Investigating the Mechanisms by which Influenza A Viruses Induce and Subsequently Dampen the Type I Interferon Response

Joshua Christian David Long

Lincoln College

Supervised by:

Professor Ervin Fodor

Sir William Dunn School of Pathology

University of Oxford

D.Phil in Pathology

Submission – September 2016

Table of Contents

Abbreviations

3′-OH = 3-prime hydroxyl

3P = Trimeric viral RNA-dependent RNA-polmerase

5′-ppp = 5-prime triphosphate

ADCC = Antibody-dependent cell-mediated cytotoxicity

AIF = Apoptosis-inducing factor

ANP32A = Acidic leucine-rich nuclear phosphoprotein family member A

ANP32B = Acidic leucine-rich nuclear phosphoprotein family member B

ANT3 = ADP/ATP translocase 3

β-gal = β-galactosidase control plasmid

CARD = Caspase-activation and recruitment domain

CBD = Cap-binding domain

cGAS = Cyclic-DMP-AMP synthase

CHX = Cycloheximide

CLIC1 = Chloride intracellular channel protein 1

CluH = Clustered mitochondria homolog

CPSF-30 = Cleavage and polyadenylation factor 30

CRM1 = Chromosome region maintenance 1

cRNA = Positive sense influenza virus genomic RNA

cRNP = cRNA-containing ribonucleoprotein complex

CTD = C-terminal domain

DI = Defective interfering

DRP1 = Dynamin-related protein 1

dsRNA = Double stranded RNA

EDTA = Ethylenediaminetetraacetic acid

eIF = Elongation initiation factor

EM = Electron microscopy

ETC = Electron transport chain

FRET = Fluorescence resonance energy transfer

GIT = Gastro-intestinal tract

HA = Haemagglutinin

HCV = Hepatitis C virus

HD = Helicase domain

HIV = Human immunodeficiency virus

HPAIV = Highly pathogenic avian influenza virus

IFITM3 = Interferon induced transmembrane protein 3

IFN = Interferon

IFN-luc = Luciferase plasmid under the control of the IFN-β promoter

IFNAR = Type I interferon receptor

IFNLR1 = Interferon lamda receptor 1

IKKε = Inhibitor-κB kinase ε

IMS = Inter-membrane space

IRF = Interferon regulator factor

ISGs = Interferon-stimulated genes

ISRE = Interferon-stimulated response elements

LGP2 = Laboratory of genetics and physiology 2

M1 = Matrix protein 1

M2 = Matrix protein 2

MAVS = Mitochondrial antiviral signalling protein

MDA-5 = Melanoma differentiation associated protein 5

Mfn = Mitofusin

MIM = Mitochondrial inner membrane

MITOC = Microtubule organising complex

MMP = Mitochondrial membrane potential

MOM = Mitochondrial outer membrane

mRNA = Messenger RNA

MTS = Mitochondrial targeting signal

Mx = Myxovirus resistance gene

NA = Neuraminidase

NEP = Nuclear export protein

NES = Nuclear export signal

NF-κB = Nuclear factor kappa-light-chain-enhancer of activated B cells

NK = Natural killer cells

NLRP3 = NACHT, LRR and PYD domains-containing protein 3

NLRX1 = NOD-like-receptor family member X1

NLS = Nuclear localisation signal

NMR = Nuclear magnetic resonance spectroscopy

NP = Nucleoprotein

NPC = Nuclear pore complex

NS1 = Non-structural protein 1

OPA1 = Optic atrophy protein 1

PA = Polymerase acidic

PABP-II = PolyA-binding protein II

PAMP = Pathogen associated molecular pattern

PB1 = Polymerase basic 1

PB2 = Polymerase basic 2

PKR = Protein kinase R

PMSF = Phenylmethylsulfonyl fluoride

Pol II = RNA polymerase II

PRR = Pattern recognition receptor

PTPC = Mitochondrial permeability transition pore complex

RE = Recycling endosome

RIG-I = Retinoic acid inducible gene 1

RNP = Ribonucleoprotein complex

ROS = Reactive oxygen species

SA = Sialic acid

SINQ = Normalised spectral index quantification

snoRNAs = nascent small nucleolar RNAs

snRNAs = short nuclear RNAs

STING = Stimulator of interferon genes

svRNA = Short viral RNA

TAP = Tandem affinity purification

TBK1 = TANK-binding kinase 1

TLR = Toll like receptor

TM = Transmembrane domain

TRAF = TNF receptor associated factor

TRIM22 = Tripartite motif containing 22

TUFM = Tu translation elongation factor, mitochondrial

URT = Upper respiratory tract

VDAC = Voltage dependent anion-selective channel

vRNA = Negative sense influenza virus genomic RNA

vRNP = vRNA-containing ribonucleoprotein complex

wt = Wild-type

Acknowledgements

I would like to express my gratitude to my supervisor, Professor Ervin Fodor, for his support over the last four years. Furthermore, I would like to thank Aartjan Te Velthuis and Marian Killip who have provided me with guidance and assistance at key junctions of my D.Phil.

I would like to acknowledge the help of the following, who took time to teach me a number of the assays which I used within this thesis: Edward Hutchinson for guidance in mass spectrometry data analysis, Marian Killip for sharing her experience in interferon assays, Errin Johnson for training in electron microscopy and fabulous company during my time in the electron microscopy facility, Itziar Serna Martin for taking the time to train me in primer extension assays, Maria Jose Oliva Martin for supervision on the flow cytometer and finally Jane Sharps, without whom none of our projects would be able to function. I would also like to acknowledge those who directly have contributed experimental time to the work within this thesis: Benjamin Thomas and Svenja Hester for processing samples by mass spectrometry and Aartjan te Velthuis for helping me to perform the preliminary work on avian adapted polymerases in the weeks before the submission of this thesis.

I would also like to thank all of those who have worked in the Fodor lab during my time in the Dunn School and made it an enjoyable place to work. Finally, I would like to thank my partner María José, whose unfaltering support through both the good times and the hard times has kept me focussed and on track, and without whom I would not have progressed nearly as far. Many thanks for taking so much time to discuss ideas, thoroughly proof read this thesis and for freely giving many weekends and holidays to help with various other parts of this project. It will always bring my great happiness remembering our time working side by side in OMPI2.

Funding for my studies was provided by the Medical Research Council (MRC), the Clarendon Fund and Lincoln College.

Investigating the Mechanisms by which Influenza A Viruses Induce and Subsequently Dampen the Type I Interferon Response

Joshua Long, Lincoln College, University of Oxford, *D.Phil in Pathology, Trinity 2016*

Abstract

The activation of antiviral innate immune responses by influenza A viruses plays a keys role in influenza virus infections. The recognition of viral RNA by the pattern recognition receptor RIG-I leads to the induction of the expression of Type I Interferons, which induce an antiviral state, and to the activation of the inflammasome, which leads to inflammation. This thesis focuses on investigating both the mechanisms by which RIG-I is activated by influenza A viruses and the mechanisms by which influenza A viruses dampen the subsequent interferon response. In this thesis, I provide evidence to support the hypothesis that the depletion of free nucleoprotein within infected cells promotes the accumulation of immunostimulatory, internal deletion viral RNAs of approximately 50 to 75 nucleotides in length, and that these may be key in the activation of RIG-I. Furthermore, I provide preliminary evidence that the polymerases of avian adapted influenza virus strains may generate higher levels of these immunostimulatory RNAs, and speculate that this may be important in the innate immune over-activation and subsequent cytokine storm which occurs during human infections with a number of avian-adapted influenza A viruses.

Furthermore, I investigate previous reports that the localisation of the PB2 subunit of the influenza virus trimeric polymerase to the mitochondria is important in dampening the expression of interferon by seeking to characterise the functions of mitochondrial localised PB2. I demonstrate that mitochondrial localising PB2 is imported into the matrix space and identify a number of matrix-localising interactors of PB2. This work demonstrates that mitochondrial PB2 is not correctly localised to interact with the mitochondrial antiviral signalling protein MAVS, which was previously believed to be the manner by which PB2 dampens the induction of interferon expression. However, I speculate that the interaction of PB2 with the matrix protein translation elongation factor Tu may mediate its dampening activity.

Chapter 1 - Introduction

Influenza viruses likely crossed into humans from an animal reservoir at the advent of farming. Many potential influenza pandemics have been identified throughout history by identifying influenza-like symptoms from historical texts [1]. Before the identification of pathogens as the causative agent, the source of influenza disease was unknown. In fact, the term influenza derives from the Latin word for 'influence', reflecting the early belief that the disease was caused by the influence of the stars. Since the discovery of the causative agent, influenza viruses, the understanding of the epidemiology and molecular biology of influenza disease has flourished. Furthermore, work into understanding host immune responses against viruses in general has underpinned an increased knowledge of the immune responses against influenza viruses. Despite this, there are still many unanswered questions in the complicated field of influenza biology, which need to be answered. It is, therefore, essential that we continue to fill these gaps to increase our ability to predict and generate treatments against a disease which is widely believed to be one of the most economically important and potentially devastating in existence. This opinion is shared by the UK government, which has placed influenza pandemics on a list of the highest risk events to 'human welfare' based on the probability of occurrence and the potential impact. Furthermore, influenza pandemics are one of only two civic emergencies (along with 'catastrophic terrorist attacks') which are ranked in the highest category of 'impact'; demonstrating the government's recognition of the potential of the disease to cause unprecedented devastation. The reason for its placement in this category is illustrated by the 1918 influenza A virus pandemic, which caused 50 million deaths world-wide in 12 months – similar to the combined death toll of from HIV since 1981 (34 million) and world war one (≈ 17 million). Furthermore, at present, infections of humans with H5N1 highly pathogenic avian influenza viruses (HPAIVs), which occur via exposure to

infected poultry, are fatal in approximately 60% of cases [2]. Thus, if a HPAIV strain gains the ability to spread robustly between humans, the potential death toll from the resulting pandemic could very realistically be higher than that of the 1918 pandemic.

The pathologies of fatal infections with the 1918 'Spanish flu' strain and HPAIV strains, in fact, proceed fairly similarly [3]. Both lead to an uncontrolled innate immune response, characterised by high levels of cytokines; termed a 'cytokine storm'. This cytokine storm causes pulmonary oedema/pneumonia, resulting in decreased lung function and patients effectively drowning in their own inflammatory exudate. Thus, understanding how influenza viruses interact with the innate immune system may be a key step in understanding how to control severe influenza pathology. In this thesis, I thus focus on how influenza viruses trigger and subsequently limit the host innate immune response.

1.1 Orthomyxoviridae Family

1.1.1 Genera

Influenza viruses are members of the *Orthomyxoviridae* family of viruses. The characteristic feature of viruses of this family is that their genomes, which are formed of negative sense single stranded RNA, are distributed over multiple distinct RNAs, termed genome segments.

The influenza viruses make up three genera of this family, namely influenza A, B and C. Despite being closely related, the influenza A, B and C virus genera infect different hosts. Viruses of the influenza A genus can infect aquatic birds, poultry and a wide range of mammalian species, including humans. Viruses of the influenza B and C genera are also human pathogens, however, are found to infect a far narrower range of host species. In this thesis, I will focus on the influenza A virus genus [3].

1.1.2 Influenza A Virus

1.1.2.1 Classification

Influenza A virus is principally a virus of aquatic birds. In this host, influenza A virus infects the digestive tract and thus is transmitted by the faecal oral route. This genus of virus, however, can infect a range of other birds, including poultry. The A genus is also capable of infecting a wide range of mammals from humans to whales [3]. In mammalian hosts, however, the virus mainly infects the respiratory tract and thus is transmitted by the airborne route.

Influenza A viruses are subdivided into subtypes based on their viral surface glycoproteins, haemagglutinin (HA) and neuraminidase (NA). These viral proteins are divided into types; with 18 known HA types and 11 known NA types. Thus, a virus which encodes a type 1 HA and a type 1 NA is termed a H1N1 virus. Despite the large number of possible HA and NA combinations, the vast majority of human influenza A disease during documented history has been caused by only three combinations. These are the H1N1, H2N2 and H3N2 subtypes, of which the H1N1 and H3N2 are the two current, widely circulating subtypes in the human population [4].

1.1.2.2 Denomination

Influenza viruses are named based on their genus, the host organism from which they were isolated, their place of isolation, their isolate number and their year of isolation (e.g. Genus/species/place/number/year). For example, A/Black Chicken/Jiangxi/10129/2014 denotes the 10129th influenza A virus strain isolated from a chicken in Jiangxi in 2014. This nomenclature varies slightly with human strains, in which the species is not stated. The virus which will be used in the majority of the experiments in this thesis is A/WSN/33; a H1N1

laboratory strain, originally isolated from a human, which is well adapted to growth in cell culture systems.

1.2 Virion

1.2.1 Introduction

Influenza disease is triggered by the successful transmission of influenza A virions from an infected host to a susceptible host by the routes described in **section 1.1.2.1**. In this section, before describing in detail the influenza A virus life cycle, the structure of the influenza A virion is detailed. These virions can be described in simple terms as an outer, viral glycoprotein-containing lipid membrane surrounding a virion core, which consists of the viral genome segments associated to a viral protein matrix. The specific structures which make up the virion will be further elaborated upon below (**section 1.2.2**) and are displayed in Figure 1 – Part A.

1.2.2 Structure

1.2.2.1 Envelope

The outer layer of the influenza A virion is formed by a lipid membrane, which is sourced from the host plasma membrane during the process of budding (**section 1.3.7**). The viral proteins HA, NA and Matrix protein 2 (M2), which localise to the plasma membrane of infected cells, are incorporated into the virion during this process [5].

The HA protein is found in the virion in homo-trimeric complexes whilst the NA protein assembles into homo-tetramers. HA and NA are integral membrane proteins due to the fact that they contain a transmembrane (TM) domain, which anchors them into the viral

membrane. The presence of a TM domain means that both HA and NA have internal and external domains; both of which appear to have roles in virion structure and function. The external region of HA and NA consists of an extended domain (termed the stem/stalk domain), which links the TM domain to a head domain [6, 7, 8] (Fig. 1 A). Due to this extended structure, the HA and NA proteins are visible via electron microscopy (EM) analysis of virions and appear as spike structures at the virion surface [9]. It is the head domain of HA and NA that contain the active site residues for receptor binding and sialic acid cleavage, respectively (activities explained in **section 1.3.1**) [8]. The M2 protein is also an integral membrane protein and is found as tetramers in the virus envelope (Fig. 1 A).

1.2.2.2 Virion Core

The viral protein M1 is found within the virion core, directly beneath the lipid membrane. M1 exists as a helical oligomer which is found around the inner face of the viral envelope and thus forms a 'net' structure under the membrane [10]. M1 directly interacts with the membrane along with the internal regions of HA, NA and M2 [11, 12]. These interactions are important during the virus budding process **(see section 1.3.7)** and likely act to give strength to the virion structure [5] (Fig. 1 A)**.**

As well as M1, the viral genomic, negative-sense RNA segments, which are termed viral RNAs (vRNAs), are found within the core of the virion. These viral RNAs, however, are not found naked, but are bound by a number of viral proteins in a complex termed the ribonucleoprotein complex (RNP). These complexes are further denoted vRNPs if the RNA contained within the complex is vRNA. The vRNP complex is formed by the viral trimeric

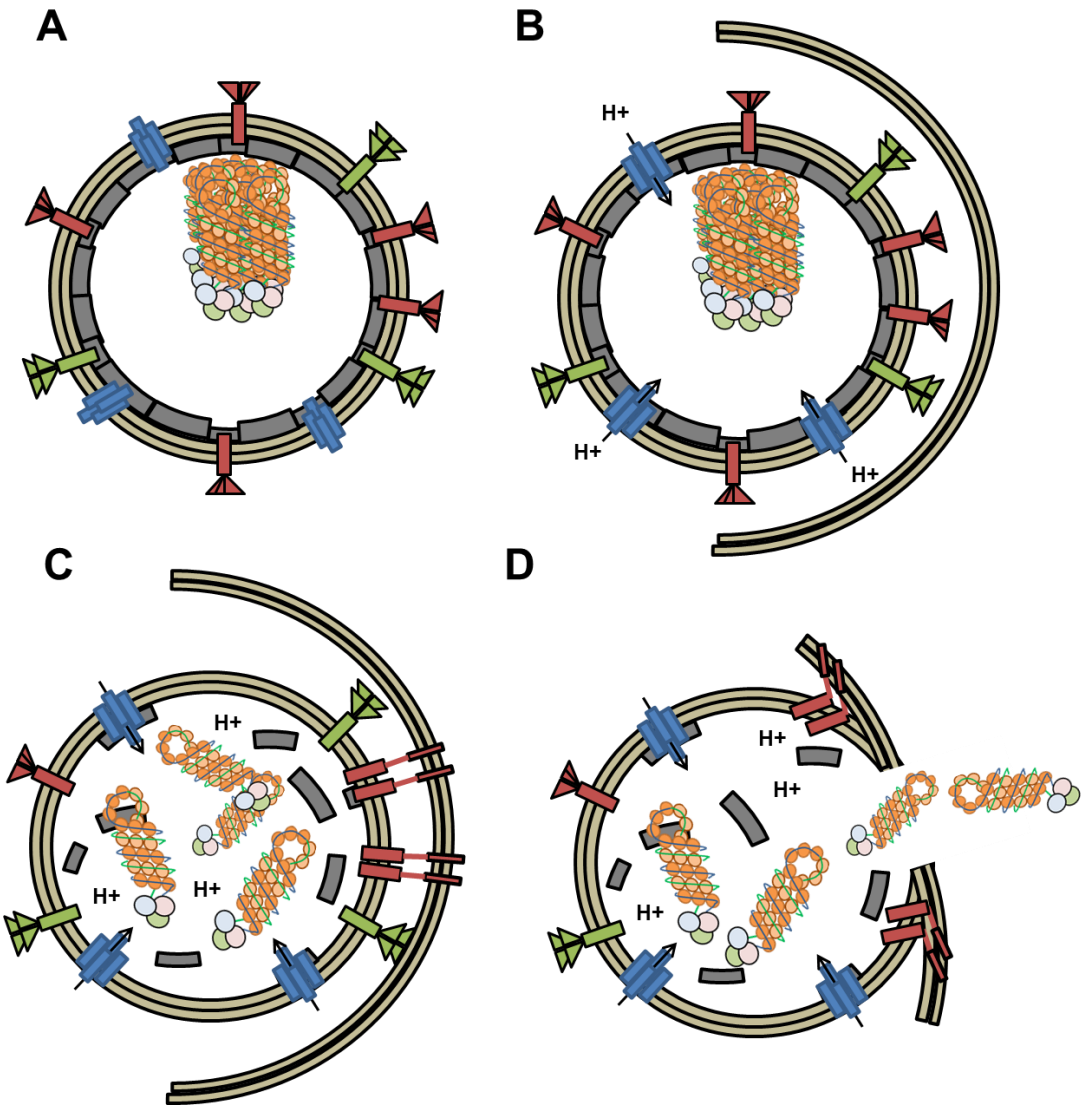


Figure 1 – (A) Structure of the influenza A virus virion - (A) HA trimer (red), NA tetramer (green) and M2 tetramer (blue) are found in the virus envelope (brown). M1 (grey) forms a helical layer under the envelope; making interactions with the membrane, HA tails and NA tails. RNPs are immobilised in an ordered octomer via binding to the M1 layer. **(B-D) RNP release from the virion and endosome** – (B) Acidification of the endosome leads to opening of the M2 channel, allowing acidification of the virion core. (C) Decreased pH induces i) dissociation of the M1 layer, ii) RNP dissociation from M1 and iii) exposure of the HA 'fusion peptide' which inserts into the endosome membrane. (D) HA pulls the endosomal and virion membranes together, leading to their fusion. This allows the release of RNPs into the cytoplasm.

RNA-dependent RNA-polymerase (3P) (which itself consists of the Polymerase Basic 1 (PB1), Polymerase Basic 2 (PB2) and Polymerase Acidic (PA) subunits), the viral nucleoprotein (NP) and the vRNA. A more detailed explanation of the RNP structure will be given in **section 1.3.3.** The RNPs within the virion interact with the M1 matrix layer and are held together in an ordered structure, potentially at one pole of the virion (Fig. 1 A) [10, 13]. The number of RNPs contained within an influenza A virion appears to be eight**;** one of each genome segment, indicating that the incorporation of vRNPs into virions is a regulated process [13].

1.2.2.3 Non-canonical Virion Structure

Influenza A viruses in cell culture are mainly found to bud as spherical virions, with the content described above. However, it is becoming increasingly evident that virions are far more complex than depicted above and can exist in a number of different states. The ways in which virions can vary from the canonical structure are explained in the following sections.

1.2.2.3.1 Morphology

Influenza A virions bud from infected cells in multiple different morphologies. The two principal morphologies are spherical virions, which are around 120 nm in size, and filamentous virions, which can be up to 30 µm in length (although intermediate length structures are also observed, which are termed to be of bacilloform morphology) [5]. The distinct roles of the different types of morphology are unknown, however, it appears that filamentous viruses have a specific role within *in vivo* infections. This hypothesis is driven by the fact that viruses from clinical isolates initially produce filamentous virions, however this ability is lost upon passage within embryonated chicken eggs or cell culture [5, 14].

Furthermore, 'spherical-only' virus strains gain the ability to produce filaments after *in vivo* infections [14]. This supports the idea that there is a specific selection for filamentous viruses in *in vivo* infections, but that producing such virions has an increased cost and thus they are selected against in cell culture [5].

1.2.2.3.2 Further Virion Content

Recent studies into virion composition have revealed that influenza A virions are also more variable in protein content than the description given in **sections 1.2.2.1 and 1.2.2.2**. Firstly, virions contain a wide range of host proteins, though the function of these proteins is still unclear [15]. Given that many of these host proteins occur at sub-stochiometric levels, there is likely host protein heterogeneity between virions. Furthermore, it has been shown that the influenza virus proteins non-structural protein 1 (NS1) and non-structural protein 2 (NS2) (otherwise termed nuclear export protein – NEP), which as their name suggests were initially believed to not be incorporated into virions, are in fact found in purified virions [15, 16]. The role of the virion NS1 is unknown, however it is possible that it acts to inhibit the anti-viral responses against 'incoming vRNPs' directly after infection (NS1 activities are explained in **section 1.6.2.1**).

1.2.2.3.2 Deficient virions

Interestingly, not all virions contain a full-complement of vRNPs, despite the apparent controlled packaging of vRNPs (**section 1.2.2.2**). Firstly, in some cases virions may package an incorrect number of vRNPs, leading to virions which do not contain the full complement of vRNPs [17]. Secondly, in some situations, incorrectly replicated vRNPs which have undergone large deletions within the vRNA they contain, are incorporated into virions. Due to these deletions, these viruses are normally deficient in their ability to express a full

complement of viral proteins. Thus, they can only productively replicate in the presence of a wild-type (wt), 'helper' virus, which provides the protein which the virus lacks [18]. Virions, which contain a short, internal-deletion vRNA, are termed defective interfering (DI) viruses due to their ability to disrupt the replication of a co-infecting wt virus [18]. DI viruses will be further explored in **section 1.6.1.4**.

1.3 Influenza Virus Life Cycle

1.3.1 Infection and Entry

The primary type of cells infected by influenza A virus are the epithelial cells of the upper respiratory tract (URT) or gastro-intestinal tract (GIT). In water fowl and poultry, influenza A virus infects principally epithelial cells of the GIT and so the virus is transmitted via the faecal-oral route. In humans, however, influenza A virus principally infects the epithelial cells of the URT and thus transmission occurs via an airborne or fomite-mediated route [3]. When an infecting virion enters the URT, it encounters the epithelial cells of the URT. The viral protein HA, which is initially in a HA0 precursor state, potentially undergoes protease mediated cleavage at this point, resulting in the formation of the HA1 and HA2 polypeptides. Without this cleavage, HA is unable to mediate virus entry [8]. This fact has been hypothesised to be key in restricting human-adapted influenza A virus strains to infections of the URT, as this is believed to be the site of localisation of the enzyme which cleaves HA0 [19, 20]. The site and timing of HA0 cleavage is controversial, however, with endosomal cleavage of HA0 being documented in some cell types [21]. Furthermore, a property of the HA0 molecules of many highly pathogenic strains is their ability to be cleaved by ubiquitously expressed, intra-cellular proteases. This allows them to infect multiple cells types and thus cause systemic infections [22].

The first role of HA in virus entry is to mediate the processes of entry receptor binding and virus internalisation. Specifically, HA recognises sialic acid (SA) on sialylated glycoproteins found on the apical face of the epithelial cells. HA binding to these glycoproteins leads to the internalisation of virions via receptor-mediated endocytosis. Once inside the cell, the internalised virions are trafficked within endosomes, which are gradually acidified. This acidification process leads to multiple conformational changes in the proteins of the virion, leading to the release of vRNPs into the host cytoplasm [3]. The various stages of this process (Fig. 1) are shown below:

- **M2-mediated acidification of the virion core**: As stated in **section 1.2.2.1**, the M2 protein is found in the virion membrane as homo-tetramers. Each of the monomers in the tetramer contributes an α-helix to form a channel across the viral membrane [23]. The channel in its basal state is found with the four α-helices tightly packed together, allowing an interaction between four key tryptophans in the channel. These four tryptophans, which are said to form a 'tryptophan gate', block the channel [23]. However, when a virion is exposed to a low external pH, as occurs in the acidifying endosome, specific histidine residues within the channel become protonated. The resultant positively charged histidines on the four helices then repel each other, disrupting both the tight helix packing and the tryptophan gate. This opens the pore, allows the passage of protons across the virion membrane and leads to acidification of the virus interior (Fig. 1 B) [23].

- **Release of vRNPs from the M1 matrix:** Within the non-acidified virion core, vRNPs are bound to the envelope-associated M1 matrix. Thus, in order for vRNPs to be released into the cytoplasm, the interaction between M1 and the vRNPs must be broken. It has been shown that an acidic environment leads to the dissociation of vRNPs from M1 (Fig. 1 C). Therefore, it appears that the opening of the M2 channel

indirectly triggers vRNP release by allowing acidification of the virion core [24, 25]. This is supported by EM analysis of virions by *Calder et* al. which shows M1 dissociation from the virion membrane under conditions of decreased pH (Fig. 1 C) [10]

- **HA-mediated fusion of the virion envelope and the endosomal membrane**: Exposure of HA to acidic conditions also leads to a drastic conformational change in the HA polypeptides. This conformational change can only occur if HA0 has been cleaved into HA1 and HA2. The conformational change exposes a hydrophobic region of HA2, which is called the fusion peptide. The fusion peptide, which is found at the tip of the HA spike under acidic conditions, inserts into the endosomal membrane (Fig. 1 C). A second conformational change then causes a constriction of the HA structure, pulling the membranes of the virion and endosome together (Fig. 1 D) [8]. This results in the fusion of the endosomal and virion membranes. The resultant pore between the virion core and the cytoplasm allows the release of the vRNPs into the cytoplasm (Fig. 1 D).

1.3.2 Nuclear Import

Once in the cytoplasm, vRNPs are trafficked to the nucleus, which is the site of vRNA replication and transcription. Nuclear localisation signals (NLS) on the NP molecules of a vRNP lead to the association of vRNPs with α-importins and the subsequent import of the vRNPs into the nucleus [26, 27]. In the nucleus, the vRNPs undergo the process of replication and transcription to generate progeny vRNPs and viral mRNAs, respectively. In order to understand the processes of transcription and replication, it is first necessary to understand the structure of the vRNP. In the following section, the structure of the vRNP complex will

be explained (**section 1.3.3**), before elaborating upon the processes of transcription (**section 1.3.4**) and genomic RNA replication (**section 1.3.5**).

1.3.3 vRNP Structure

As stated in **section 1.2.2.2**, vRNPs consist of a vRNA bound by a single 3P and multiple NP molecules. In this section, I firstly detail the structure of the vRNA and how this is affected by the binding of 3P and NP. Secondly, I detail the specific structure of 3P.

1.3.3.1 3P Bound vRNA

vRNAs are of negative sense and, unlike host mRNAs, are neither capped nor poly-adenylated at their 5′ and 3′ ends, respectively. Instead, the 5′ end of vRNA contains a 5′ triphosphate (5′-ppp) whilst the 3′ end contains a 3′ hydroxyl (3′-OH) group. The terminal 12 nt and 13 nt of the 3′ and 5′ regions of influenza virus vRNAs, respectively, are highly conserved. This is the case between the different vRNA segments and between the different influenza virus genera [28]. Furthermore, the termini of the 3′ and 5′ regions of the vRNA are partially complementary, which has led to the hypothesis that free vRNA in solution forms a double stranded (dsRNA) structure, termed the ‘panhandle’ (Fig. 2). This idea is supported by nuclear magnetic resonance spectroscopy (NMR) structures of RNAs containing the vRNA 5′ and 3′ regions, which form panhandle duplexes in solution [29, 30].

The viral RNAs, however, are not normally found naked within the cell, but are bound at their 5′ and 3′ ends by 3P. The binding of vRNA to 3P has been previously hypothesised to lead to the separation of the terminal 10/11nt of the 5′ and 3′ ends. Recent data using Fluorescence Resonance Energy Transfer (FRET) had indicated that the termini may adopt a ‘corkscrew’ conformation [31]; however, a more recently solved structure

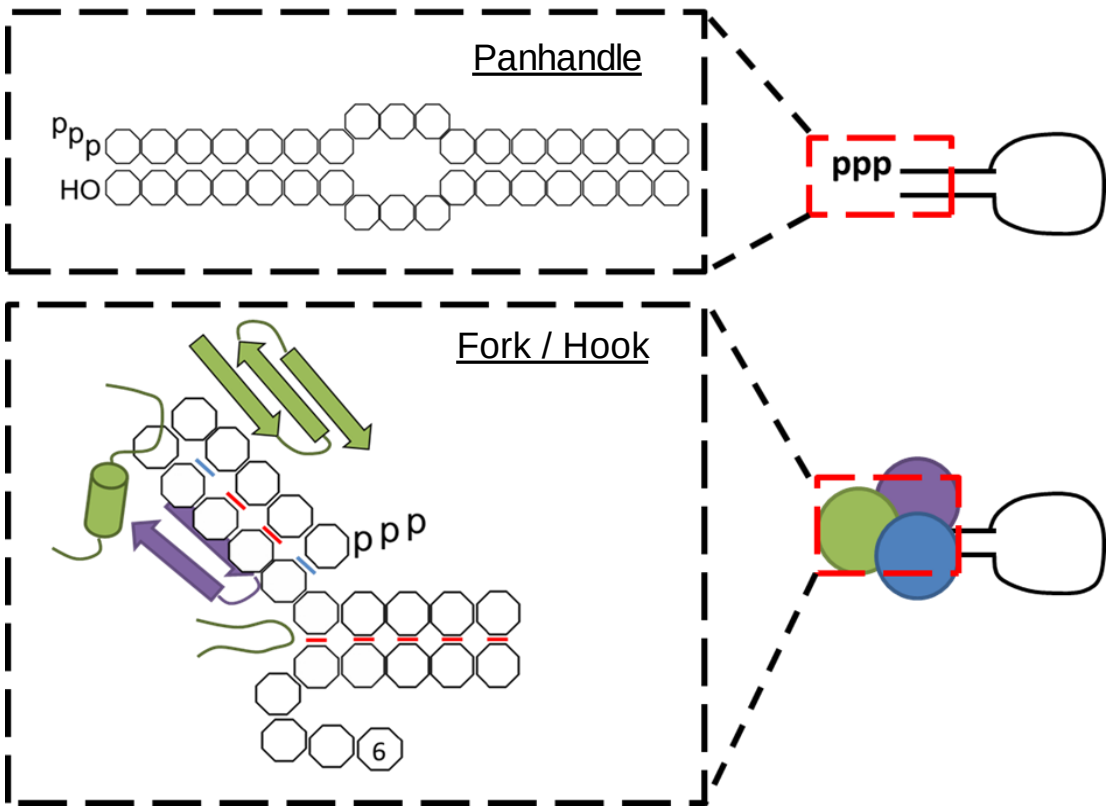


Figure 2 – vRNA 5' / 3' termini conformations – The extreme 5' and 3' termini of influenza vRNAs and cRNAs are partial complementary. These termini are believed to have different conformations in their 3P bound and 3P free form. 3P is denoted as three circles consisting of the three subunits (PB2 – blue, PB1 – green, PA – purple). **Panhandle -** In their free form, vRNAs have been seen by NMR to form 'panhandle' structures, with a blunt ended termini containing a 5'ppp. The presence of mismatches with the sequence leads to bulges within this terminal region. **Fork / Hook –** Crystal structures of vRNA bound 3P have shown that, when bound to 3P, the 5' and 3' termini of the vRNA are separated into two binding pockets on the surface of 3P. Base pairing within the 5' termini forms a hairpin structure. This structure is stabilised by multiple interactions between the RNA and regions of PA and PB1. The terminal nucleotides of the 3' termini are not resolved in the structure; the number 6 denotes the 6th nucleotide from the 3' termini. Red lines represent Watson-Crick interactions whilst blue lines represent non-Watson-Crick interactions.

of vRNA bound 3P shows that the RNA in fact adopts a 'fork/hook' conformation [32] (Fig. 2). In this conformation, the terminal 10 and 9 residues of the 5′ and 3′ ends respectively are separated (Fig. 2). The 5′ terminus forms a hairpin-like structure which is stabilised by Watson-Crick and non-Watson-Crick residues base pairs as well as by interactions with a 5′ binding pocket formed by regions of PA and PB1 [32] (Fig. 2). In this structure, the 3′ terminus of the vRNA is not completely visible [32]. It is, however, visible in a second structure in which it binds a shallow pocket on the surface of 3P instead of being inserted into the 3P active site [33].

1.3.3.2 NP Bound vRNA

Figure 3A shows a vRNA in which the 3′ proximal and 5′ proximal halves of the vRNA coloured have been coloured differently (Fig. 3 A). Alongside 3P binding to the vRNA termini (Fig. 3 B), vRNA is bound by NP. NP binds to the vRNA 'loop region', which in this thesis is used to denote the vRNA region which is not bound by 3P. The NP molecules within the vRNP make a number of different interactions which contribute to the final vRNP structure [34]. Figure 3B gives a diagrammatic representation of these interactions:

- NP is believed to make an interaction with components of 3P [35, 36]. This interaction may be important for the formation of the RNP structure.
- The first NP:NP interaction is a linear interaction between neighbouring NP molecules on the v/cRNA (Fig. 3 B). Specifically, the 'tail-loop' structure of one vRNA-bound NP binds the 'insertion groove' of the neighbouring NP molecule [34, 37].
- NP molecules have a positively charged region on their surface, which is essential for RNA binding [34, 38].The binding of vRNA to NP is likely mediated by electrostatic interactions between the negatively charged RNA backbone and the positively

charged residues in this region (Fig. 3 B). The position of vRNA in the vRNP has not been structurally determined, due to limitations in the resolution of present cryo-EM structures of vRNPs, however, it likely wraps around the outside of the NP structure [39, 40]. This hypothesis is supported by the fact that the vRNA in vRNPs is sensitive to RNase degradation [41].

- The second NP:NP interaction in the vRNP occurs between NP molecules on different regions of the vRNA. Given that the 3′ and 5′ regions of the vRNA are both bound by 3P, the two termini of the vRNA are in close apposition. The NP molecules on the 3′ proximal half of an NP-bound vRNA form an 'inter-strand' interaction with the NP molecules on the 5′ proximal half in a tail loop/insertion groove independent manner (Fig 3. B). This causes NP-bound vRNAs to form into a double helical structure (Fig. 3 C).

Together, these interactions contribute to the final structure of the vRNP. Multiple groups have determined the structure of the vRNP by cryo-EM studies [39, 40]. It has been found that the majority of the NP bound-region of vRNA forms into a double helical filament, with two antiparallel strands of NP/vRNA running alongside each other (Fig. 3 C). At the end of this rod, a loop of NP-bound vRNA can be seen, while the other end contains 3P. The structure of the vRNP, and its importance in the induction of interferon, is further explored in chapter 5.

