velthuis

7.8.3 *Phosphatase activity assay*

250nM purified nsp13 with an N-terminal His6-Zbasic tag (a kind gift from Yuliana Yosaatmadja and Opher Gileadi) was incubated with 4.75µM ATP, 0.25µM γ-^{32}P-ATP and 5mM $MgCl_2$ in a total reaction volume of 2µl for 0.5-5 mins at 30°C, then reactions were stopped by addition of 3µl loading dye (80% formamide, 1mM EDTA, bromophenol blue, xylene cyanol) and heating to 95°C for 3 mins. 0.5U/µl FastAP Thermosensitive Alkaline Phosphatase was incubated with 4.75µM ATP and 0.25µM γ-^{32}P-ATP for 1 hour at 37°C as a positive control. Reaction products were resolved by 20% denaturing PAGE.

7.8.4 *Guanylyltransferase activity assay*

A 5µM mixture of di- and triphosphorylated 20-mer RNA (**Table 7.7**) was treated with 250nM purified nsp13 in the presence of 5mM $MgCl_2$ for 5 mins at 30°C, then nsp13 was heat inactivated at 70°C for 5 mins. The resulting diphosphorylated 20-mer RNA stock was used for all

guanylyltransferase reactions. Where indicated, this 5µM diphosphorylated 20mer RNA was then treated with 0.3U/µl FastAP Thermosensitive Alkaline Phosphatase or 0.5U/µl RNA 5′ Pyrophosphohydrolase (NEB) in 1xNEBuffer 2 (NEB) for 1 hour at 37°C. Enzymes were heat inactivated at 75°C for 5 mins before the resulting RNA substrates were used in GTase reactions.

To run GTase reactions 1µM diphosphorylated 20-mer RNA was incubated with 500nM nsp12 or purified nsp12 NiRAN domain (a kind gift from Haitian Fan), 5mM $MgCl_2$, 0.15µl α-^{32}P-GTP (3000Ci/mmol), 10mM KCl, 1U RNasin and 1mM DTT in 3µl reaction volume for 1-4 hours at 30°C. Unless stated otherwise, reactions were run for 4 hours. Where indicated, purified SARS-CoV-2 nsp7, nsp8 and nsp12 proteins were pre-mixed at a molar ratio of 5:5:1 and used in place of nsp12. GTase reactions involving vaccinia capping enzyme were run for 1 hour under the same conditions, using 0.01U/µl vaccinia capping enzyme (NEB) instead of SARS-CoV-2 protein.

Where indicated, the completed GTase reactions were then treated with 0.3U/µl FastAP Thermosensitive Alkaline Phosphatase or 0.5U/µl RNA 5′ Pyrophosphohydrolase in 1xNEBuffer 2 for 1 hour at 37°C. All reactions were stopped by addition of 3µl loading dye (80% formamide, 1mM EDTA, bromophenol blue, xylene cyanol), followed by heating to 95°C for 3 mins, and products were resolved by 20% denaturing PAGE.

7.8.5 *RNA ligase activity assay*

LS1 template and radiolabelled LS2 primer or radiolabelled tLS2 primer RNAs (**Table 7.7**) were mixed at a concentration of 150nM, then annealed by heating to 70°C for 5 mins and cooling to room temperature. 50nM annealed RNA was incubated with 500nM nsp12, 5mM $MgCl_2$, 0.5mM of each ATP, UTP, GTP and CTP (unless stated otherwise), 10mM KCl, 1U RNasin and 1mM DTT in 3µl reaction volume for 30-120 mins at 30°C. Unless stated otherwise, reactions were run for 120 mins. Reactions were stopped by adding 3µl loading dye (80% formamide, 1mM EDTA,

bromophenol blue, xylene cyanol), followed by heating to 95°C for 3 mins, and products were resolved by 20% denaturing PAGE.

7.8.6 *Product purification and sequencing*

LS2 primer (**Table 7.7**) was 5′ phosphorylated using T4 polynucleotide kinase, then used in place of radiolabelled LS2 to carry out an RNA ligase activity assay as described above. Reaction products were resolved by 20% denaturing PAGE, then the 60nt product was excised and eluted from the gel overnight in 0.75mL dH_2O at 4°C. Eluted RNA was mixed with 20μg glycogen and precipitated in an equal volume of isopropanol, then reverse transcribed with SuperScript™ III according to the manufacturer's instructions using LS1-3′RT or LS2-3′RT primers (**Table 7.8**). cDNA was amplified by PCR using LS1-3′ PCR and LS1-3′RT primers, or LS2-3′ PCR and LS2-3′RT primers (**Table 7.8**). PCR products were separated by 1% agarose gel electrophoresis and purified using a QIAquick Gel Extraction Kit, then phosphorylated using T4 polynucleotide kinase and ligated into pcDNA3A at the EcoRV restriction site using Mighty Mix DNA Ligation Kit (Takara Bio). Plasmids were transformed into chemically competent *E. coli*, then single clones were purified using a QIAprep Spin Miniprep Kit and were Sanger sequenced using an SP6 promoter primer (**Table 7.8**).

References

1. (WHO) WHO. 2018. Influenza (Seasonal). World Health Organization.

2. Taubenberger JK, Morens DM. 2006. 1918 Influenza: The mother of all pandemics. Emerging Infectious Diseases 12:15–22.

3. Wei CJ, Crank MC, Shiver J, Graham BS, Mascola JR, Nabel GJ. 2020. Next-generation influenza vaccines: opportunities and challenges. Nature Reviews Drug Discovery 19:239–252.

4. Lowen AC. 2017. Constraints, Drivers, and Implications of Influenza A Virus Reassortment. Annual Review of Virology 4:105–121.

5. Heneghan CJ, Onakpoya I, Thompson M, Spencer EA, Jones M, Jefferson T. 2014. Zanamivir for influenza in adults and children: Systematic review of clinical study reports and summary of regulatory comments. BMJ (Online) 348.

6. Jefferson T, Jones M, Doshi P, Spencer EA, Onakpoya I, Heneghan CJ. 2014. Oseltamivir for influenza in adults and children: Systematic review of clinical study reports and summary of regulatory comments. BMJ (Online) 348.

7. Uehara T, Hayden FG, Kawaguchi K, Omoto S, Hurt AC, De Jong MD, Hirotsu N, Sugaya N, Lee N, Baba K, Shishido T, Tsuchiya K, Portsmouth S, Kida H. 2020. Treatment-Emergent Influenza Variant Viruses With Reduced Baloxavir Susceptibility: Impact on Clinical and Virologic Outcomes in Uncomplicated Influenza. The Journal of infectious diseases 221:346–355.

8. Furuta Y, Komeno T, Nakamura T. 2017. Favipiravir (T-705), a broad spectrum inhibitor of viral RNA polymerase. Proceedings of the Japan Academy Series B: Physical and Biological Sciences 93:449–463.

9. Rossman JS, Lamb RA. 2011. Influenza virus assembly and budding. Virology 411:229–236.

10. Peukes J, Xiong X, Erlendsson S, Qu K, Wan W, Calder LJ, Schraidt O, Kummer S, Freund SMV, Kräusslich HG, Briggs JAG. 2020. The native structure of the assembled matrix protein 1 of influenza A virus. Nature 587:495–498.

11. Hutchinson EC, Charles PD, Hester SS, Thomas B, Trudgian D, Martínez-Alonso M, Fodor E. 2014. Conserved and host-specific features of influenza virion architecture. Nature Communications 5:1–11.

12. Nakatsu S, Murakami S, Shindo K, Horimoto T, Sagara H, Noda T, Kawaoka Y. 2018. Influenza C and D Viruses Package Eight Organized Ribonucleoprotein Complexes. Journal of Virology 92.

13. Eisfeld AJ, Neumann G, Kawaoka Y. 2015. At the centre: Influenza A virus ribonucleoproteins. Nature Reviews Microbiology 13:28–41.

14. Noda T, Sagara H, Yen A, Takada A, Kida H, Cheng RH, Kawaoka Y. 2006. Architecture of ribonucleoprotein complexes in influenza A virus particles. Nature 439:490–492.

15. Limburg H, Harbig A, Bestle D, Stein DA, Moulton HM, Jaeger J, Janga H, Hardes K, Koepke J, Schulte L, Koczulla AR, Schmeck B, Klenk H-D, Böttcher-Friebertshäuser E. 2019. TMPRSS2 Is the Major Activating Protease of Influenza A Virus in Primary Human Airway Cells and Influenza B Virus in Human Type II Pneumocytes. Journal of Virology 93:649–668.

16. Edinger TO, Pohl MO, Stertz S. 2014. Entry of influenza A virus: Host factors and antiviral targets. Journal of General Virology 95:263–277.

17. Miyake Y, Keusch JJ, Decamps L, Ho-Xuan H, Iketani S, Gut H, Kutay U, Helenius A, Yamauchi Y. 2019. Influenza virus uses transportin 1 for vRNP debundling during cell entry. Nature Microbiology 4:578–586.

18. Walker AP, Fodor E. 2019. Interplay between Influenza Virus and the Host RNA Polymerase II Transcriptional Machinery. Trends in Microbiology 27:398–407.

19. Vreede FT, Jung TE, Brownlee GG. 2004. Model Suggesting that Replication of Influenza Virus Is Regulated by Stabilization of Replicative Intermediates. Journal of Virology 78:9568–9572.

20. Te Velthuis AJW, Fodor E. 2016. Influenza virus RNA polymerase: Insights into the mechanisms of viral RNA synthesis. Nature Reviews Microbiology 14:479–493.

21. Dadonaite B, Gilbertson B, Knight ML, Trifkovic S, Rockman S, Laederach A, Brown LE, Fodor E, Bauer DLV. 2019. The structure of the influenza A virus genome. Nature Microbiology 4:1781–1789.

22. Moeller A, Kirchdoerfer RN, Potter CS, Carragher B, Wilson IA. 2012. Organization of the influenza virus replication machinery. Science 338:1631–1634.

23. Arranz R, Coloma R, Chichón FJ, Conesa JJ, Carrascosa JL, Valpuesta JM, Ortín J, Martín-Benito J. 2012. The structure of native influenza virion ribonucleoproteins. Science 338:1634–1637.

24. Coloma R, Arranz R, de la Rosa-Trevín JM, Sorzano COS, 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 5:727–734.

25. Gallagher JR, Torian U, McCraw DM, Harris AK. 2017. Structural studies of influenza virus RNPs by electron microscopy indicate molecular contortions within NP supra-structures. Journal of Structural Biology 197:294–307.

26. Chan W-H, Ng AK-L, Robb NC, Lam MK-H, Chan PK-S, Au SW-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 84:7337–7345.

27. York A, Hengrung N, Vreede FT, Huiskonen JT, 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:E4238–E4245.

28. Turrell L, Lyall JW, Tiley LS, Fodor E, Vreede FT. 2013. The role and assembly mechanism of nucleoprotein in influenza A virus ribonucleoprotein complexes. Nature Communications 4:1591.

29. te Velthuis AJW, Grimes JM, Fodor E. 2021. Structural insights into RNA polymerases of negative-sense RNA viruses. Nature Reviews Microbiology 1–16.

30. Guilligay D, Tarendeau F, Resa-Infante P, Coloma R, Crepin T, Sehr P, Lewis J, Ruigrok RWH, Ortin J, Hart DJ, Cusack S. 2008. The structural basis for cap binding by influenza virus polymerase subunit PB2. Nature Structural and Molecular Biology 15:500–506.

31. Tarendeau F, Crepin T, Guilligay D, Ruigrok RWH, Cusack S, Hart DJ. 2008. Host determinant residue lysine 627 lies on the surface of a discrete, folded domain of influenza virus polymerase PB2 subunit. PLoS Pathogens 4:e1000136.

32. Subbarao EK, London W, Murphy BR. 1993. A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. Journal of Virology 67:1761–1764.

33. Dias A, Bouvier D, Crépin T, McCarthy AA, Hart DJ, Baudin F, Cusack S, Ruigrok RWH. 2009. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. Nature 458:914–918.

34. Pflug A, Guilligay D, Reich S, Cusack S. 2014. Structure of influenza A polymerase bound to the viral RNA promoter. Nature 516:355–360.

35. Reich S, Guilligay D, Pflug A, Malet H, Berger I, Crepin T, Hart D, Lunardi T, Nanao M, Ruigrok RWH, Cusack S. 2014. Structural insight into cap-snatching and RNA synthesis by influenza polymerase. Nature 516:361–366.

36. Serna Martin I, Hengrung N, Renner M, Sharps J, Martínez-Alonso M, Masiulis S, Grimes JM, Fodor E. 2018. A Mechanism for the Activation of the Influenza Virus Transcriptase. Molecular Cell 70:1101-1110.e4.

37. Fan H, Walker AP, Carrique L, Keown JR, Serna Martin I, Karia D, Sharps J, Hengrung N, Pardon E, Steyaert J, Grimes JM, Fodor E. 2019. Structures of influenza A virus RNA polymerase offer insight into viral genome replication. Nature 573:287–290.

38. Thierry E, Guilligay D, Kosinski J, Bock T, Gaudon S, Round A, Pflug A, Hengrung N, El Omari K, Baudin F, Hart DJ, Beck M, Cusack S. 2016. Influenza Polymerase Can Adopt an Alternative Configuration Involving a Radical Repacking of PB2 Domains. Molecular Cell 61:125–137.

39. Hengrung N, El Omari K, Serna Martin I, Vreede FT, Cusack S, Rambo RP, Vonrhein C, Bricogne G, Stuart DI, Grimes JM, Fodor E. 2015. Crystal structure of the RNA-dependent RNA polymerase from influenza C virus. Nature 527:114–117.