1.3.3.3 Viral Trimeric Polymerase

The structure of 3P has been recently solved by multiple groups and has provided important insights into the mechanism of replication and transcription [32, 33, 42]. In this section, the overall structure of 3P will be briefly described, whilst the importance of key domains in the polymerase subunits will be explained, along with their corresponding functions, in later sections.

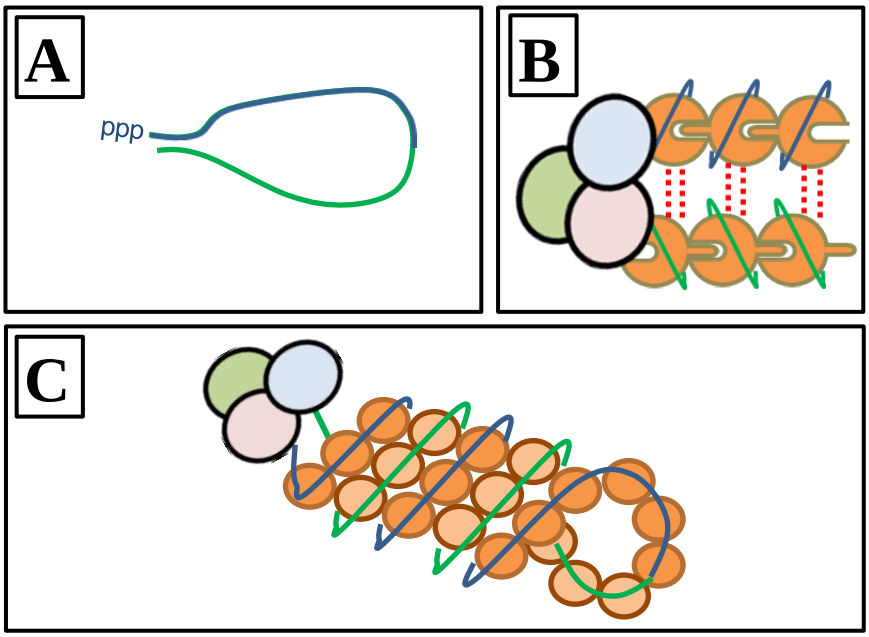


Figure 3 – The structure of RNPs and the interactions between the RNP components – (A) Free v/cRNA forms a panhandle structure due to the association of the 5′ and 3′ termini. The RNA has been coloured in blue and green to indicate the 5′ and 3′ halves of the RNA, respectively. (B) Interactions between 3P, NP and RNA within an RNP. 3P is denoted as three circles representing the three subunits (PB2 – blue, PB1 – green, PA – pink). NP is denoted as orange/cream circles. 3P binds to the 5′ and 3′ termini of the RNA. RNA is wrapped around the outside of the NP molecules. The terminal NP molecules are believed to make interactions with 3P. NPs also interacts linearly with neighbouring NPs via the binding of a 'tail loop' structure into an insertion groove. NPs on the 5′ proximal RNA region interact with NP molecules on the 3′ proximal RNA region. (C) The NP bound RNA within an RNP is found in a helical filament arrangement, with the 5′ and 3′ proximal RNA regions in close apposition. At one end of the helical region, there is a loop of NP-bound RNA whilst 3P is bound at the other.

The polymerase core is formed by PB1, the PB2 N-terminus and the C-terminus of PA, and is found in a similar arrangement in all of the resolved influenza virus 3P structures [32, 33, 42, 43]. Within this core, the extreme N-terminal region of PB2 (which is further studied within the first part of this thesis) mediates a specific interaction with the C-terminus of PB1. The active site of 3P resides in this core region and is formed by PB1.

Despite having an apparently constant core region, the arrangement of other domains in 3P varies between the different structures. Specifically, the PB2 cap-binding domain (CBD), PB2 627-domain and PA endonuclease domain are differently orientated in the different structures [32, 33, 42]. This has led to the suggestion that conformational rearrangements of these 'flexible' domains are important for 3P function (explored further in **section 1.3.4**).

1.3.4 Transcription

Each of the 8 segments of vRNA encodes at least one viral protein. The proteins encoded by each genome segment are: 1/ **PB2** and *PB2-S1*, 2/ **PB1,** *PB1-F2* and PB1-N40, 3/ **PA**, *PA-X*, PA-N155 and PA-N182, 4/ **HA**, 5/ **NP**, 6/ **NA**, 7/ **M1**, **M2** and M42, 8/ **NS1, NEP** and NS3 (**bold** denotes the main proteins whilst *italics* denotes the auxiliary proteins which will be discussed in this thesis). In order to express these viral proteins, mRNAs encoding the viral protein ORFs must first be transcribed from vRNAs (Fig. 4). This process of transcription is performed by 3P. The binding of 3P to both the 5′ and 3′ vRNA termini is required for 3P's transcriptional activity and so these vRNA termini together are termed the viral 'promoter' [44]. Transcription starts with 3P sourcing short 7-methyl-guanosine capped primers from host nascent transcripts of not just host nascent-mRNAs but also nascent short nuclear RNAs (snRNAs) and nascent small nucleolar RNAs (snoRNAs) [45] via a process termed 'cap snatching'. To achieve this, vRNPs localise to host RNA polymerase II (Pol II) via binding to

its C-terminal domain (CTD) [46]. This interaction is selective for Ser-5 phosphorylated CTD, which is a marker of initiating Pol II, and thus likely brings 3P in close apposition to capped, short nascent RNAs [46, 47]. The process of cap snatching occurs initially by the binding of the cap-structure on the nascent transcript to a cap binding site found on the CBD of PB2. The nascent RNA is then cleaved by 3P in a process mediated by the PA endonuclease domain [48, 49]. This generates short primers of 8-14 nucleotides in length containing a terminal cap structure. These short primers are then used by 3P to prime transcription.

The different structures of 3P, referred to in **section 1.3.3.3**, support a model in which the process of cap-snatching is regulated by conformational changes, with RNA binding acting as a key regulator of these changes [32, 33, 42, 43]. In the influenza C virus 3P 'apo' structure, where 3P is not bound to promoter RNA, the cap-binding site is occluded [42]. However, in another structure, which is of a vRNA-bound influenza A virus 3P, the CBD can be seen closely localised to and orientated in the direction of the PA endonuclease domain [32]. This state might thus allow cleavage of a CBD-bound nascent mRNA. Finally, in the influenza B virus RNA-bound 3P structure, the PB2 CBD is rotated, orientating the capped-primer in the direction of the PB1 active site via the RNA product exit channel of 3P [33]. Overall, these observations suggest that the activation of 3P's transcriptional activity depends on vRNA promoter binding, which is supported by observations that 3P is unable to cap snatch without being bound to the 5′ region of the vRNA promoter [50].

The next stage of transcription requires the transition of the 3′ vRNA promoter end from its surface binding site (**section 1.3.3.3**) into the active site of 3P. Here, it interacts with the capped primer, which is likely inserted into the active site via the RNA product exit channel. Free nucleotides enter the active site via a positively charged 'nucleotide entry channel' in the

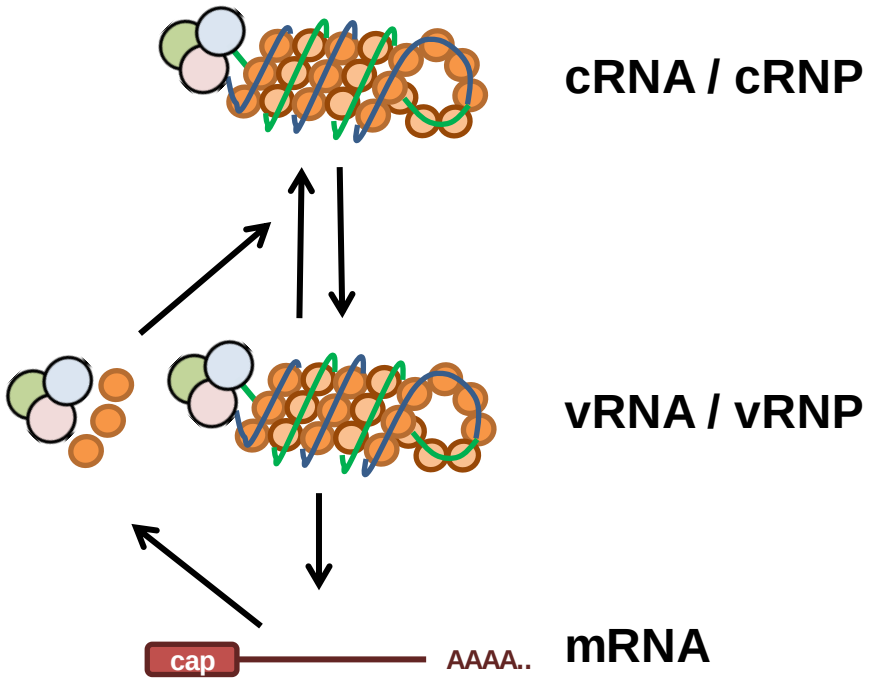


Figure 4 – Viral replication and transcription – The structure of vRNPs, cRNPs and viral mRNAs are shown. 3P is denoted as three circles representing the three subunits (PB2 – blue, PB1 – green, PA – pink). NP is denoted as orange/cream circles. 3P binds to the 5′ and 3′ termini of the RNA. vRNAs, within the context of a vRNP, act as template for transcription of viral mRNAs and generation of cRNAs. Viral mRNAs encode the viral proteins which are required to bind vRNAs and cRNAs. Without 3P and NP, cRNAs are unstable and thus cRNA accumulation requires the expression of viral proteins. cRNAs, within the context of a cRNP, act as a template for the generation of progeny vRNAs, which are packaged into vRNPs.

3P structure [32]. These free nucleotides are then used to extend the capped primer, using the 3′ vRNA termini as a template. The active site residues involved in the nucleotide condensation are found in the PB1 subunit [32]. The template vRNA is fed through the 3P structure as it is copied into viral mRNA. As 3P reaches the 5′ end of the vRNA, 3P maintains its interaction with the 5′ terminus. This creates steric hindrance which causes 3P to stutter on a run of U residues in the vRNA sequence, which leads to the generation of a poly-A tail on the viral mRNA [51]. The resultant mRNAs thus contain a 5′ cap structure and a 3′ poly-A tail. Due to this, viral mRNAs have been found to bind host mRNA binding proteins [52], which are likely both important for mRNA export and mRNA stability.

Viral mRNAs are in some cases spliced, leading to the generation of multiple viral ORFs from the same vRNA. The two principal examples of this are in segment 7, which encodes M1 and M2, and in segment 8, which encodes NS1 and NEP. Splicing has been shown to lead to the generation of other auxiliary proteins, however the only one of these which will be discussed in this thesis is termed PB2-S1 [53].

The viral mRNAs are then exported from the nucleus and translated by the host machinery. Interestingly, it has been found that viral mRNA production is not constant throughout infection but predominates early in infections. On the other hand, replication products appear to accumulate more constantly over time from early to late in infections [54]. This early peak of mRNA generation is hypothesised to occur due to the fact that Pol II is degraded during infection, which likely leads to a lack of capped-primers [55]. After translation, PB1, PB2, PA and NP are imported into the nucleus to allow the generation of progeny RNPs [26] (Fig. 4)

1.3.5 vRNA Replication

1.3.5.1 Introduction

The process of vRNA replication occurs via the generation of a positive sense intermediate termed a cRNA. This cRNA is generated by replication of vRNA (within a vRNP) and thus i) is complementary to vRNA and ii) contains partially complementary 5′ and 3′ promoter ends. The resultant cRNA is then packaged, like vRNA, into RNP structures which are termed cRNPs. These cRNPs have a similar structure to that of vRNPs (described above in **section 1.3.3.2)** [56]. The cRNAs, within the context of a cRNP, are then used as a template to generate progeny vRNA (Fig. 4). Despite accumulating to far lower levels than vRNAs and mRNAs in infection, cRNAs are the source of high levels of progeny vRNPs.

As with transcription, the process of replication is catalysed by 3P, however the mechanism of replication is distinct in many ways. The mechanism of replication and the differences relative to transcription are further explained below.

1.3.5.2 The Requirement for Non-RNP Associated Polymerase

The first major difference between transcription and replication is that, in replication, free non-RNP associated 3P has an essential role. As shown above, in transcription, it is the 3P molecule of an RNP (termed the resident 3P) which transcribes mRNA from the vRNA of said RNP. In replication, however, *Jorba et al*. have provided data that shows that free 3P is required for replication of RNPs containing an inactive resident 3P; however, it was not determined if it was required for vRNA replication, cRNA replication or both. They therefore hypothesised that free 3P, not the resident 3P, catalyses the generation of v/cRNA [57]. The free 3P which carries out the replication of the v/cRNA is termed in their model a ‘trans-acting’ polymerase. This idea is supported by cryo-EM observations of forked RNP structures, which

are hypothesised to represent a trans-acting polymerase (and the associated nascent replication product) replicating a template RNP [39]. Furthermore, the authors provide data which indicates that 3P is found at the branch points of the fork, however, a lack of resolution along with the lack of clear labelling of the resident 3P makes it hard to determine whether this, in fact, represents a trans-acting 3P [39].

Work by *York et al.,* on the other hand, has demonstrated that inactive free 3P appears to be able to stimulate the replication of purified cRNPs, leading to the hypothesis that the role of free 3P is to stimulate the activity of the resident 3P [56]. Thus, they state that a '*trans-activating'* polymerase stimulates the cRNA bound 3P to carry out replication. It is at present unclear which of the two apparently contradicting models is correct.

The host proteins acidic leucine-rich nuclear phosphoprotein family members A and B (ANP32A and ANP32B) are important in stimulating the generation of vRNA from cRNA [58]. Although the mechanism by which ANP32A and B promote this activity is unknown, it has been found that ANP32A and B specifically binds to free 3P and thus could potentially be involved in promoting the recruitment or activity of a trans-acting or trans-activating polymerase [58].

1.3.5.3 Initiation

A second major difference between transcription and replication is the mechanism of initiation. As stated above, transcription initiates using a host-derived primer. Replication, on the other hand, is initiated *de novo. De novo* initiation refers to the initiation of replication by the polymerisation of two single dNTPs instead of using a primer RNA. Despite both vRNA to cRNA and cRNA to vRNA initiation occurring via *de novo* initiation, there are also differences in the mechanisms between the two. Firstly, for cRNA synthesis by a vRNP, the *de novo* initiation occurs by terminal initiation, which involves the generation of a

dinucleotide at positions 1 and 2 of the 3′ vRNA promoter. 'Terminal' *de novo* initiation by viral polymerases is normally linked to the presence of a structure in the polymerase active site called a 'priming loop', which acts to stabilise the first dNTP at position 1 of the template. Such a loop has been observed in the recently solved 3P structure of influenza viruses [32, 33, 42], and has been shown to be functionally important [59]. It was observed that without this priming loop structure, cRNA synthesis from a vRNA template was compromised [59], likely due to the lack of stabilisation of the initial nucleotide.

In contrast, in the process of vRNA synthesis from a cRNP, the *de novo* initiation occurs internally at position 4 and 5 of the 3′ cRNA promoter [60]. This process does not require the priming loop structure, which has been explained by *te Velthuis et al.* to be due to the fact that the activation energy of internal *de novo* initiation is 11x lower than that of terminal *de novo* initiation [59]. After this internal initiation, the generated dinucleotide realigns to position 1 and 2 of the 3′ promoter and elongation continues using the 'dinucleotide primer' [60].

1.3.5.4 Elongation

The process of elongation in vRNA and cRNA synthesis likely proceeds similarly to that of transcription, with the template RNA being fed through the replicating 3P's active site. Interestingly, it has been found that steric hindrance of the priming loop may limit the rate of elongation and thus that a conformational change in the priming loop may be required to allow elongation to occur [59]. This is supported by observations that elongation occurs more efficiently in mutant 3Ps which have a deleted priming loop [59].

Once initiation has occurred, the presence of NP on the vRNA or cRNA template appears essential for 3P processivity and the generation of an elongated product. It has been shown that short vRNAs/cRNAs can be replicated efficiently in the absence of NP, however, longer RNAs cannot [37, 61]. It is possible that the role of NP is to inhibit the formation of secondary

structures in vRNA and cRNA, which might decrease the processivity of 3P [62]. The same role of NP applies for the extension of viral mRNAs during 3P-mediated transcription.

1.3.5.5 RNA Encapsidation

As explained in **section 1.3.4**, viral mRNAs are bound by host mRNA-binding proteins. Both vRNAs and cRNAs, in contrast, are encapsidated by 3P and NP (**section 1.3.3**). 3P and NP likely bind to nascent v/cRNAs during their synthesis, however the exact sequence of events is not fully understood. 3P likely initially binds onto the 5′ termini of nascent v/cRNAs. Loading of NP is then potentially nucleated by the binding of NP to 3P at the nascent RNA termini [35, 36], though this has not been definitively proven. Loading proceeds as free NP molecules interact with the terminal v/cRNA-bound NP in a growing NP chain. Specifically, the tail-loop of a free NP molecule binds into the insertion groove of the terminal NP of the v/cRNA bound NP-chain [34] (the tail-loop/insertion groove interaction is shown in Figure 3 B). This model is supported by data which shows that NP molecules with a mutated insertion groove, but a wt tail loop (NP-E339A), appear to be able to block the loading of wt NP onto vRNPs [37, 63]. This is likely as the mutant NP can insert into a growing wt NP chain, however, does not allow insertion of the following NP molecule [37]. Interestingly, phosphorylation of NP may control the transition of NP from a monomeric to an oligomerised state, thus blocking inappropriate oligomerisation of NP [64, 65, 66]. Furthermore, the loading of NP onto nascent v/cRNAs appears to be facilitated by the host enzymes UAP56 [67], Tat-SF1 [68] and Fragile X Mental Retardation Protein [69].

1.3.5.6 Copying of the 5′ Terminus of vRNA and cRNA

The final major difference between transcription and replication relates to the mechanism of copying the 5′ end of the template. When the replicating 3P reaches the 5′ terminus of the template vRNA or cRNA, unlike in transcription, where a poly-A tail is synthesised by

stuttering (**section 1.3.4**), the replicating 3P successfully copies the 5′ terminal region [3]. This is clearly important for the maintenance of the full genome sequence. The mechanism by which the 5′ terminus is released from 3P to allow its replication is unknown; however, it may be linked to the binding of short viral RNAs (svRNAs) to 3P [70, 71] (discussed in **section 1.3.5.7**).

1.3.5.7 Regulation of Transcription and Replication

Much work has focussed on the mechanisms which control the transition of 3P from a transcriptase to a replicase. The idea of this switch stemmed from observations that viral mRNA could be seen to accumulate in infections at earlier time points than cRNA/vRNA, indicating that transcription precedes replication [54].

One hypothesis to explain this observation has been provided by *Vreede et al.,* who demonstrated that 3P is able to stochastically initiate via capped-primer dependent or via *de novo* initiation, and that the lack of early cRNA accumulation only occurs due to its instability in the host cell [72]. This could be countered, however, by pre-expressing 3P, which binds to the cRNAs (Fig. 4). This shelters cRNAs from degradation and thus reveals the generation of cRNAs at early time points. In this model, therefore, replication only occurs productively when viral transcription has led to 3P (and NP) accumulation [72].

A second explanation has been provided by work looking at the effect of the NEP protein on the balance of transcription and replication. NEP appears to be important in limiting transcription at late time points post-infection and promoting 3P's replicase activity [73]. It has been found that NEP has a role in generating a species of short viral RNAs (svRNAs) which correspond to the first 22-25 nt of vRNA [71]. These svRNA are generated by 3P using cRNA as a template. The accumulation of these svRNAs correlates with the switch from transcription to replication [71]. svRNAs appear to promote genome replication by promoting

the generation of vRNAs from cRNPs. This occurs in a segment dependent manner, e.g. HA derived svRNAs promote the replication of the HA segment; however, the mechanism by which svRNAs act is still unclear [70]. Interestingly, the role of NEP in this process appears to be solely in stimulating the initial generation of cRNA, which is required for svRNA synthesis [70]. Thus, in this model, two factors regulate the switch; (i) NEP by stimulating the generation of cRNA by vRNPs, (ii) svRNAs by stimulating the generation of vRNA from cRNPs.

In summary, it is possible that multiple factors regulate the apparent switch in conjunction, with the levels of host mRNAs available for cap snatching, the levels of free 3P, the levels of NEP, and the levels of 5′ vRNA (likely svRNAs) all having a role in the process [74].

1.3.6 vRNP Export and Transport

In order for the formation of new virions at the plasma membrane to occur, progeny vRNPs must be trafficked out of the nucleus, transported across the cytoplasm and then incorporated into the budding virion.

The major cellular pathway which controls vRNP nuclear export, is the Chromosome Region Maintenance 1 (CRM1) mediated export pathway, which targets cargo for export through nuclear pore complexes (NPC). This can be experimentally demonstrated by the use of leptomycin-B (a CRM1 inhibitor), which blocks the export of vRNPs during infection [75, 76]. However, it is unclear which viral factor is bound by the export factor CRM1; with NEP, M1 and NP all containing nuclear export signals (NES) [75]. Nuclear export appears to be promoted by caspase-3 activity, which is associated to the induction of apoptotic signalling [77]. It has been hypothesised that this occurs due to caspase-3 leading to an increase in the diameter of the NPC pore; thus facilitating the export of the bulky vRNP complexes [77].

Once in the cytoplasm, vRNPs are trafficked to the apical membrane for the budding of new virions. This initiates with the targeting of vRNPs to the microtubule organising complex (MITOC) [78], where they associate to Rab11 found on recycling endosomes (RE) [78, 79, 80]. These REs traffic vRNPs to the plasma membrane where new virion formation occurs [79].

1.3.7 Budding

At the plasma membrane, or during the trafficking of vRNPs to the plasma membrane on REs, vRNPs appear to associate into hetero-octomers (with one of each of the eight vRNPs present). The association of these vRNPs in an ordered manner, which allows correct virion formation, is controlled by the presence of 'packing signals' in each of the vRNAs. These packaging signals are encoded within the vRNA sequence. There are many documented regions which appear to act as packaging sequences. These are generally found in the vRNA termini, often in the non-translated regions (previously reviewed [81, 82]). The budding process appears to be driven by the viral proteins HA, NA, M1, M2 and the vRNPs, though the exact mechanism has not been fully elucidated. This is supported by the observation that many of these components can drive budding to some degree when expressed by themselves in cells [5, 83]. The budding observed, however, is higher when these viral factors are expressed together, demonstrating their cooperative action [5, 83]. The hypothesised role of each of these proteins in the process is discussed below.

In infected cells, HA [84, 85] and NA [5] localise to certain regions of the membrane, termed lipid rafts, which are enriched in lipids such as cholesterol and sphingolipids. Recruitment of M1 to the forming virion is principally mediated by its interaction with the HA and NA intracellular tails. The octamer vRNP complexes then interact with (or are already interacting with) M1 and thus are recruited to the forming virion. This daisy chain mechanism of

recruiting RNPs to the forming virion is supported by the fact that mutations in the HA/NA cytoplasmic tails cause decreased levels of M1 and vRNP incorporation into the forming virions [86, 87, 88]. The interaction between M1/RNPs with the cytoplasmic tails of HA and NA then possibly stimulates the ability of HA and NA to cause localised membrane curvature and thus the formation of the virion shape. The idea that vRNP binding to M1 promotes virion formation is supported by the observation that in filamentous virions, RNPs are found localised to the apical virion pole [10].

The final process in the budding of the forming virions is termed scission and is the mechanism by which the 'virion' membrane separates from the plasma membrane. This process is mediated by the M2 protein. Thus, M2 deficient viruses are unable to release forming virions from the plasma membrane surface [5, 89]. It is hypothesised that under conditions of low cholesterol, which would occur when the lipid raft has been incorporated into the forming virion, M2 promotes membrane curvature and scission of the virus [89].

Post-scission, however, the SA-binding activity of HA can inhibit virion release from the infected cell, due to HA-mediated binding of the virion to SA-containing receptors on the cell surface or on other virions. NA acts to cleave the SA and thus stops virion aggregation and allows release from the cell membrane. Thus, infections with viruses encoding mutant NAs with low activity results in the retention of virion aggregates at the plasma membrane of infected cells [3].

1.4 Innate Host Anti-viral Responses

1.4.1 Introduction

Host cells possess a range of different innate anti-viral responses, which they activate upon detection of a viral infection. These responses are aimed principally at limiting the replication of the virus, limiting the spread of the virus from the infected cell and stimulating the upregulation of anti-viral proteins in neighbouring cells in preparation for a subsequent infection. The three principal stages of the innate anti-viral responses are: i) the detection of the viral infection, which is mediated by a range of cellular receptors, ii) the induction of a number of intra-cellular and inter-cellular signalling pathways and iii) the expression of a wide range of anti-viral factors or the induction of a number of anti-viral processes. In this thesis, I will focus on the interferon (IFN) response. There are three types of IFN, types I, II and III. This thesis will focus mainly on Type I IFN and thus, apart from when stated, IFN will refer to Type I IFN. Despite this, the conclusions drawn are likely relevant for Type III IFN, which is induced and acts in a similar manner to Type I IFN in influenza virus infections [90].

1.4.2 Viral Detection - Pattern Recognition Receptors

Cells are able to recognise a wide range of pathogens using Pattern Recognition Receptors (PRR). PRRs recognise a large variety of molecules which are generated by multiple different types of pathogens. These molecules are termed Pathogen Associated Molecular Patterns (PAMPs), and are mainly molecules that are not found in the uninfected host but are found widely in infections with different pathogens. Beyond this, some PAMPs are molecules that are found in host cells, but PRRs detect their abnormal localisation within infected cells. PAMPs encompass molecules such as viral RNAs, cytoplasmic DNA and bacterial liposaccharides. Given that different PRRs detect only a subset of PAMPs, and that a certain

pathogen will only produce a small range of PAMPs, this means that distinct PRRs contribute differentially to the detection of different types of pathogens [91].

The main types of PRRs are listed below:

- **Toll Like Receptors (TLRs):** The TLR family of receptors is made up of thirteen different receptors denoted TLR1-13. These receptors are mainly found on the plasma membrane and in the endosomes of immune cells, where they survey the extracellular regions for PAMPs. The TLR family recognises a wide range of ligands from viral RNA to bacterial liposaccharide and flagellin. In terms of influenza A virus infection, the TLR3 and TLR7 receptors are the most relevant as they are able to recognise viral RNA species in the endosome [92]. In response to TLR receptor ligand binding, MyD88/TRIF mediate the activation of IRF3/7 and the induction of IFN expression [93].
- **STING:** The stimulator of interferon genes (STING) protein is just one of the many proteins which have been documented to be a viral DNA sensor [94]. The ER-associated STING protein uses the presence of DNA in the cytoplasm, which does not normally occur in uninfected cells, to detect DNA virus infections. This can occur either via the direct detection of DNA by STING or via the activation of STING by the DNA receptor cyclic-DMP-AMP synthase (cGAS). Examples of viruses detected by STING are herpes simplex virus 1 and Kaposi's sarcoma-associated herpesvirus [95]. When activated, STING signals downstream to induce IFN expression [96].
- **RIG-I Like Receptors:** The RIG-I receptor family consists of Retinoic acid-inducible gene 1 (RIG-I), Melanoma-associated differentiation protein 5 (MDA-5) and Laboratory of genetics and physiology 2 (LGP2). Both RIG-I and MDA-5 localise to the cytoplasm and detect the presence of cytoplasmic viral RNAs. RIG-I principally

recognises short regions of blunt-ended dsRNAs with a terminal 5′-ppp or 5′-pp [97, 98]. Furthermore, recent work has shown that 5′-ppp-less blunt-ended dsRNA which contain a terminal bend in the double stranded region are also able to activate RIG-I [99]. MDA-5 on the other hand is activated by longer stretches of dsRNA, independent of the presence of a 5′-ppp/pp [98]. The two receptors have similar structures; two N-terminal Caspase-Activation and Recruitment Domains (CARD), a central helicase region and a C-terminal Regulatory Domain (CTD). The helicase region is formed by two helicase domains (HDs) that are termed Hel1 and Hel2, and a region termed Hel2i [100]. In this thesis, I will focus on RIG-I, as it has been shown to be the principal receptor which detects influenza A virus infection and triggers the IFN response [98]. This is demonstrated by the fact that the knockout of RIG-I almost completely abrogates the IFN response against influenza A virus.

1.4.3 Induction of Interferon Expression - RIG-I/MAVS

In its ligand-free form, RIG-I adopts a conformation which is incompatible with downstream signalling (Fig. 5). In this conformation, CARD2 is bound by Hel2i, likely sequestering the former and thus inhibiting downstream signalling [101]. 5′-ppp-dsRNA is likely initially captured by binding to free CTD. It is hypothesised that ATP hydrolysis is then required to disrupt the CARD2:Hel2i interactions, allowing a conformational change and the binding of the dsRNA into a pocket formed by the CTD and helicase region. Specifically, in the resultant structure, the helicase region forms a hollow ‘disk structure’ which surrounds the dsRNA. One end of this disk is capped by the CTD domain, which makes interactions with the dsRNA terminus. The CTD recognises the 5′-ppp by interacting with the α-phosphate (closest to ribose ring) and the β-phosphate (second closest to ribose ring), with further electrostatic interactions with the γ-phosphate (terminal) hypothesised [100]. The RNA-bound RIG-I is able to hydrolyse ATP, which is hypothesised to cause the stochastic release of

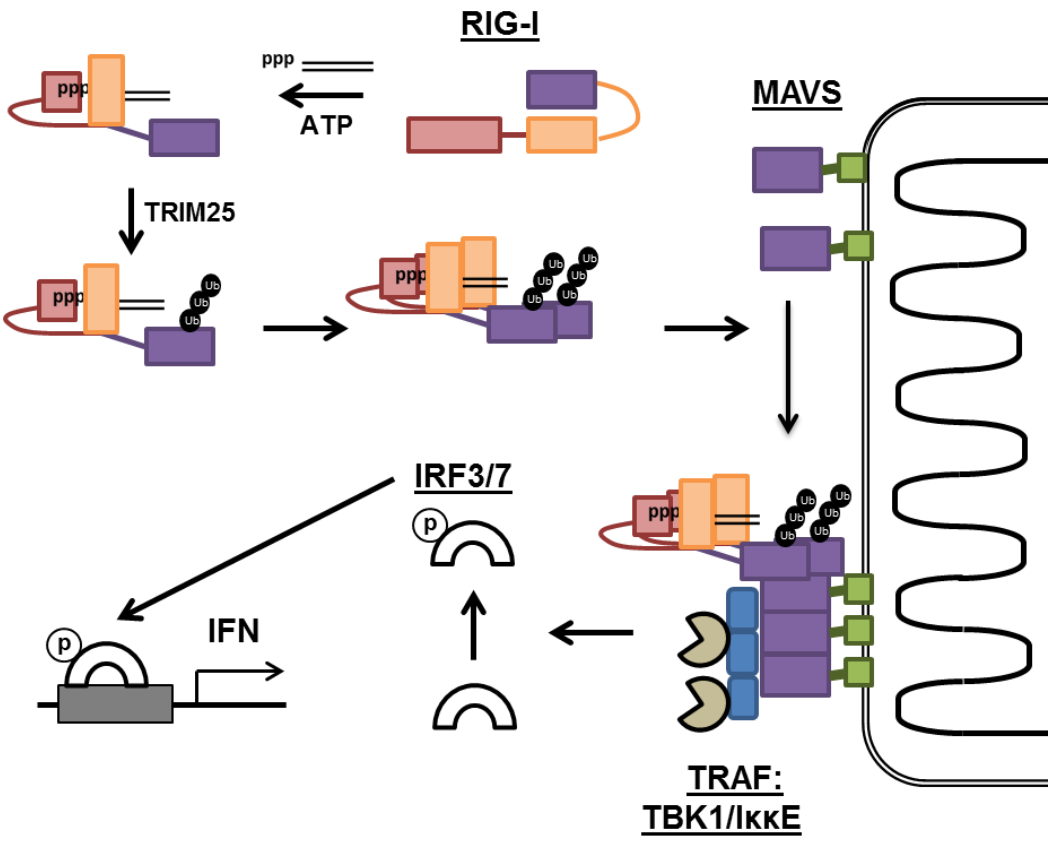


Figure 5 – The RIG-I/MAVS signalling pathway - RIG-I CARD domains (purple) are held in an unexposed state by the RIG-I helicase domain (orange). Binding RIG-I to stimulatory dsRNA, along with ATP hydrolysis by the helicase domain, leads to a conformational change in the receptor which exposes the RIG-I CARD domains. TRIM25 then ubiquitinates the CARD domains. Ubiquitinated RIG-I oligomerises on the stimulatory RNA and interacts with the MAVS (purple/green). RIG-I binding stimulates the homoligomerisation of MAVS into mitochondrial associated filaments. These filaments are able to interact with members of the TRAF family of adaptor proteins which in turn recruit and activate TBKI and IkkE. These activated kinases phosphorylate the transcription factors IRF3, IRF7 and NFkB. In their phosphorylated state, these factors translocate into the nucleus where they mediate the induction of IFN expression.

RNAs [100, 102]. This release likely allows RIG-I to survey multiple RNA molecules and stops the inappropriate activation of RIG-I. Indeed, RIG-I mutants which cannot hydrolyse ATP are constitutively active [103]. Subsequent ubiquitination of the CARD domains by tripartite motif containing 25 (TRIM-25) then occurs, blocking a return of the CARD domains to their auto-repressive conformation [104] (Fig. 5). This poly-ubiquitination allows the CARD domains of RIG-I molecules to oligomerise, leading to the formation of RIG-I oligomers [100]. It is controversial, however, as to whether these oligomers are required for downstream signalling (Fig. 5) [100, 105]. Activated RIG-I interacts with the outer-mitochondrial membrane protein MAVS (mitochondrial anti-viral signalling protein), which, like RIG-I, is normally found in a signalling incompetent conformation [106, 107, 108, 109]. The interaction of RIG-I with MAVS leads to a conformational change in the latter, allowing MAVS to form homo-oligomers (Fig. 5) [110]. These homo-oligomers act in a prion-like manner, stimulating the incorporation of further monomeric MAVS molecules into the forming complexes [110]. The MAVS complexes consist of long helical filaments, with the MAVS CARD domains directed into the filament core [111]. These helical filaments have been seen to form on mitochondria in cells, with an average length of 400 nm [111]. The mitochondrial associated homo-oligomeric MAVS complexes are then able to bind members of the TNF receptor associated factor (TRAF) family of proteins which in turn recruit the kinases TANK-binding kinase 1 (TBK1) and inhibitor-κB kinase ε (IKKε). This is potentially due to the high avidity of interaction between the TRAF proteins and the homo-oligomeric MAVS filaments. These kinases in turn activate the transcription factors interferon regulator factors 3 and 7 (IRF3/7) and nuclear factor kappa-light-chain-enhancer of activated B cells (NF-KB) , allowing their import into the nucleus where they promote the expression of cytokines including Type I and III IFN (Fig. 5) [105].

1.4.4 Anti-viral Activities – Interferon Stimulated Genes

IFN signalling occurs in an autocrine and paracrine manner, leading to the upregulation of anti-viral genes in both the IFN-expressing cell and its neighbours, respectively. Type I IFN and Type III IFN bind and activate the Type I Interferon Receptor (IFNAR) and the Interferon Lamda Receptor 1 (IFNLR1), respectively [112]. Both receptors stimulate the JAK/STAT signalling pathway leading to the activation of genes which contain IFN-

stimulated response elements (ISRE). These genes are termed IFN stimulated genes (ISGs). ISGs have a wide range of functions and can be divided into two main groups:

- **ISGs that sensitise cells to detect infection** – Many of the proteins listed in section 1.4.3, which are involved in RIG-I/MAVS signalling, are themselves ISGs. For example, the expression of both RIG-I and MAVS are controlled by an ISRE. Thus, IFN stimulated cells express higher levels of PRRs and components of the IFN pathway. This means that, when an infected cell produces IFN, it causes its neighbours to become more sensitised to detect infection [113]. For example, treatment of cells with Type I IFN increases the sensitivity of their detection of influenza A virus vRNAs [114].
- **ISGs that control infection** – Many ISGs have direct roles in controlling the replication and spread of viruses. Despite the fact that each ISG is not necessarily active against every virus, there are multiple ISGs which target the replication and spread of influenza A virus. Some of the ISGs which are known to act to control influenza A virus infection are shown below [92]:

  (i) ***IFITM3* – Interferon Induced Transmembrane Protein 3** is a protein which acts in the endosomal compartment. It is believed that this protein blocks the ability of influenza virus to infect cells by inhibiting the fusion of the viral and endosomal membranes, which is required for the release of vRNPs into the cytoplasm (**section 1.3.1**) [115, 116].