40. Fodor E, Pritlove DC, Brownlee GG. 1994. The influenza virus panhandle is involved in the initiation of transcription. Journal of Virology 68:4092–4096.

41. Pritlove DC, Fodor E, Seong BL, Brownlee GG. 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 76:2205–2213.

42. Cianci C, Tiley L, Krystal M. 1995. Differential activation of the influenza virus polymerase via template RNA binding. Journal of virology 69:3995–3999.

43. Fodor E, Pritlove DC, Brownlee GG. 1995. Characterization of the RNA-fork model of virion RNA in the initiation of transcription in influenza A virus. Journal of virology 69:4012–4019.

44. Wandzik JM, 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 181:877-893.e21.

45. 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 45:3353–3368.

46. Tomescu AI, Robb NC, Hengrung N, Fodor E, Kapanidis AN. 2014. Single-molecule FRET reveals a corkscrew RNA structure for the polymerase-bound influenza virus promoter. Proceedings of the National Academy of Sciences of the United States of America 111:E3335–E3342.

47. Pritlove DC, Poon LLM, Devenish LJ, Leahy MB, Brownlee GG. 1999. A Hairpin Loop at the 5′ End of Influenza A Virus Virion RNA Is Required for Synthesis of Poly(A)+ mRNA In Vitro. Journal of Virology 73:2109–2114.

48. Leahy MB, Pritlove DC, Poon LLM, Brownlee GG. 2001. Mutagenic Analysis of the 5′ Arm of the Influenza A Virus Virion RNA Promoter Defines the Sequence Requirements for Endonuclease Activity. Journal of Virology 75:134–142.

49. Robb NC, Te Velthuis AJW, Fodor E, Kapanidis AN. 2019. Real-time analysis of single influenza virus replication complexes reveals large promoter-dependent differences in initiation dynamics. Nucleic acids research 47:6466–6477.

50. Robb NC, Te Velthuis AJW, Wieneke R, Tampe R, Cordes T, Fodor E, Kapanidis AN. 2016. Single-molecule FRET reveals the pre-initiation and initiation conformations of influenza virus promoter RNA. Nucleic Acids Research 44:10304–10315.

51. Pflug A, Lukarska M, Resa-Infante P, Reich S, Cusack S. 2017. Structural insights into RNA synthesis by the influenza virus transcription-replication machine. Virus Research 234:103–117.

52. Peng R, Xu X, Jing J, Wang M, Peng Q, Liu S, Wu Y, Bao X, Wang P, Qi J, Gao GF, Shi Y. 2020. Structural insight into arenavirus replication machinery. Nature 579:615–619.

53. 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 GF, Shi Y. 2019. Structural insight into RNA synthesis by influenza D polymerase. Nature Microbiology 4:1750–1759.

54. Gerlach P, Malet H, Cusack S, Reguera J. 2015. Structural insights into bunyavirus replication and its regulation by the vRNA promoter. Cell 161:1267–1279.

55. Engelhardt OG, Smith M, Fodor E. 2005. Association of the Influenza A Virus RNA-Dependent RNA Polymerase with Cellular RNA Polymerase II. Journal of Virology 79:5812–5818.

56. 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 90:6014–6021.

57. 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 541:117–121.

58. Koppstein D, Ashour J, Bartel DP. 2015. Sequencing the cap-snatching repertoire of H1N1 influenza provides insight into the mechanism of viral transcription initiation. Nucleic Acids Research 43:5052–5064.

59. Gu W, Gallagher GR, Dai W, Liu P, Li R, Trombly MI, Gammon DB, Mello CC, Wang JP, Finberg RW. 2015. Influenza A virus preferentially snatches noncoding RNA caps. Rna 21:2067–2075.

60. Pflug A, Gaudon S, Resa-Infante P, Lethier M, Reich S, Schulze WM, Cusack S. 2018. Capped RNA primer binding to influenza polymerase and implications for the mechanism of cap-binding inhibitors. Nucleic Acids Research 46:956–971.

61. te Velthuis AJW, Oymans J. 2017. Initiation, Elongation, and Realignment during Influenza Virus mRNA Synthesis. Journal of Virology 92:e01775-17.

62. Li X, Palese P. 1994. Characterization of the polyadenylation signal of influenza virus RNA. Journal of Virology 68:1245–1249.

63. Robertson JS, Schubert M, Lazzarini RA. 1981. Polyadenylation sites for influenza virus mRNA. Journal of Virology 38:157–163.

64. Luo GX, Luytjes W, Enami M, Palese P. 1991. The polyadenylation signal of influenza virus RNA involves a stretch of uridines followed by the RNA duplex of the panhandle structure. Journal of Virology 65:2861–2867.

65. Poon LLM, Pritlove DC, Fodor E, Brownlee GG. 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 73:3473–3476.

66. Amorim MJ, Read EK, Dalton RM, Medcalf L, Digard P. 2007. Nuclear export of influenza A virus mRNAs requires ongoing RNA polymerase II activity. Traffic 8:1–11.

67. Read EKC, Digard P. 2010. Individual influenza A virus mRNAs show differential dependence on cellular NXF1/TAP for their nuclear export. Journal of General Virology 91:1290–1301.

68. York A, Fodor E. 2013. Biogenesis, assembly and export of viral messenger ribonucleoproteins in the influenza A virus infected cell. RNA Biology 10:1274–1282.

69. Dubois J, Terrier O, Rosa-Calatrava M. 2014. Influenza viruses and mRNA splicing: Doing more with less. mBio 5:1–13.

70. Perez JT, Varble A, Sachidanandam R, Zlatev I, Manoharan M, García-Sastre A, TenOever BR. 2010. Influenza A virus-generated small RNAs regulate the switch from transcription to replication. Proceedings of the National Academy of Sciences of the United States of America 107:11525–11530.

71. Vreede FT, Brownlee GG. 2007. Influenza Virion-Derived Viral Ribonucleoproteins Synthesize both mRNA and cRNA In Vitro. Journal of Virology 81:2196–2204.

72. Deng T, Vreede FT, Brownlee GG. 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 80:2337–2348.

73. Te Velthuis AJW, Robb NC, Kapanidis AN, Fodor E. 2016. The role of the priming loop in influenza A virus RNA synthesis. Nature Microbiology 1:16029.

74. González S, Ortín J. 1999. Distinct regions of influenza virus PB1 polymerase subunit recognize vRNA and cRNA templates. EMBO Journal 18:3767–3775.

75. Leahy MB, Zecchin G, Brownlee GG. 2002. Differential Activation of Influenza A Virus Endonuclease Activity Is Dependent on Multiple Sequence Differences between the Virion RNA and cRNA Promoters. Journal of Virology 76:2019–2023.

76. Oymans J, te Velthuis AJW. 2017. A Mechanism for Priming and Realignment during Influenza A Virus Replication. Journal of Virology 92:e01773-17.

77. Huet S, Avilov S V., Ferbitz L, Daigle N, Cusack S, Ellenberg J. 2010. Nuclear Import and Assembly of Influenza A Virus RNA Polymerase Studied in Live Cells by Fluorescence Cross-Correlation Spectroscopy. Journal of Virology 84:1254–1264.

78. Jorba N, Area E, Ortín J. 2008. Oligomerization of the influenza virus polymerase complex in vivo. Journal of General Virology 89:520–524.

79. Jorba N, Coloma R, Ortín J. 2009. Genetic trans-complementation establishes a new model for influenza virus RNA transcription and replication. PLoS Pathogens 5:e1000462.

80. World Health Organization (WHO). 2020. Cumulative Number of Confirmed Human Cases of Avian Influenza A/(H5N1). WHO. World Health Organization.

81. Long JS, Mistry B, Haslam SM, Barclay WS. 2019. Host and viral determinants of influenza A virus species specificity. Nature Reviews Microbiology 17:67–81.

82. Taft AS, Ozawa M, Fitch A, Depasse J V., Halfmann PJ, Hill-Batorski L, Hatta M, Friedrich TC, Lopes TJS, Maher EA, Ghedin E, Macken CA, Neumann G, Kawaoka Y. 2015. Identification of mammalian-adapting mutations in the polymerase complex of an avian H5N1 influenza virus. Nature Communications 6:7491.

83. Chin AWH, Li OTW, Mok CKP, Ng MKW, Peiris M, Poon LLM. 2014. Influenza A viruses with different amino acid residues at PB2-627 display distinct replication properties in vitro and in vivo: Revealing the sequence plasticity of PB2-627 position. Virology 468:545–555.

84. Gabriel G, Dauber B, Wolff T, Planz O, Klenk HD, Stech J. 2005. The viral polymerase mediates adaptation of an avian influenza virus to a mammalian host. Proceedings of the National Academy of Sciences of the United States of America 102:18590–18595.

85. Shinya Y, Hatta M, Staker BL, Watanabe S, Imai M, Shinya K, Sakai-Tagawa Y, Ito M, Ozawa M, Watanabe T, Sakabe S, Li C, Kim JH, Myler PJ, Phan I, Raymond A, Smith E, Stacy R, Nidom CA, Lank SM, Wiseman RW, Bimber BN, O'Connor DH, Neumann

G, Stewart LJ, Kawaoka Y. 2010. Biological and structural characterization of a host-adapting amino acid in influenza virus. PLoS Pathogens 6:15–16.

86. Mehle A, Doudna JA. 2009. Adaptive strategies of the influenza virus polymerase for replication in humans. Proceedings of the National Academy of Sciences of the United States of America 106:21312–21316.

87. Mänz B, Brunotte L, Reuther P, Schwemmle M. 2012. Adaptive mutations in NEP compensate for defective H5N1 RNA replication in cultured human cells. Nature Communications 3:802.

88. Long JS, Giotis ES, Moncorgé O, Frise R, Mistry B, James J, Morisson M, Iqbal M, Vignal A, Skinner MA, Barclay WS. 2016. Species difference in ANP32A underlies influenza A virus polymerase host restriction. Nature 529:101–104.

89. Reilly PT, Yu Y, Hamiche A, Wang L. 2014. Cracking the ANP32 whips: Important functions, unequal requirement, and hints at disease implications. BioEssays 36:1062–1071.

90. Zhang H, Zhang Z, Wang Y, Wang M, Wang X, Zhang X, Ji S, Du C, Chen H, Wang X. 2019. Fundamental Contribution and Host Range Determination of ANP32A and ANP32B in Influenza A Virus Polymerase Activity. Journal of Virology 93:e00174-19.

91. 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:e1008989.

92. Staller E, Sheppard CM, Neasham PJ, Mistry B, Peacock TP, Goldhill DH, Long JS, Barclay WS. 2019. ANP32 Proteins Are Essential for Influenza Virus Replication in Human Cells. Journal of Virology 93:e00217-19.

93. Baker SF, Ledwith MP, Mehle A. 2018. Differential Splicing of ANP32A in Birds Alters Its Ability to Stimulate RNA Synthesis by Restricted Influenza Polymerase. Cell Reports 24:2581-2588.e4.

94. 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 4:1–19.

95. Mistry B, Long JS, Schreyer J, Staller E, Sanchez-David RY, Barclay WS. 2019. Elucidating the Interactions between Influenza Virus Polymerase and Host Factor ANP32A. Journal of Virology 94:e01353-19.

96. Domingues P, Hale BG. 2017. Functional Insights into ANP32A-Dependent Influenza A Virus Polymerase Host Restriction. Cell Reports 20:2538–2546.

97. Long JS, Idoko-Akoh A, Mistry B, Goldhill D, Staller E, Schreyer J, Ross C, Goodbourn S, Shelton H, Skinner MA, Sang H, McGrew MJ, Barclay W. 2019. Species specific differences in use of ANP32 proteins by influenza A virus. eLife 8:696–697.

98. Zhang H, Li H, Wang W, Wang Y, Han GZ, Chen H, Wang X. 2020. A unique feature of swine ANP32A provides susceptibility to avian influenza virus infection in pigs. PLoS Pathogens 16:e1008330.

99. Ma W, Kahn RE, Richt JA. 2009. The pig as a mixing vessel for influenza viruses: Human and veterinary implications. Journal of Molecular and Genetic Medicine 03:158–166.

100. Kerry PS, Willsher N, Fodor E. 2008. A cluster of conserved basic amino acids near the C-terminus of the PB1 subunit of the influenza virus RNA polymerase is involved in the regulation of viral transcription. Virology 373:202–210.

101. Hara K, Schmidt FI, Crow M, Brownlee GG. 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 80:7789–7798.

102. Killip MJ, Fodor E, Randall RE. 2015. Influenza virus activation of the interferon system. Virus Research 209:11–22.

103. Kodama Y, Hu CD. 2012. Bimolecular fluorescence complementation (BiFC): A 5-year update and future perspectives. BioTechniques 53:285–298.

104. Shyu YJ, Liu H, Deng X, Hu CD. 2006. Identification of new fluorescent protein fragments for bimolecular fluorescence complementation analysis under physiological conditions. BioTechniques 40:61–66.

105. Gettemans J, de Dobbelaer B. 2021. Transforming nanobodies into high-precision tools for protein function analysis. American Journal of Physiology - Cell Physiology 320:C195–C215.

106. Chen KY, Afonso EDS, Enouf V, Isel C, Naffakh N. 2019. Influenza virus polymerase subunits co-evolve to ensure proper levels of dimerization of the heterotrimer. PLoS Pathogens 15:e1008034.