  (ii) ***Mx1/MxA* – Myxovirus resistance gene (Mx) 1 and A** are homologues from the mouse and human, respectively. Both ISGs have been shown to protect mice against influenza virus infection [117]. Despite this, the mechanism of action of Mx1 and MxA are not fully understood. Interestingly, MxA localises primarily to the cytoplasm whilst Mx1 localises to the nucleus,

indicating the two may act in distinct manners [118]. Both ISGs have been shown to bind to components of the influenza virus RNP and inactivate RNPs in an unknown manner [119, 120]. Mx1 has been hypothesised to act by inhibiting the interaction of 3P and NP, and thus blocking the encapsidation of progeny vRNAs into vRNPs; however, this has not been conclusively proven [119]. Furthermore, work has demonstrated that MxA may act in influenza virus infections to retain vRNPs in the cytoplasm [121].

(iii) ***TRIM22*** **– Tripartite Motif Containing 22** is a protein which inhibits the replication of influenza A virus. This is caused by the ability of TRIM22 to target NP for degradation [122]. This would limit the levels of NP within an infected cell, limit the formation of v/cRNPs and thus the ability of vRNA and cRNAs to be replicated.

(iv) ***Viperin*** **–** Viperin disrupts the composition of lipid rafts, which, as stated above, are important in virion formation [123]. In support of this, viperin overexpression specifically inhibited virion release by apparently blocking the final scission process. This is demonstrated by the presence of 'daisy chain like structures', which are formed by a chain of budding virions that are unable to complete the budding process [123].

(v) ***PKR*** **– Protein Kinase R** is a receptor which recognises dsRNA produced by viruses during infection. When activated by dsRNA, PKR phosphorylates Eukaryotic Initiation Factor (eIF) 2α. This factor is important in translation initiation where it cycles between a GDP and GTP bound state to complete its function. The conversion from the GDP to GTP bound state is controlled by the GTP exchange factor EIF-2B. In its phosphorylated form, eIF2α becomes

locked in a non-functional interaction with EIF-2B, blocking the function of eIF2α and thus inhibiting translation [124]. Given influenza A virus proteins are translated by the host translational machinery, this also leads to an inhibition of viral protein synthesis.

1.5 Mitochondria and Viral Infections

1.5.1 Introduction

Mitochondria are multifunctional organelles with roles in metabolism, cell death and immune responses. They are formed by two membranes; the inner membrane (MIM), which is highly folded, and the outer-membrane (MOM). This double membrane causes mitochondria to be subdivided into two compartments. The first, termed the matrix, is the region surrounded by the MIM, whilst the second, termed the intermembrane space (IMS), is the region between the two membranes.

As stated above, mitochondria are the site of MAVS signalling and thus are an important organelle in the response against virus infections. This is highlighted by the fact that many influenza A virus strains encode the viral protein PB1-F2, which disrupts mitochondrial function (**section 1.6.2.2**). Furthermore, in some influenza A virus strains, the viral protein PB2 also targets the mitochondria (**section 1.6.2.2**). Given that the first part of this thesis aims at further understanding the effects of PB2 on mitochondrial function, the main roles of mitochondria are discussed below.

1.5.2 Electron Transport Chain

The proteins of the electron transport chain (ETC) are found within the MIM. The ETC drives the formation of a mitochondrial membrane potential (MMP) across the MIM by

pumping protons into the matrix space. The best known role of the MMP is to drive ATP synthase, however, it is important for many other functions. The ETC also generates reactive oxygen species (ROS), which have important roles in many processes such as autophagy and IFN induction [125, 126].

1.5.3 Apoptosis

Mitochondria are the key mediator of the intrinsic apoptosis pathway. This is as they contain multiple apoptosis regulators, including cytochrome-c and apoptosis-inducing factor (AIF), within the IMS. Many different types of apoptotic stimuli can lead to the permeabilisation of the MOM and the release of these components into the cytoplasm, leading to the induction of apoptotic cell death. The mechanism of this permeabilisation is unknown, however, opening of a mitochondrial permeability transition pore complex (PTPC) is hypothesised to be key [127]. It has also been hypothesised that the MOM protein VDAC (voltage dependent anion-specific channel) also has a role in this permeabilisation process [127].

1.5.4 Fusion and Fission

Mitochondria are highly dynamic organelles which are able to undergo fusion and fission. This process has been hypothesised to be important in mixing mitochondrial contents [128]. Mitochondrial fusion and fission are mediated by the activity of multiple mitochondrial proteins. Specifically, the mitofusin (MFN) family and Optic atrophy protein 1 (OPA1) are important in driving fusion, and DRP1 (dynamin-related protein 1) is important for driving fission [128, 129]. Interestingly, mitochondrial fusion and fission have been shown to be important for both mitochondrial metabolism [129] and RIG-I/MAVS signalling [130] (**sections 1.5.2 and 1.5.5**).

1.5.5 Innate Immune System

As stated above in **section 1.4.3**, the MOM is the site of formation of MAVS oligomers, which are essential in the RIG-I and MDA-5 mediated IFN response. The mitochondrial association of MAVS is essential for its signalling capacity. This is demonstrated by the fact that the hepatitis C virus (HCV) protein NS3/4A inactivates MAVS by cleaving it from the MOM [131]. Interestingly, mitochondria have been found to modulate MAVS signalling in multiple ways. Firstly, the matrix proteins NOD-like-receptor family member X1 (NLRX1) and Tu translation elongation factor mitochondrial (TUFM) are both regulators of MAVS signalling, though their mechanisms of action are unknown [132, 133]. Secondly, fusion and fission of mitochondria are important in regulating the RIG-I/MAVS pathway; with fusion promoting RIG-I/MAVS signalling and fission inhibiting it [134]. Finally, the MMP is required for correct RIG-I/MAVS signalling.

Mitochondria are not only important in the production of IFN. Detection of PAMPs also leads to the activation inflammasome complex. This complex catalyses the generation of inflammatory cytokines, such as IL-1β, from their inactive pro-forms. The inflammasome associates with mitochondria via an interaction with MAVS [135, 136], and this association is important for inflammasome activity [135, 136]. This may be linked to the role of mitochondria in generating ROS, which promotes inflammasome activity [137].

1.6 Influenza A Virus and the Innate Immune Response

1.6.1 Interferon Induction by Influenza A Virus

Influenza A virus infection can lead to the induction of IFN production in infected cells. Work by *Kato et al*. has demonstrated that the IFN response against influenza A virus is mainly mediated by the RIG-I receptor [98, 138]. The most likely candidate RNAs detected by

RIG-I in influenza A virus infected cells are the 5′-ppp-containing vRNA and cRNAs. As stated in **section 1.4. 2**, in order to activate RIG-I, the 5′-ppp of vRNAs or cRNAs must be found at the end of a region of blunt-ended dsRNA. This could be envisaged to occur in three different ways: (i) free vRNA forming a panhandle structure, (ii) free cRNA forming a panhandle structure or (iii) given their complementarity, vRNA and cRNA together forming a perfect dsRNA. On the other hand, viral mRNAs do not contain a 5′-ppp, as they are capped, and thus do not have the canonical RIG-I ligand structure. The role of vRNA and cRNA in IFN induction is consistent with the data of *Rehwinkel et al.*, who generated RNPs from plasmid-expressed components in order to test the stimulatory activity of the different viral RNA products. They found that wt RNPs and transcription-incompetent RNPs were comparably able to induce IFN expression; however, replication-incompetent RNPs induced IFN more weakly [114]. Furthermore, treatment of the product RNAs with calf intestinal phosphatase, which removes 5′ phosphates from RNAs, abolishes the IFN-inducing capacity of the viral RNAs [114]. This demonstrates that, within this experimental system, viral replication, which generates vRNA and cRNA, is essential for RIG-I ligand generation whilst viral transcription, which generates capped poly-adenylated transcripts, is not [114]. This led to the hypothesis that in influenza A virus infected cells, RIG-I detects genomic vRNAs and/or antigenomic cRNAs. However, despite this apparently simple explanation of RNA detection, there are many unanswered questions in the field of influenza A virus RNA detection.

1.6.1.1 The Importance of vRNA and cRNA Synthesis

Despite the work by *Rehwinkel et al.* supporting the hypothesis that progeny v/cRNAs are the principal RIG-I inducer, there is controversy about the requirement of v/cRNA replication for IFN induction. Firstly, it has been hypothesised that incoming vRNPs (those released from virions) can activate RIG-I [139]. This hypothesis has been disputed by *Killip et al.*, who

demonstrate that inhibitors of viral transcription and nuclear RNA export (which inhibit v/cRNA replication and subsequent v/cRNA export) effectively inhibit influenza virus mediated IFN induction [140]. This effect would not occur if incoming vRNPs were the principal activators, as the above inhibitors do not affect vRNP release from virions. Interestingly, this work also shows that both the drug cycloheximide (CHX) and knockdown of NP, despite efficiently blocking v/cRNA synthesis, do not completely inhibit IFN induction in infection [140]. In fact, in infections with certain viral strains, CHX increases the IFN response. Based on this data, the authors hypothesise that replication of full length v/cRNAs is not required for IFN induction in this system, but that the major RIG-I ligands may be aberrant RNA products which can be generated under CHX treatment [140]

1.6.1.2 The Importance of Protein Shielding of RNAs

As stated in **section 1.3.3**, vRNAs and cRNAs are not found naked in the cell, but are bound at their 3′ and 5′ termini by 3P and along their length by NP (Fig. 3). This binding has been hypothesised to block the association of host factors to the RNA by steric hinderance [138]. Furthermore, when bound to 3P, the termini of vRNA are separated into a 'fork/hook' structure and thus the vRNA does not have a blunt-dsRNA termini [32]. In this thesis, the inhibition of RIG-I binding to viral RNA species in this manner will be termed as 'shielding'. Shielding of stimulatory RNAs/DNAs by viral proteins is not undocumented, with human immunodeficiency virus (HIV) capsid protein believed to be involved in shielding HIV cDNA from cGAS in the cytoplasm [141]. Due to the potential effect of shielding, it has been hypothesised that a free viral RNA species must be generated during infection in order for RIG-I activation to occur [138]. However, it is at present unclear by which mechanism such free viral RNA could be generated and transported into the cytoplasm of an infected cell.

1.6.1.3 The Importance of vRNA and cRNA Promoter Mismatches

Interestingly, vRNA and cRNA are relatively poor inducers of the RIG-I/MAVS response. *Anchisi et al.* recently demonstrated that mismatches between the 5′ and 3′ termini in the v/cRNA panhandle structure (Fig. 2) act to decrease the stability of viral RNAs in complex with RIG-I and thus heavily limit their ability to induce IFN [142]. Thus, point mutations that remove these mismatches in the double stranded termini lead to a profound (multiple log-fold) increase in the inducing activity of viral RNAs [142]. These mismatches have been hypothesised to have evolved to allow influenza A virus to more effectively evade the host innate immune response [142]. In addition to this, work by *Rehwinkel et al.* has shown that not all vRNPs containing full-length vRNA induce a robust IFN response when generated within cells [114]. It is, however, unclear whether the poor induction by these vRNA species is due to i) the shielding of the vRNAs within the vRNP (**section 1.6.1.2**), ii) by the imperfect double stranded nature of the RNA termini or iii) by both. *Rehwinkel et al.* could, however, cause these vRNPs (containing full length vRNAs) to trigger RIG-I activation in conditions where cells had been pre-treated with IFN [114]. This was hypothesised to occur due to the fact that IFN pre-treatment leads to over-expression of RIG-I within cells, which increases the sensitivity of cells to RIG-I ligands [114] (**section 1.4.4**).

1.6.1.4 The Importance of Aberrant Infections

It has been observed that within a population of influenza A infected cells, only a small percentage of cells express IFN [143]. This same observation has been made in paramyxovirus infections, even in strains where the principal IFN antagonist has been deleted [144]. This led to the hypothesis that IFN induction is associated with infrequent, erroneous cell infections. However, the nature of the erroneous influenza A virus infection that causes IFN induction is

poorly understood. One possibility is given by work focussing on DI virus (**section 1.2.2.3**) infections [138, 145, 146].

Influenza DI viruses are generated when, during the replication of a vRNP or cRNP, 3P makes an error and only copies the termini of the vRNA or cRNA. This error thus leads to a large internal deletion in the v/cRNAs, however the 5′ and 3′ termini are preserved [18, 138]. This process appears to occur most commonly in influenza virus genome segments 1, 2 and 3 [138, 147]. The mechanism by which 3P makes this error is unknown, however, interestingly, 3P appears to copy similar lengths of the 3′ and 5′ regions during these internal deletion events (Fig. 6) [147]. Furthermore, the final 2 or 3 nucleotides of the v/cRNA before the 'jump-site' are often the same as the final 2 or 3 nucleotides which are not copied (Fig. 6 A) [147]. This implies that 3P loses association to the template v/cRNA, but, maintains the nascent v/cRNA in its active site, and then somehow allows a re-association between the nascent RNA and a complementary sequence at the 3′ side of the template (Fig. 6 A) [147]. The fact that similar lengths of the 5′ and 3′ template termini are copied has been explained by the idea that the polymerase may dissociate from the 3′ proximal region of the v/cRNA and re-associate to the nearby 5′ proximal region of the v/cRNA in a 'template jumping' event [148] (Fig. 6 B).

These resultant internal deletion RNAs can be packaged into vRNPs or cRNPs, due to the presence of the 5′ and 3′ RNA termini. However, these RNPs are shorter in nature and often have large deletions in or even completely lack the protein coding sequence of the vRNA [147, 149]**.** These RNAs can replicate alongside their full-length counterparts within the cell and, as long as they contain intact packaging sequences (**section 1.3.7**) [82], can be incorporated into new virions. The internal deletion RNAs contained by DI viruses will be termed as DI RNAs within this thesis. Given that these viruses lack the coding capacity for one or more functional viral proteins, they can only successfully replicate within a cell in the presence of a wild-type 'helper

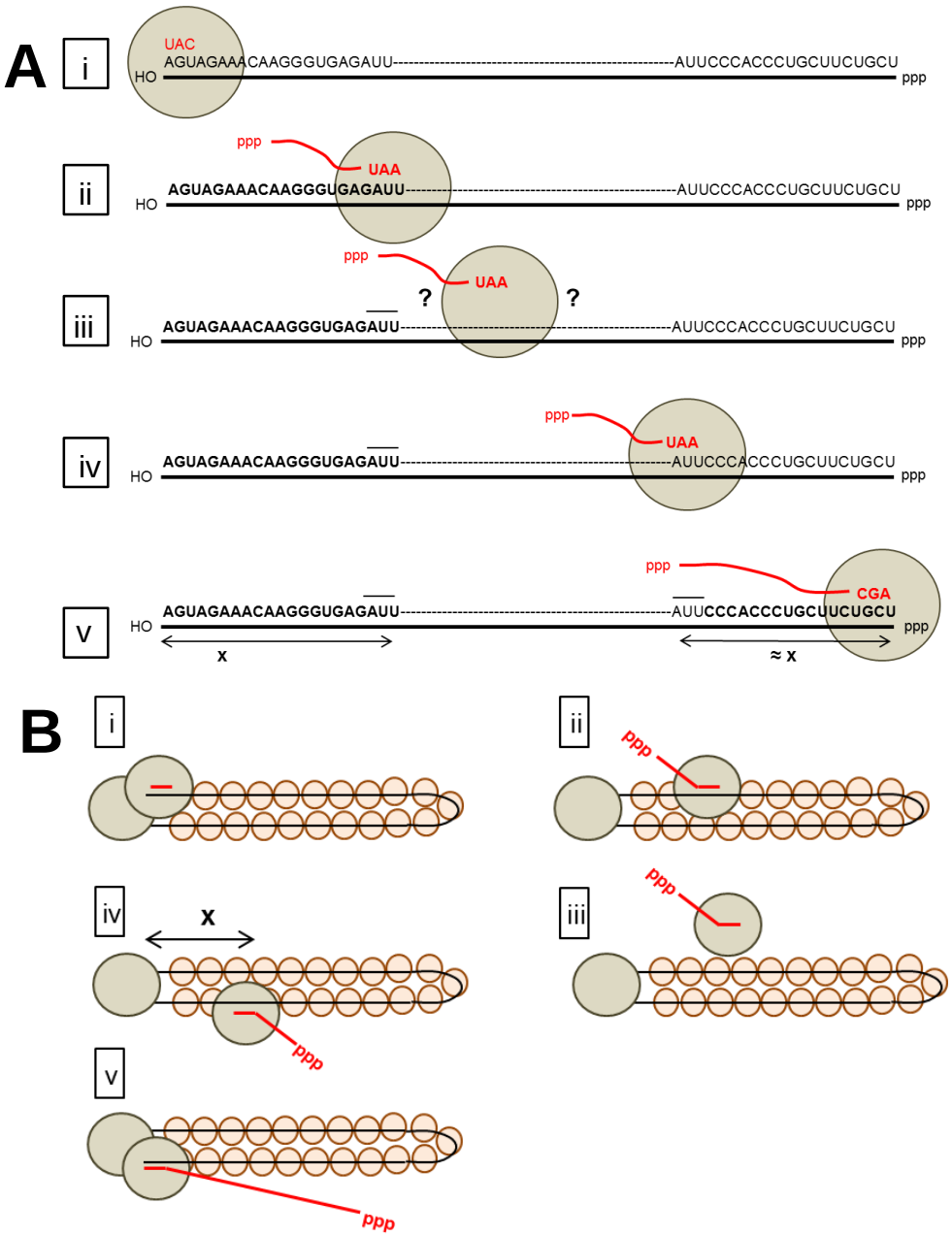


Figure 6 – A model of DI RNA / internal deletion RNA formation – (A) (i) During the replication of v/cRNA (black), *de novo initiation* by 3P (grey) occurs at the 3' end of the template. (ii) Elongation proceeds until 3P likely stalls. The base pairing between the template and product at this point is shown (UAA:AUU). (iii) The polymerase dissociates from the vRNA by an unknown mechanism. (iv) 3P re-associates to a downstream region on the vRNA by base pairing between the template and nascent RNA within the 3P active site (UAA:AAU). The final nucleotides before the 'jump start site' (short line) are often identical to last nucleotides before the 'jump end site' (short line). (v) Elongation continues until the end of the template. The replicated region of the template is highlighted in bold. The length of RNA (x) replicated before the 'jump start site' is generally similar to the length replicated after then 'jump end site' (B) The stages of (i) initiation, (ii) elongation and (iii) 3P dissociation are shown on an illustrative RNP. (iv) The dissociated 3P is hypothesised to re-associate onto a nearby region of the RNA in the RNP structure. If this region of RNA is of the reverse polarity, then an internal deletion would occur. (v) 3P then completes the replication of the template.

virus', which provides the DI virus with the missing viral protein(s) [18, 149]. Interestingly, however, in these dually infected cells, the DI RNAs outcompete their full length counterparts and replicate to far higher levels, which causes the generation of sequentially higher levels of DI virions relative to wt virions [150]. It is unclear what causes this effect, but it is believed that DI RNAs, due to their shorter length, may out-compete the full length viral RNAs for a limiting viral or host protein that is essential for replication. Interestingly, it has been found that mutations in NEP and PA can promote the formation (or accumulation) of DI RNAs [151, 152].

In terms of IFN induction, it has been shown that influenza A virus stocks that contain high levels of DI viruses are very strong inducers of IFN [138, 145, 153]. This idea is supported by work by *Perez-Cidoncha et al.* in which influenza A virus was passaged through non-IFN-responsive cells in order to remove the pressure against the accumulation of mutant viruses which induce a strong IFN response. A number of the resultant IFN inducing virus strains were found to accumulate high levels of DI RNAs during infection [146]. Finally, infection of cells with a reverse genetics generated influenza DI virus protects the cells from subsequent influenza virus infection in an IFN dependent manner [154, 155].

The mechanism by which influenza DI viruses induce a strong IFN response, however, is still unclear. *Baum et al.* have found that the purification of RIG-I from influenza A virus infected cells leads to co-purification of higher levels of DI RNAs than equivalent full-length RNAs [156]. However, it is not clear whether this is because RIG-I has a preference for DI RNAs, or simply because DI RNAs are found at higher levels than their full length counterparts. The idea of RIG-I (or MDA-5) activation by DI RNAs is not new, with paramyxoviruses infections triggering RIG-I activation due to the production of a type of DI RNA called a 'copy-back' [144, 156, 157, 158]. Copy-back RNAs appear to occur when the replicating paramyxovirus polymerase dissociates from the template strand and starts using its nascent

strand as a template. This process results in a perfectly dsRNA with a 5′-ppp, which has all the features of a strong RIG-I/MDA-5 ligand. However, influenza virus DI RNAs are only internal deletion RNAs, and thus have identical 5′ and 3′ termini to their full length counterparts [147, 156]. It is unclear why changes to the internal region of a vRNA or cRNA, which are on average more than 100 nts downstream of the v/cRNA termini, would affect the ability of the RNA to activate RIG-I [18, 147]. Finally, it has been hypothesised that DI RNAs may encode truncated viral proteins which are able to induce the IFN response [159].

1.6.2 Dampening – Suppressing Interferon Induction

In the previous section, I looked at the mechanisms by which influenza viruses might induce IFN. However, this is not the only factor that determines the magnitude of the IFN response against influenza virus, as the virus is, in fact, able to decrease the magnitude of the IFN response in cells in which RIG-I has been activated [92]. In this thesis, I will term this activity as 'dampening'. Influenza A virus does this in multiple ways using a range of different proteins.

1.6.2.1 NS1 – The Principal Dampener

The NS1 protein consists of an N-terminal RNA Binding Domain, an Effector Domain and a short unstructured C-terminus [160]. NS1 acts in multiple manners to both inhibit the detection of viral RNAs by RIG-I and to dampen the innate immune response in cells where RIG-I has been activated. These activities are listed below, though, interestingly, it appears that the NS1 proteins of different viral strains do not always possess every one of these activities [160].

- **RNA Sequestration -** The first activity of NS1 is to inhibit the detection of viral RNAs by RIG-I. This is mediated by NS1's N-terminal RNA Binding Domain which allows NS1 to sequester double stranded regions of viral RNAs [92, 160]. Position 38 is key for the

RNA binding function of NS1 and its mutation abolishes NS1's RNA binding capacity [160]. NS1-mediated sequestration of viral dsRNAs likely limits their ability to be recognised by the RIG-I family of receptors. Indeed, *Rehwinkel et al.* have demonstrated that if NS1 is purified from infected cells, RNAs co-purify which are able to induce IFN when re-transfected back into cells [114].

- **TRIM-25 Inhibition –** As stated in **section 1.4.3**, when RIG-I binds a 5′-ppp-dsRNA ligand, it undergoes a conformational change which leads to exposure of its CARD domains. NS1 can inhibit the ability of RIG-I to signal downstream at this stage and thus dampen the IFN response. This occurs by NS1 binding and inhibiting the activity of the TRIM25 ubiquitin ligase (Fig 5) [161], which is important in RIG-I signalling.
- **CPSF-30/PABPII Binding -** Finally, NS1 can dampen the IFN response by binding Cleavage and polyadenylation factor 30 (CPFS-30) and PolyA-binding protein II (PABPII). These activities are mediated by the effector domain and unstructured C-terminus, respectively [160]. NS1 binding to CSPF-30 inhibits cleavage and poly-adenylation of cellular mRNAs [162]. Beyond this, NS1 binding to PABPII inhibits correct poly-A tail formation and nuclear export of host mRNAs [163]. In this manner, NS1 inhibits host mRNAs export to the cytoplasm for translation [162, 163]. Thus, NS1 leads to a general shut-down of cellular protein expression, which subsequently dampens the IFN response in multiple ways. Firstly, since IFN itself is expressed from an mRNA, a general reduction in protein expression reduces IFN expression [160]. Secondly, the shut-down of IFN expression would limit the expression of ISGs in neighbouring cells, which would otherwise prime the cells to detect influenza virus infection (**section 1.4.4**).

1.6.2.2 Alternative Dampeners

Despite the fact that influenza A virus encodes such a strong IFN inhibitor as NS1, it also expresses other proteins which dampen IFN induction. These proteins, and their dampening activities, are listed below.

- **PB1-F2 -** The PB1-F2 protein is encoded in a +2 frame of the PB1-encoding mRNA [164] and has been shown to possess dampening activity in IFN induction assays [165, 166]. The PB1-F2 of the vast majority of avian influenza virus strains is imported into the mitochondria [165, 167], where it disrupts mitochondrial function by promoting their depolarisation. This is likely mediated by the opening of pores / channels in the MIM. This could occur either by the direct formation of mitochondrial membrane pores by PB1-F2, or by the interaction of PB1-F2 with the ADP/ATP translocase 3 (ANT3) and/or VDAC, which are mitochondrial membrane channels [168, 169]. Given that the MMP is required for mitochondrial morphology and the correct signalling of MAVS on mitochondria (**sections 1.5.4 and 1.5.5**) [130], it is likely that PB1-F2 inhibits IFN induction in this manner. However, there are reports that PB1-F2 may act by directly binding and inhibiting MAVS at the MOM [170, 171].
- **PA-X -** The PA-X protein is expressed from the same mRNA as the PA protein but arises when, during the process of PA translation, a frameshift event leads to a different C-terminal sequence [172]. The PA-X protein contains the same endonuclease domain as PA (**section 1.3.4**), but acts in a distinct manner. PA-X has been shown to degrade cellular RNAs [173]. Via this mechanism, PA-X is hypothesised to limit the IFN response by degrading IFN mRNAs or ISG mRNAs [174].
- **PB2** - In some influenza A virus strains, the PB2 protein also localises to the mitochondria [175, 176]. Following the observation that PB2 is able to bind MAVS, it was observed that overexpression of PB2 dampens IFN induction [176, 177, 178]. This led to the hypothesis that PB2 might inhibit the activation or downstream signalling of MAVS.

Inhibition of IFN signalling at the level of MAVS has already been demonstrated for HCV, whose NS3/4A protein cleaves MAVS from the mitochondrial membrane [131], and rotavirus, whose NSP1 protein mediates MAVS degradation [179].

1.7 Adaptive Immune Responses, Antigenic Drift and Antigenic Shift

1.7.1 Introduction

The innate branch of the immune system is important in controlling the initial infection, however, viral clearance and protection against re-infection are mediated by the adaptive immune response. Though not the main focus of this thesis, I will here briefly consider the various adaptive immune responses against influenza A virus, and the manner by which influenza A virus circumvents them.

1.7.2 Cell-Mediated Adaptive Immune Responses

As previously mentioned, the adaptive immune system is essential in controlling influenza A virus infection. The adaptive immune system is made up of several cell types, including lymphocytes and natural killer (NK) cells. T lymphocytes play a key role in controlling influenza virus spread; demonstrated by the fact that animals deficient in T lymphocytes die of influenza virus infection [3]. T lymphocytes are categorised into two major subsets: CD8+ lymphocytes, which mediate direct elimination of viruses or virus-infected cells, and CD4+ lymphocytes, which clear viruses via activating B lymphocytes. CD8+ lymphocytes, also termed cytotoxic lymphocytes, can eliminate viruses either via killing infected cells or via producing pro-inflammatory cytokines. Although most CD8+ lymphocytes disappear after an infection is cleared, memory CD8+ lymphocytes have been detected *in vivo* [180]. Interestingly, these memory cells have shown cross-reactivity between different influenza virus strains,

which has important roles in immunity [3]. A potential example of this is the low rate of severe infections in the elderly in the 2009 pandemic, which is likely due to the similarities between the pre-1950 and 2009 H1N1 virus antigens [181].

1.7.3 Antibody Mediated Adaptive Immune Responses

The major influenza virus antigens recognised by antibodies are the surface glycoproteins HA and NA. Antibodies to these proteins mainly detect the highly antigenic head domains of both proteins (**section 1.2.2.1**). A number of antibodies against the HA head are termed to be 'neutralising', as they block the ability of virions to bind to SA and thus infect host cells. On the other hand, antibodies against NA are not neutralising, but, can inhibit the sialidase activity of NA. This does not block the process of virion internalisation directly, but, inhibits virion spread by blocking release of virions from the surface of infected cells (**section 1.3.7**) [182]. Furthermore, both HA and NA specific antibodies are able to recognise the HA and NA present on the plasma membrane of infected cells and initiate antibody-dependent cell-mediated cytotoxicity (ADCC) [3]. ADCC is the process by which antibodies can attract cells of the innate immune system and cause targeted killing of infected cells.

The antibody response generated against a certain virus strain, or influenza vaccine, provides protection against reinfection. However, this protection is only mediated against viruses with the same HA and/or NA types, thus antibodies raised against an H1N1 strain will likely provide poor protection against a subsequent H3N2 infection [3].

1.7.4 Influenza Evasion of Adaptive Immunity

Virus strains encoding the same subtype of HA and/or NA can cause influenza disease in a human multiple times within said human's life [3]. This is a clear demonstration that the adaptive immune response generated against a certain virus subtype does not provide lifelong

protection against all strains of that subtype. In fact, influenza A viruses are very capable of evading the host adaptive immune response. This occurs via two main mechanisms; antigenic drift and antigenic shift.

1.7.4.1 Antigenic Drift

The first mechanism by which influenza viruses evade previously generated adaptive immune responses is termed antigenic drift. This process is driven by the high variability between influenza virions within a population. This variability is caused by the high error rate of 3P, which lacks proof-reading activity and thus is not able to correct the mis-incorporation of a base during v/cRNA replication [3]. Due to this, the population of influenza A virions generated during an infection is, in fact, made up of a large array of variants, termed quasispecies, which all derive from the original infecting strain. The majority of these quasispecies will replicate worse or equally as well as the original infecting strain, so they will be maintained as a low proportion of the population. However, if a quasispecies has a replicative advantage it will start to increasingly become over-represented in the population. For instance, some quasispecies within the population will contain mutations in NA or HA that decrease the binding affinity between the mutant HA/NA and the pre-exisiting antibodies. Thus, any virion or infected cell which expresses these HA or NA molecules will be less effectively targeted by the host adaptive immune response and will have a replicative advantage. Overtime, this replicative advantage process leads to further, optimising mutations which lead to a 'drift' away from the original virus strain that was strongly recognised by the HA and NA antibodies. The new antigenic drift strains are thus increasingly able to evade previous adaptive immune responses and to re-infect previously infected humans. This gives influenza viruses the ability to cause yearly epidemics in the winter months, as a large proportion of the human population each year will have an increased susceptibility to

infection. The process of antigenic drift explains the requirement for frequent updates of the virus strains used in influenza vaccines in order to maintain immunity against the present circulating strains [3, 4]. In this thesis, the strains that circulate in the human population over long periods will be termed 'seasonal strains'. The pathogenesis and impact of seasonal influenza disease is explained further in **section 1.8.2**.

Interestingly, the stem region of HA, which is less variable than the head domain, is a promising target for antibody targeting. This is as the high conservation in this region will likely resist the accumulation of mutations. Thus, antigenic drift against anti-stem antibodies may occur less readily than against the anti-head antibodies. Furthermore, due to the high conservation in the stem region, anti-stem antibodies have been found to be active against a wide range of different virus subtypes [183].

1.7.4.2 Antigenic Shift

The second process by which influenza viruses can evade pre-existing adaptive immune responses in a population is termed antigenic shift. This process also results in the presence of HA and/or NA proteins that are antigenically distinct from the circulating strains; however, the mechanism by which they are acquired is different.

Antigenic shift occurs when a cell is infected by two virus strains which are adapted to two different hosts, such as an avian and a human adapted strain. Swine have been termed to be an effective 'mixing vessel' (site for this process) as swine URT cells can be infected efficiently by both human and avian adapted strains. After replication of the vRNAs of both infecting viruses, the process of vRNP packaging into virions occurs. If, during this process, vRNAs from the two strains have compatible packaging sequences [82], hybrid virions can form which contain some genome segments from the first virus strain and some genome segments from the second. This results in 'reassortment viruses', which encode a subset of

viral proteins adapted to one host and a subset of proteins adapted to another. If the HA and/or NA encoded by a reassortment virus are not derived from a human-adapted stain, the virus will likely be antigenically distinct from the circulating seasonal strains. This can occur via the acquisition of a different type of HA or NA (e.g. HxNx → HyNy) or by the acquisition of antigenically distinct HAs and NAs of the same type (e.g. HxNx → HxNx). The resultant virus will have the potential to infect a large proportion of the human population, due to the likely low levels of immunity.

1.8 Influenza Epidemiology and Pathogenesis

1.8.1 Introduction

As stated above, influenza A virus infects the cells of the URT in humans and thus causes respiratory illness. The severity of influenza disease is highly variable, and is dependent on the virus strain and the host. The severity varies from lowly pathogenic seasonal infections to highly pathogenic pandemic and avian virus infections [3]. In the sections below, the virus source, pathogenesis and global importance of different 'types' of influenza disease will be further explained.

1.8.2 Seasonal Epidemics

Most years, the vast majority of human influenza A virus infections occur during the winter months. These winter epidemics are caused by the circulating seasonal influenza virus strains. At present, the circulating seasonal virus strains are of the H1N1 and H3N2 subtypes [4]. Furthermore, influenza B virus seasonal epidemics occur during the winter months and, in some years, influenza B virus infections are more predominant than those of influenza A virus [3]. Seasonal infections are generally of low virulence in healthy adults and are

characterised by the onset of fever, malaise, joint pain, coughing, sneezing etc. These symptoms can persist for 1-2 weeks with approximately 5-10% of adults becoming infected every year [3, 4]. Thus, the economic burden of seasonal influenza A virus infection is high due to lost work hours and decreased productivity. Despite the normally low severity of seasonal infections in healthy adults, seasonal strains should not be considered to be non-dangerous, as they cause significant mortality in certain groups. These 'at risk' groups include the young, old, pregnant and immunocompromised. In fact, seasonal epidemics result in 250,000 – 500,000 deaths globally per year [4]. Thus, vaccination is currently recommended for the young, old, pregnant and other at risk groups such as health workers [4].

1.8.3 Influenza A Virus Pandemics

As well as causing seasonal epidemics, influenza A viruses are capable of causing pandemics. These pandemics can have far higher death tolls than seasonal epidemics. The pandemics which have occurred in the last century are the 1918 H1N1, 1957 H2N2, 1968 H3N2, 1977 H1N1 and 2009 H1N1 pandemics [3]. The strains which caused many of the above pandemics arose via the process of antigenic shift (**section 1.7.4.2**, NB/ the 1977 pandemic is hypothesised to have occurred due to the laboratory release of a stored H1N1 strain, whilst the causative strain of the 1918 pandemic is thought to be a virus of avian origin [184]). The process of antigenic shift can result in viruses to which a large proportion of the human population have low pre-existing immunity. This likely allowed the rapid spread of pandemic strains through the global population. Unlike human seasonal influenza virus infections, which as stated above are generally mild in their symptoms, pandemic strains can exhibit highly variable virulence. The 2009 pandemic strain for example caused no higher mortality than average seasonal epidemics. Furthermore, the H2N2 pandemic strain was not

more pathogenic than average seasonal strains and thus the cause of the increased mortality was likely the low level of pre-exisiting immunity to H2N2 subtype viruses [3]. On the other hand, the 1918 H1N1 pandemic is the archetypal example of the unpredictability of influenza pandemics. As with other pandemics, factors such as the likely low levels of pre-existing immunity allowed the rapid spread of the virus over three waves, causing the infection of approximately 20% of the global population. The distinguishing factor of this pandemic, however, was the mortality. The virus had a 2.5% mortality rate (25x higher than standard epidemics) and was abnormally pathogenic in healthy populations [3]. In total, the pandemic is believed to have caused 50 – 100 million deaths worldwide. Infected individuals displayed a rapid disease progression, underlined by a number of the normal symptoms as well as acute pulmonary pneumonia/oedema and haemorrhage. This oedema obstructed gas-exchange in the lungs, caused patients to rapidly become cyanotic, leading to eventual death.