107. Garcin D, Kolakofsky D. 1992. Tacaribe arenavirus RNA synthesis in vitro is primer dependent and suggests an unusual model for the initiation of genome replication. Journal of Virology 66:1370–1376.

108. Garcin D, Kolakofsky D. 1990. A novel mechanism for the initiation of Tacaribe arenavirus genome replication. Journal of Virology 64:6196–6203.

109. Garcin D, Lezzi M, Dobbs M, Elliott RM, Schmaljohn C, Kang CY, Kolakofsky D. 1995. The 5’ ends of Hantaan virus (Bunyaviridae) RNAs suggest a prime-and-realign mechanism for the initiation of RNA synthesis. Journal of virology 69:5754–5762.

110. Taubenberger JK, Kash JC. 2010. Influenza virus evolution, host adaptation, and pandemic formation. Cell Host and Microbe 7:440–451.

111. Carrique L, Fan H, Walker AP, Keown JR, Sharps J, Staller E, Barclay WS, Fodor E, Grimes JM. 2020. Host ANP32A mediates the assembly of the influenza virus replicase. Nature 587:638–643.

112. Nilsson BE, te Velthuis AJW, Fodor E. 2017. Role of the PB2 627 Domain in Influenza A Virus Polymerase Function. Journal of Virology 91:e02467-16.

113. Gabriel G, Herwig A, Klenk HD. 2008. Interaction of polymerase subunit PB2 and NP with importin α1 is a determinant of host range of influenza A virus. PLoS Pathogens 4:e11.

114. Corman VM, Muth D, Niemeyer D, Drosten C. 2018. Hosts and Sources of Endemic Human Coronaviruses, p. 163–188. *In* Advances in Virus Research. Academic Press Inc.

115. Gorbalenya AE, Baker SC, Baric RS, de Groot RJ, Drosten C, Gulyaeva AA, Haagmans BL, Lauber C, Leontovich AM, Neuman BW, Penzar D, Perlman S, Poon LLM, Samborskiy D V., Sidorov IA, Sola I, Ziebuhr J. 2020. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nature Microbiology 5:536–544.

116. Cheng VCC, Lau SKP, Woo PCY, Kwok YY. 2007. Severe acute respiratory syndrome coronavirus as an agent of emerging and reemerging infection. Clinical Microbiology Reviews 20:660–694.

117. World Health Organization (WHO). 2019. Middle East respiratory syndrome coronavirus (MERS-CoV). WHO.

118. Ke Z, Oton J, Qu K, Cortese M, Zila V, McKeane L, Nakane T, Zivanov J, Neufeldt CJ, Cerikan B, Lu JM, Peukes J, Xiong X, Kräusslich HG, Scheres SHW, Bartenschlager R, Briggs JAG. 2020. Structures and distributions of SARS-CoV-2 spike proteins on intact virions. Nature 588:498–502.

119. Neuman BW, Adair BD, Yoshioka C, Quispe JD, Orca G, Kuhn P, Milligan RA, Yeager M, Buchmeier MJ. 2006. Supramolecular Architecture of Severe Acute Respiratory Syndrome Coronavirus Revealed by Electron Cryomicroscopy. Journal of Virology 80:7918–7928.

120. Klein S, Cortese M, Winter SL, Wachsmuth-Melm M, Neufeldt CJ, Cerikan B, Stanifer ML, Boulant S, Bartenschlager R, Chlanda P. 2020. SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. Nature Communications 11:1–10.

121. Perlman S, Netland J. 2009. Coronaviruses post-SARS: Update on replication and pathogenesis. Nature Reviews Microbiology 7:439–450.

122. V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. 2021. Coronavirus biology and replication: implications for SARS-CoV-2. Nature Reviews Microbiology 19:155–170.

123. Muth D, Corman VM, Roth H, Binger T, Dijkman R, Gottula LT, Gloza-Rausch F, Balboni A, Battilani M, Rihtarič D, Toplak I, Ameneiros RS, Pfeifer A, Thiel V, Drexler JF, Müller MA, Drosten C. 2018. Attenuation of replication by a 29 nucleotide deletion in SARS-coronavirus acquired during the early stages of human-to-human transmission. Scientific Reports 8:1–11.

124. He JF, Peng GW, Min J, Yu DW, Liang WJ, Zhang SY, Xu RH, Zheng HY, Wu XW, Xu J, Wang ZH, Fang L, Zhang X, Li H, Yan XG, Lu JH, Hu ZH, Huang JC, Wan ZY, Hou JL, Lin JY, Song HD, Wang SY, Zhou XJ, Zhang GW, Gu BW, Zheng HJ, Zhang XL, He M, Zheng K, Wang BF, Fu G, Wang XN, Chen SJ, Chen Z, Hao P, Tang H, Ren SX, Zhong Y, Guo ZM, Liu Q, Miao YG, Kong XY, He WZ, Li YX, Wu CI, Zhao GP, Chiu RWK, Chim SSC, Tong YK, Chan PKS, Tam JS, Lo YMD. 2004. Molecular Evolution of the SARS Coronavirus, during the Course of the SARS Epidemic in China. Science 303:1666–1669.

125. Raj VS, Mou H, Smits SL, Dekkers DHW, Müller MA, Dijkman R, Muth D, Demmers JAA, Zaki A, Fouchier RAM, Thiel V, Drosten C, Rottier PJM, Osterhaus ADME,

Bosch BJ, Haagmans BL. 2013. Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. Nature 495:251–254.

126. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, Schiergens TS, Herrler G, Wu NH, Nitsche A, Müller MA, Drosten C, Pöhlmann S. 2020. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. Cell 181:271-280.e8.

127. Letko M, Marzi A, Munster V. 2020. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. Nature Microbiology 5:562–569.

128. Gierer S, Bertram S, Kaup F, Wrensch F, Heurich A, Kramer-Kuhl A, Welsch K, Winkler M, Meyer B, Drosten C, Dittmer U, von Hahn T, Simmons G, Hofmann H, Pohlmann S. 2013. The Spike Protein of the Emerging Betacoronavirus EMC Uses a Novel Coronavirus Receptor for Entry, Can Be Activated by TMPRSS2, and Is Targeted by Neutralizing Antibodies. Journal of Virology 87:5502–5511.

129. Matsuyama S, Nagata N, Shirato K, Kawase M, Takeda M, Taguchi F. 2010. Efficient Activation of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein by the Transmembrane Protease TMPRSS2. Journal of Virology 84:12658–12664.

130. Simmons G, Gosalia DN, Rennekamp AJ, Reeves JD, Diamond SL, Bates P. 2005. Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. Proceedings of the National Academy of Sciences of the United States of America 102:11876–11881.

131. Kawase M, Shirato K, van der Hoek L, Taguchi F, Matsuyama S. 2012. Simultaneous Treatment of Human Bronchial Epithelial Cells with Serine and Cysteine Protease Inhibitors Prevents Severe Acute Respiratory Syndrome Coronavirus Entry. Journal of Virology 86:6537–6545.