The genome sequence of the causative strain of the 1918 pandemic has been recovered from samples extracted from formalin-fixed and permafrost-buried tissue [185, 186]. These sequences were used to reconstruct a strain that encoded the viral proteins of the 1918 virus [187]. The resultant virus strain was highly virulent in mammals, demonstrating that the high virulence observed during the pandemic was a feature of the virus and was not solely caused by a lack of immunity in the infected humans [187]. Infection in a mouse model led to an uncontrolled innate immune response, termed a cytokine storm, which caused extensive inflammation and pulmonary oedema. This matches with the observations in humans during the pandemic. Furthermore, it was found that there was no single gene that caused the high virulence, with many of the 1918 viral proteins conferring increased virulence in the mouse model [187]. For example, the 1918 PB1-F2 protein has been demonstrated to be important in increasing pulmonary inflammation and immune cell recruitment [188].

In the years following a pandemic, derivatives of pandemic strains cause yearly epidemics in the same manner as seasonal strains. This is also often accompanied by the pandemic subtype virus out-competing the pre-existing seasonal viral strains, leading to their eventual disappearance (e.g. the H2N2 pandemic lead to disappearance of the H1N1 human lineage, the H3N2 pandemic lead to disappearance of H2N2 human lineage) [3]. Furthermore, in cases where the pandemic virus had high virulence, such as the 1918 strain, this process is accompanied by a decrease in the virulence of the virus strain, representing the adaptation of the virus to the human host. This can likely be explained by the fact that the 1918 viruses encoded non-human adapted proteins, which were acquired from an avian adapted virus [186, 189]. As the virus proceeded to replicate in the human host, however, host adaptation occurred, leading to a gradual decrease in virulence of the virus in the human host.

1.8.4 High Pathogenic Avian Influenza Viruses

A final type of influenza A infection, which is of particular importance at present, is the infection of humans with highly pathogenic avian influenza viruses (HPAIV) of the H5N1 subtype. These infections are contracted directly from infected poultry, and despite occurring at low frequency, exhibit high mortality rates [2]. This high mortality, like that caused by the 1918 pandemic strain, appears to stem from high degrees of viral replication and high cytokine release, leading to decreased gas exchange and cyanosis [3]. Beyond this, human infection with H7N9 avian viruses has been documented, also with high mortality rates [190]. The causative highly pathogenic H5N1 and H7N9 strains, despite inducing high mortality in humans, are at present unable to complete human to human transmission; however, recent work has demonstrated that only a small number of mutations are required for ferret to ferret

transmission (used as a model for mammal to mammal transmission) of a H5N1 to be achieved [191, 192].

The low rates of avian to human and human to human transmission stem from the poor adaptation of H5N1 and H7N9 virus to the human host. Furthermore, as stated in **section 1.8.3**, a lack of adaptation to the human host could be a cause for the high levels of cytokines induced and the high mortality in HPAIV infections. It is thus of great interest to understand the changes that occur in avian viruses during the process of human adaptation. Further information on host adaptation and host range is given in section 1.9, and is a focus of the first part of this thesis.

1.9 Host Range and Host Adaptation

1.9.1 Introduction

As stated in **section 1.1.2.1**, influenza A virus is able to infect a wide range of different hosts. Whilst replicating in a certain host, a given influenza A virus strain adapts to the specific environment and life-cycle of the host organism in a process termed 'host adaptation'. This adaptation can, however, render the strain incapable of productively infecting a different host, thus limiting its 'host range'. In order for a virus strain to become zoonotic (infect a different host species), it needs to accumulate specific polymorphisms in its genome that allow it to successfully complete the stages of the life-cycle (listed in **section 1.3**) in the new species.

Host adaptation of non-human adapted strains of influenza viruses is a potential source of future outbreaks of influenza disease in humans. Thus, it is important to understand the barriers that determine host range. At present, given its high mortality rate in human infections, the potential adaptation of HPAIVs to the human host in a manner which allows

human to human transmission is of grave concern. Many of the stages of the virus life cycle represent points of host-range restriction. To demonstrate the wide range of factors which control host range, a number of examples will be shown chronologically from across the virus life cycle.

1.9.2 Virus Entry – SA Type

HA mediates cellular receptor binding and virus internalisation by recognising SA moieties on sialylated receptors. However, there are multiple different types of SA groups whose distribution varies between hosts. For example, in birds, the epithelial cells of the GIT express one variant of SA (α-2,3-linked), whilst the epithelial cells of the URT of mammals express another (α-2,6-linked) [3]. Due to this, avian strains, which are adapted to recognising and binding α-2,3-linked SAs, are generally poorly capable of binding α-2,6-linked SA, and vice versa. This thus limits the ability of avian viruses to infect human hosts. In fact, one of the adaptive changes which allowed a H5N1 HPAIV strain in the laboratory to transmit between mammals resulted in an increased affinity of HA for α-2,6-linked SAs [191, 192]. Interestingly, pigs express both types of SA, and thus can be infected by both human and avian viruses [3]. This point is important for the process of 'antigenic shift', as pig cells can be infected simultaneously with pig-adapted, human-adapted and avian-adapted viruses.

1.9.3 Genome Replication By the Viral Polymerase

The replicative activity of the majority of avian adapted 3Ps in human cells is very low. This is caused principally by a difference in position 627 of PB2 between avian and human adapted PB2s [193]. This difference is taken as one of the model examples of host adaptation as pre-2009 there was an almost complete conservation of both the avian (Glutamic Acid – 627E) associated amino acid in avian strains and the human (Lysine – 627K) amino acid in human strains. Post-2009 human strains were found to lack this 627K and exhibit the avian-

associated 627E. However, compensatory mutations in the PB2s from viruses from the post-2009 H1N1 human lineage were soon found which are found spatially close to amino acid 627 in the 3P structure [194, 195]. It has been hypothesised that amino acid 627 of PB2 is important in the interaction between NP and 3P within RNPs [196, 197] (**section 1.3.3.2**). Furthermore, it has been shown recently that there are differences between the human and avian homologues of ANP32A and ANP32B (**section 1.3.5.2**), and that expression of the avian ANP32A homologue or ANP32B homologue can recover the activity of avian-adapted 3P in human cells [198]. This supports earlier work which hypothesised that human cells lacked an avian host factor which was required for the replicative activity of avian adapted 3P (encoding aspartic acid at position 627) in human cells [199].

1.9.4 Dampening – PB1-F2 and PB2

As stated in **section 1.6.2.2**, the viral protein PB1-F2 is targeted into the mitochondria of cells and thus acts to disrupt the induction of IFN. This localisation is controlled by a C-terminal mitochondrial targeting sequence (MTS). Interestingly, this mitochondrial localisation of PB1-F2, which is highly conserved in avian influenza A virus strains, has been lost from multiple human lineages [200, 201]. This has occurred due to the truncation of the PB1-F2 sequence by the introduction of premature stop codons, thus removing the C-terminal MTS. These truncated PB1-F2s do not disrupt mitochondrial function in the same manner as full length PB1-F2s [202]. This indicates that there are differing pressures to disrupt mitochondrial function between human and avian cells. This idea is supported by work which focusses on the dampening of IFN induction by PB2 (**section 1.6.2.2**). *Graef et al.* and *Miotto et al.* showed that there was an apparent selection for an amino acid at position 9 of PB2 in seasonal human (pre-2009 pandemic) influenza A virus strains which drives mitochondrial localisation of PB2 [176, 203] and that these mitochondrial localising PB2s were almost completely absent from avian strains. Interestingly, *Graef et al.* have shown that

mitochondrial PB2s are stronger IFN dampeners [176]. The effects of PB2 at the mitochondria will be further investigated in the first part of this thesis.

AIMS

In this thesis, I seek to further understand the mechanisms by which influenza A virus induces IFN expression in cells and how it subsequently dampens the response.

In chapters 3 and 4, I firstly investigate the role of mitochondrial PB2, which has been hypothesised to dampen RIG-I/MAVS signalling. Specifically, within this section, I attempt to further characterise the effects of PB2 on mitochondrial function and investigate the hypothesis that mitochondrial PB2 might be inhibiting MAVS signalling indirectly by affecting other mitochondrial functions; as appears to be the case with PB1-F2.

In the chapter 5, I seek to explain some of the unanswered questions in the field of RIG-I mediated detection of influenza A virus infection. These include: i) how RIG-I is able to detect 3P/NP coated viral RNAs and ii) how DI virus infections induce IFN. In this chapter, I set out to test the initial hypothesis that DI RNA-containing RNPs are less stable, due to adopting an aberrant structure, and thus are more prone to expose the v/cRNA they contain to RIG-I binding.

Chapter 2 -

Materials and Methods

2.1 Plasmids

2.1.1 PB2 Plasmids

Expression plasmids for PB2 and C-terminal GFP-tagged PB2 with N9 or D9 have been described previously [176]. The expression plasmid for GFP-tagged PB2 with T9 was generated via site directed mutagenesis of the expression plasmid for GFP-tagged PB2 with N9 (Primers used = A + B – Oligonucleotide Table – **Section 6.1.1**). The expression plasmid for C-terminally tandem affinity purification (TAP)-tagged PB2 has been described previously [204]. To generate expression plasmids for C-terminal APEX-tagged PB2s, the APEX sequence was PCR amplified from the MTS-APEX plasmid [205] using Phusion® HF polymerase and primers containing the XbaI or NotI restriction sites (Oligonucleotide Table; PCR primers = C +D). The resultant PCR product was digested with XbaI and NotI. The expression plasmids for PB2-GFP with N9, T9 or D9 were also digested with XbaI and NotI. The digested plasmids and PCR product were ligated using DNA Ligation Kit Mighty Mix (Takara), as per manufacturer's instructions.

2.1.2 Other Influenza Protein and RNA Expression Plasmids

2.1.2.1 Protein Expression Plasmids

The expression plasmids for the influenza A virus proteins NP, PB1 and PA from A/WSN/33 (H1N1) have been described previously [37]. The expression plasmid for D445A/D446A-PB1

(polymerase replication and transcription inactive mutant) has been described previously [37].

2.1.2.2 vRNA Expression Plasmids

The expression plasmids for segment 5 derived vRNAs (pPolI vector) used in section **5.2.1** have been described previously [37]. The expression plasmids for segment 5 derived vRNAs (pPolI vector) used in **sections 5.2.2 – 5.2.10** were a kind gift of Aartjan te Velthuis. The expression plasmids (pPolI vector) for segment 4 and segment 6 derived vRNAs were a kind gift of Aartjan te Velthuis. The pPolI GFP expression plasmid encodes a vRNA (vGFP) which, when transcribed by 3P, generates an mRNA which encodes GFP.

2.1.3 Other Plasmids

The MTS-APEX plasmid was purchased from Addgene and has been described previously [205]. To measure IFN-β expression levels, a luciferase plasmid (IFN-luc) under the control of the IFN-β promoter was used. To control for cell numbers during the IFN assays, a β-galactosidase control plasmid (β-gal) was used. The reporter plasmids were a kind of Marian Killip.

2.2 Cell Lines

2.2.1 Non-SILAC Labelled Cell Lines

Human Embryonic Kidney 293 T (HEK 293T), Madin-Darby Bovine Kidney epithelial (MDBK), monkey kidney epithelial (Vero) and human adenocarcinoma alveolar basal epithelial (A549) cells were from the Cell Bank of the Sir William Dunn School of Pathology.

The IFN-GFP HEK 293T cell line, which encodes GFP under the control of the IFN promoter, was a kind gift of Marian Killip.

The $MAVS^{+/+}$ and $MAVS^{-/-}$ mouse embryonic fibroblasts (MEFs) were a kind gift of Caetano Reis e Sousa.

2.2.2 SILAC Labelled Cell Lines

The SILAC cell lines were generated using A549 cells. The cells were cultured in Dulbecco's Modified Eagle Medium (lacking lysine and arginine) supplemented with i) fetal calf serum (FCS; to 10%) and ii) either light (K0, R0), medium (K4, R6) or heavy (K8, R10) lysine (to 0.8 mM) or arginine (to 0.4 mM). Cells were cultured in one of the media types for 2 weeks. To test for complete incorporation of the relevant amino acid type, samples of the medium and heavy labelled cells were lysed and analysed by mass spectrometry (**section 2.10.1**). The presence of light, medium and heavy labelled peptides within the samples was analysed by Svenja Hester of the Sir William Dunn School of Pathology Mass Spectrometry Facility.

2.3 Viruses

2.3.1 Virus Infections

The wt A/WSN/33 (N9-WSN) and N9D-PB2 A/WSN/33 (D9-WSN) viruses are as previously described [176]. The wt A/WSN/33 and PB2-strep A/WSN/33 viruses are as previously described [206]. The mutant NS1 virus (Δ99R38A) was a kind gift of Marian Killip.

Virus infections of HEK 293T and A549 were performed with DMEM supplemented to 0.5% FCS (0.5% FCS/DMEM), at the stated MOI. The cells were then left for the stated length of time before the cells were lysed or samples of the media were taken.

For specific infection conditions, see the following sections:

Plaque assay (Section 2.3.2)

Infection of SILAC cell lines for NP immunofluorescence (Section 2.5.2)

Infection of SILAC cell lines for mitochondrial isolation (Section 2.8.1)

Infection of A549 cells with Strep-PB2-WSN or A/WSN/33 (Section 2.8.2)

Infection of TUFM knockdown cells (Section 2.12.1)

Infection of cells transfected for RNP reconstitutions (Section 2.13.1.2)

Infection of cells expressing NP and/or 3P (Sections 2.13.1.3, 2.14.3 and 2.15)

2.3.2 Plaque Assays

Serial ten-fold dilutions of virus-containing samples were performed in Minimum Essential Media supplemented to 0.5% FCS (0.5% FCS/MEM). Confluent monolayers of MDBK cells in the wells of a 6 well plate were washed with phosphate buffered saline (PBS). 200 µl of each dilution of the virus-containing sample was added to a well of the 6 well plate. The plates were incubated for 1 h at room temperature before the inoculum was removed and the cells were overlaid with 0.5 % FCS/MEM containing 1% agarose. The cells were incubated at 37^{o}C for 72 h. The number of plaques was counted in the lowest dilution well which had clearly separated plaques. This number of plaques was used to calculate the viral titre in the original virus-containing sample, as per standard protocols.

2.4 Transfections

2.4.1 Plasmid Transfection

2.4.1.1 General Protocol

Plasmids were diluted in 37°C optiMEM (50µl for a 24 well plate, 500 µl 6 well plate and 1.5 ml for 10 cm dish). Lipofectamine 2000 (Invitrogen) (1.5 µl per 1 µg of plasmid) was added to optiMEM (in the same volumes as above) and incubated for 5 min. Equal volumes of the optiMEM/DNA and optiMEM/Lipofectamine mixes were combined and incubated for 20 min. Except where stated, cells in suspension were added into the wells of plates or dishes in DMEM supplemented to 10% FCS (10% FCS/DMEM; $1/50^{th}$ of a confluent T75 in 0.1 ml for a 24 well plate, $1/12^{th}$ of a confluent T75 in 0.5 ml for a 6 well plate, 1/2 of a confluent T75 in 1.5 ml for a 10cm dish). The OptiMEM/DNA/lipofectamine mix was added to the suspension cells and incubated for 5 min. 10% FCS/DMEM was added, and the cells were incubated at 37°C for the stated period of time.

For specific transfection conditions, see the following sections:

PB2-GFP Localisation in MEFs (Section 2.5.1)

PB2-APEX Immunofluorescence (Section 2.5.2)

PB2-APEX Electron Microscopy (Section 2.7.1)

RNP Reconstitutions in HEK 293T cells (Section 2.13.1.2)

RNP Reconstitutions in IFN-GFP HEK 293T cells (Section 2.13.2)

Expression of 3P, or 3P and NP in HEK 293T cells prior to infection (Section 2.14.3.1)

2.4.2 siRNA Transfections

Transfections of siRNAs were performed using the protocol from **section 2.4.1**, however, using siRNAs instead of plasmid DNA. siRNAs were transfected using 2 µl of Lipofectamine 2000.

For specific transfection conditions, see the following sections:

TUFM knockdown (Section 2.12)

RIG-I knockdown (Section 2.13.1.2)

NP knockdown (Section 2.14.3.1)

2.5 Fluorescence Microscopy

2.5.1 PB2 Localisation in MAVS$^{+/+}$ and MAVS$^{-/-}$ MEFs

MEFs generated from either wild-type or MAVS$^{-/-}$ mice, were grown on glass coverslips in 6-well plates. Cells were transfected with 2.5 µg of expression plasmids for PB2 or PB2-GFP. 24 h post transfection, cells were stained for 30 min with 10 nM Mitotracker Red (Molecular Probes) in 10% FCS/DMEM, washed for 2 min with 10% FCS/DMEM and then fixed with 3.5% formaldehyde in 10% FCS/DMEM. Coverslips were washed three times for 5 min in PBS and mounted in Mowiol. Images were obtained with an Axioplan 2 microscope. Image analysis was carried out using ImageJ software.

2.5.2 Immunofluorescence

To determine the localisation of PB2-APEX, vero cells grown on glass-coverslips in 6-well plates were transfected with expression plasmids for PB2-APEX. Cells were stained with

Mitotracker and fixed as in **section 2.5.1**. Cells were then blocked with 2% FCS in PBS (blocking buffer) overnight at 4°C and then permeabilised for 7.5 min using 0.2% Triton X-100 in PBS. Cells were stained with a rabbit polyclonal anti-PB2 antibody and a Cy2-conjugated anti-rabbit secondary antibody (Antibody table – **Section 6.1.2**). Coverslips were mounted in Mowiol and images were obtained with an Axioplan 2 microscope. Image analysis was carried out using ImageJ software.

To detect NP expression in infected SILAC amino acid labelled A549 cells, cells were infected with N9-WSN or D9-WSN (MOI=1). Immunofluorescence was performed as above using a rabbit polyclonal anti-NP antibody and a Cy2-conjugated anti-rabbit secondary antibody (Antibody Table; **Section 6.1.2**).

2.6 SDS-PAGE and Western Blot

Polyacrylamide gel electrophoresis (PAGE) was performed as per standard protocols. Western blots from polyacrylamide gels were performed as per standard protocols using a wet-transfer onto a nitrocellulose membrane**.** The membranes were blocked overnight in 5 % milk powder in PBS (milk). The membrane was then incubated for 2 h in milk containing the relevant primary antibodies (see Antibody Table – **Section 6.1.2**). The membrane was washed four times for 5 min in 0.1% Tween 20-supplemented PBS. The membrane was then incubated for 1 h in milk containing the relevant Horseradish peroxidase (HRP)-conjugated secondary antibody (see Antibody Table – **Section 6.1.2**). The membrane was then washed four times for 5 min in 0.1% Tween 20-supplemented PBS before washes were completed as above. The HRP was detected using one of two chemiluminescence reagents; Immobilon™ Western Chemiluminescent HRP Substrate Kit (Millipore) (High Sensitivity) or

Amersham™ ECL Start Western Blotting Detection Reagent (GE Healthcare) (Low Sensitivity). Densitometry of western blots was performed using the Aida software.

2.7 APEX and Electron Microscopy

2.7.1 Transfection, Staining and Section Preparation

HEK 293T cells ($1/50^{th}$ of a confluent T75), transfected in suspension with 1 µg of expression plasmids for either PB2 (with N9 or D9), PB2-APEX (with N9, D9 or T9) or MTS-APEX, were grown on plastic coverslips (Thermonox) in 24-well plates. Transfected cells were incubated for the indicated period of time (1 or 2 days) and then fixed in pre-warmed fixative solution (2.5% formaldehyde, 2% paraformaldehyde, 0.1 M sodium cacodylate, pH 7.4, and 2mM $CaCl_2$) for 1 h. Cells were washed five times with wash buffer (0.1 M sodium cacodylate, pH 7.4, 2 mM $CaCl_2$), once with wash buffer containing 1.5 mg/ml glycine, followed by five washes in wash buffer (all at 4°C). Cells were incubated with wash buffer containing 0.5 mg/ml diaminobenzidine and 0.03% H_2O_2 for the indicated period of time (15 min or 4 h, room temperature). Cells were washed five times with wash buffer, stained with 2% osmium tetroxide, washed three times in distilled water and then stained with 0.5% uranyl acetate. The samples were dehydrated in a cold ethanol series of increasing ethanol concentration up to 100%. Resin infiltration was then completed with a graded resin series of increasing resin concentration. The resin was left to polymerise at 60°C for 48 h. Sections were generated using a Leica UC7 microtome, mounted on 200 mesh copper grids and post stained with Reynold's lead citrate. Sections were imaged using a Tecnai 12 transmission electron microscope (TEM) with a Gatan US100 CCD camera. Densitometry and mitochondrial size analysis were carried out using ImageJ.

2.8 Mitochondrial Isolation

2.8.1 Mitochondrial Isolation for SILAC Studies

Confluent T175 flasks of light, medium and heavy-labelled SILAC cells were washed twice with PBS. A virus inoculum containing N9-WSN or D9-WSN (MOI = 5), or an inoculum containing no virus, was added to the cells on ice. 1 h post addition, cell monolayers were washed twice with ice cold PBS before 20 ml of the relevant, pre-warmed SILAC media was added. At 12 h post infection, the cells were detached by trypsinisation and 10 ml of the relevant SILAC media was added. The light, medium and heavy labelled cells were pelleted independently (800xg, 5 min) and each pellet was resuspended in 2 ml of PBS. The cell concentration in each of the cell suspensions was calculated using a haemocytometer. Equal numbers of cells from each of the suspensions were combined together. The resultant sample was centrifuged (800xg, 5 min) and the mass of the resultant cell pellet was measured. The volume of the pellet was calculated using the formula: Pellet Volume (ml) = Pellet Weight (g) x 0.8. The pellet was resuspended in 9 volumes of Hypotonic Buffer (20 mM HEPES pH 7.8, 5 mM KCl, 1.5 mM $MgCl_2$, 2 mM (dithiothreitol) DTT, 1 mM phenylmethylsulfonyl fluoride (PMSF)) and lysed with a dounce homogeniser (90 strokes). The lysate was mixed with 2 ml of 2.5x MSH buffer (525 mM Mannitol, 175 mM Sucrose, 20 mM HEPES pH 7.8, 5 mM ethylenediaminetetraacetic acid (EDTA), 2 mM DTT, 1 mM PMSF) for every 3 ml of hypotonic buffer. Large insoluble material was removed by centrifugation (800xg, 10 min) and mitochondria were isolated by centrifugation (9800xg, 10 min). The resultant pelleted mitochondria, termed the crude mitochondria, were resuspended in 0.5 volumes of 1xMSH (2.5x MSH diluted 2.5-fold with water). Two sucrose step-gradients were produced in ultracentrifuge tubes. The gradients consisted of 3 ml of 1 M sucrose solution layered carefully on top of 1.5 ml of 1.5 M sucrose solution. The resuspended mitochondria were

layered on top of one of the gradients and the second was used as a balance. The gradients were centrifuged (135,00xg, 4^{o}C) on a Beckmann Ultracentrifuge using a SW55 Ti rotor. The material at the interface between the 1.5 M and 1 M sucrose solutions was collected in 400 µl of sucrose solution. The 400 µl of sucrose solution was then diluted in 1.1 ml of 1xMSH to decrease its viscosity and the mitochondria were pelleted (9800xg, 15 min). The resultant pellet was resuspended in resuspension solution (8M Urea, 1% Sodium Deoxycholate) and snap frozen before submission to the Central Proteomics Facility of the Sir William Dunn School of Pathology Mass Spectrometry Facility for analysis (**Section 2.10.1**).

2.8.2 Mitochondrial Isolation for StrepTactin Purifications

Four confluent T175 flasks of A549 cells were infected with either wt A/WSN/33 or with PB2-Strep A/WSN/33 (MOI 1.5). At 12 h post infection, the cells were detached by trypsination and 16 ml of 10% FCS/DMEM was added per flask. The cells were pelleted (800xg, 5 min) and washed in PBS. The cells were then resuspended in hypotonic buffer (20 mM HEPES pH 7.8, 5 mM KCl, 1.5 mM $MgCl_2$, 2mM DTT, 1mM PMSF, 1x protease inhibitors (Roche)) and lysed with a dounce homogeniser (90 strokes). The lysate was mixed with 2 ml of 2.5x MSH buffer (525 mM Mannitol, 175 mM Sucrose, 20mM HEPES pH 7.8, 5 mM EDTA, 2mM DTT, 1mM PMSF) for every 3 ml of hypotonic buffer. Large insoluble material was removed (700xg, 10 min) by centrifugation and then mitochondria isolated (9800xg, 10 min) by centrifugation. The resultant pelleted mitochondria were resuspended in a 2:3 mix of 2.5xMSH and Hypotonic buffer, and the sample was centrifuged (9800xg, 10 min) to pellet mitochondria. The process of resuspension and pelleting was repeated a third time. The mitochondria were then lysed and a StrepTactin purification was performed (see **section 2.9**).

2.9 StrepTactin Affinity Purification

Mitochondria isolated from wt-WSN and Strep-PB2-WSN infected cells (**section 2.8.2**) were lysed in 0.5 ml of lysis buffer (50 mM Tris-HCl (pH 8.0), 200mM NaCl, 33% glycerol, 0.5% Igepal, 1mM DTT, 1x Protease Inhibitors (Roche)) for 1 h at 4^{o}C. The insoluble material was then removed by centrifugation (17,000xg, 10 min). The remaining lysate (400 µl) was diluted in binding buffer (800µl; 20mM Tris-HCl (pH 8.0), 200mM NaCl, 1x Protease Inhibitor). The diluted lysate was incubated overnight with 50 µl of pre-washed (in wash buffer – 100mM Tris-HCl (pH 8.0), 150 mM NaCl, 1mM EDTA, 0.1 % Igepal, 1 mM DTT, 1 x Protease Inhibitor) StrepTactin® Superflow beads (IBA). The beads were pelleted, and washed four times for 10 min in 1 ml of wash buffer. The beads were then incubated in 200 µl of 1x biotin elution buffer (IBA) for 2 h at 4^{o}C. The beads were then pelleted and the eluate collected to be submitted to mass spectrometry for protein identification (**section 2.10.1**).

2.10 Mass Spectrometry

2.10.1 Sample Submission and SINQ Analysis

Elutions from the StrepTactin purifications, and purified mitochondria from the SILAC experiments, were submitted to the Central Proteomics Facility (Sir William Dunn School of Pathology). The proteins were digested with trypsin and analysed using a Thermo Q Exactive Orbitrap mass spectrometer (by Svenja Hester or Benjamin Thomas of the Central Proteomics Facility). The peptide spectral matches were analysed using the central proteomics facility pipeline against a database containing the proteome of *Homo sapiens*¸ the proteome of A/WSN/33 and common contaminant proteins (by Svenja Hester or Benjamin Thomas of the Central Proteomics Facility) with a 1% false discovery rate. Label free

quantitation of protein levels was performed using the spectral index normalised quantification (SINQ) method [207], as previously described [206].

2.10.2 Mass Spectrometry Data Analysis

2.10.2.1 SILAC Experiments – Analysis of the Mitochondrial Proteome

The level of a protein from the D9-WSN or N9-WSN infected cells (quantified by SINQ analysis) was divided by its level in the uninfected cells to calculate the fold change in level upon infection (relative protein level). Filtering of proteins by their mitochondrial or non-mitochondrial ontology was completed using the Proteome-Discoverer software.

In order to approximate the level of NP within the N9-WSN and D9-WSN samples, the SINQ value for NP in each sample was divided by the SINQ value of the lowest level protein within the sample (taken as the background level).

2.10.2.2 Strep-PB2 Experiments – Analysis of the Matrix PB2 Interactome

The level of proteins within the Strep-PB2 WSN and WSN samples were quantified by SINQ as in **section 2.10.1.1**. Proteins which were likely purifying non-specifically were identified as described previously [206]. These proteins were assumed to be mainly contaminants and thus to be at constant levels between samples. The median protein level of these proteins was calculated for each sample, and all data points in that sample were divided by this median. The normalised data was then used to calculate fold differences in each protein's level between the Strep-PB2 WSN and WSN samples (relative protein level), and an average relative protein level for each protein across the three repeats was calculated. The detected proteins were cross referenced against two previous PB2 interaction screens [206, 208].

2.11 Tandem Affinity Purification

HEK 293T cells were transfected with 15 µg of PB2 expression plasmids and 5 µg of Flag-TUFM expression plasmid in a 10cm dish. After 24 h, cells were incubated at 4^{o}C in 500 µl TAP-lysis buffer (5mM HEPES (pH 8.0), 200 mM NaCl, 25% glycerol, 0.5% Igepal, 1 mM β-mercaptoethanol, 1x protease inhibitors (Roche)) for 10 min. Insoluble material was removed by centrifugation (17,000xg, 10 min). 350 µl of lysate was diluted with 1.4 ml of 140 mM NaCl solution. The diluted lysate was then applied to 50 µl of IgG Sepharose beads which had been pre-washed in binding buffer (10 mM HEPES (pH 8.0), 150 mM NaCl, 0.1 % Igepal, 10% glycerol, 1x protease inhibitors), and incubated for 2h at 4^{o}C. The beads were washed twice for 10 min in binding buffer and once for 10 min in cleavage buffer (binding buffer supplemented to 1 mM DTT). TAP-tagged PB2 was released from the beads using cleavage buffer containing ACTEV protease (1 µl ACTEV per 100 µl of cleavage buffer) at 4^{o}C overnight. Cell lysates and eluates were analysed by SDS-PAGE (**section 2.6**).

2.12 TUFM Knockdown

2.12.1 Cell Viability and Viral Growth Kinetics

Three siRNAs specific to the TUFM gene were generated using the Dharmacon siDesign Centre to target different positions on the TUFM mRNA (N, O and P; Oligonucleotide Table, **Section 6.1.1**). An siRNA specific to the GFP sequence was used as a negative control (Q; Oligonucleotide Table, **Section 6.1.1**). HEK 293T cells (1/12th of a confluent T75) were transfected with 100 pmol of either an siRNA against TUFM or an siRNA against GFP, using 2 µl of Lipofectamine 2000.

At 24 h post transfection, cells were either: i) Lysed and their viability determined using the CellTiter-Glo luminescent assay according to the manufacturer's instructions (Promega). ii)

Lysed and the expression levels of TUFM and β-actin were measured by Western Blot. iii) Infected at an MOI of 0.01 with either N9-WSN or D9-WSN. Samples of the supernatant were collected at various time points and the viral titre determined by plaque assay (**Section 2.3.2**).

2.12.2 Interferon Assay

For IFN analyses, HEK 293T cells ($1/50^{th}$ of a confluent T75) were transfected with 33 pmol siRNA specific to TUFM or GFP, as in **section 2.12.1**, along with 100 ng of the IFN-luc and β-gal reporter plasmids (2.13.1.1). 24 h post transfection, cells were either mock infected or infected with the Δ99R38A-NS1 virus strain (MOI = 10). 15 h post infection, cells were lysed and the reporter activities were measured (**Section 2.13.1.1**).

2.13 Interferon Analyses

2.13.1 Interferon Analysis Using Luciferase and β-Galactosidase Reporter Plasmids

2.13.1.1 Measuring Luciferase and β-galactosidase Activities

Reporter Lysis buffer (Promega) was applied to transfected and/or infected cell monolayers, which were then placed at -20°C. Samples were then thawed, the lysates collected and the insoluble material removed by centrifugation (17,000xg, 10 min). The luciferase signal was then measured using Luciferase Assay Reagent (Promega), as per manufacturer's instructions using a spectrophotometer. The β-galactosidase activity was calculated, as per standard protocols. In short, samples were incubated with ortho-nitrophenyl-β-galactosidase (ONPG)/Phosphate Buffer (60 mM Na_2HPO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, 38.5 mM β-mercaptoethanol, 0.8 mg/ml ONPG) until the sample became yellow, and then the absorbance was measured at 420 nm.

2.13.1.2 RNP Reconstitutions

HEK 293T cells ($1/50^{th}$ of a confluent T75) were transfected with 100 ng of IFN-luc and β-gal reporter plasmids alongside 0.25 µg of PB1/PB2/PA/NP expression plasmids and 0.25 µg of a pPolI (vRNA expression) plasmid. Cells transfected with a control plasmid (pcDNA3A) were used to calculate the background luciferase signal, which was subtracted from all samples, except when stated. For the majority of experiments, the cells were then incubated at 37^{o}C for 24 h. The protocol for measuring the luciferase and β-galactosidase reporter activities is shown in **section 2.13.1.1**.

For experiments where infections were performed post RNP reconstitution (**Section 5.2.1**), HEK 293T cells were transfected as above. Cells were either mock infected or infected with A/WSN/33 (MOI=5) at 15 h post transfection. At 12 h post infection, the cells were lysed and the reporter activities were measured (**Section 2.13.1.1**).

For the RIG-I knockdown experiments, HEK 293T cells ($1/50^{th}$ of a confluent T75) were transfected with 33 pmol of either an siRNA specific to TUFM or an siRNA specific to GFP (R and Q; Oligonucleotide Table – **Section 6.1.1**), alongside 100 ng of the IFN-luc and β-gal reporter plasmids, using 2 microlitres of Lipofectamine 2000. 24 h post transfection, cells were transfected with the expression plasmids for an RNP reconstitution, as shown above. 24 h post second transfection, the cells were lysed and the reporter activities were measured (**Section 2.13.1.1**).

For RNP reconstitutions involving 'limiting NP', an RNP reconstitution was performed as per above. However, the amount of NP expression plasmid that was transfected was varied. The total amount of DNA transfected was controlled using pcDNA3A (empty vector). 24 h

post transfection, the cells were lysed and the reporter activities were measured (**Section 2.13.1.1**).

For RNP reconstitutions involving the 3P of the 1918 pandemic influenza virus strains, the RNP reconstitution protocol for 'limiting NP' was used, as above, however, expression plasmids for the PB1, PB2 and PA subunits of the 1918 pandemic influenza virus 3P were used. 24 h post transfection, the cells were lysed and the reporter activities were measured (**Section 2.13.1.1**).

2.13.1.3 Infection of 3P, or 3P and NP Expressing Cells

For experiments where cells expressing exogenous 3P and/or NP were infected, HEK 293T cells were transfected with 0.25 μg expression plasmids for NP and/or PB1, PB2, PA (all of A/WSN/33), and 100 ng of the IFN-luc and β-gal reporter plasmids. 24 h post transfection, cells were infected with A/WSN/33 (MOI = 5). Cells were lysed 12 h post transfection and reporter activities were measured (**Section 2.13.1.1**).

2.13.1.4 Transfecting Extracted RNAs

RNP reconstitutions were performed as per the standard protocol in **section 2.13.1.2,** however, the reporter plasmids were not transfected. 24 h post transfection, the cells were lysed in trizol and the total RNA extracted by chloroform extraction, as per standard protocols. The concentration of the total RNA was measured by nanodrop and was found to be similar in all samples. HEK 293T cells ($1/50^{th}$ of a confluent T75) were transfected with 400 ng of RNA, and 100 ng of the IFN-luc and β-gal reporter plasmids, using 2 μl of lipofectamine. 24 h post transfection, the cells were lysed and the reporter activities measured (**Section 2.13.1.1**).