132. Shang J, Wan Y, Luo C, Ye G, Geng Q, Auerbach A, Li F. 2020. Cell entry mechanisms of SARS-CoV-2. Proceedings of the National Academy of Sciences of the United States of America 117:11727–11734.

133. Cantuti-Castelvetri L, Ojha R, Pedro LD, Djannatian M, Franz J, Kuivanen S, van der Meer F, Kallio K, Kaya T, Anastasina M, Smura T, Levanov L, Szirovicza L, Tobi A, Kallio-Kokko H, Österlund P, Joensuu M, Meunier FA, Butcher SJ, Winkler MS, Mollenhauer B, Helenius A, Gokce O, Teesalu T, Hepojoki J, Vapalahti O, Stadelmann

C, Balistreri G, Simons M. 2020. Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. Science 370:856–860.

134. Daly JL, Simonetti B, Klein K, Chen KE, Williamson MK, Antón-Plágaro C, Shoemark DK, Simón-Gracia L, Bauer M, Hollandi R, Greber UF, Horvath P, Sessions RB, Helenius A, Hiscox JA, Teesalu T, Matthews DA, Davidson AD, Collins BM, Cullen PJ, Yamauchi Y. 2020. Neuropilin-1 is a host factor for SARS-CoV-2 infection. Science 370:861–865.

135. Finkel Y, Mizrahi O, Nachshon A, Weingarten-Gabbay S, Morgenstern D, Yahalom-Ronen Y, Tamir H, Achdout H, Stein D, Israeli O, Beth-Din A, Melamed S, Weiss S, Israely T, Paran N, Schwartz M, Stern-Ginossar N. 2021. The coding capacity of SARS-CoV-2. Nature 589:125–130.

136. Bhatt PR, Scaiola A, Loughran G, Leibundgut M, Kratzel A, McMillan A, O' Connor KM, Bode JW, Thiel V, Atkins JF, Ban N. 2020. Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome. bioRxiv 2020.10.26.355099.

137. Schubert K, Karousis ED, Jomaa A, Scaiola A, Echeverria B, Gurzeler LA, Leibundgut M, Thiel V, Mühlemann O, Ban N. 2020. SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. Nature Structural and Molecular Biology 27:959–966.

138. Knoops K, Kikkert M, Van Den Worm SHE, Zevenhoven-Dobbe JC, Van Der Meer Y, Koster AJ, Mommaas AM, Snijder EJ. 2008. SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. PLoS Biology 6:1957–1974.

139. Ulasli M, Verheije MH, de Haan CAM, Reggiori F. 2010. Qualitative and quantitative ultrastructural analysis of the membrane rearrangements induced by coronavirus. Cellular Microbiology 12:844–861.

140. Snijder EJ, Decroly E, Ziebuhr J. 2016. The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing. Advances in Virus Research 96:59–126.

141. Sawicki SG, Sawicki DL. 1995. Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. Advances in experimental medicine and biology 380:499–506.

142. Sola I, Almazán F, Zúñiga S, Enjuanes L. 2015. Continuous and Discontinuous RNA Synthesis in Coronaviruses. Annual Review of Virology 2:265–288.

143. Kim D, Lee JY, Yang JS, Kim JW, Kim VN, Chang H. 2020. The Architecture of SARS-CoV-2 Transcriptome. Cell 181:914-921.e10.

144. Kopecky-Bromberg SA, Martínez-Sobrido L, Frieman M, Baric RA, Palese P. 2007. Severe Acute Respiratory Syndrome Coronavirus Open Reading Frame (ORF) 3b, ORF 6, and Nucleocapsid Proteins Function as Interferon Antagonists. Journal of Virology 81:548–557.

145. Shi C-S, Qi H-Y, Boularan C, Huang N-N, Abu-Asab M, Shelhamer JH, Kehrl JH. 2014. SARS-Coronavirus Open Reading Frame-9b Suppresses Innate Immunity by Targeting Mitochondria and the MAVS/TRAF3/TRAF6 Signalosome. The Journal of Immunology 193:3080–3089.

146. Jiang H wei, Zhang H nan, Meng Q feng, Xie J, Li Y, Chen H, Zheng Y xiao, Wang X ning, Qi H, Zhang J, Wang PH, Han ZG, Tao S ce. 2020. SARS-CoV-2 Orf9b suppresses type I interferon responses by targeting TOM70. Cellular and Molecular Immunology 17:998–1000.

147. Siu YL, Teoh KT, Lo J, Chan CM, Kien F, Escriou N, Tsao SW, Nicholls JM, Altmeyer R, Peiris JSM, Bruzzone R, Nal B. 2008. The M, E, and N Structural Proteins of the Severe Acute Respiratory Syndrome Coronavirus Are Required for Efficient Assembly, Trafficking, and Release of Virus-Like Particles. Journal of Virology 82:11318–11330.

148. Escors D, Ortego J, Laude H, Enjuanes L. 2001. The Membrane M Protein Carboxy Terminus Binds to Transmissible Gastroenteritis Coronavirus Core and Contributes to Core Stability. Journal of Virology 75:1312–1324.

149. Miller S, Krijnse-Locker J. 2008. Modification of intracellular membrane structures for virus replication. Nature Reviews Microbiology 6:363–374.

150. Snijder EJ, van der Meer Y, Zevenhoven-Dobbe J, Onderwater JJM, van der Meulen J, Koerten HK, Mommaas AM. 2006. Ultrastructure and Origin of Membrane Vesicles Associated with the Severe Acute Respiratory Syndrome Coronavirus Replication Complex. Journal of Virology 80:5927–5940.

151. Angelini MM, Akhlaghpour M, Neuman BW, Buchmeier MJ. 2013. Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles. mBio 4:e00524-13.

152. Wolff G, Limpens RWAL, Zevenhoven-Dobbe JC, Laugks U, Zheng S, de Jong AWM, Koning RI, Agard DA, Grünewald K, Koster AJ, Snijder EJ, Bárcena M. 2020. A molecular pore spans the double membrane of the coronavirus replication organelle. Science 369:1395–1398.

153. Lehmann KC, Gulyaeva A, Zevenhoven-Dobbe JC, Janssen GMC, Ruben M, Overkleeft HS, Van Veelen PA, Samborskiy D V., Kravchenko AA, Leontovich AM, Sidorov IA, Snijder EJ, Posthuma CC, Gorbalenya AE. 2015. Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. Nucleic Acids Research 43:8416–8434.

154. Kirchdoerfer RN, Ward AB. 2019. Structure of the SARS-CoV nsp12 polymerase bound to nsp7 and nsp8 co-factors. Nature Communications 10:1–9.

155. Gao Y, Yan L, Huang Y, Liu F, Zhao Y, Cao L, Wang T, Sun Q, Ming Z, Zhang L, Ge J, Zheng L, Zhang Y, Wang H, Zhu Y, Zhu C, Hu T, Hua T, Zhang B, Yang X, Li J, Yang H, Liu Z, Xu W, Guddat LW, Wang Q, Lou Z, Rao Z. 2020. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. Science 368:779–782.

156. Subissi L, Posthuma CC, Collet A, Zevenhoven-Dobbe JC, Gorbalenya AE, Decroly E, Snijder EJ, Canard B, Imbert I. 2014. One severe acute respiratory syndrome coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proceedings of the National Academy of Sciences of the United States of America 111:E3900–E3909.

157. te Velthuis AJW, Arnold JJ, Cameron CE, van den Worm SHE, Snijder EJ. 2009. The RNA polymerase activity of SARS-coronavirus nsp12 is primer dependent. Nucleic Acids Research 38:203–214.

158. Gao Y, Yan L, Huang Y, Liu F, Zhao Y, Cao L, Wang T, Sun Q, Ming Z, Zhang L, Ge J, Zheng L, Zhang Y, Wang H, Zhu Y, Zhu C, Hu T, Hua T, Zhang B, Yang X, Li J, Yang H, Liu Z, Xu W, Guddat L, Wang Q, Lou Z, Rao Z. 2020. Structure of RNA-dependent RNA polymerase from 2019-nCoV, a major antiviral drug target. bioRxiv 2020.03.16.993386.

159. Hillen HS, Kokic G, Farnung L, Dienemann C, Tegunov D, Cramer P. 2020. Structure of replicating SARS-CoV-2 polymerase. Nature 584:154–156.

160. Wang Q, Wu J, Wang H, Gao Y, Liu Q, Mu A, Ji W, Yan L, Zhu Y, Zhu C, Fang X, Yang X, Huang Y, Gao H, Liu F, Ge J, Sun Q, Yang X, Xu W, Liu Z, Yang H, Lou Z, Jiang B, Guddat LW, Gong P, Rao Z. 2020. Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase. Cell 182:417-428.e13.

161. Yin W, Mao C, Luan X, Shen DD, Shen Q, Su H, Wang X, Zhou F, Zhao W, Gao M, Chang S, Xie YC, Tian G, Jiang HW, Tao SC, Shen J, Jiang Y, Jiang H, Xu Y, Zhang S, Zhang Y, Xu HE. 2020. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. Science 368:1499–1504.

162. Imbert I, Guillemot JC, Bourhis JM, Bussetta C, Coutard B, Egloff MP, Ferron F, Gorbalenya AE, Canard B. 2006. A second, non-canonical RNA-dependent RNA polymerase in SARS coronavirus. EMBO Journal 25:4933–4942.

163. Te Velthuis AJW, Van Den Worm SHE, Snijder EJ. 2012. The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. Nucleic Acids Research 40:1737–1747.

164. Xiao Y, Ma Q, Restle T, Shang W, Svergun DI, Ponnusamy R, Sczakiel G, Hilgenfeld R. 2012. Nonstructural Proteins 7 and 8 of Feline Coronavirus Form a 2:1 Heterotrimer That Exhibits Primer-Independent RNA Polymerase Activity. Journal of Virology 86:4444–4454.

165. Zhai Y, Sun F, Li X, Pang H, Xu X, Bartlam M, Rao Z. 2005. Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer. Nature Structural and Molecular Biology 12:980–986.

166. Tanner JA, Watt RM, Chai YB, Lu LY, Lin MC, Peiris JSM, Poon LLM, Kung HF, Huang JD. 2003. The severe acute respiratory syndrome (SARS) coronavirus NTPasefhelicase belongs to a distinct class of 5′ to 3′ viral helicases. Journal of Biological Chemistry 278:39578–39582.

167. Singleton MR, Dillingham MS, Wigley DB. 2007. Structure and Mechanism of Helicases and Nucleic Acid Translocases. http://dx.doi.org/101146/annurev.biochem76052305115300 76:23–50.

168. Jia Z, Yan L, Ren Z, Wu L, Wang J, Guo J, Zheng L, Ming Z, Zhang L, Lou Z, Rao Z. 2019. Delicate structural coordination of the Severe Acute Respiratory Syndrome coronavirus Nsp13 upon ATP hydrolysis. Nucleic acids research 47:6538–6550.

169. Chen J, Malone B, Llewellyn E, Grasso M, Shelton PMM, Olinares PDB, Maruthi K, Eng ET, Vatandaslar H, Chait BT, Kapoor TM, Darst SA, Campbell EA. 2020. Structural Basis for Helicase-Polymerase Coupling in the SARS-CoV-2 Replication-Transcription Complex. Cell 182:1560-1573.e13.

170. Hao W, Wojdyla JA, Zhao R, Han R, Das R, Zlatev I, Manoharan M, Wang M, Cui S. 2017. Crystal structure of Middle East respiratory syndrome coronavirus helicase. PLoS Pathogens 13:e1006474.

171. Ivanov KA, Thiel V, Dobbe JC, van der Meer Y, Snijder EJ, Ziebuhr J. 2004. Multiple Enzymatic Activities Associated with Severe Acute Respiratory Syndrome Coronavirus Helicase. Journal of Virology 78:5619–5632.

172. Adedeji AO, Marchand B, Te Velthuis AJW, Snijder EJ, Weiss S, Eoff RL, Singh K, Sarafianos SG. 2012. Mechanism of nucleic acid unwinding by SARS-CoV helicase. PloS one 7:e36521.

173. Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LLM, Guan Y, Rozanov M, Spaan WJM, Gorbalenya AE. 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. Journal of Molecular Biology 331:991–1004.

174. Minskaia E, Hertzig T, Gorbalenya AE, Campanacci V, Cambillau C, Canard B, Ziebuhr J. 2006. Discovery of an RNA virus 3′→5′ exoribonuclease that is critically involved in coronavirus RNA synthesis. Proceedings of the National Academy of Sciences of the United States of America 103:5108–5113.

175. Bouvet M, Imbert I, Subissi L, Gluais L, Canard B, Decroly E. 2012. RNA 3′-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. Proceedings of the National Academy of Sciences of the United States of America 109:9372–9377.

176. Ma Y, Wu L, Shaw N, Gao Y, Wang J, Sun Y, Lou Z, Yan L, Zhang R, Rao Z. 2015. Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. Proceedings of the National Academy of Sciences of the United States of America 112:9436–9441.

177. Eckerle LD, Lu X, Sperry SM, Choi L, Denison MR. 2007. High Fidelity of Murine Hepatitis Virus Replication Is Decreased in nsp14 Exoribonuclease Mutants. Journal of Virology 81:12135–12144.

178. Smith EC, Blanc H, Vignuzzi M, Denison MR. 2013. Coronaviruses Lacking Exoribonuclease Activity Are Susceptible to Lethal Mutagenesis: Evidence for Proofreading and Potential Therapeutics. PLoS Pathogens 9:e1003565.