2.13.2 RNP Reconstitutions in IFN-GFP HEK 293T Cells

IFN-GFP HEK 293T cells were transfected as per **section 2.13.1.2**, however, the reporter plasmids were not transfected. A further condition was added in which an RNP reconstitution was performed using the pPolI GFP plasmid. The pPolI GFP plasmid produces a vRNA (vGFP) which, when transcribed by 3P, generates a GFP-encoding mRNA. At 48 h post transfection, the cell monolayers were imaged by fluorescence microscopy. The cells were then trypsinised until the monolayers had been separated into single cells. The cells were then fixed using 3.7% formaldehyde in 10% FCS/DMEM overnight at 4°C**.** The cells were washed twice in PBS and analysed using a FACS Calibur Flow Cytometer. Data analysis was performed using the FlowJo Software, using the pcDNA3A transfected cells to gate cells as GFP positive or GFP negative.

2.14 Primer Extension Analysis of Viral RNAs

2.14.1 Primer Labelling

Primers for analysing the levels of vRNA or 5S rRNA within samples were labelled with [γ-^{32}P]-ATP using polynucleotide kinase (PNK; Roche). In short, 1 µl of primer (10 µM) was mixed with 1 µl of [γ-^{32}P]-ATP and PNK in 1x PNK kinase buffer A (Roche), and incubated for 1 h at 37°C. The primer was then purified using the QiaQUICK® nucleotide removal kit (Qiagen), following manufacturer's instructions.

2.14.2 vRNA Analyses

RNA was extracted from either cell lysates or directly from cell monolayers using the standard Trizol/chloroform extraction protocol. The extracted RNA was analysed using the [γ-^{32}P]-labelled primers [73]**.** In short, RNA was incubated with 0.25 µl of the relevant [γ-^{32}P]-

labelled primer to detect vRNA (Primer K or L for vRNA; Oligonucleotide Table – **Section 6.1.1**), 0.025 µl of [γ-^{32}P]-labelled control primer to detect 5S rRNA (Primer M; Oligonucleotide Table – **Section 6.1.1**) and 0.475 µl of a non-[γ-^{32}P]-labelled control primer. The total volume of the RNA : primer solution was then adjusted to 5 µl using RNase free H_2O. The solution was heated to 95^{o}C for 2 min to allow RNA denaturation and then cooled on ice for 2 min. The RNA was then combined with 5 µl of a primer extension master mix (2x first-strand buffer, 1 mM mixed dNTP, 20 mM DTT and 50 U SuperScript® III Reverse Transcriptase (Thermofisher)) and incubated for 1 h at 50^{o}C. The samples were then incubated at 70^{o}C for 10 min. 10 µl of formamide was added prior to incubation at 95^{o}C for 3 min. The products were analysed on polyacrylamide gels (12% - 16% acrylamide depending on the product size) containing 7 M Urea in tris-borate-EDTA solution, as previously described [37]. Quantitation was performed using a FUJI phosphorimager. Densitometry was performed using the AIDA software.

2.15 Reverse Transcription PCR

2.15.1 Quantitative Reverse Transcription PCR - MAVS$^{+/+}$ and MAVS$^{-/-}$ MEFs

MAVS$^{+/+}$ and MAVS$^{-/-}$ MEFs were lysed and total RNA extracted. Reverse transcription (rt) was performed using superscript III (Thermofisher), as per manufacturer's instruction, with a dT20 primer (Primer T; Oligonucleotide Table – **Section 6.1.1**). Quantitative PCR (qPCR) analysis was performed using the QuantiTect SYBR® Green PCR Kit (Qiagen), following manufacturer's instructions, and either primers specific to the MAVS mRNA (Primers E and F; Oligonucleotide Table – **Section 6.1.1**) or primers specific to the β-actin mRNA (Primers G and H; Oligonucleotide Table – **Section 6.1.1**).

2.15.2 Reverse Transcription PCR - RNP Reconstitutions

RNP reconstitutions were performed as described in **section 12.2.1.2**. Cells were lysed and the total RNA was extracted at 24 h post transfection. Contaminant DNA was degraded using RQ1 DNAse (Promega; 15 minute incubation), as per manufacturer's instructions. Reverse transcription PCR was performed using SuperScript III (Thermofisher), as per manufacturer's instruction, using a primer specific to the 3' terminal region of the segment 5 vRNA (Primer N; Oligonucleotide table – **Section 6.1.1**). PCR amplification was performed using GoTaq® G2 DNA polymerase and primers specific to the 3'- and 5'-terminal regions of the segment 5 vRNA (Primers N and O; Oligonucleotide Table – **Section 6.1.1**). PCR was performed as per standard protocols, with an annealing temperature of 47^{o}C and an extension time of 1 min.

2.15.3 Reverse Transcription PCR - Virus Infections

HEK 293T cells (1/50th of a confluent T75) were transfected and infected as per **section 2.13.1.3**. However, in a number of conditions, the cells were co-transfected with 33 pmol of an siRNA specific to NP, which has been previously documented[209], or an siRNA specific to GFP (J and I – Oligonucleotide Table – **Section 6.1.1**). Total RNA was extracted at 24 h post infection. Reverse transcription was performed using SuperScript® III (Thermofisher), as per manufacturer's instructions, with a primer specific to the 3'-terminal region of either the segment 1 (Primer P; Oligonucleotide Table – **Section 6.1.1**) or segment 3 (Primer R; Oligonucleotide Table – **Section 6.1.1**) vRNA. PCR amplification was performed using the GoTaq® G2 (Promega) DNA polymerase, following manufacturer's instructions, and primers against the 3'- and 5'-terminal regions of either the segment 1 (Primers P and Q; Oligonucleotide Table – **Section 6.1.1**) or segment 3 (Primers R and S; Oligonucleotide Table – **Section 6.1.1**) vRNA. PCR was performed as per standard protocols, with an annealing temperature of 49^{o}C and an extension time of 4 min.

2.16 Topo-TA Cloning

HEK 293T cells ($1/50^{th}$ of a confluent T75) were transfected with expression plasmids for PB1, PB2 and PA (0.25 µg each). 24 h post transfection, cells were infected with A/WSN/33 (MOI=5). Cells were lysed at 12 h post infection and total RNA was extracted. Reverse transcription was performed using SuperScript® III, as per manufacturer's instructions, with a primer specific to the 3'-terminal region of the segment 1 (Primer P; Oligonucleotide table – **Section 6.1.1**). PCR amplification was performed using GoTaq® G2 (Promega) DNA polymerase and primers specific to the 3'- and 5'-terminal regions of the segment 1 vRNA (Primers P and Q; Oligonucleotide table – **Section 6.1.1**) as per standard protocols, with an annealing temperature of 49°C and an extension time of 0.5 min. The resultant cDNAs were purified using a QiaQUICK® nucleotide removal kit (Qiagen). The cDNAs were ligated into the vector of the TOPO TA® Cloning® Kit for Sequencing and plated onto Ampicillin-containing (50 µg/ml) plates supplemented to 20 µg/ml X-gal and 100 µM IPTG. The plates were incubated for 24 h. The DNA from white colonies was analysed by PCR using GoTaq® G2 (Promega) DNA polymerase and primers specific to the 3'- and 5'-terminal regions of the segment 1 vRNA (Primers P and Q; Oligonucleotide table – **Section 6.1.1**). Purified TOPO-TA® plasmids were sequenced by Sanger Sequencing (Source Bioscience). The length of the vRNA- or cRNA-derived cDNA inserts was then calculated.

Part I

Chapter 3

Mitochondrial Localisation of PB2

3.1 Introduction

The influenza virus polymerase carries out the transcription of viral genes and replication of genomic RNA in the nucleus. Thus, the principal site of localisation of PB1, PB2 and PA is the nucleus [26, 210]. However, PB2 of some influenza virus strains has also been found to localise to the mitochondria [176]. This mitochondrial localisation of PB2 occurs independently of PB1 and PA [175, 176]. The localisation of PB2 to the mitochondria has been hypothesised to be driven by the N-terminus of PB2 [175, 176]. Despite high sequence conservation in the N-terminal region of PB2, position 9 shows variation in influenza A viruses of different hosts and has been shown to be a determinant of the mitochondrial localisation of PB2 [176, 203]. Specifically, sequence analysis of PB2 by *Miotto et al.* showed that human seasonal influenza virus pre-2009 H1N1, H2N2 and H3N2 strains encode asparagine or threonine at position 9 in over 99% of cases [176, 203]. PB2s with an asparagine at position 9 (N9-PB2) localise to the nucleus and mitochondria, whereas the localisation of PB2s with a threonine at position nine (T9-PB2) is not known [176]. On the other hand, avian influenza A viruses encode PB2 with either aspartic acid or glutamic acid at position 9 in greater than 98% of cases [176, 203]. PB2s with an aspartic acid at position 9 (D9-PB2) localise to the nucleus but not to the mitochondria, whilst the localisation of PB2s with a glutamic acid at position 9 is not known [176]. The swine origin pandemic H1N1 influenza A viruses that were first detected in the human population in 2009 encode predominantly non-mitochondrial D9-PB2. This is in agreement with the avian origin of the PB2 segment of these viruses.

The mechanism by which PB2 localises to the mitochondria, however, has not been proven. As stated in **section 1.6.2.2**, PB2 interacts with the outer mitochondrial membrane protein MAVS [53, 159, 176, 177, 211]. The N-terminal 37 amino acids of PB2 (within which position nine is found) are hypothesised to mediate the interaction between PB2 and MAVS [211]. Furthermore, *Boergeling et al.* have provided experimental data that siRNA-mediated knockdown of MAVS decreases the association between PB2 and mitochondria [159]. Together, these results suggest that the mitochondrial association of PB2 might be mediated via its interaction with MAVS at the MOM, and that changes at position nine might affect this association. However, a separate study has found that both mitochondrial and non-mitochondrial PB2 can interact with MAVS, suggesting that the mitochondrial targeting of PB2 occurs independently of its interaction with MAVS [176]. In this study, it was hypothesised that the N-terminus of PB2 contains a MTS that causes the mitochondrial localisation of PB2, and that the presence of aspartic acid at position 9 disrupts the function of this MTS. In summary, it has been hypothesised that PB2 localisation to mitochondria is mediated by either i) the interaction between PB2 and MAVS at the MOM and/or ii) a MTS within the N-terminus of PB2.

Thus, the first aim of this chapter is to investigate the mechanisms of localisation of PB2 to the mitochondria. The second aim of this chapter is to determine the sub-mitochondrial localisation of PB2. This information would act as a first stage in understanding the mitochondrial roles of PB2, as it would reveal which mitochondrial proteins (e.g. matrix, intermembrane space or outer-membrane localised) would be able to interact with mitochondrial PB2.

3.2 Results

3.2.1 Human Seasonal Pre-2009 Influenza A Virus Strains Have a Conservation of Mitochondrial Localising PB2s

Previous work has demonstrated that the majority of human seasonal influenza virus pre-2009 H1N1, H2N2 and H3N2 strains encode N9-PB2 (90% of strains), which localises to the mitochondria, or T9-PB2 (9% of strains). On the other hand, the majority of avian influenza viruses encode D9-PB2 (95% of strains), which does not localise to the mitochondria [176, 203]. Therefore, it was hypothesised that there may be a selection for mitochondrial localising PB2s in human seasonal influenza virus strains.

To further investigate the conservation of mitochondrial PB2s in human seasonal influenza A virus strains, the localisation of N9-, T9- and D9-PB2s was investigated. Wild-type (wt) mouse embryonic fibroblasts (MEFs) were transfected with either the expression plasmids for GFP-tagged N9- or D9-PB2, which have been previously described [176], or with the expression plasmid for GFP-tagged T9-PB2, which was generated by point mutagenesis of the expression plasmid for GFP-tagged N9-PB2. The localisation of PB2 was determined by fluorescence microscopy. The MEFs were stained with Mitotracker to determine the co-localisation of PB2 with the mitochondria. Furthermore, PB2-expressing MEFs were scored as having either nuclear/mitochondrial or nuclear only localisation of PB2. Given that PB2 contains a nuclear localisation sequence, all three PB2s showed nuclear localisation, as expected (Fig. 7 A). As previously reported, there was no detectable GFP signal outside of the nucleus of cells expressing GFP-tagged D9-PB2 (Fig. 7 A and B) (Graef et al., 2010). However, both GFP-tagged N9- and T9-PB2 showed mitochondrial localisation in the vast majority of cells (Fig. 7 A and B).

In order to investigate host species specific differences in the amino acid at position 9 of PB2, the amino acid at position 9 of PB2s of pre-2009 avian influenza virus strains and seasonal human influenza H1N1, H2N2 and H3N2 virus strains was analysed using the Influenza Research Database. The percentage of strains which encode each amino acid at position nine of PB2 is shown. In agreement with *Graef et al.* and *Miotto et al.*, it was found that over 99% of human seasonal influenza virus pre-2009 H1N1, H2N2 and H3N2 strains encoded a mitochondrial localising (N9 or T9) PB2 whereas at least 95% of avian strains encoded a non-mitochondrial localising (D9) PB2 (Fig. 7 C) [176, 203]. Furthermore, this analysis showed that over 99% of avian strains encoded a PB2 with an acidic amino acid (aspartic or glutamic acid) at position 9 whilst over 99% of human seasonal strains encoded a PB2 with a non-acidic acid at position 9 (Fig. 7 C). These observations support the hypothesis that there might be differential evolutionary pressures for the mitochondrial localisation of PB2 between different viral strains and between human and avian hosts.

3.2.2 Mitochondrial Localisation Is Not Mediated By Binding of PB2 to MAVS

Previous work has demonstrated that the localisation of PB2 to the mitochondria may be dependent on the interaction of PB2 and MAVS at the MOM [159]. To test whether the interaction between PB2 and MAVS is required for the mitochondrial localisation of PB2, GFP-tagged N9-, T9- or D9-PB2 was expressed in $MAVS^{-/-}$ MEFs (which had been previously generated from $MAVS^{-/-}$ mice). The MEFs were stained with Mitotracker and the localisation of PB2 was determined as in **section 3.2.1**. As before, GFP-tagged D9-PB2 localised to the nucleus but not to the mitochondria (Fig. 7 D and E) whereas GFP-tagged N9- and T9-PB2 localised to both the nucleus and mitochondria (Fig. 7 D and E). The identity of the $MAVS^{-/-}$ MEFs was tested by qPCR analysis of the levels of

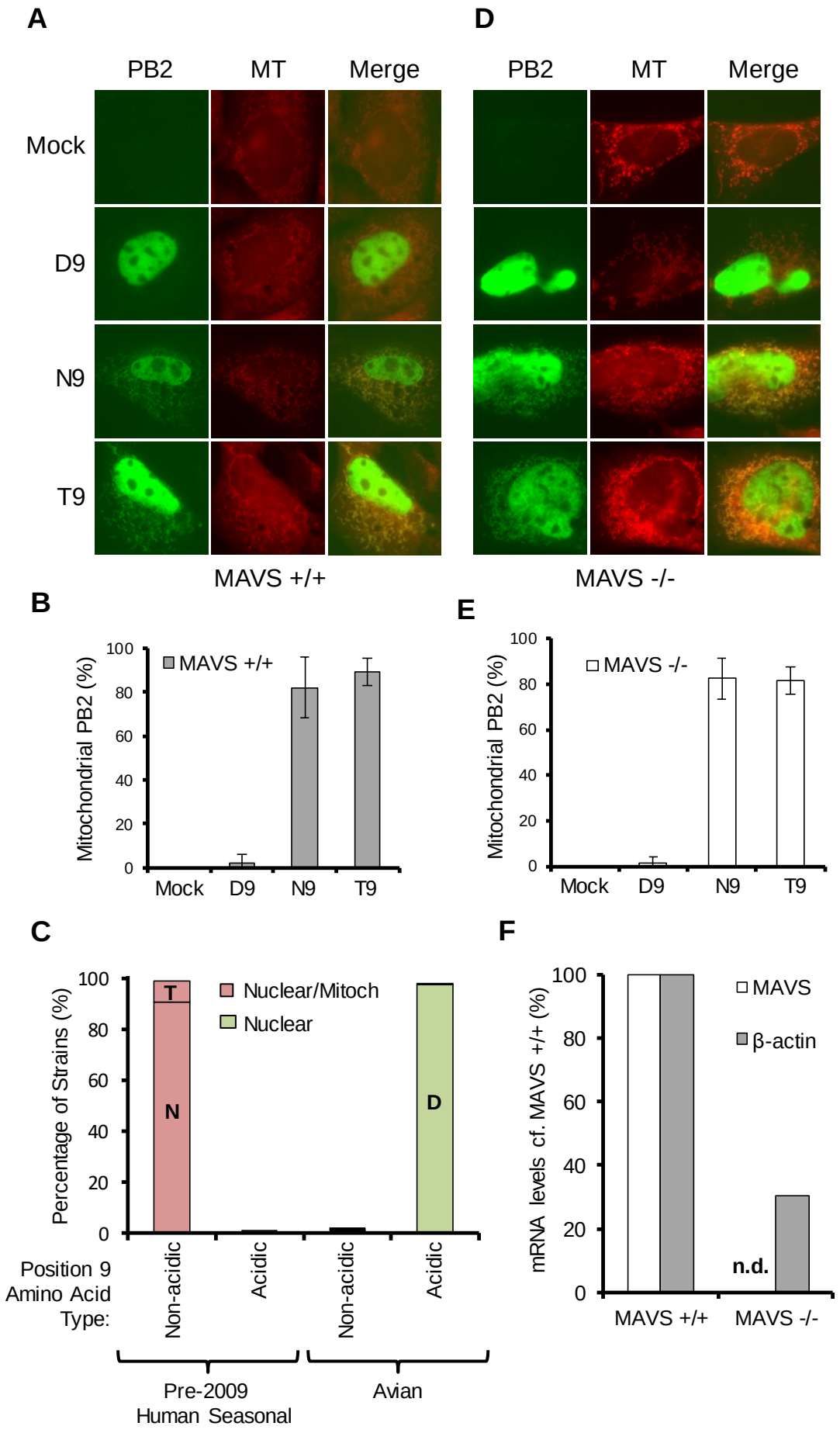

A
PB2
MT
Merge
Mock
D9
N9
T9
MAVS +/+
D
PB2
MT
Merge
MAVS -/-
B
Mitochondrial PB2 (%)
MAVS +/+
Mock
D9
N9
T9
E
Mitochondrial PB2 (%)
MAVS -/-
Mock
D9
N9
T9
C
Percentage of Strains (%)
Nuclear/Mitoch
Nuclear
T
N
D
Position 9 Amino Acid Type:
Non-acidic
Acidic
Non-acidic
Acidic
Pre-2009 Human Seasonal
Avian
F
mRNA levels cf. MAVS +/+ (%)
MAVS
β-actin
n.d.
MAVS +/+
MAVS -/-

Figure 7 – (A-C) Seasonal human influenza virus pre-2009 H1N1, H2N2 and H3N2 strains encode a mitochondrial-localising PB2 whereas avian influenza virus strains encode a non-mitochondrial PB2 (A) Cellular distribution of GFP-tagged N9-, T9- and D9-PB2 in transfected wildtype MEF cells. Cells were stained with the fluorescent mitochondrial marker Mitotracker (MT) and imaged by fluorescence microscopy. Representative images of the GFP and MT signals, as well as a merged image, are shown. (B) PB2-expressing cells were scored for mitochondrial localisation of PB2. Scoring was performed by eye by identifying cells with mitochondrial GFP signal above background. Columns represent the percentage of PB2-expressing cells with mitochondrial PB2 signal. Bars represent standard deviations of average data from three independent experiments (n = 17 to 21 cells/experiment). (C) Analysis of the amino acid encoded at position 9 of the PB2 of avian influenza virus strains and human seasonal pre-2009 H1N1, H2N2 and H3N2 virus strains (using the *fludb* Influenza Research Database). **(D-F) The interaction between PB2 and MAVS is not required for the mitochondrial localisation of PB2** (D) The localisation experiments from 'A' were repeated in MAVS -/- MEFs. Cells were stained, fixed and imaged as in Panel A. Representative images of the GFP and MT signals, as well as a merged image, are shown. (E) PB2-expressing cells were scored for mitochondrial localisation of PB2. Columns represent the percentage of PB2-expressing cells with mitochondrial PB2 signal. Bars represent standard deviations of average data from three independent experiments (n = 17 to 21 cells/experiment). (F) Total RNA was extracted from wt and MAVS -/- MEFs. Reverse transcription was completed using a dT20 primer and qPCR was used to quantify the levels of IFN-β and β-actin mRNA. mRNA levels are normalised against those from the wt cells. n.d. = not detected.

MAVS mRNA and β-actin mRNA in the lysates of wt and MAVS$^{-/-}$ MEFs. With the wt MEFs, both MAVS mRNA and β-actin mRNA were detected (Fig. 7 F). On the other hand, with the MAVS$^{-/-}$ MEFs, despite detecting β-actin mRNA, no MAVS mRNA was detected (Fig. 7 F). This demonstrates that the MAVS$^{-/-}$ cells are not generating detectable levels of MAVS mRNA and thus likely do not express MAVS. The fact that both GFP-tagged N9- and T9-PB2 showed the same localisation pattern in wt MEFs and MAVS$^{-/-}$ MEFs indicates that the interaction between MAVS and PB2 is not required for the localisation of PB2 to the mitochondria. It cannot be excluded that some mitochondrial PB2 is associated to the mitochondria via binding to MAVS, however, the data above demonstrates that a majority of PB2 is localised to the mitochondria in a MAVS-independent manner.

3.2.3 Mitochondrial Localising PB2s Are Imported Into Mitochondria

As determined in the previous section, the localisation of PB2 to mitochondria is not solely mediated by the interaction of PB2 and MAVS. Therefore, this section focusses on determining which other mechanisms mediate the mitochondrial localisation of PB2. Previous work by Carr et al. demonstrated that the extreme N-terminal region of PB2 mediates the mitochondrial localisation of PB2 [175]. The N-terminal region of PB2 is predicted to fold into an α-helix which is amphipathic in nature due to the presence of multiple leucines, isoleucines and charged residues [175]. Amphipathic N-terminal α-helices can act as MTSs and cause the import of proteins into mitochondria [212]. This involves the interaction of the N-terminal amphipathic helix with the Tom20 component of the Translocase of the Mitochondrial Outer Membrane Complex (TOM) [213]. This binding occurs in a sequence independent manner, as long as amphipathicity is maintained. However, Tom5, which

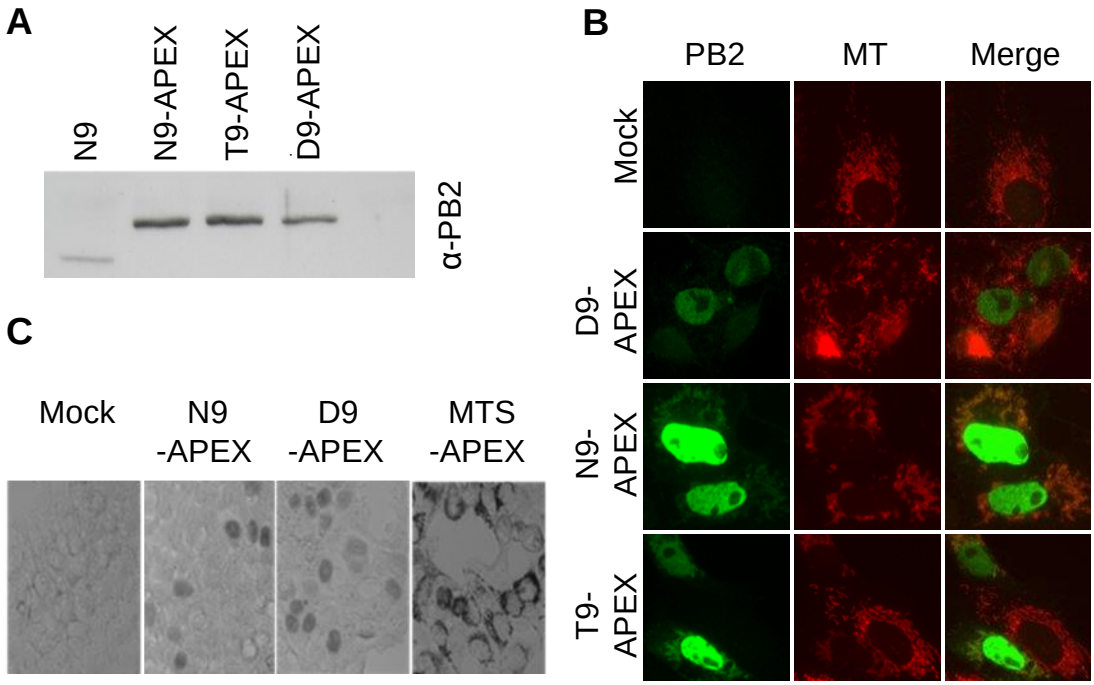


Figure 8 – **APEX is active when fused to PB2, and does not affect the localisation of PB2 –** (A) 293T cells were transfected with an expression plasmid for untagged N9-PB2 or with an expression plasmid for APEX-tagged N9-, T9- or D9-PB2. The cells were lysed 24 h post transfection and the levels of PB2 or PB2-APEX were determined by Western Blot using a polyclonal α-PB2 antibody. (B) Vero cells were transfected with empty vector (mock) or with an expression plasmid for APEX-tagged N9-, T9- or D9-PB2. 24 h post transfection, cells were stained with the fluorescent mitochondrial marker Mitotracker (MT). PB2-APEX was visualised by staining with a polyclonal α-PB2 antibody and a cy2-conjugated secondary antibody. Cells were imaged by fluorescence microscopy. Representative images of the cy2 (PB2) and MT signal are shown as well as a merged image. (C) Cells were transfected with empty vector (mock) or an expression plasmid for N9-PB2-APEX, D9-PB2-APEX or MTS-APEX. 24 h post transfection, cells were fixed and treated for 15 min with DAB and H_2O_2. Cells were imaged by light microscopy and representative images are shown.

facilitates translocation of proteins into the mitochondria through the Tom40 channel, has specificity for basic N-terminal sequences [213]. It was thus hypothesised that the presence of a negatively charged amino acid at position nine of PB2, such as in D9-PB2, might disrupt the binding of the MTS to Tom5 and thus the mitochondrial import of PB2.

In order to determine whether N9- and T9-PB2 are imported into the mitochondria, PB2 was tagged with the APEX (Enhanced Ascorbate Peroxidase) tag. The APEX tag is derived from the ascorbate peroxidase enzyme and, upon addition of H2O2, catalyses the oxidative polymerisation of diaminobenzidine (DAB) to generate a cross-linked and locally deposited precipitate [205]. Furthermore, the precipitate is membrane impermeant and thus is restricted to the cellular compartment in which the APEX tag is localised. The DAB precipitate can be detected directly by light microscopy (LM) or, after staining with heavy metals, by electron microscopy (EM). Thus, fusion of APEX to a protein of interest allows the sub-organelle localisation of the protein to be determined.

In order to test the expression and cellular localisation of N9-, T9- and D9-PB2-APEX, cells were transfected with a PB2-APEX expression plasmid. The cells were then lysed and the levels of PB2 or PB2-APEX were determined by Western Blot using an antibody against PB2. All three PB2-APEX fusions were found to be expressed (Fig. 8 A). The localisation of PB2 was then analysed by immunofluorescence in cells expressing PB2-APEX using a polyclonal α-PB2 antibody. The cells were stained with Mitotracker to determine the co-localisation of PB2 with the mitochondria. All three PB2s showed nuclear localisation, whilst N9-PB2-APEX and T9-PB2-APEX also localised to the mitochondria (Fig. 8 B). This demonstrates that the APEX-tag does not disrupt mitochondrial localisation of N9- and T9-PB2.

Next, in order to test whether the APEX-tag in the PB2-APEX fusion was active, cells were transfected with either a PB2-APEX expression plasmid, an empty vector or with an MTS-APEX expression plasmid. MTS-APEX consists of the APEX-tag fused to an N-terminal MTS, and was chosen as a positive control for APEX-activity as it has been shown previously to be active [205]. The transfected cells were fixed and treated with DAB and H_2O_2 for 15 min. The generation of DAB precipitate was then analysed by LM. No apparent DAB-precipitate was seen in the empty vector (mock) transfected cells. On the other hand, cells expressing PB2-APEX displayed an increased nuclear contrast whilst cells expressing MTS-APEX displayed punctate cytoplasmic regions of high contrast that likely represent the mitochondria (Fig. 8 C). However, the nuclear contrast in the PB2-APEX expressing cells was relatively weak compared to the punctate contrast in MTS-APEX expressing cells. Furthermore, no punctate cytoplasmic contrast could be detected by LM in N9-PB2-APEX transfected cells (Fig. 8 C). The weak signal in the PB2-APEX conditions could have multiple explanations such as: i) Given its larger size, PB2-APEX was likely expressed at lower levels than MTS-APEX. ii) The majority of PB2-APEX is targeted to the nucleus whilst the vast majority of MTS-APEX is targeted to the mitochondria. Given that the nucleus is a larger structure than the mitochondria, DAB precipitate in the nucleus is likely more dispersed and thus results in a lower contrast. iii) PB2-APEX may be intrinsically less active than the MTS-APEX. Overall, despite these issues, all PB2-APEX fusion proteins localised as expected and were active.

For analysis by electron microscopy, cells expressing the untagged PB2s or PB2-APEX fusion proteins were treated with DAB and H_2O_2 and then stained with osmium tetroxide and uranyl acetate. Due to the low contrast generated in PB2-APEX expressing cells, the expression period for PB2-APEX was increased from 24 h to 48 h and the incubation period with DAB and H_2O_2 was also increased. No high contrast was observed in cells expressing

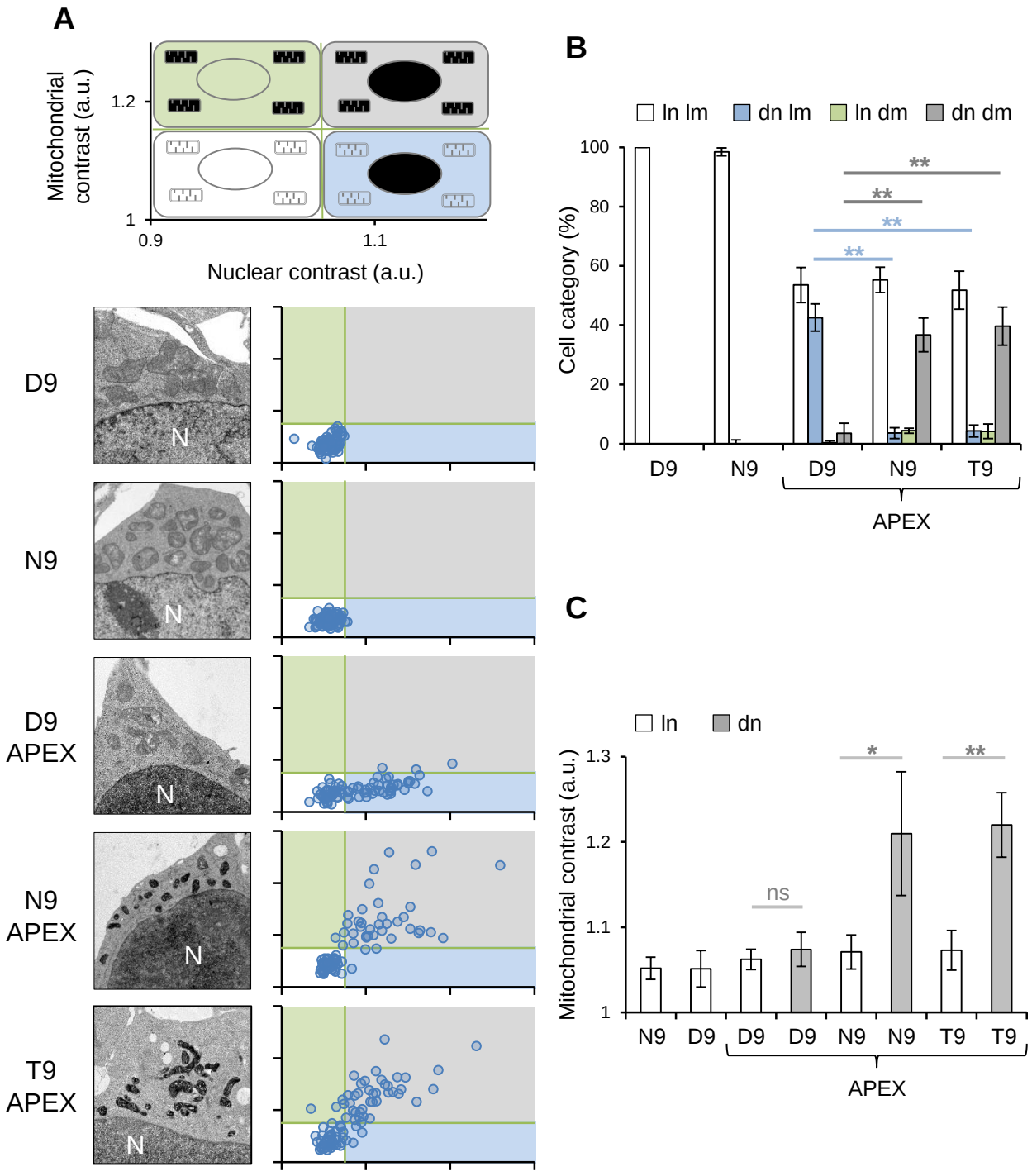

A
Mitochondrial contrast (a.u.)
1.2
1
0.9
1.1
Nuclear contrast (a.u.)
D9
N9
D9 APEX
N9 APEX
T9 APEX
N
B
ln lm
dn lm
ln dm
dn dm
Cell category (%)
100
80
60
40
20
0
**
**
**
**
D9
N9
D9
N9
T9
APEX
C
ln
dn
Mitochondrial contrast (a.u.)
1.3
1.2
1.1
1
ns
*
**
N9
D9
D9
D9
N9
N9
T9
T9
APEX

Figure 9 - N9-PB2 and T9-PB2 are imported into the mitochondria. (A) HEK 293T cells expressing untagged PB2 with D9 or N9 or APEX-tagged PB2 with D9, N9 or T9 were analysed by electron microscopy. Scatter plots show densitometry analysis of the mitochondrial and nuclear contrasts of individual cells relative to their cytoplasmic contrast. Each point represents a single cell. The data was gated into categories of low or high mitochondrial/cytoplasmic (y-axis) and low or high nuclear/cytoplasmic contrast (x-axis) based on the contrast of mitochondria and nuclei of cells expressing untagged PB2. Diagrammatic key displaying axis titles and values for scatter plots is shown at the top. Staining phenotypes of cells for each quadrant are also shown (white, light mitochondria/light nucleus; blue, light mitochondria/dark nucleus; green, dark mitochondria/light nucleus; grey, dark mitochondria/dark nucleus). N, position of the nucleus; a.u., arbitrary units (B) Quantitative analysis of cells expressing untagged PB2 with D9 or N9 or APEX-tagged PB2 with D9, N9 or T9, categorised as having either light nuclei (ln) or dark nuclei (dn) and light mitochondria (lm) or dark mitochondria (dm), based on the gating in Panel A. Columns represent percentage of cells expressing the indicated PB2 belonging to the indicated staining categories. Bars represent standard deviations of average data from three independent experiments. Asterisks indicate a significant difference between samples (one sample t-test; **, $p < 0.01$) (C) Comparison of mitochondrial signals of cells with light (ln) or dark (dn) nuclei expressing the indicated PB2, based on the gating in Panel A. Columns represent average mitochondrial contrast of each population. Bars represent the standard deviations of averages of contrast data from three independent experiments. Asterisks indicate a significant difference between samples (one sample t-test; **, $p < 0.01$; *, $p < 0.05$; ns, $p > 0.05$).

untagged D9- and N9-PB2 (Fig. 9 A). All cells expressing PB2-APEX displayed a strong nuclear staining (Fig. 9 A). However, only cells expressing N9-PB2-APEX and T9-PB2-APEX displayed strong mitochondrial staining (Fig. 9 A).

For quantitative evaluation, imaged cells were analysed by calculating the contrast of each cell's nucleus and mitochondria relative to its cytoplasm. Given that PB2-APEX was not seen to localise to detectable levels in the cytoplasm (Fig. 7 A and D), the assumption was made that the cytoplasmic contrast would be comparable in all cells. The resultant data point from each cell was then plotted onto a scatter plot of increasing nuclear/cytoplasmic (N/C) contrast (x-axis) against increasing mitochondrial/cytoplasmic (M/C) contrast (y-axis) (Fig. 9 A). Cells expressing untagged D9-PB2 and N9-PB2 exhibited low M/C and N/C contrasts (Fig. 9 A). When cells were transfected with the D9-PB2-APEX expression plasmid, a population of cells with low N/C contrast and M/C contrast was seen, which likely represents untransfected cells (Fig. 9 A). However, a second population of cells was observed which had increased N/C contrast with no associated increased in M/C contrast. A similar population of cells with increased N/C was observed in N9- and T9-PB2-APEX expressing cell populations; however, these cells also had an increase in M/C contrast. In order to combine the data from three independent repeats, two gates were produced to sort cells into four categories based on their M/C contrast (low or high) and N/C contrast (low or high). These gates were placed so that the data points for cells expressing untagged N9-PB2 and D9-PB2 were found in the low/ low contrast quadrant. A schematic key of the gating strategy and which type of cell is found in each category is shown (Fig. 9 A, top). The resultant quantitative data showed that the low/low contrast cells occurred at a similar frequency when cells were transfected with the D9-, N9- and T9-PB2-APEX expression plasmids. This supports the idea that they represent untransfected cells (Fig. 9 B). Furthermore, the quantitative data revealed that D9-PB2-APEX expression only led to increased N/C contrast whilst N9- and T9-PB2-APEX

expression caused both increased N/C and M/C contrasts (Fig. 9 B). Together, this demonstrates that N9-PB2 and T9-PB2, as well as containing a nuclear localisation signal, also contain an MTS, which drives their import into mitochondria.

Finally, to quantitatively evaluate whether or not there was any import of D9-PB2-APEX into mitochondria, cells from the three independent experiments in each condition were separated into groups of cells with light (low N/C contrast) or dark (high N/C contrast) nuclei using the gating parameters shown in Figure 9 A, and the average mitochondrial contrast for each group was calculated. As expected, the cells with dark nuclei expressing T9-PB2-APEX and N9-PB2-APEX had a higher average mitochondrial contrast than cells with light nuclei (Fig. 9 C). However, the cells with dark nuclei expressing D9-PB2-APEX did not have significantly different mitochondrial contrast than cells with light nuclei. Therefore, there was no detectable import of D9-PB2 into the mitochondria.

Taken together, these data show that the vast majority of human seasonal influenza virus pre-2009 H1N1, H2N2 and H3N2 strains encode a PB2 which is imported into the mitochondria and that position 9 is a determinant of the mitochondrial import.

3.2.4 PB2 Is Imported into the Mitochondrial Matrix

In order to determine the sub-mitochondrial localisation of mitochondrial PB2, high contrast mitochondria from N9-PB2-APEX, T9-PB2-APEX or MTS-APEX expressing cells were imaged at high magnification. Mitochondria are formed by a pair of membrane bilayers which separate two distinct regions. The first of these, termed the matrix space, is the area surrounded by the MIM and contains the mitochondrial DNA and mitochondrial ribosomes. The MIM, which is highly folded, contains the proteins of the ETC. The MIM is surrounded by the MOM, resulting in a space between the two membranes termed the IMS. The IMS

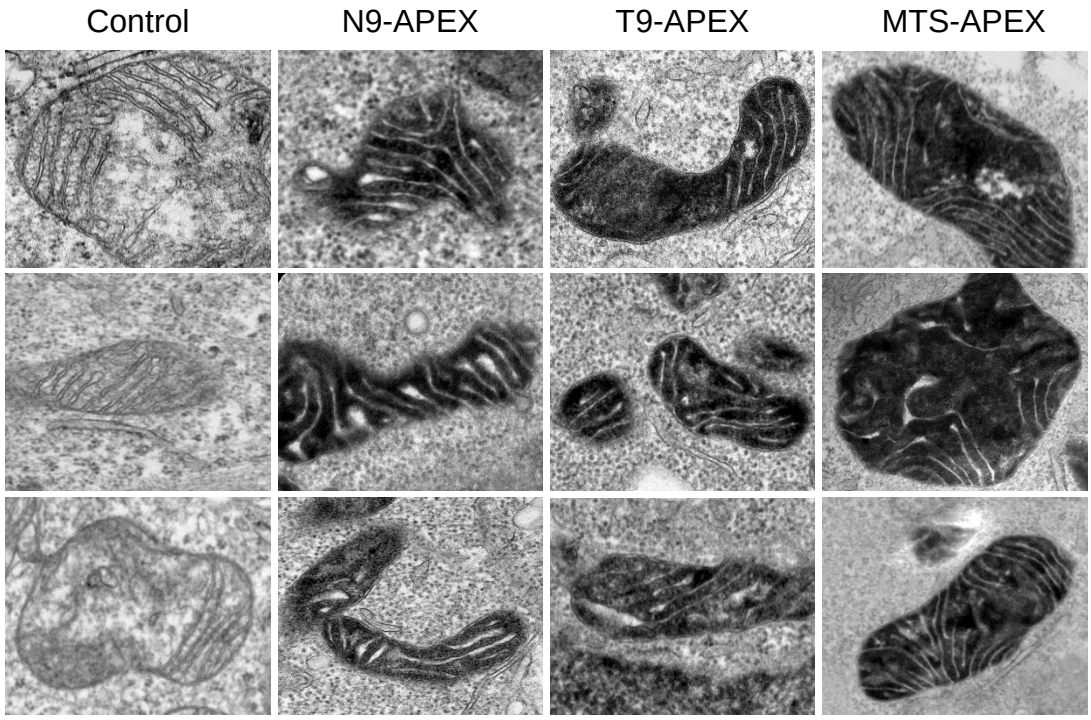


Figure 10 - PB2 is imported into the mitochondrial matrix. HEK 293T cells expressing APEX-tagged N9- or T9-PB2, or MTS-APEX were analysed by electron microscopy. Three representative images of mitochondria are shown per condition.

includes the volume formed by the invaginations in the MIM called cristae. High contrast mitochondria from N9-PB2-APEX and T9-PB2-APEX expressing cells were found to show high contrast in the mitochondrial matrix but not in the IMS, as indicated by the presence of low contrast cristae invaginations (Fig. 10). A similar staining pattern was observed for the MTS-APEX construct, which is known to be targeted into the mitochondrial matrix space [205]. This demonstrates that PB2 is found within the mitochondrial matrix space.

3.2.5 Mitochondrial PB2 Causes Changes in Mitochondrial Morphology

It was hypothesised that matrix localising PB2 might disrupt normal mitochondrial function. For example, it was possible that the import of PB2 into the matrix leads to changes in mitochondrial morphology. In order to test whether this was the case, the EM images from Figure 9 were used to measure the maximum length and area of mitochondria from cells expressing D9-, N9- or T9-PB2-APEX. Furthermore, the number of mitochondria in the cells was also counted. The analysis revealed a significant reduction in the average size of high contrast mitochondria from N9- or T9-PB2-APEX expressing cells relative to the low contrast mitochondria from D9-PB2-APEX expressing cells (Fig. 11 A - left). This decrease in mitochondrial size was accompanied by an increase in average mitochondrial numbers per cell (Fig. 11 B - left). On the other hand, these differences were not observed in cells expressing MTS-APEX relative to untransfected cells, indicating that they were not caused by the import of the APEX tag into mitochondria (Fig. 11 A and B - right). Together, these observations suggest that the matrix localisation of PB2-APEX might lead to fragmentation of mitochondria, potentially by disrupting the mitochondrial fusion/fission equilibrium.

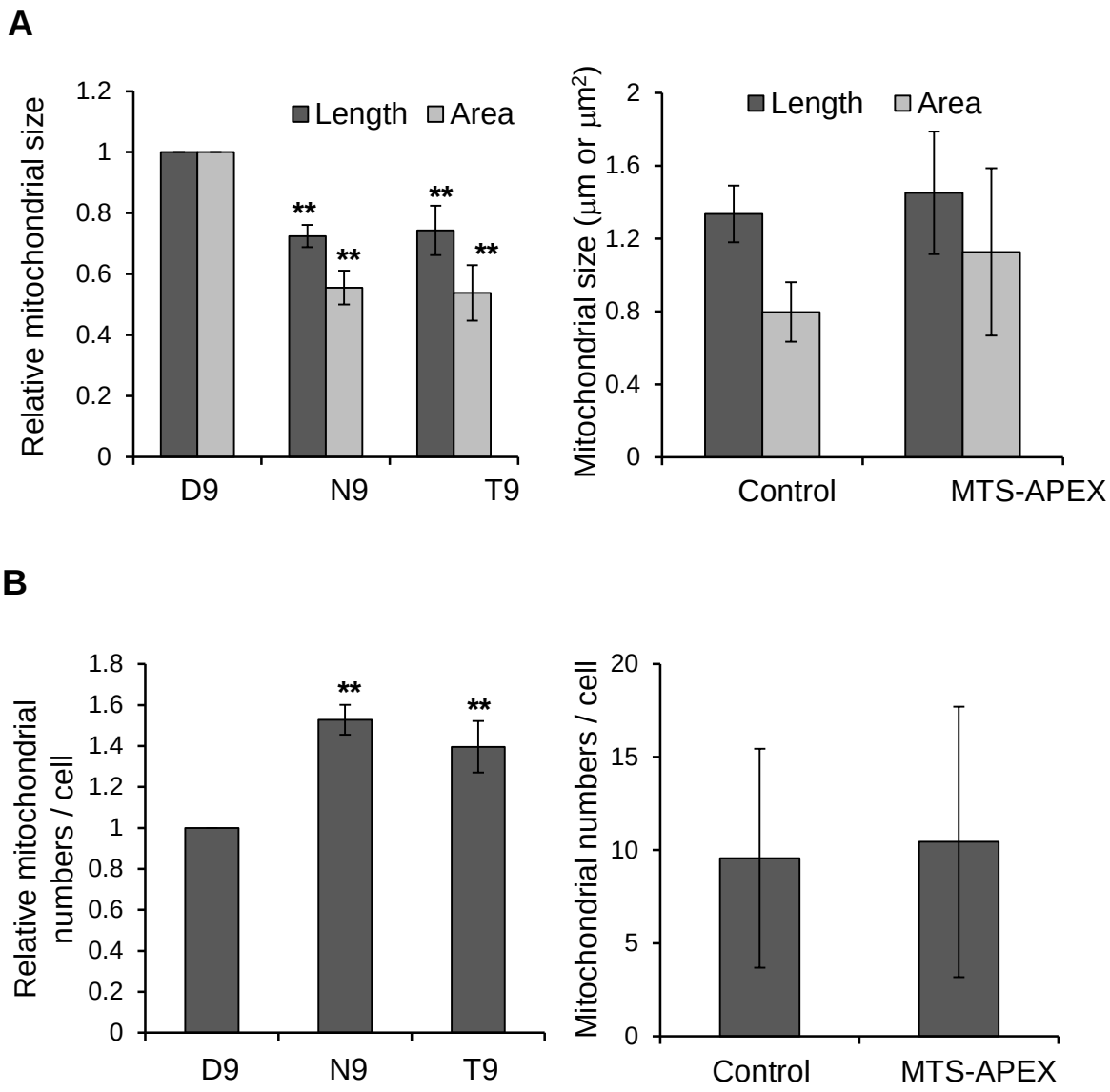


Figure 11 - Mitochondria of cells expressing mitochondrial PB2 show reduced size and increased numbers compared to cells expressing non-mitochondrial PB2. Mitochondria from cells expressing APEX-tagged PB2 with D9, N9 or T9, MTS-APEX or untransfected cells (control) were analysed by measuring their longest diameter and largest cross sectional area (A) and by counting their numbers per cell (B). For APEX-tagged PB2 with D9, N9 or T9, 75 to 200 mitochondria from 13 to 15 cells per condition were analysed (left panels). Columns represent relative mitochondrial length or area (A) or mitochondrial numbers per cell (B) from three independent experiments with data expressed relative to D9-PB2-APEX. Bars represent standard deviations. Asterisks indicate a significant difference between the marked sample and the corresponding D9 sample (one sample t-test; **, $p < 0.01$). For MTS-APEX and untransfected cells (Control), 94 and 67 mitochondria from 9 and 7 cells, respectively, were analysed (right panels). Columns represent mitochondrial length or area (A) or mitochondrial numbers per cell (B). Bars represent standard deviations.

3.5 Discussion

The overall aim of this chapter was to further understand the mitochondrial localisation of PB2. Firstly, this chapter further investigated the difference in conservation of mitochondrial localising PB2s in avian and human-adapted influenza A viruses. Previous work by *Graef et al*. demonstrated firstly that position 9 is a determinant of the mitochondrial localisation of PB2 and, secondly, that there are host specific differences in the conservation of position 9 between human seasonal influenza A virus pre-2009 H1N1, H2N2 and H3N2 strains, and avian strains [176]. Specifically, it was shown that the vast majority of human seasonal pre-2009 H1N1, H2N2 and H3N2 strains encode a mitochondrial localising PB2 (N9-PB2) whilst the vast majority of avian strains encode a non-mitochondrial PB2. Here, the mitochondrial and non-mitochondrial localisation of N9-PB2 and D9-PB2, respectively, were confirmed. Furthermore, T9-PB2, which was found in approximately 9% of human influenza A virus pre-2009 H1N1, H2N2 and H3N2 strains, was also shown to be compatible with the mitochondrial localisation of PB2. Given the prevalence of these amino acid residues in influenza virus isolates (as shown by sequence analysis), these results suggest that the PB2 of over 99% of all human seasonal pre-2009 H1N1, H2N2 and H3N2 influenza viruses was targeted into the mitochondrial matrix. On the other hand, the vast majority (95%) of avian influenza viruses encode non-mitochondrial D9-PB2. This implies that there may be a selection for viruses encoding mitochondrial localising PB2 during influenza A virus infections in the human host, which is not present in the avian host. Interestingly, the localisation of PB1-F2 also differs between influenza A virus strains. Most influenza A viruses encode full length 87 to 90 amino acid long PB1-F2s that associate to the mitochondria due to the presence of the C-terminal MTS and act to disrupt mitochondrial function (further discussed in **section 4.1**). However, all human pre-2009 H1N1 viruses isolated after 1947 encoded a truncated, 57 amino acid long, non-mitochondrial PB1-F2 due

to the presence of a stop codon, while pandemic 2009 H1N1 viruses and the resultant H1N1 human seasonal lineage encode a truncated 11 amino acid, non-mitochondrial version [165, 201]. Overall, therefore, these observations support the idea that mitochondria are actively targeted by influenza A viruses, that this might be mediated by both PB1-F2 and PB2, and that there might be differential evolutionary pressures for modulating mitochondrial function between different viral strains and between human and avian hosts.

The second focus of this chapter was to determine the mechanism by which PB2 localises to the mitochondria. It was found that, in contrast to previous reports [159], the mitochondrial localisation of PB2 is not dependent upon its interaction with the MOM-protein MAVS. This is in agreement with previous data that both mitochondrial and non-mitochondrial PB2 can interact with MAVS [176]. If PB2 is not recruited to the mitochondria through its binding to MAVS, the question arose as to which other mechanism could be targeting PB2 to the mitochondria. It has previously been proposed that the N-terminal α-helix of PB2 acts as an MTS. This is supported by the fact that this helix contains a number of hydrophobic and positively charged amino acid residues, hallmarks of amphipathic helices implicated in mitochondrial targeting. Indeed, the mutation of two hydrophobic residues within this helix (L7 and L10) has been shown to disrupt the mitochondrial localisation of PB2 [175]. MTS-containing proteins are recognised by the Tom20 component of the translocase of the outer membrane (TOM) complex, which acts as a general entry gate for protein import into the mitochondria by mediating protein transfer into the IMS [213]. The Tom5 component of the TOM complex, which is acidic in nature, is hypothesised to be involved in promoting the translocation of proteins which contain a basic MTS [213]. Thus, *Graef et al.* hypothesised that the presence of an acidic amino acid at position 9 might disrupt the interaction between PB2 and Tom5, and thus disrupt mitochondrial import of PB2 [176]. In order to test whether position 9 is a determinant of the mitochondrial import of PB2, the sub-mitochondrial localisation of

APEX-tagged PB2 was determined by EM. It was found that N9- and T9-PB2s are imported into the mitochondrial matrix, whilst D9-PB2s are not. Therefore, it is likely that, after passage through the TOM channel, N9- and T9-PB2 are imported into the matrix space by the translocase of the inner membrane TIM23 complex [214]. In support of this hypothesis, our group recently identified a range of chaperones and components of the TIM23 and TOM complexes, including Tom22 and Tim50, that co-purify with PB2 from influenza virus infected cells [206]. In summary, it is likely that PB2 is delivered to the mitochondria with the assistance of the Hsp70 and Hsp90 chaperones [206, 215, 216, 217] and then its translocation into the mitochondrial matrix is mediated by the TOM and TIM23 complexes of the mitochondrial import pathway.

The finding that PB2 is imported into the mitochondrial matrix raises further questions about the functional implications of the mitochondrial association of PB2. As PB2 is imported into the mitochondrial matrix, mitochondrial PB2 is unlikely to interfere with the function of MAVS by directly binding to it as previously suggested [176]. We have not definitively proven here, however, that MAVS binding does not have a role in matrix import. Given that matrix localising PB2 leads to a decrease in mitochondrial size and an increase in their numbers, it is possible that matrix PB2 disrupts the mitochondrial fusion/fission equilibrium. Mitochondrial fusion and fission is controlled by the mitofusin proteins and Opa1, respectively. Mitofusins and mitochondrial dynamics are important for many processes, including mitochondrial metabolism, IFN induction and clearance of damaged mitochondria [218, 219]. Furthermore, mitochondrial fragmentation occurs during the process of programmed cell death [220].

In summary, over 99% of human seasonal influenza A virus pre-2009 H1N1, H2N2 and H3N2 strains encode a mitochondrial localising PB2. This localisation is not mediated by the interaction between PB2 and the MOM-protein MAVS but occurs due to the import of PB2

into the mitochondrial matrix. Although the effects of matrix PB2 import have not been investigated, the import of PB2 appears to result in mitochondrial fragmentation.

3.6 Summary

A) Both PB2s with asparagine or threonine at position 9 localise to mitochondria.

B) Over 99% of pre-2009 H1N1, H2N2 and H3N2 human seasonal influenza virus strains encode a mitochondrial localising PB2, whilst over 95% of avian strains encode a non-mitochondrial localising PB2. This highlights a potential selection for mitochondrial localising PB2s in human seasonal influenza virus strains.

C) Mitochondrial PB2 is not recruited to the mitochondria by its interaction with MAVS, as previously hypothesised.

D) The presence of asparagine or threonine at position 9 of PB2 leads to the import of PB2 into mitochondria.

D) Mitochondrial PB2 is found in the mitochondrial matrix space.

E) Cells expressing mitochondrial PB2 contain a higher number of mitochondria which are on average smaller than the mitochondria of cells expressing non-mitochondrial PB2.

Chapter 4

Mitochondrial Functions of PB2

4.1 Introduction

Viruses target mitochondrial processes during infection to promote their own replication (Anand and Tikoo, 2013). Given that mitochondria act as central hubs for energy production and distribution, apoptosis, and innate immune responses this is perhaps not surprising. Mitochondria are best known for their role in oxidative ATP synthesis, which requires the generation of the MMP across the mitochondrial inner membrane. However, they also contain cell death mediators and are involved in the induction of cytokine expression [136, 221, 222].

One example of a process which is widely modulated by viruses is the induction of apoptotic cell death. By inhibiting cell death, viruses can increase their ability to replicate or induce latency. On the other hand, viruses can increase virion release and dissemination by promoting cell death [223]. The induction of apoptotic cell death is normally triggered by the permeabilisation of the mitochondrial membranes; a process controlled by the Bcl-2 family of proteins. Viruses such as hepatitis B virus and human adenovirus encode apoptosis regulating proteins with Bcl-2 homology which can be either pro- or anti-apoptotic in nature [224, 225]. Another targeted process is the induction of IFN expression in response to viral infection, such as that mediated by the RIG-I/MAVS signalling pathway. Multiple viruses act to dampen the signalling in this pathway. One example is hepatitis C virus, which expresses a protein that causes cleavage of MAVS and this limits IFN release [107]. Beyond this,

mitochondrial processes such as calcium homeostasis, the maintenance of the MMP and the generation of reactive oxygen species have also been found to be affected by viruses [223].

In terms of influenza viruses, the PB1-F2 protein has been shown to disrupt mitochondrial function [167, 168, 169, 202]. The import of PB1-F2 into mitochondria is driven by a MTS in the C-terminal half of PB1-F2 and is mediated by the TOM40 channel [202]. This leads to its targeting to the IMS where it is tightly associated to the MIM [167, 202]. Here, the protein leads to mitochondrial membrane permeabilisation, disrupts the MMP and triggers the induction of apoptosis [164, 168] either by directly forming membrane pores or by interacting with ANT3/VDAC [168, 169]. As well as disrupting the MMP and inducing cell death, the import of PB1-F2 into mitochondria disrupts RIG-I/MAVS signalling, NLRP3 inflammasome activation and leads to mitochondrial fission [202]. However, it has been hypothesised that the disruption of all of these functions is linked to the loss of the MMP [130, 170, 202]. This demonstrates how, by disrupting a single mitochondrial process, PB1-F2 can potentially indirectly interfere with multiple anti-viral processes such as the induction of IFN expression and the generation of pro-inflammatory cytokines.

In terms of the effects of PB2 on mitochondrial function, very little is known. Recombinant influenza A/WSN/33 viruses expressing PB2 with mutations at L7 and L10 in the MTS, rendering PB2 non-mitochondrial, showed reduced growth in cell culture and animal models [175]. It has also been show that a recombinant influenza A/WSN/33 virus encoding wild-type mitochondrial N9-PB2 (N9-WSN) replicates to higher titres in mice than an equivalent virus encoding a non-mitochondrial D9-PB2 (D9-WSN). Finally, a D9N mutation in the PB2 of an avian H5N1 virus resulted in increased virulence in mice [226]. Together, these observations indicate that the mitochondrial localisation of PB2 may have an important function in increasing virus replication in the mammalian host. In terms of the cellular effects of PB2, it has been proposed that mitochondrial PB2 acts to preserve the MMP and acts as a dampener

of IFN induction [175, 176]. The latter has been hypothesised to occur via the binding of PB2 to MAVS at the MOM, however, the matrix localisation of PB2 seems incompatible with this model.

This chapter, therefore, focuses on attempting to further understand the mitochondrial functions of matrix-localised PB2.

4.2 Results

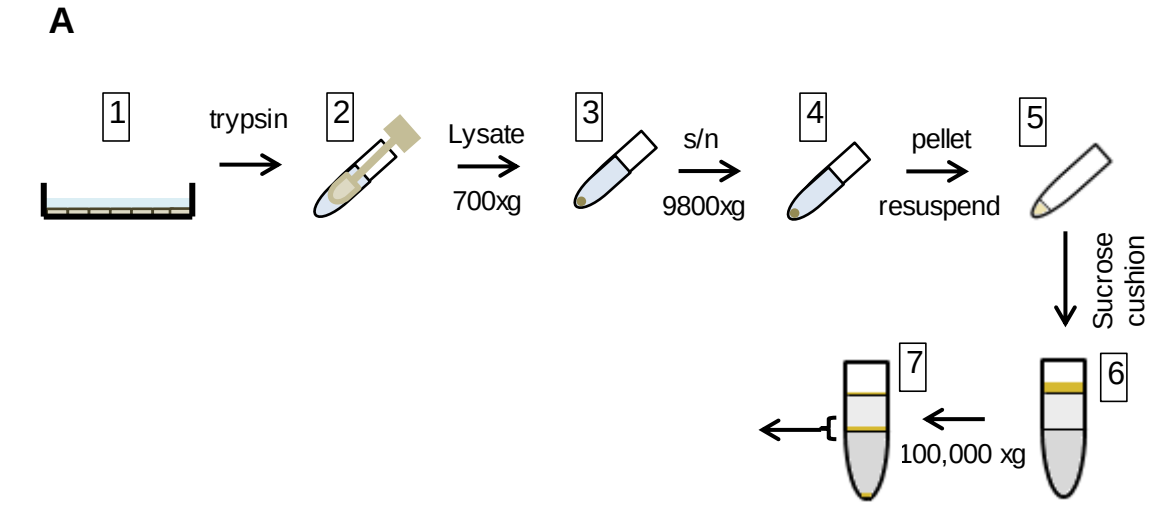
4.2.1 Optimising the Isolation of Mitochondria From Infected Cells

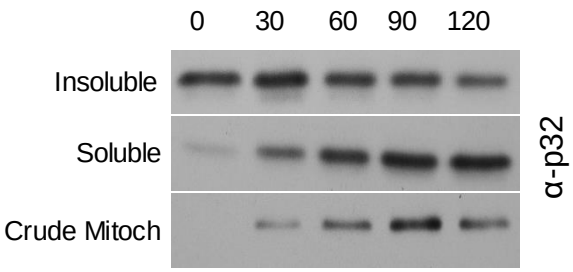
Given i) that mitochondrial PB2 has previously been reported to be a dampener of IFN induction and ii) that the mitochondrial localisation of PB2 has been reported to promote viral replication in mice, the rest of this chapter focusses on using mass spectrometry (MS) to identify the potential roles of matrix PB2 [176]. This MS analysis specifically aims to i) determine the effect of mitochondrial PB2 on the mitochondrial proteome and ii) characterise the interactome of mitochondrial PB2.

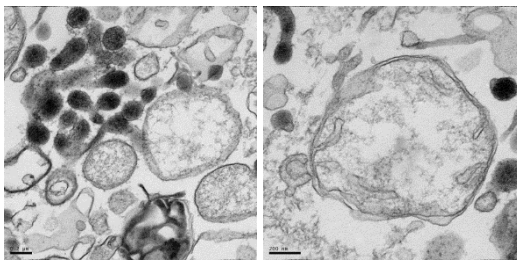
In order to complete these analyses, mitochondria needed to be isolated from cells for multiple reasons: i) to attempt to decrease the complexity of the samples submitted for MS analysis by removing cytoplasmic and nuclear proteins, ii) to allow the quantification of the mitochondrial levels of a protein which is also found in a secondary cellular location, and iii) to allow the specific purification of mitochondrial PB2. For this purpose, a protocol for the isolation of mitochondria was designed. This protocol consisted of i) the lysis of cells by dounce homogenisation in order to liberate mitochondria (Fig. 12 A – 1 to 2), ii) the removal

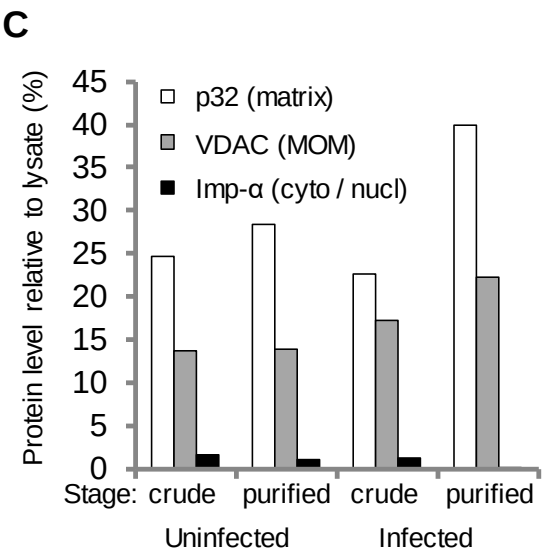
of unlysed cells and insoluble material from the lysate by centrifugation (Fig. 12 A – 2 to 3), iii) an initial purification of mitochondria from the lysate by centrifugation (the resultant sample in this thesis is termed the 'crude mitochondria') (Fig. 12 A - 3 to 5) and iv) a secondary purification of mitochondria by ultracentrifugation on a sucrose cushion (the resultant sample in this thesis is termed the 'purified mitochondria') (Fig. 12 A – 5 to 7). This section focusses on optimising this isolation protocol and evaluating the purity of the isolated mitochondria.

Firstly, in order to determine the optimal level of lysis to liberate intact mitochondria, A549 cells were subjected to 0, 30, 60, 90 or 120 strokes of a dounce homogeniser. The isolation protocol described above was then performed on the resultant lysates, and samples were collected of three different fractions: i) the unlysed cells and insoluble material from the lysate (shown in Fig. 12 A - 3), ii) the soluble material from the lysate (shown in Fig. 12 A - 3) and iii) the crude mitochondria (Fig. 12 A - 5). The level of the matrix protein p32 in each fraction was then determined by Western Blot. As the number of strokes was increased, the level of p32 in the insoluble fraction decreased whereas the level in the soluble fraction increased (Fig. 12 B). The level of p32 in the crude mitochondria sample increased between 0 and 90 strokes, but decreased between 90 and 120 strokes (Fig. 12 B). The decrease and increase in p32 levels in the insoluble and the soluble fractions of the lysate, respectively, is likely explained by the release of mitochondria from cells. Thus, a higher number of strokes appeared to lead to an increased liberation of mitochondria. This also likely explained the increased level of p32 in the crude mitochondrial sample up to 90 strokes. The decrease in p32 levels between 90 and 120 strokes likely was due to the damage to mitochondria due to excessive homogenisation, which would lead to the release of p32 from the mitochondria.


A
1
trypsin
2
Lysate
700xg
3
s/n
9800xg
4
pellet
resuspend
5
Sucrose
cushion
6
7
100,000 xg


B
0
30
60
90
120
Insoluble
Soluble
Crude Mitoch
α-p32


D


C
Protein level relative to lysate (%)
45
40
35
30
25
20
15
10
5
0
p32 (matrix)
VDAC (MOM)
Imp-α (cyto / nucl)
Stage: crude
purified
crude
purified
Uninfected
Infected

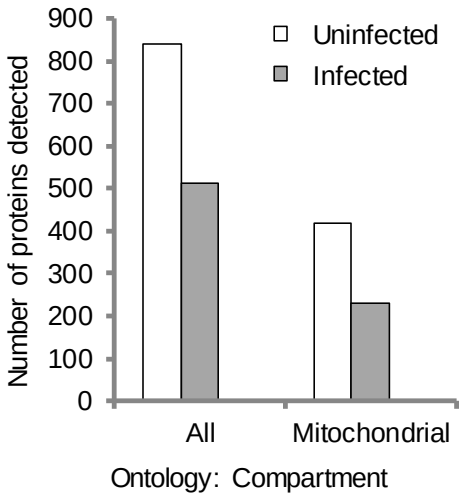

E
Number of proteins detected
900
800
700
600
500
400
300
200
100
0
Uninfected
Infected
All
Mitochondrial
Ontology: Compartment

Figure 12 – Optimisation of a protocol for the isolation of mitochondria – (A) Protocol for mitochondrial isolation. Infected or uninfected A549 cells were trypsinised [1] and lysed using a dounce homogeniser [2]. A 700xg centrifugation was used to pellet 'unlysed cells and large insoluble material' [3]. The supernatant (s/n) was collected and a 9800xg centrifugation was used to pellet mitochondria [4]. The resultant pellet was resuspended [5] to form the 'crude mitochondria' sample. The crude mitochondria were then loaded onto a sucrose cushion [6]. The sucrose cushion was then subjected to ultracentrifugation at 100,000xg and the material at the cushion interface was collected [7]. The collected material is termed the 'purified mitochondria'. (B) A549 cells were subjected to different degrees of homogenisation (0 – 120 strokes with a dounce homogeniser) and then the mitochondrial isolation protocol shown in Panel A was performed. Samples were taken at stage 3 (insoluble and soluble material) and at stage 5 (crude mitochondria). The samples were analysed by Western Blot using an anti-p32 antibody. (C) Mitochondria were purified from uninfected or infected A549 cells, and samples were taken of the lysate at stage 2 (lysate), of the material at stage 5 (crude mitochondria), and of the material at the cushion interface at stage 7 (purified mitochondria). The samples were analysed by Western Blot using an anti-Importin-α (a cytoplasmic and nuclear protein; cyto / nucl) antibody, an anti-p32 (matrix protein) antibody and an anti-VDAC (mitochondrial outer membrane protein; MOM) antibody. Densitometry analysis of the Western Blots are shown. Values represent the level of a protein in a given sample relative to the protein level in the lysate. (D) The 'purified mitochondria' from the optimised isolation protocol was analysed by EM (F) A549 cells were infected with A/WSN/33 or were mock infected. 12 h post infection, cells were lysed and their mitochondria isolated. The purified mitochondria samples were analysed by MS. The number of proteins detected by mass spectrometry analysis was calculated. Bars shows the total number of detected proteins (all) and the number of detected proteins with mitochondrial ontology (mitochondria).

Based on this analysis, 90 strokes with a dounce homogeniser was determined to lead to the most efficient liberation of intact mitochondria.

Next, in order to determine the purity of the mitochondria generated by the isolation procedure, the levels of proteins in the crude and purified mitochondria samples relative to the lysate were analysed. Specifically, levels of the matrix protein p32, MOM protein VDAC and the nuclear/cytoplasmic protein importin α1 were analysed by Western blot (data not shown). Densitometry was then used to quantify the protein levels in each sample and to calculate the levels of the proteins in the crude and purified mitochondrial samples relative to the lysate (Fig. 12 C). It was found that the levels of VDAC and p32 in the crude and purified mitochondrial samples were approximately 25-35% of that in the lysate, whilst the levels of importin α1 in the crude and purified mitochondrial samples were approximately 1-2% of that in the lysate (Fig. 12 C). The fact that the isolation procedure leads to enrichment in VDAC and p32, but not in importin α1, indicates that mitochondria are isolated. Furthermore, the fact that VDAC and p32 are enriched comparably indicates that the purified mitochondria do not contain high levels of i) ruptured mitochondria, which would release p32, or ii) mitoplasts, which are mitochondria which have lost the MOM. Overall, therefore, it appears that the isolation protocol leads to a semi-pure sample of intact mitochondria.

Given that the mitochondria were isolated based on their density using the sucrose step gradient, it was hypothesised that the mitochondrial isolation procedure would also purify any organelle which has a similar density to mitochondria. To test the purity of the isolated mitochondria, the purified mitochondria samples were analysed by EM. This analysis showed that the purified mitochondria samples contained a range of membranous structures around the size of mitochondria (Fig. 12 D - left). This indicates that multiple organelles were being isolated along with mitochondria. Furthermore, structures containing double membranes, consistent with mitochondria, were observed (Fig. 12 D - right). It was difficult, however, to

definitively identify mitochondria in the samples, likely due to changes to the mitochondrial ultrastructure upon fixing and exposure to high sucrose concentrations. These images, however, are consistent with EM analysis of isolated mitochondria from other studies and therefore the isolation protocol was determined to be suitable [227].

Finally, to analyse the protein content of mitochondria isolated from either infected or uninfected cells, purified mitochondria samples from said conditions were submitted for MS analysis. It was found that approximately half of the detected proteins had mitochondrial ontology (Fig. 12 E). This data indicated that the purified mitochondria sample is enriched in mitochondria.

Overall, therefore, it appeared that the optimised mitochondrial isolation protocol described above yields a sample which is highly enriched in mitochondrial proteins, and that is comparable to the purified mitochondria samples from previous MS-based mitochondrial proteome studies [227].

4.2.2 Analysing the Effect of Influenza Virus Infection on the Mitochondrial Proteome

4.2.2.1 Generation of SILAC Cell Lines

Section 4.2.3 focusses on investigating the effects of matrix-PB2 on mitochondrial function by quantifying the proteome of mitochondria extracted from cells which have been either: i) mock infected, ii) infected with an influenza A virus encoding a mitochondrial PB2 (N9-WSN), or iii) infected with an influenza A virus encoding a non-mitochondrial PB2 (D9-WSN). It was hypothesised that this could potentially give information on: i) the effects of influenza A virus infection on mitochondria, and ii) the specific effects of mitochondrial PB2

on mitochondria. A similar MS based strategy has been used by other groups to investigate the effects of virus infections on mitochondria [227, 228].