179. Agostini ML, Andres EL, Sims AC, Graham RL, Sheahan TP, Lu X, Smith EC, Case JB, Feng JY, Jordan R, Ray AS, Cihlar T, Siegel D, Mackman RL, Clarke MO, Baric RS, Denison MR. 2018. Coronavirus susceptibility to the antiviral remdesivir (GS-5734) is mediated by the viral polymerase and the proofreading exoribonuclease. mBio 9:221–239.

180. Decroly E, Ferron F, Lescar J, Canard B. 2012. Conventional and unconventional mechanisms for capping viral mRNA. Nature Reviews Microbiology 10:51–65.

181. Yan L, Ge J, Zheng L, Zhang Y, Gao Y, Wang T, Huang Y, Yang Y, Gao S, Li M, Liu Z, Wang H, Li Y, Chen Y, Guddat LW, Wang Q, Rao Z, Lou Z. 2021. Cryo-EM Structure of an Extended SARS-CoV-2 Replication and Transcription Complex Reveals an Intermediate State in Cap Synthesis. Cell 184:184-193.e10.

182. Chen Y, Cai H, Pan J, Xiang N, Tien P, Ahola T, Guo D. 2009. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. Proceedings of the National Academy of Sciences of the United States of America 106:3484–3489.

183. Decroly E, Imbert I, Coutard B, Bouvet M, Selisko B, Alvarez K, Gorbalenya AE, Snijder EJ, Canard B. 2008. Coronavirus Nonstructural Protein 16 Is a Cap-0 Binding Enzyme Possessing (Nucleoside-2′O)-Methyltransferase Activity. Journal of Virology 82:8071–8084.

184. Decroly E, Debarnot C, Ferron F, Bouvet M, Coutard B, Imbert I, Gluais L, Papageorgiou N, Sharff A, Bricogne G, Ortiz-Lombardia M, Lescar J, Canard B. 2011. Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2′-o-methyltransferase nsp10/nsp16 complex. PLoS Pathogens 7:e1002059.

185. Bouvet M, Debarnot C, Imbert I, Selisko B, Snijder EJ, Canard B, Decroly E. 2010. In vitro reconstitution of sars-coronavirus mRNA cap methylation. PLoS Pathogens 6:1–13.

186. Chen Y, Su C, Ke M, Jin X, Xu L, Zhang Z, Wu A, Sun Y, Yang Z, Tien P, Ahola T, Liang Y, Liu X, Guo D. 2011. Biochemical and structural insights into the mechanisms of sars coronavirus RNA ribose 2′-O-methylation by nsp16/nsp10 protein complex. PLoS Pathogens 7:e1002294.

187. Schlee M, Hartmann G. 2016. Discriminating self from non-self in nucleic acid sensing. Nature Reviews Immunology 16:566–580.

188. Te Velthuis AJW, Zubkova TG, Shaw M, Mehle A, Boltz D, Gmeinwieser N, Stammer H, Milde J, Müller L, Margitich V. 2021. Enisamium Reduces influenza virus shedding and improves patient recovery by inhibiting viral RNA polymerase activity. Antimicrobial Agents and Chemotherapy 65:e02605-20.

189. Van Vliet ALW, Smits SL, Rottier PJM, De Groot RJ. 2002. Discontinuous and non-discontinuous subgenomic RNA transcription in a nidovirus. EMBO Journal 21:6571–6580.

190. Lai MM, Stohlman SA. 1981. Comparative analysis of RNA genomes of mouse hepatitis viruses. Journal of Virology 38:661–670.

191. Lai MM, Stohlman SA. 1981. Genome structure of mouse hepatitis virus: comparative analysis by oligonucleotide mapping. Advances in experimental medicine and biology 142:69–82.

192. Shuman S. 1990. Catalytic activity of vaccinia mRNA capping enzyme subunits coexpressed in Escherichia coli. Journal of Biological Chemistry 265:11960–11966.

193. Deana A, Celesnik H, Belasco JG. 2008. The bacterial enzyme RppH triggers messenger RNA degradation by 5′ pyrophosphate removal. Nature 451:355–358.

194. Chakravarty AK, Subbotin R, Chait BT, Shuman S. 2012. RNA ligase RtcB splices 3′-phosphate and 5′-OH ends via covalent RtcB-(histidinyl)-GMP and polynucleotide-(3′)pp(5′)G intermediates. Proceedings of the National Academy of Sciences of the United States of America 109:6072–6077.

195. Ahn DG, Choi JK, Taylor DR, Oh JW. 2012. Biochemical characterization of a recombinant SARS coronavirus nsp12 RNA-dependent RNA polymerase capable of copying viral RNA templates. Archives of Virology 157:2095–2104.

196. Ogino T, Green TJ. 2019. RNA synthesis and capping by nonsegmented negative strand RNA viral polymerases: Lessons from a prototypic virus. Frontiers in Microbiology 10:1490.

197. Baddock HT, Brolih S, Yosaatmadja Y, Ratnaweera M, Bielinski M, Swift LP, Cruz-Migoni A, Morris GM, Schofield CJ, Gileadi O, McHugh PJ. 2020. Characterisation of the SARS-CoV-2 ExoN (nsp14ExoN-nsp10) complex: implications for its role in viral genome stability and inhibitor identification. bioRxiv 2020.08.13.248211.

198. Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, Shi Z, Hu Z, Zhong W, Xiao G. 2020. Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. Cell Research 30:269–271.

199. Gorbalenya AE, Enjuanes L, Ziebuhr J, Snijder EJ. 2006. Nidovirales: Evolving the largest RNA virus genome. Virus Research 117:17–37.

200. Fodor E, Crow M, Mingay LJ, Deng T, Sharps J, Fechter P, Brownlee GG. 2002. A Single Amino Acid Mutation in the PA Subunit of the Influenza Virus RNA Polymerase Inhibits Endonucleolytic Cleavage of Capped RNAs. Journal of Virology 76:8989–9001.

201. Ng AKL, Chan WH, Choi ST, Lam MKH, Lau KF, Chan PKS, Au SWN, Fodor E, Shaw PC. 2012. Influenza polymerase activity correlates with the strength of interaction between nucleoprotein and PB2 through the host-specific residue K/E627. PLoS ONE 7:e36415.

202. Deng T, Sharps J, Fodor E, Brownlee GG. 2005. In Vitro Assembly of PB2 with a PB1-PA Dimer Supports a New Model of Assembly of Influenza A Virus Polymerase Subunits into a Functional Trimeric Complex. Journal of Virology 79:8669–8674.

203. Fodor E, Devenish L, Engelhardt OG, Palese P, Brownlee GG, García-Sastre A. 1999. Rescue of Influenza A Virus from Recombinant DNA. Journal of Virology 73:9679–9682.

204. Kodama Y, Hu CD. 2010. An improved bimolecular fluorescence complementation assay with a high signal-to-noise ratio. BioTechniques 49:793–803.