To be able to quantify the effects of influenza virus infection on the mitochondrial proteome during infection, it was necessary to isolate mitochondria from infected and uninfected cells using the technique described in the previous section. However, it was hypothesised that it would not be possible to repeat the mitochondrial isolation procedure identically on sequential samples as the dounce homogenisation and collection of mitochondria from the sucrose cushion are manual processes. Thus, it was necessary to combine the infected and uninfected cells before cell lysis and mitochondrial isolation.

In order to distinguish proteins from the infected and uninfected cells in the combined sample, the stable isotope labelling by amino acids in cell culture (SILAC) technique was used. SILAC labelled cell lines were generated by growing A549 cells for approximately 30 doubling times in medium containing either 'light' (K0, R0), 'medium' (K4, R6) or 'heavy' (K8, R10) arginine and lysine. The resultant cell lines will be termed the light, medium and heavy cell lines. The human A549 cell line was chosen as it is derived from alveolar epithelial cells and thus can be used as a model of the cells of the URT, which are infected by influenza A viruses. To determine whether the medium and heavy cell lines contained any residual light amino acid containing proteins, samples of the medium and heavy cell lines were lysed and the lysates were submitted for MS analysis. Due to their different masses, peptides containing either light, medium or heavy arginine or lysine can be distinguished. The cells which were grown in medium amino acid containing media contained peptides with 'medium' arginine and lysine, but non-detectable levels of peptides with light arginine and lysine (data not shown). The cells which were grown in heavy amino acid containing media contained peptides with 'heavy' arginine and lysine, but non-detectable levels of peptides with light arginine and lysine (data not shown). Overall, this demonstrates that the medium

and heavy cell lines did not contain detectable levels of light lysine and arginine, and thus were suitable for use.

4.2.2.2 Testing the Infection of SILAC Cell Lines

In order to characterise the effects of influenza A virus infection on mitochondria, it was important that the generated SILAC cell lines could be infected with influenza A virus.

Therefore, in order to test whether the SILAC cell lines could be infected by and support the replication of influenza A virus, the light, medium and heavy SILAC cell lines were either infected with D9-WSN, mock infected or infected with N9-WSN, respectively. The infection of the SILAC cells lines was analysed in two ways: i) the cells were fixed at 12 h post infection and expression of NP was analysed by immunofluorescence using an α-NP antibody, or ii) the replication of the virus in the SILAC cell lines was measured by plaque assay. It was found that NP was expressed and localised as expected in infected cells (Fig. 13 A), however, no NP expression was seen in the mock infected cells (data not shown). Furthermore, the replication of influenza A virus in the infected light and heavy labelled SILAC cell lines proceeded with the expected growth kinetics and led to virus titres of approximately 5 x 10^6 pfu/ml (Fig. 13 B). Overall, this data demonstrates that the tested SILAC cell lines can be infected by and support the replication of influenza A virus.

4.2.2.3 Quantifying the Effect of Influenza Virus Infection on the Levels of Mitochondrial Proteins

In order to quantify the effect of infection on the mitochondrial proteome, the SILAC cell lines were infected with either N9-WSN or D9-WSN, or were mock infected. In each of the three completed repeats, the combination of cell line and condition was changed (e.g. the

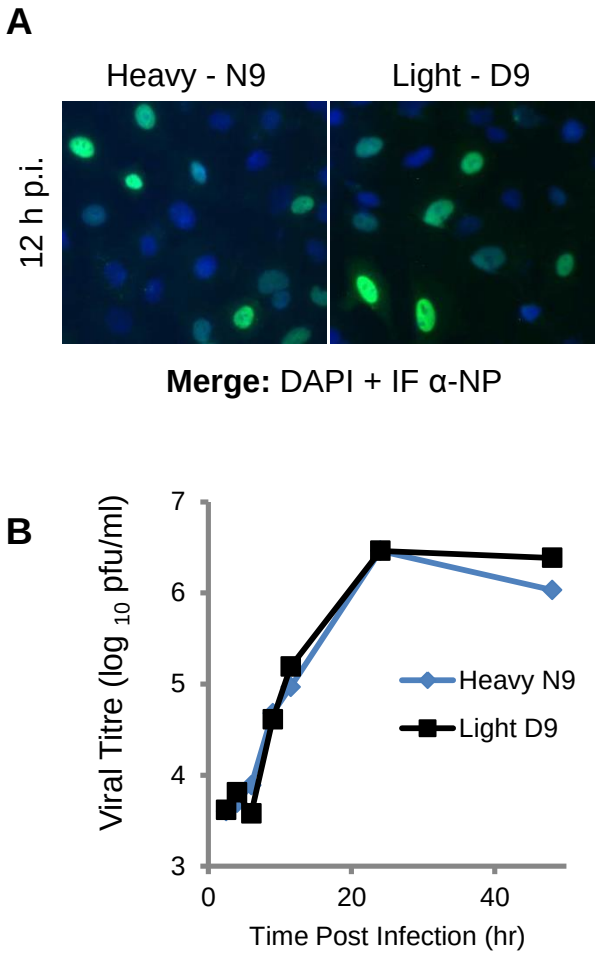


Figure 13 – Optimising infections of SILAC cell lines – (A) The light, medium and heavy-labelled SILAC cell lines were either infected with a virus encoding a non-mitochondrial PB2 (D9), mock infected or infected with a virus encoding a mitochondrial PB2 (N9), respectively. (A) The cells were infected at an MOI of 1. At 12 h post infection, cells were fixed and immunofluorescence was completed using an α-NP antibody and a cy2-conjugated secondary antibody. The cells were then stained with DAPI. Representative merged images of the cy2 (green) and DAPI (blue) signals from the N9 and D9 infected cells are shown. (B) The cells were infected at an MOI of 0.01. Samples of the media were taken at various time points post infection and the virus titres were quantified by plaque assay.

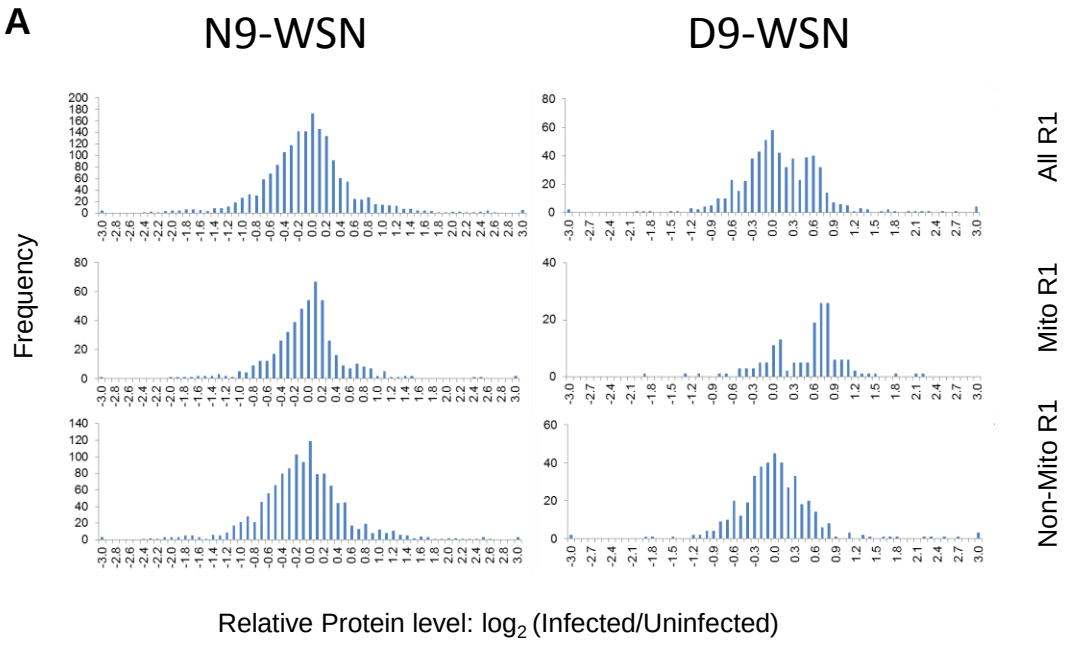
light labelled cells were mock infected in repeat 1, N9-WSN infected in repeat 2 and D9-WSN infected in repeat 3). 12 h post infection, the cells were trypsinised and equal numbers of cells from the three conditions were combined. The mitochondrial isolation procedure was then performed and the 'purified mitochondria' were submitted for MS analysis. After MS analysis of the isolated mitochondria, detected peptides were assigned to the relevant protein in the relevant labelled state using the Central Proteomics Facilities Pipeline (CPFP). This yielded a list of detected proteins from each of the different SILAC conditions. Quantitative data on the levels of each protein in each labelled state (light, medium or heavy labelled) were generated by SINQ analysis. The resultant data was used to generate protein level fold changes upon infection (relative protein levels), by dividing the level of a protein from each of the infected samples by the level of the same protein in the uninfected sample. Therefore, each repeat produced: i) a list of proteins detected in both the N9-WSN and the mock infected samples, and a relative protein level for each protein, and ii) a list of proteins detected in both the D9-WSN and the mock infected samples, and a relative protein level for each protein. These lists will, from herein, be referred to as 'datasets' i.e. the N9- dataset.

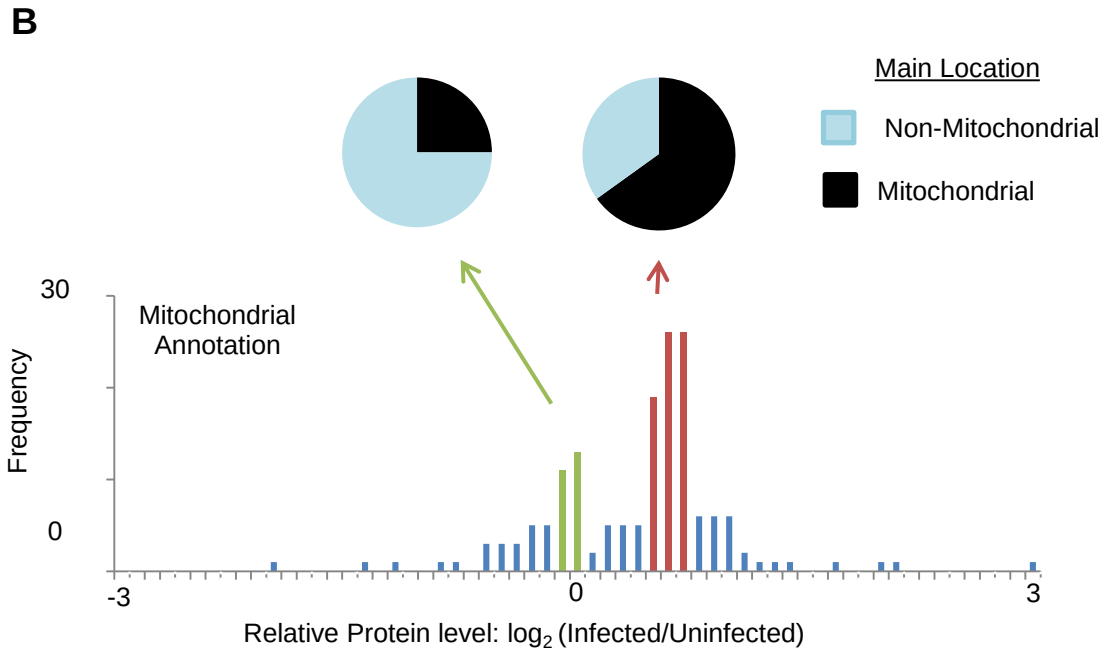
4.2.2.3.1 Analysing the Distribution of Relative Protein Levels of Mitochondrial Proteins

It was hypothesised that the infection of cells would only have a substantial effect on the levels of a small number of mitochondrial proteins whilst the levels of the vast majority would remain unchanged. In order to determine whether this was the case, the relative protein levels of the proteins in each data set were analysed to see if they followed such a trend. This was achieved by placing all of the relative protein levels from a given dataset into groups or 'bins' based on their magnitude. The size of the bins were 0.1 on a log 2 scale. Therefore, for example, a protein whose level increased 2.1 fold upon infection (relative protein level = 1.07 on a log 2 scale) was placed into the 1 – 1.1 bin. A histogram was then produced which plots

the magnitude of the relative protein level (x-axis, divided into bins) against the frequency (y-axis). Interestingly, the D9-data set from the first repeat did not follow a normal distribution but showed an apparent bimodal distribution; resembling two normal distributions with distinct medians/centres (Fig. 14 A, Top). This consisted of a major peak (centring around $\log_2$ relative protein level = 0) and a minor peak (centring around $\log_2$ relative protein level = 0.7). When proteins in this data-set were separated by ontology into mitochondrial (Fig. 14 A, Middle) and non-mitochondrial (Fig. 14 A, Bottom) proteins, it was found that the minor peak was formed by mainly mitochondrial-annotated proteins whilst the major peak was formed mainly by non-mitochondrial-annotated proteins. Interestingly, however, even with the mitochondrial annotated proteins, two peaks were still observed in the D9 data set (Fig. 14 A, Middle). When an ontology search for the 'main location' of proteins from these two peaks was performed, it was found that the majority of proteins from the minor peak had a 'main location' which was distinct from the mitochondria, whilst the majority of those in the major peak had mitochondria as their 'main location' (Fig. 14 B). This demonstrates that in the D9-WSN data set from repeat 1 the relative protein levels of mitochondrial and non-mitochondrial proteins both follow normal distributions, but centre around different points. Furthermore, in the D9-WSN data set from repeat 1, by filtering the detected proteins via ontology to select only the mitochondrial annotated proteins, the vast majority of non-mitochondrial proteins were removed. This resulted in a major peak which consisted of mitochondrial annotated proteins and resembled a normal distribution.

Overall, this data supports the conclusion that the vast majority of mitochondrial proteins are equally affected by infection with only a small number of mitochondrial proteins showing a larger or smaller fold change.


A
N9-WSN
D9-WSN
All R1
Mito R1
Non-Mito R1
Frequency
Relative Protein level: log2 (Infected/Uninfected)


B
Main Location
Non-Mitochondrial
Mitochondrial
Mitochondrial Annotation
Frequency
30
0
-3
0
3
Relative Protein level: log2 (Infected/Uninfected)

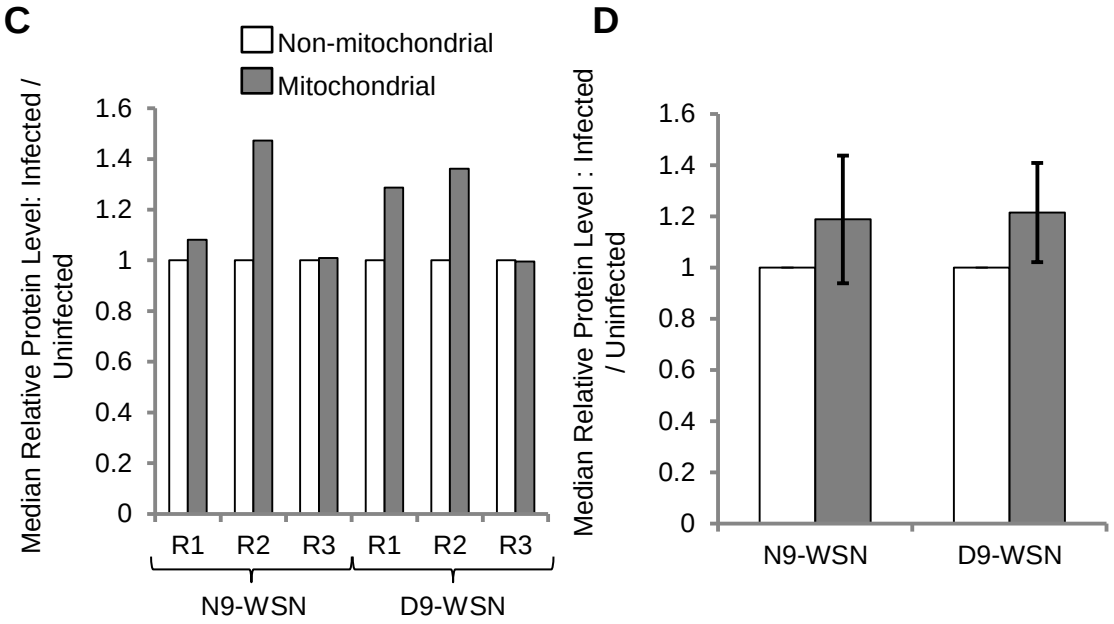


Figure 14 – Analysing the fold-changes in protein levels upon infection – Light, medium and heavy labelled SILAC cells lines were infected with either a virus encoding a mitochondrial PB2 (N9-WSN; MOI=5), a virus encoding a non-mitochondrial PB2 (D9-WSN; MOI=5) or mock infected. 12 h post infection, mitochondria were isolated and submitted for mass spectrometry analysis. The level of each protein in the infected (N9-WSN or D9-WSN) sample relative to the level of the same protein in the uninfected sample (termed the relative protein level) was calculated. The detected proteins were placed into 'bins' based on the magnitude of their relative protein level. (A) The relative protein level distributions for all proteins (all), mitochondrial annotated (mito), non-mitochondrial annotated (non-mito) from repeat 1 are shown (B) The relative protein level distribution of the 'mitochondrial annotated proteins' from the D9-WSN data set in Panel A is shown. The proteins from the peak which centres around a relative protein level of 0 (highlighted in green) and then proteins from the peak which centres around a relative protein level of 0.7 (highlighted in red) were selected and ontology analysis was used to identify their main location. Pie-charts display the main location of the proteins found in each peak. The 'black' wedge represents the proteins which have mitochondria as their main location. (C) Median relative protein levels of the mitochondrial (grey) and non-mitochondrial (white) proteins from N9-WSN and D9-WSN data sets from repeats 1, 2 or 3 (R1-R3). Data is normalised against the median relative protein level for non-mitochondrial proteins. (D) Average relative protein levels values from Panel D for the N9-WSN and D9-WSN data sets.

4.2.2.3.2 The Levels of Mitochondrial and Non-Mitochondrial Proteins are Differentially Affected by Influenza Virus Infection

It was next tested as to whether the bimodal distribution observed in the D9-WSN data set from repeat 1 was reproducible between the three repeats. In order to determine this, the median relative protein level in each data set was determined for either mitochondrial-annotated or non-mitochondrial annotated proteins. The median was chosen as it can be used as a model of the centre of a normal distribution. The median relative protein level of the mitochondrial annotated proteins was found to be higher than that of the median relative protein level of non-mitochondrial proteins in 4 of the 6 data sets (Fig. 14 C). On average, however, neither the average median relative protein level of mitochondrial-annotated proteins from N9-WSN or D9-WSN was significantly higher than the corresponding average median relative protein level of non-mitochondrial proteins (Fig. 14 D). Retrospective analysis of the levels of viral proteins in the six samples, however, showed that the viral protein NP was at lower levels in repeat 3 than in repeats 1 and 2 (Supplementary Figure 1). This is likely due to the infection in repeat 3 being less robust and thus could potentially explain why the bimodal distribution was not observed in the third repeat.

Therefore, together this data supports the conclusion that the infection of cells with influenza virus differentially affects the levels of mitochondrial and non-mitochondrial proteins which are isolated. This result could be explained: i) if infected cells have different levels of mitochondrial proteins than mock infected cells, however similar levels of non-mitochondrial proteins. This could for example be explained by differing levels of mitophagy between mock infected and infected cells. Or ii) if mitochondria are isolated at different efficiencies from infected and uninfected cell lysates, however non-mitochondrial proteins are isolated at similar efficiencies. This could occur if there was a difference in the density of mitochondria

between infected and mock infected cells, which would likely change the purification efficiency. This might happen, for example, if changes in mitochondrial structure occurred, e.g. due to differences in mitochondrial fusion and fission.

4.2.2.3.3 Influenza Virus Infection Does Not Lead To Changes in the Mitochondrial Proteome by 12 Hours Post Infection

As shown above, the infection of cells with influenza virus led to a general increase in the levels of isolated mitochondrial proteins compared to isolated non-mitochondrial proteins. In order to determine whether the infection of cells with either N9-WSN or D9-WSN also led to statistically significant changes in the levels of any mitochondrial proteins on top of this general increase, the data sets from the three repeats needed to be combined. However, as shown above, relative protein level distribution of mitochondrial proteins of the different repeats centred around different points (Fig. 14 C). Therefore, it was firstly necessary to normalise the data from all data-sets so that the relative protein levels of mitochondrial proteins centred around a relative protein level of zero. With normal distributions, this can be achieved by dividing all data points within the distribution by the median (demonstrated in Fig. 15 A). In order to perform this normalisation on the data sets, the proteins in the data sets were first filtered to select only mitochondrial annotated proteins (as per Fig. 14) so to focus the analysis specifically on the effects of infection on the levels of mitochondrial proteins. The resultant filtered data sets were normalised by the median relative protein level of each data set. The normalised data sets were all found to centre approximately around a relative protein level of 0 and thus the normalisation technique was deemed to be appropriate.

In order to determine whether the infection of cells with either N9-WSN or D9-WSN led to statistically significant changes in the levels of any mitochondrial proteins on top of the general increase seen in **section 4.2.3.3.2**, the average relative protein level of each protein in either the normalised N9-WSN data or the D9-WSN data sets was then calculated. The

averages were only calculated for those mitochondrial proteins which were detected in all three of the repeats. Furthermore, a p-value for each protein was calculated by comparing the protein's relative protein level to that of a hypothetical protein whose levels are not affected by infection (e.g. relative protein level = 0 on a $\log_2$ scale, standard deviation = 0). The resultant data was then plotted on two volcano plots (the first for the N9-WSN and the second for the D9-WSN data sets). A volcano plot displays the average relative protein level on a $\log_2$ scale on the *x*-axis against the P-value of this change on a $\log_{10}$ scale on the *y*-axis. The majority of points on a volcano plot normally centre around (0,0); as most proteins will have a low magnitude average relative protein level which will in general have low statistical significance. The points of interest are those in the top right and top left quadrants of the graph, which represent those proteins with a high magnitude average relative protein level which is highly significant. The statistical threshold used in this analysis was calculated using the Bonferonni Correction, which takes into account the number of proteins in the analysis. It was found that with infections with both the N9-WSN and D9-WSN influenza viruses, the majority of detected mitochondrial proteins had a relative protein level of less than 0.5 or more than -0.5 (less than a 2 fold change in levels between infected and uninfected samples). Of those which exhibited a higher magnitude relative protein level, none had a P-value higher than the Bonferonni corrected significance limit (Fig. 15 B). The full list of detected proteins and the corresponding relative protein levels from this analysis are found in Supplementary Figures 2 and 3.

It was possible, however, that the lack of robust infection in repeat 3 (Supplementary Figure 1) led to low magnitude relative protein levels in this repeat, masking any differences. It was not possible, however, to complete a T-test using only 2 independent repeats. Therefore, the four data sets from repeats 1 and 2 (both N9-WSN and D9-WSN) were analysed together to

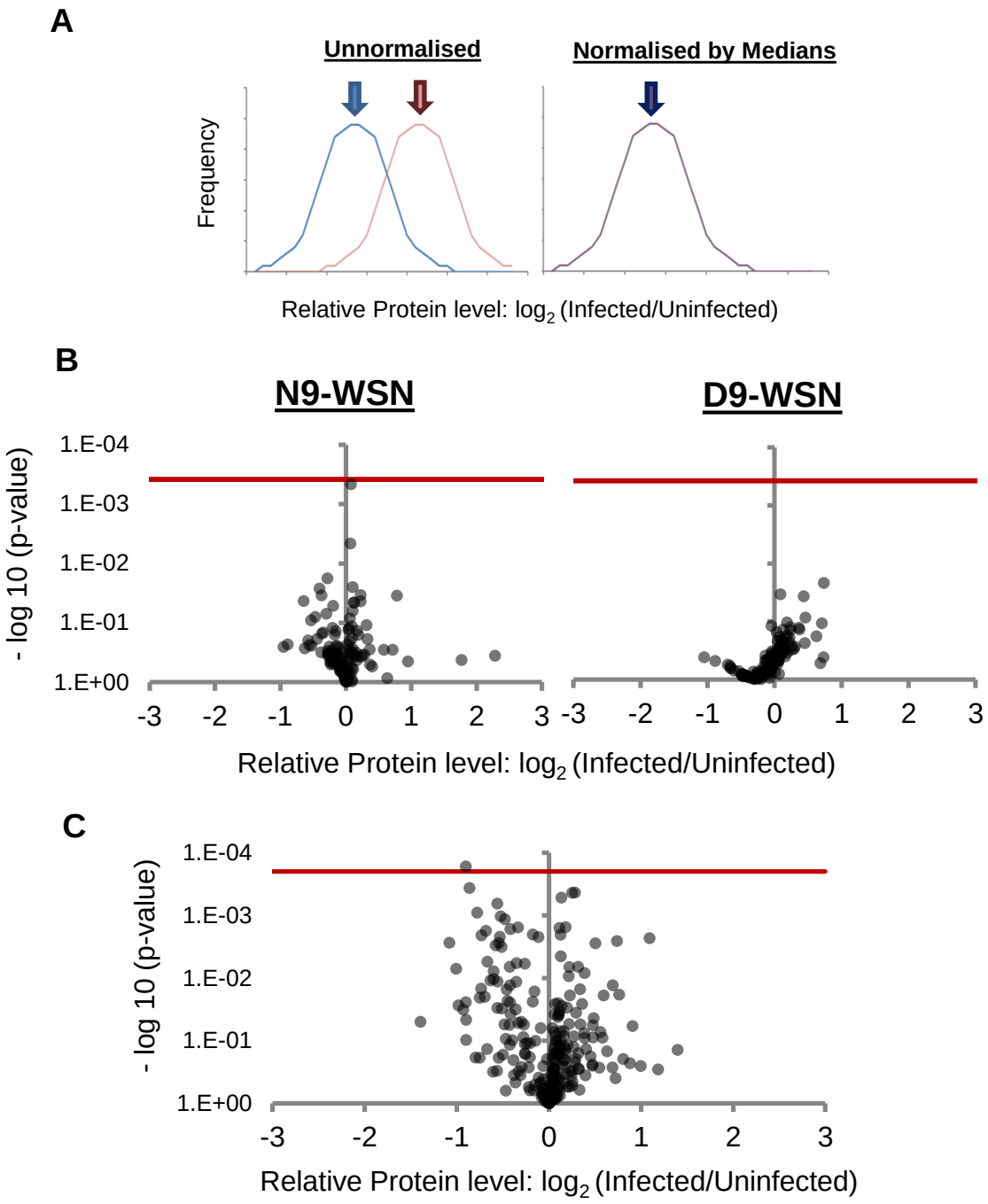


Figure 15– Influenza A virus infection does not have detectable effects on the mitochondrial proteome by 12 h p.i. – (A) Representative image of the normalisation of two normal distributions using their respective medians (arrows). (B) Volcano plot of the relative protein levels on infection with N9-WSN or D9-WSN. Only mitochondrial annotated proteins which were detected and quantified in all the repeats (n=3) are displayed. (C) Volcano plot of the relative protein levels on infection from repeats 1 and 2 with N9-WSN and D9-WSN. Only mitochondrial annotated proteins which were detected and quantified in at least 3 data sets (n = 3 or 4) are displayed.

determine whether influenza virus infection, independent of the expression of mitochondrial PB2, led to any changes in the mitochondrial proteome. It was found that only chloride intracellular channel protein 1 (CLIC1) had a relative protein level of greater than 0.5 (on a $\log_2$ scale) and a P-value which was higher than that of the Bonferroni corrected significance limit (Fig. 15 C). However, no documented mitochondrial function is present in the literature for this protein and thus it is not clear how changes in CLIC1 might affect mitochondrial function. The full list of detected proteins and the corresponding relative protein levels from this analysis are found in Supplementary Figure 4.

In conclusion, it was found that mitochondrial proteins in general are isolated at higher levels than non-mitochondrial proteins from influenza A virus infected cells (relative to mock infected cells). However, on top of this effect, only one mitochondrial annotated protein, CLIC1, was found to be significantly affected by influenza virus infection. Given that CLIC1 has no documented mitochondrial function, this analysis therefore did not identify any specific mitochondrial pathways which might be affected by influenza virus infection. Furthermore, this analysis did not identify any mitochondrial pathways which might be specifically affected by the presence of mitochondrial PB2.

4.2.3 Mapping the Mitochondrial Interactors of PB2

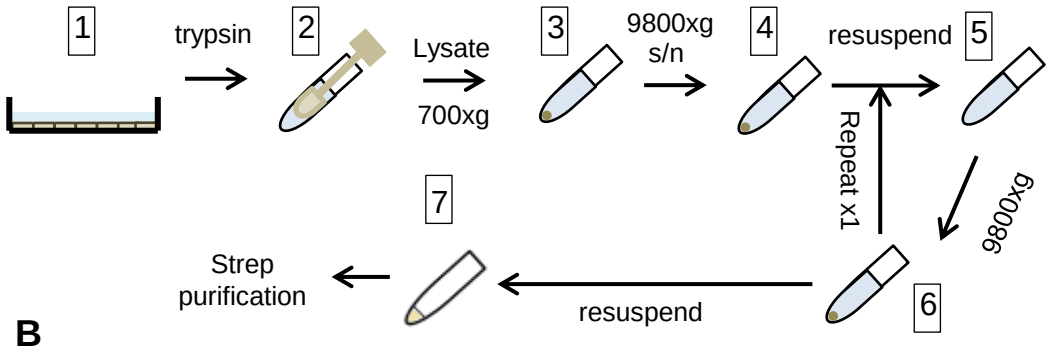
4.2.3.1 Mitochondrial PB2 Interacts with a Number of Matrix Localising Factors

It was hypothesised that the mitochondrial localisation of PB2 might allow PB2 to bind a matrix localised host factor and thus affect mitochondrial function. Therefore, section 4.2.3 aimed to analyse the interactome of mitochondrial PB2 by the affinity purification of PB2 from isolated mitochondria.

In order to perform the interactome analysis, a higher number of mitochondria were required than for the SILAC analysis. The limiting stage in the purification of mitochondria from above was the ultracentrifugation stage. However, previous analysis had determined that the isolation of mitochondria by differential centrifugation (used to generate the crude mitochondria) also enriched efficiently for mitochondria (Fig. 12 C). Therefore, the protocol from Figure 12 A was modified such that the crude mitochondria were twice resuspended in buffer and re-pelleted (Fig. 16 A). It was hypothesised that this would further enrich for mitochondrial proteins and decrease the levels of nuclear and cytoplasmic contaminants.

In order to characterise the interactome of mitochondrial PB2, cells were infected with either wt A/WSN/33 (WSN) or an A/WSN/33 derived strain which encodes a strep-tagged mitochondrial PB2 (Strep-PB2 WSN). At 12 h. p.i, the mitochondria were isolated from cells using the mitochondrial isolation protocol shown in Figure 16 A. The mitochondria were then lysed and a StrepTactin affinity purification was performed. The resultant samples were analysed by MS and SINQ analysis, as before. The protein levels in the samples were normalised in a manner similar to that previously used by *York et al.* [206]. In short, proteins were identified which were present in all three of the repeats in both the WSN and Strep-PB2-WSN conditions, given that the majority of these proteins would likely be contaminant proteins which were co-purifying non-specifically. Thus, it was assumed that these contaminant proteins would be present in comparable levels in both the WSN and Strep-PB2-WSN samples. The median fold-change of the 'contaminant proteins' in each of the six samples (three repeats of two conditions) was calculated and the protein levels in each sample were normalised by the relevant median. After this normalisation, the fold difference in the level of each protein between the WSN sample and Strep-PB2-WSN sample was

A



B

Protein	Sub-mitochondrial Location	Relative Protein Level (Strep-PB2 WSN / WSN)				Watanabe	York
		R1	R2	R3	Average		
HSP60	Matrix	12.34	10.00	46.85	23.06	PA PB1 PB2	PB2
TUFM	Matrix	5.42		12.02	8.72	PA PB2	PB2
OGDC-E2	Matrix	11.73	2.79	7.09	7.20		
HSPA9		6.43	1.56	10.68	6.22	PA PB1 PB2	PB2
C1QBP	Matrix	2.82	10.00	4.72	5.85	PB1 PA	
ARALAR2	MIM	2.91		7.02	4.97	PB2	PB2

C

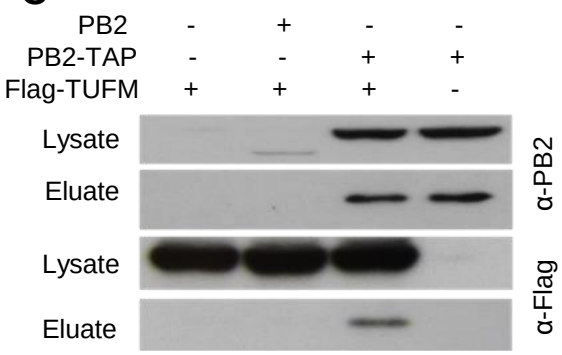


Figure 16 – Characterising the interacting partners of matrix PB2 – (A) Modified protocol for mitochondrial isolation. Cells were infected with either A/WSN/33 (WSN) or Strep-PB2-A/WSN/33 (Strep-PB2-WSN). Cells were trypsinised [1 to 2] and lysed using a dounce homogeniser [2 to 3]. A crude mitochondrial pellet was generated by differential centrifugation; a 700xg centrifugation to pellet 'cell debris' [2 to 3] and a 9800xg centrifugation to pellet mitochondria (termed the crude mitochondria) [3 to 4]. The resultant pellet was resuspended and a 9800xg centrifugation was performed to pellet mitochondria. This resuspension and pelleting process was repeated a second time. [4 to 6]. The mitochondria were then resuspended and lysed [6 to 7], and a strep-purification was completed. (B) SINQ analysis was used to quantify the levels of mitochondrial proteins in the elution. These were then normalised against the median level of the 'contaminant proteins' in the elution. A fold difference in the level of each protein between the PB2-Strep-WSN and WSN samples (relative protein level) was calculated for repeat 1, 2 and 3 (R1/2/3) and an average relative protein level was calculated. If a protein was only detected in the Strep-PB2-WSN sample then the relative protein level was arbitrarily set to 10. Those mitochondrial proteins which were found consistently at higher levels in the Strep-PB2-WSN sample are shown. Furthermore, detected proteins were cross-referenced against two previous interactome screens. (C) Flag-TUFM was expressed in HEK 293T cells alongside TAP-tagged or non-TAP-tagged PB2. A TAP-purification was performed. Western blot analysis of the lysates and purified proteins was then performed using an anti-PB2 and an anti-Flag antibody.

calculated (termed the relative protein level). These values were averaged to identify proteins which were consistently at higher levels in the Strep-PB2-WSN sample and thus were candidate interactors of mitochondrial PB2. If a protein was detected in the Strep-PB2 WSN sample but not in the WSN sample, its relative protein level was arbitrarily set at 10. If a protein was detected in the WSN sample but not in the Strep-PB2 WSN sample, its relative protein level was arbitrarily set at 0.1. Furthermore, the identified proteins were cross referenced against two previous interaction screens: i) *York et al*. which characterised the total interactome of the Strep-PB2 from Strep-PB2-WSN infected cells and ii) *Watanabe et* al., which characterised the interactome of individually expressed viral proteins [206, 208]. Finally, the Uniprot database was used to determine the sub-mitochondrial location of the identified proteins. The full list of detected proteins and the associated relative protein levels are shown in Supplementary Figure 5. From the detected proteins, a number of mitochondrial proteins were identified which were consistently at higher levels in the Strep-PB2 WSN samples than the WSN samples (relative protein level greater than 1). These were: i) heat shock protein 60 (Hsp60), ii) elongation factor tu, mitochondrial (TUFM), iii) dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex (DLST), iv) stress-70 protein (HSPA9), v) complement component 1 Q subcomponent-binding protein (C1QBP) and vi) mitochondrial aspartate glutamate carrier 2 (ARALAR2). Of these proteins, HSP60, TUFM, HSPA9 and ARALAR2 have been previously identified as PB2 interactors in previous screens (Fig. 16 B). Furthermore, all four of these proteins are either i) localised to the matrix space or ii) are MIM proteins which have matrix localised regions. Therefore, this data supports the hypothesis that the mitochondrial matrix localisation of PB2 may allow it to interact with a number of matrix localising proteins.

4.2.4.2 PB2 Interacts with the Mitochondrial Matrix Protein TUFM

The protein TUFM was selected for further investigation given its matrix localisation, its detection in the PB2 purifications shown above, and the fact it was found in two previous PB2 interactome screens [206, 208]. From the experimental data above, it was not clear whether TUFM interacted with monomeric PB2, or with PB2 within the context of 3P. In order to confirm whether PB2 interacts with TUFM, a tandem affinity purification (TAP) was performed using the lysate of cells expressing TAP-tagged or non-tagged PB2, and Flag-TUFM. The cell lysates and the affinity purified proteins were analysed by Western Blot using an anti-PB2 and an anti-Flag antibody. It was found that TUFM was purified from cells expressing TAP-tagged PB2 but not from those expressing untagged PB2 (Fig. 16 C). This demonstrates that monomeric PB2 is able to interact with TUFM. Given that PB1 and PA have not been observed to be imported into the matrix, this finding is consistent with the hypothesis that matrix PB2 interacts with TUFM.

4.2.4 Investigating the Effect of TUFM on Influenza Virus Replication and the Interferon Response against Influenza Virus

The host protein TUFM is a nuclear encoded protein which contains an N-terminal MTS and is trafficked into the mitochondrial matrix space. Here, its principal role is documented to be the loading of amino-acyl-tRNAs into ribosomes during translation. It has also be shown, however, to act as a regulator of RIG-I/MAVS mediated IFN induction [132]. It has been hypothesised that TUFM works in concert with NLRX1 by binding to MAVS and inhibiting the RIG-I:MAVS interaction [132, 133]. However, this result has been disputed by papers which point out that NLRX1 is, in fact, a matrix localised protein and thus is not correctly localised for MAVS binding [229, 230]. A second, conflicting study by *Jaworska et al.*, concluded that, in

macrophages, NLRX1 promotes the IFN response specifically against PB1-F2 containing viruses [231]. This led to the hypothesis that NLRX1 acts to maintain mitochondrial viability, which is reduced by PB1-F2, and thus safeguards the ability of cells to efficiently induce IFN expression.

In order to test whether TUFM is important for IFN induction in response to influenza A virus infection, the induction of IFN by influenza A virus in TUFM knockdown or control cells was tested. Cells were transfected with a control siRNA (GFP) or with one of three siRNAs against TUFM, which target different regions of the TUFM mRNA. The cells were lysed 24 h post transfection and the levels of TUFM and β-actin were analysed by Western Blot. The levels of TUFM and β-actin were quantified by densitometry and the level of TUFM was normalised to that of β-actin. The transfection of cells with each of the siRNAs against TUFM led to a reduction of the levels of TUFM to approximately 20% of that of the control siRNA transfected cells (Fig. 17 A). Furthermore, in order to test the effects of the knockdown of TUFM on cell viability, the levels of ATP in the lysates of the siRNA transfected cells were measured. It was found that there were comparable levels of ATP in all of the lysates (Fig. 17 B). Together, this demonstrates that the transfection of cells with the TUFM specific siRNAs led to an 80% knockdown of TUFM but had no significant effect on cellular viability. In order to test the effect of TUFM knockdown on IFN induction by influenza A virus infection, cells were transfected with the siRNAs, as above, as well as two reporter plasmids to measure IFN induction. These reporter plasmids encode i) luciferase under the control of the IFN promoter (IFN-luc; used to measure IFN induction) and ii) β-galactosidase under the control of a constitutive promoter (β-gal). At 24 h post transfection, the cells were either mock infected or infected with N9-WSN. At 12 h post infection, the cells were lysed and the expression of IFN (measured as the β-galactosidase normalised luciferase activity) was quantified. It was found that the infection of cells with N9-WSN did not lead to

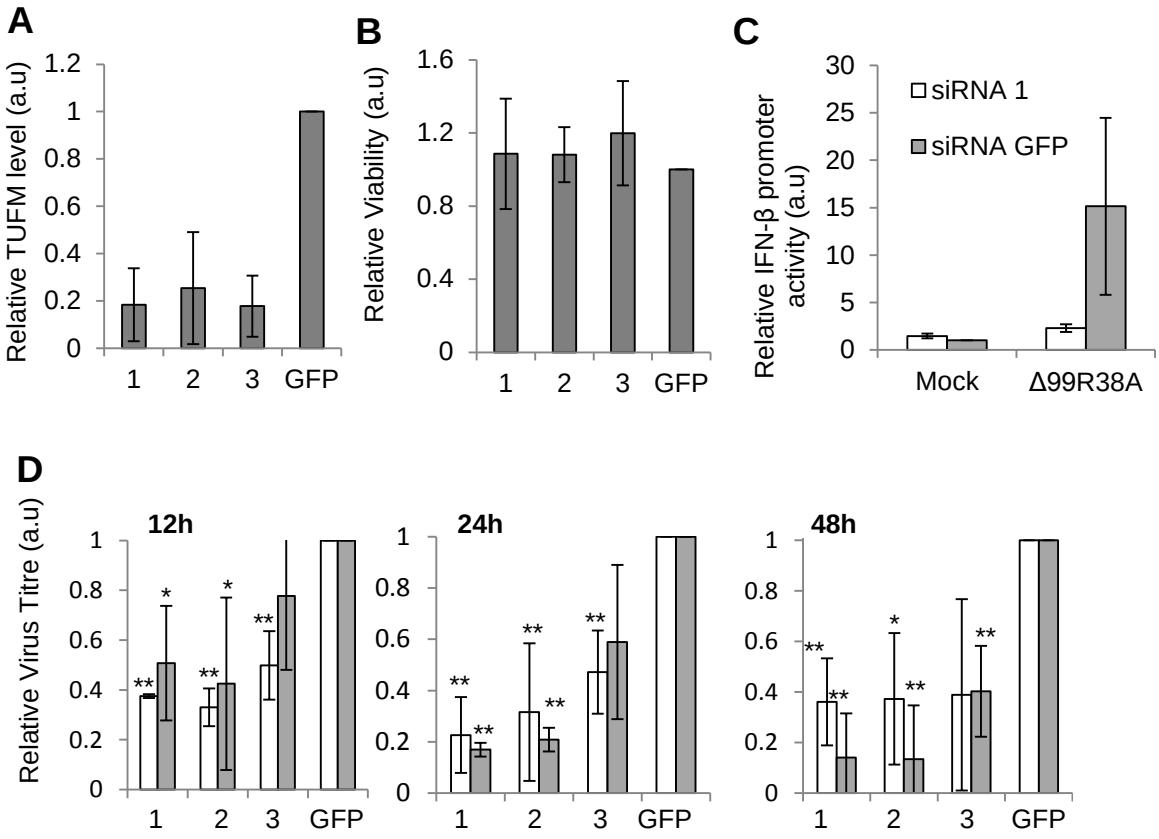


Figure 17 – TUFM is important in the induction of interferon in response to influenza virus infection – (A) HEK 293T cells were transfected with either a TUFM-targeting siRNA (1, 2 or 3) or a GFP-targeting control siRNA (GFP). Cells were lysed and the levels of β-actin and TUFM were analysed by Western Blot using β-actin and TUFM-specific antibodies. TUFM levels were normalised to the level of actin, and expressed relative GFP condition. a.u = arbitrary units. Bars represent standard deviations of average data from three independent experiments (B) Cells were transfected as in 'A' and, 24 h post transfection, viability was determined using the CellTiter-Glo luminescent assay which measures ATP levels. The viability levels are expressed relative to the GFP condition. Bars represent standard deviations of average data from three independent experiments. (C) Cells were co-transfected with either siRNA 1 or with siRNA GFP, along with an IFN-luciferase (luc) and a β-galactosidase (β-gal) reporter plasmid. At 24 h post transfection, cells were either mock infected or infected at an MOI of 10 with a virus strain encoding a mutant NS1 (Δ99R38A). At 12h post transfection, cells were lysed and the luc and β-gal activities were measured. The β-gal normalised luciferase activities (relative IFN-β promoter activities) are shown relative to that of the Mock/GFP condition. Bars represent standard deviations of average data from three independent experiments (D) Cells were transfected as in 'A' and infected at an MOI of 0.01 with either N9-A/WSN/33 (white) or D9-A/WSN/33 (grey) at 24 h post transfection. At 12, 24 and 48h post infection, supernatants were collected and virus titres were determined by plaque assay. The virus titres at each time point are expressed relative to those from the GFP condition. Bars represent standard deviations of averages of the viral titres from three independent experiments. Asterisks indicate a significant difference between the marked sample and the corresponding GFP sample (one sample *t* test; * = P <0.05, ** = P < 0.01)

a detectable increase in IFN expression (data not shown). It was hypothesised that this was due to the expression of NS1 by the virus. Therefore, the experiment was repeated using an influenza A virus strain which has a truncated, RNA-binding deficient NS1 protein (Δ99R38A). It was found that the knockdown of TUFM decreased the average induction of IFN in response to Δ99R38A infection, though this effect was not significant (Fig. 17 C). This highlights the possibility that TUFM is a regulator of the IFN response against influenza virus infection.

To determine whether the interaction of matrix PB2 and TUFM was important for virus replication, the growth of N9-WSN and D9-WSN in control or TUFM-knockdown cells was analysed. In short, siRNA transfected cells were infected at 24 h post infection with either N9-WSN and D9-WSN at an MOI of 0.01, and virus replication was analysed by plaque assay. It was found that at 12, 24 and 48 h post infection, both the N9-WSN and D9-WSN replicated to lower titres in TUFM-specific siRNA transfected cells than in control transfected cells (Fig. 17 D). However, there were no significant differences in the effect of TUFM knockdown between the N9-WSN and D9-WSN viruses (Fig. 17 D). This indicates that TUFM knockdown does not differentially affect N9-WSN and D9-WSN replication. However, as observed above, infection of cells with N9-WSN (and likely D9-WSN) does not lead to a robust induction of IFN. Thus, if the role of mitochondrial PB2 is to dampen the IFN response by binding TUFM, no effect would be expected in this system.

4.3 – Discussion

The focus of this chapter was to identify potential roles of the matrix-localised PB2 identified in Chapter 3. Previous data has shown that an influenza A virus encoding a mitochondrial PB2 replicated to higher titres in infected mice than an isogenic virus encoding a non-mitochondrial PB2 [176]. Furthermore, mitochondrial PB2 has been previously hypothesised to

act to preserve the MMP and dampen IFN induction [175, 176]. Given that the viral protein PB1-F2, when imported into mitochondria, has been documented to bind IMS localised proteins and thus disrupt multiple mitochondrial functions, it was hypothesised that matrix PB2 might interact with matrix localised proteins and thus modulate mitochondrial processes.

In order to determine whether the expression of matrix localising PB2 by influenza A virus affects mitochondrial function, mitochondria were isolated from i) mock infected cells, ii) cells infected with a virus which encodes a non-mitochondrial PB2 and iii) cells infected with a virus which encodes a mitochondrial PB2, and the mitochondrial proteomes were analysed by mass spectrometry. Analysis of mitochondrial proteomes has been previously successfully used to analyse the effects of viral infections upon the mitochondria and represents a high-throughput technique to identify pathways or processes which are specifically affected. An example of this is a recent study on the effects of respiratory syncytial virus infection on mitochondria [228]. In this study by *Munday et al.*, the level of a number of proteins changed specifically in the virus infected cells. Ontological analysis of these proteins demonstrated that infection affected the recruitment of antiviral signalling proteins to the mitochondria and changed the association of mitochondria with other cellular membrane bound organelles.

Here, it was found that the infection of cells had a differential effect on the levels of mitochondrial proteins isolated relative to non-mitochondrial proteins. It is possible that this effect is due to differences in the purification efficiency of mitochondria, potentially due to changes in mitochondrial structure upon infection. This could for example occur if infection of cells with influenza virus promoted the fragmentation of mitochondria. Beyond this, it was found that, when comparing the protein fold-changes of mitochondrial annotated proteins, the infection of cells with influenza A virus led only to a significant decrease in the levels of CLIC1. This protein, however, has no documented mitochondrial functions and thus it is not clear how this difference would affect mitochondrial function.

It is possible that 12 h post infection was not a late enough time point to detect further changes in the mitochondrial proteome caused by virus infection or by the expression of matrix PB2. Furthermore, it possible that matrix PB2 may only lead to mitochondrial proteome changes in a small proportion of infected cells. For example, given that it has been shown that only a small proportion of cells express IFN within an influenza A virus infected cell population, if matrix PB2 was inhibiting IFN expression, this would only be apparent in these cells [143]. However, the effect of matrix PB2 on the mitochondrial proteome of this small subset of cells would be masked by the lack of effect in the majority of infected cells. Furthermore, it is possible that the effect of mitochondrial PB2 may be less apparent when using cell culture lines, which frequently lack functional antiviral pathways, or when using lab-adapted influenza virus, such as WSN. Therefore, any effects of mitochondrial PB2 might only be observed when using cells from animals or primary cell lines which have been infected with a recent field isolate (that has not undergone extensive passaging through cell lines).

Due to the fact that the above analysis did not identify any mitochondrial pathways which might be affected by the presence of mitochondrial PB2, a second technique was employed to attempt to characterise the effects of matrix PB2 at mitochondria. It was hypothesised that by characterising the interactome of matrix PB2, it would be possible to determine which mitochondrial functions might affected by matrix PB2. The interactome of PB2 has been characterised by multiple groups [206, 208]. However, it was hypothesised that given that matrix PB2 likely only constitutes a small proportion of the total PB2 in an infected cell, the interactors of matrix PB2 might have been masked by the more highly abundant interactors of nuclear PB2 in these studies. To analyse the interactome of matrix PB2, cells were infected with either a virus which encoded a strep-tagged PB2 or a non-tagged PB2, the mitochondria were isolated and a streptactin affinity purification was completed. It was found that a

number of mitochondrial proteins were consistently found to co-purify at higher levels from cells expressing strep-tagged PB2 than from those expressing non-tagged PB2. The first was the chaperone Hsp60, which has been shown to bind to proteins which have been newly imported into the matrix space of mitochondria [232]. This chaperone may therefore be involved in the folding of imported PB2 in the mitochondria. Two of the other detected proteins are involved in mitochondrial metabolic pathways. Firstly, ARALAR2 is an aspartate-glutamate antiporter which is important in ETC function [233]. Secondly, OGDC-E2 is a component of the 2-oxoglutarate dehydrogenase complex which is part of the citric acid cycle and is thus important for the generation of $NADH/FADH_2$, which are substrates for ETC function [234]. It could therefore be of future interest to test the effects of mitochondrial and non-mitochondrial PB2s on the respiratory function of mitochondria in order to determine whether mitochondrial PB2 impacts upon the function of this system. In this thesis, however, only the matrix localising protein TUFM was further investigated. TUFM was selected for further analysis due to the fact that it has been identified as a PB2 interactor in two previous PB2 interaction screens and that it has documented roles in the induction of IFN [132, 206, 208].

The knockdown of TUFM led to a decrease in the expression of IFN in response to infection with an influenza A virus strain encoding a mutated NS1 protein. It is not clear whether this was due to TUFM directly regulating the RIG-I/MAVS pathway or due to the importance of TUFM in general mitochondrial function. This observation, however, highlighted the possibility that if mitochondrial PB2s are in fact stronger IFN dampeners than non-mitochondrial PB2s, as postulated by *Graef et al.* [176], the interaction of PB2 with TUFM may be mediating this effect.

Interestingly, it has recently been found that influenza A virus can generate a number of other proteins which are derived from the segment 1 (PB2) vRNA. The first of these, termed PB2Δ, is encoded by an internal deletion, DI RNA derived from the segment 1 [159]. Given this protein

contains the same N-terminal sequence as full length PB2, it can be targeted to the mitochondria and is likely also imported into the matrix. The second protein is termed PB2-S1 and is generated from a spliced form of the PB2 mRNA [53]. Given that the N-terminal MTS of PB2-S1 leads to its mitochondria localisation, it is likely that PB2-S1 is also imported into the mitochondrial matrix. It is thus possible that mitochondrial localising PB2-S1 and PB2Δ could also have yet undocumented functions associated to their matrix localisation.

Taken together, in this chapter, it has been demonstrated that the import of PB2 into the matrix space of the mitochondria may be important in the interaction of PB2 with a number of matrix localising host factors. These factors have been documented to be involved in processes such as ETC function and the induction of IFN [132, 233, 234]. Finally, it was found that one of the interactors, TUFM, is important for the induction of IFN in response to influenza A virus infection. Given that previous work by *Graef et al.* showed that mitochondrial PB2 is a stronger dampener of IFN expression than non-mitochondrial PB2 [176], this highlights the possibility that matrix PB2 might dampen IFN expression by binding and modulating the activity of TUFM.

4.4 Summary

A) The infection of cells with N9-WSN or D9-WSN differentially affected the levels of mitochondrial and non-mitochondrial proteins which purified with isolated mitochondria.

B) HSP60, DLST, HSPA9, C1QBP and ARALAR2 are candidate interactors of matrix localised PB2.

C) PB2 interacts with TUFM

D) The knockdown of TUFM causes a decrease in IFN expression in response to influenza A virus infection.

4.5 Part I – Final Discussion

The first two results chapters of this thesis focussed on investigating the roles mitochondrial localising PB2. It was demonstrated that over 99% of pre-2009 human seasonal influenza A virus strains encoded a mitochondrial localising PB2s. It was shown that this PB2 is imported into the mitochondrial matrix, where it potentially interacts with a number of host factors including TUFM. Interestingly, the knockdown of TUFM was found to impact upon the induction of IFN in response to influenza A virus infection, though the reason for this is unclear. Given that previous work by *Graef et al*. demonstrated that mitochondrial localising PB2s may be stronger dampeners of IFN expression that non-mitochondrial localising PB2s, it is possible that the interaction of the dampening activity of mitochondrial PB2 is related to its binding to TUFM [176]. Furthermore, a number of proteins involved in mitochondrial metabolic pathways were identified as potential interactors of matrix PB2, however, these interactions were not further investigated. Thus, it is possible that matrix PB2s may also have a function related to mitochondrial metabolism which should be further investigated.

Interestingly, as of yet, no increase of asparagine or threonine usage has been observed in the H1N1 virus strains derived from the 2009 H1N1 pandemic strains, which encoded a PB2 of avian origin. It is possible that there are compensatory mutations in the PB2, or another viral protein, of these viral strains which circumvents the requirement for asparagine or threonine at position 9. This is not undocumented as certain polymorphisms in the PB2s of post-2009 H1N1 influenza viruses compensate for the fact that these PB2s encode the avian-associated aspartic acid at position 627 [194, 195]. It will be of interest, however, to observe whether aspargine and threonine, or other non-negatively charged amino acid residues, occur with

increased frequency at position 9 of PB2s of post-2009 H1N1 influenza A virus strains in the future. If this occurs, this would appreciably strengthen the case that mitochondrial localisation of PB2 has a function in human and/or mammalian infections.

Part II

Chapter 5

The Role of RNP Structure in Interferon Induction

5.1 Introduction

The focus of the second part of this thesis is on attempting to further understand how RIG-I detects influenza A virus infection, which, as shown in **section 1.6.1**, is still poorly understood.

Given that RIG-I has only been documented to localise to the cytoplasm, it seems likely that this is the site of viral RNA detection in influenza A virus infected cells. As stated in **section 1.6.1.1**, vRNA and cRNAs are the only viral RNA species which contain a 5′-ppp and thus are the most likely candidates to be detected by RIG-I in influenza A virus infected cells. However, it is unlikely that RIG-I would be able to bind to 3P-bound v/cRNAs due to the effects of 'shielding' (explained in **section 1.6.1.2**). Thus, for the detection of vRNA or cRNA to occur, it is likely that RIG-I must encounter naked (or at least 3P-less) vRNA and/or cRNA in the cytoplasm. It is, however, unclear how this would occur. One simple explanation could be that RIG-I detects vRNAs and/or cRNAs which are not encapsidated (by 3P and NP) post-synthesis, although this seems unlikely for two reasons:

- Firstly, given that vRNA replication occurs in the nucleus, unencapisated vRNA or cRNA would need to reach the cytoplasm before they could be detected by RIG-I.

However, the only known mechanism of viral RNA export from the nucleus is that of vRNA in the context of a vRNP with the NESs of viral proteins driving export (**section 1.3.6**) [75]. Cellular RNA species are exported to the cytoplasm via a number mechanisms; involving the binding of adaptor proteins which recruit exportins [235]. However, there is at present no known mechanism for the export of free vRNA (or free cRNA) without 3P and NP binding. Due to this, it is unlikely that vRNAs or cRNAs reach the cytoplasm in a completely naked, unshielded form.

- Secondly, given vRNAs and cRNAs contain a 5′-ppp instead of a cap-structure and lack a poly-A tail, even if there was a mechanism for free vRNA or cRNA transport to the cytoplasm, the RNA would likely be rapidly degraded. The reason for this is that cap structures protect RNAs from 5′-3′ exonucleases such as XRNI/II, whilst RNAs lacking a poly-A tail are subject to 3′-5′ degradation by the exosome [236]. This rapid degradation can be seen by the instability of cRNA in the absence of encapsidation by 3P and NP [72].

Therefore, it is likely that, once an RNP reaches the cytoplasm, the dissociation of 3P is required before RIG-I can detect an influenza virus viral RNA. However, it is unknown whether this is a stochastic process or driven by a host protein. Interestingly, it has been recently shown that vRNAs released from influenza B virions during infections are able to induce IFN [237]. Given it seems improbable that vRNAs could be incorporated into virions without being encapsidated into a vRNP (**section 1.3.7**), this work supports the idea that vRNAs which have been packaged into vRNPs are able to induce IFN. Furthermore, 'incoming' influenza B virus vRNPs induce IFN more strongly than 'incoming' influenza A virus vRNPs [237]. This highlights the interesting possibility that 3P dissociation from influenza B virus vRNPs may occur more readily than from influenza A virus vRNPs. Finally, *Weber et al*. have shown that incoming RNPs containing PB2-627E (avian adapted PB2) are

stronger activators of RIG-I than those containing PB2-627K (human adapted PB2, section **1.9.4**) [238]. This highlights the possibility that human-adapted 3Ps are more efficient at shielding viral RNAs from RIG-I.

On the other hand, *Rehwinkel et al.* found that full length RNPs (RNPs containing full length vRNAs) containing segment 3, 4, 5, 6, 7 and 8 vRNAs, when generated within transfected cells, are unable to induce IFN expression [114]. This effect could potentially be explained by the shielding of the v/cRNA from RIG-I by 3P or by the imperfect double stranded nature of the v/cRNA panhandle (**section 1.6.1.3**). Interestingly, however, these full length RNPs were able to induce IFN in cells which had been pre-treated with IFN [114]. Therefore, this means that the priming of cells with IFN (**section 1.4.4**) can potentially allow them to detect the presence of wt influenza virus RNAs in the context of RNPs. This leads to the conclusion that, once one cell has started to produce IFN, all neighbouring cells will be more able to detect the presence of a vRNA in the context of a vRNP, induce IFN and thus propagate the innate immune response. However, how the first cell manages to detect the influenza A virus infection without pre-priming is unknown. One possible clue is given by the fact that, in the work by *Rehwinkel et al.*, the only RNPs which robustly induced IFN without priming were those containing segment 1 and 2 vRNAs [114]. Interestingly, segments 1-3 have been shown to be the most common source of DI-RNAs [147]. This implies that the generation of DI-RNA within a cell may represent the mechanism of RIG-I activation in unprimed cells. This hypothesis is supported by observations that the presence of high levels of DI viruses during an infection correlates to a strong induction of IFN [138, 145, 153]. Based on the information provided above, it was hypothesised that in influenza DI virus infected cells, the formation of DI RNAs may be connected to the generation of a strong RIG-I ligand which can be detected in unprimed cells. This chapter thus focusses on understanding the reasons for the difference in IFN induction between wt influenza virus and influenza DI virus infected cells.

As stated in the **section 1.6.1.4**, DI viruses are likely formed by internal deletion events during the replication of vRNA or cRNA. The observed deletions that lead to DI RNAs are normally at least 100 nt away from the 5′ and 3′ RNA termini [147, 148] (Fig. 18 A). Thus, DI RNAs have identical 5′ and 3′ termini to full length wt v/cRNAs (Fig. 18 A). Therefore, it seems logical that the 5′-ppp blunt-ended dsRNA regions formed by DI RNAs, whether they are panhandle structures or v/cRNA hybrids, would be structurally identical to those formed by full length RNAs (Fig. 18 A). Thus, it is hard to conceive how structural changes so distal to the RIG-I binding site could have an effect on RIG-I binding and/or activation (Fig. 18 A). Despite this, DI rich stocks of influenza A virus have been found to be strong inducers of IFN expression [138, 145, 153]. Furthermore, affinity purification of RIG-I from influenza A virus-infected cells leads to a sample enriched in short, viral RNAs containing internal deletions, likely representing DI RNAs [156]. To attempt to understand this disparity, instead of focussing on the structure of DI RNAs specifically, it was hypothesised that the ability of DI viruses to induce IFN might be explained by structural differences between RNPs containing DI RNAs and those containing full length viral RNAs. Specifically, it was hypothesised that DI-RNA containing RNPs (due to the short length of some DI-RNAs) might contain only a limited number of NP molecules and, thus, might be unable to fold into the canonical double-helical structure (Fig. 18 B). Without the ability to form the canonical RNP structure, these RNPs might be less stable and more likely to release 3P (Fig. 18 B). Given, that the binding of 3P likely shields vRNA and cRNA from RIG-I binding, a decrease in RNP stability might translate into an increased propensity of RNPs to expose the vRNA or cRNA they contain to RIG-I (Fig. 18 B). Thus, the starting hypothesis could be most simply explained as: **RNPs containing short viral RNAs less effectively shield the vRNA and/or cRNA they contain from RIG-I binding**.

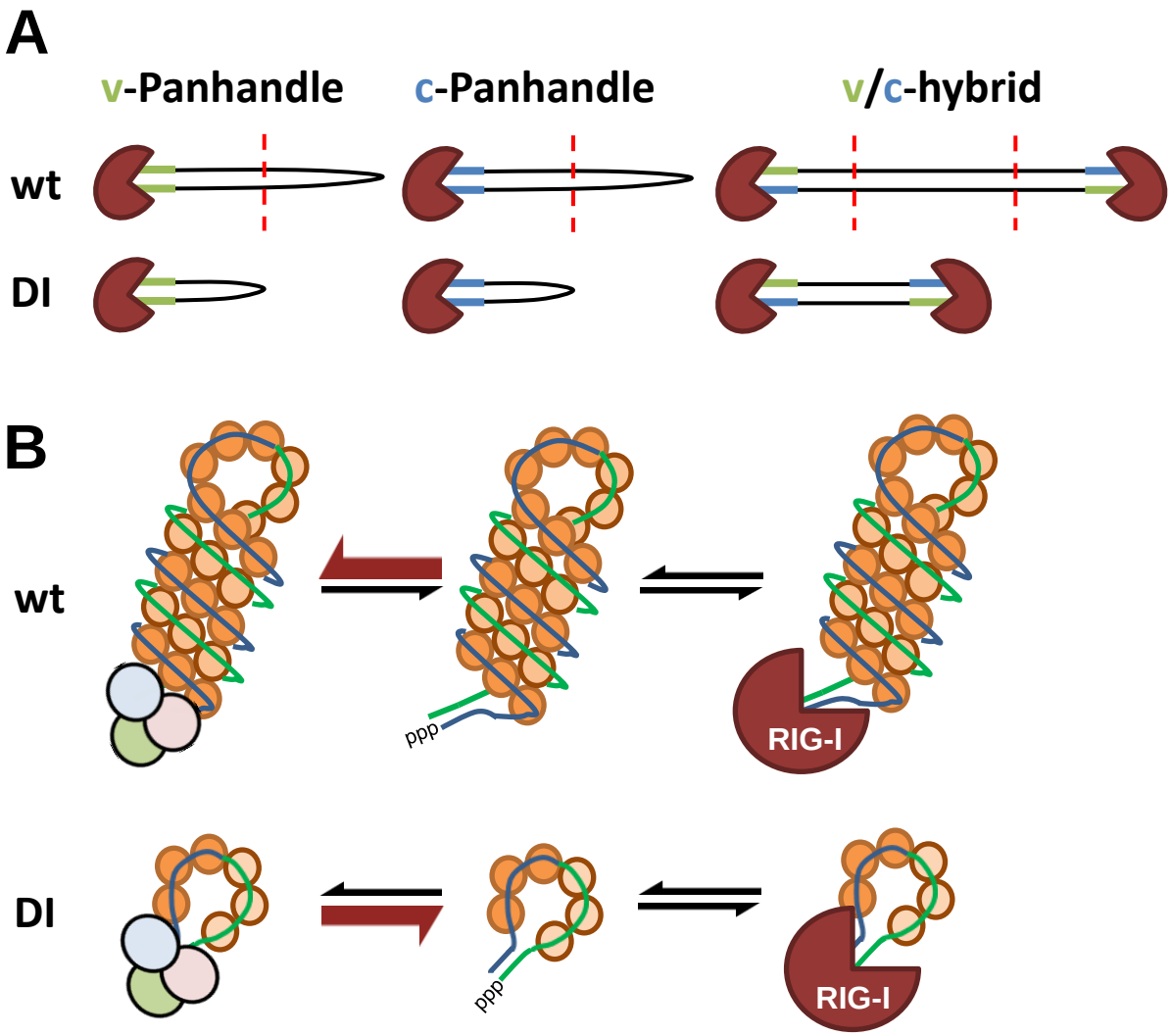


Figure 18 – Model of RIG-I activation by DI RNAs and DI RNPs – (A) Model of RIG-I binding to either a vRNA panhandle, cRNA panhandle or a vRNA:cRNA hybrid formed from full length (wt) or defective interfering (DI) RNAs. Red dotted lines are representative of the start and end points of an internal deletion in full length vRNA or cRNA to demonstrate their distal nature relative to the RNA-termini. (B) Model of RIG-I binding to the RNA contained within wt RNPs and DI RNPs. Full length RNPs have been shown to fold into a 'filament and loop structure' with potential interactions between 3P and the RNA bound NP. In this model, the canonical NP:RNA filament does not occur in DI RNPs, leading to a lower stability of 3P on vRNA and cRNA. This results in a higher propensity for 3P dissociation and a higher probability of RIG-I binding to the vRNA or cRNA panhandle.

5.2 Results

5.2.1 Replication of 76 Nucleotide Long vRNAs by the Viral Polymerase Leads to the Induction of Interferon

In order to test the hypothesis that RNP structure is important in the shielding of vRNA and cRNA from RIG-I, a range of different RNP structures were generated within cells to measure how strongly they induced IFN. To achieve this, a technique called an 'RNP reconstitution' was used in which vRNA, 3P and NP are expressed from plasmids. These factors associate together to form vRNPs, which replicate within the cell. This technique has previously been used to test the IFN induction by full length RNPs [114]. Previous work by *Turrell et al*. showed that active RNPs can be generated using this technique using vRNAs of a wide range of lengths, as long as the 3′ and 5′ termini of the vRNA are maintained (Fig. 19 A) [37]. Thus, by varying the length of the vRNA expressed, it should be possible to see how varying the length of the RNA within an RNP affects its ability to induce IFN. Furthermore, this technique can be used to generate 'NP-less RNPs'; that is 'RNPs' which consist of only RNA and 3P. As stated in **section 1.3.5.4**, RNPs are only active in RNP reconstitutions in the absence of NP if they contain RNAs shorter than approximately 100 nt. Thus, by expressing vRNAs that are shorter than 100 nt in the presence of 3P but the absence of NP, NP-less RNPs can be generated (Fig. 19 A). These NP-less RNPs could then be used to determine whether the presence of NP in a RNP is important in the shielding of the RNA.

For the RNP reconstitutions, cells were transfected with a vRNA expression plasmid and the expression plasmids for the subunits of 3P (termed herein as the '3P expression plasmids'). Beyond this, either the expression plasmid for NP or a control plasmid (3A) was included in the transfection. In this section, for clarity, the combination of plasmids used in each RNP reconstitution will be described as follows: i) the transfected plasmids will be listed in the

order: *vRNA / 3P / NP*, ii) a dash will be used to denote the absence of a plasmid, iii) in later sections, where the amount of a plasmid transfected is varied, the amount of plasmid will also be included. For example, if a cell is transfected with the 76 nt long RNA and 3P expression plasmids, however, not the NP expression plasmid, then the transfection will be denoted as *76 nt / 3P / -* . Furthermore, in this section, an RNP containing, for example, a 100 nt long RNA will be termed a '100 nt long RNP'. Diagrammatic representations of the predicted structures of the resultant RNPs in each of the RNP reconstitutions are shown (Fig. 19 A) and are described below:

- In RNP reconstitutions performed with RNAs of lengths far over 100 nt in the presence of NP (*e.g. 707 nt / 3P / NP*), it was hypothesised that RNA within the resultant RNPs would be bound by both 3P and NP. In this thesis, the central region of the RNA, which is not bound by 3P, will be termed the 'loop region'. Thus, in this condition, the RNA loop would be 'NP-coated' Due to the length of the RNA, these RNPs might be able to fold into a canonical, double-helical RNP structure (Fig. 19 A – Pink).

- In RNP reconstitutions performed with RNAs of lengths slightly over 100 nt in the presence of NP (*e.g. 246 nt / 3P / NP*), RNPs would likely be generated with NP-coated loops. However, due to the length of the RNA, these RNPs might not be able to fold into the canonical RNP structure (Fig. 19 A – Grey).

- In RNP reconstitutions performed with RNAs of lengths over 100 nt in the absence of NP (*e.g. both 246 nt / 3P / - and 707 nt / 3P / -*), based on previous observations [37], there would not be detectable levels of vRNA or cRNA, and thus RNPs, generated (Fig. 19 A – Grey and Pink). This is as '246 and 707 nt long' RNPs, in the absence of NP, are not active in replication.

- In RNP reconstitutions performed with RNAs of lengths slightly less than 100 nt in the presence of NP (*e.g. 76 nt / 3P / NP)*, it is not clear whether the resultant RNPs would contain NP-coated loops, NP-less loops or a mixed population (Fig. 19 A – Green).

- In RNP reconstitutions performed with RNAs of lengths slightly less than 100 nt in the absence of NP (*e.g. 76 nt / 3P /* -), 'NP-less' RNPs would be generated that would have 'NP-less loops' (Fig. 19 A – Green).

- In RNP reconstitutions completed with very short RNAs (e.g. *37 nt / 3P / NP* or *37 nt / 3P /* -), it was hypothesised that the RNA within the resultant RNPs would be almost completely shielded by 3P and that there would be no RNA accessible for NP binding (Fig. 19 A – Blue).

To test the IFN inducing ability of these multiple RNP types, cells were transfected with the IFN-luc and β-gal reporter plasmids alongside the vRNA, NP and 3P expression plasmids. The transfection of cells with *37 nt / 3P / NP* or *37 nt / 3P /* - did not induce high levels of IFN (Fig. 19 B – Mock). This is despite the fact that 37 nt long RNPs are likely able to replicate vRNA both in the presence and absence of NP [37]. This indicates that 37 nt long RNAs, in conditions where 3P is available to allow RNA encapsidation, are poor inducers of IFN (Fig. 19 A and B). The transfection of cells with *246 nt / 3P / NP* and *707 nt / 3P / NP* also did not lead to high levels of IFN (Fig. 19 B – Mock). Interestingly, cells transfected with *246 nt / 3P / NP* or *707 nt / 3P / NP* did not induce significantly more IFN than cells transfected with *246 nt / 3P /* - or *707 nt / 3P /* -. This is despite the fact that, in the latter conditions, there would not be expected to be replication of vRNA and cRNA to detectable

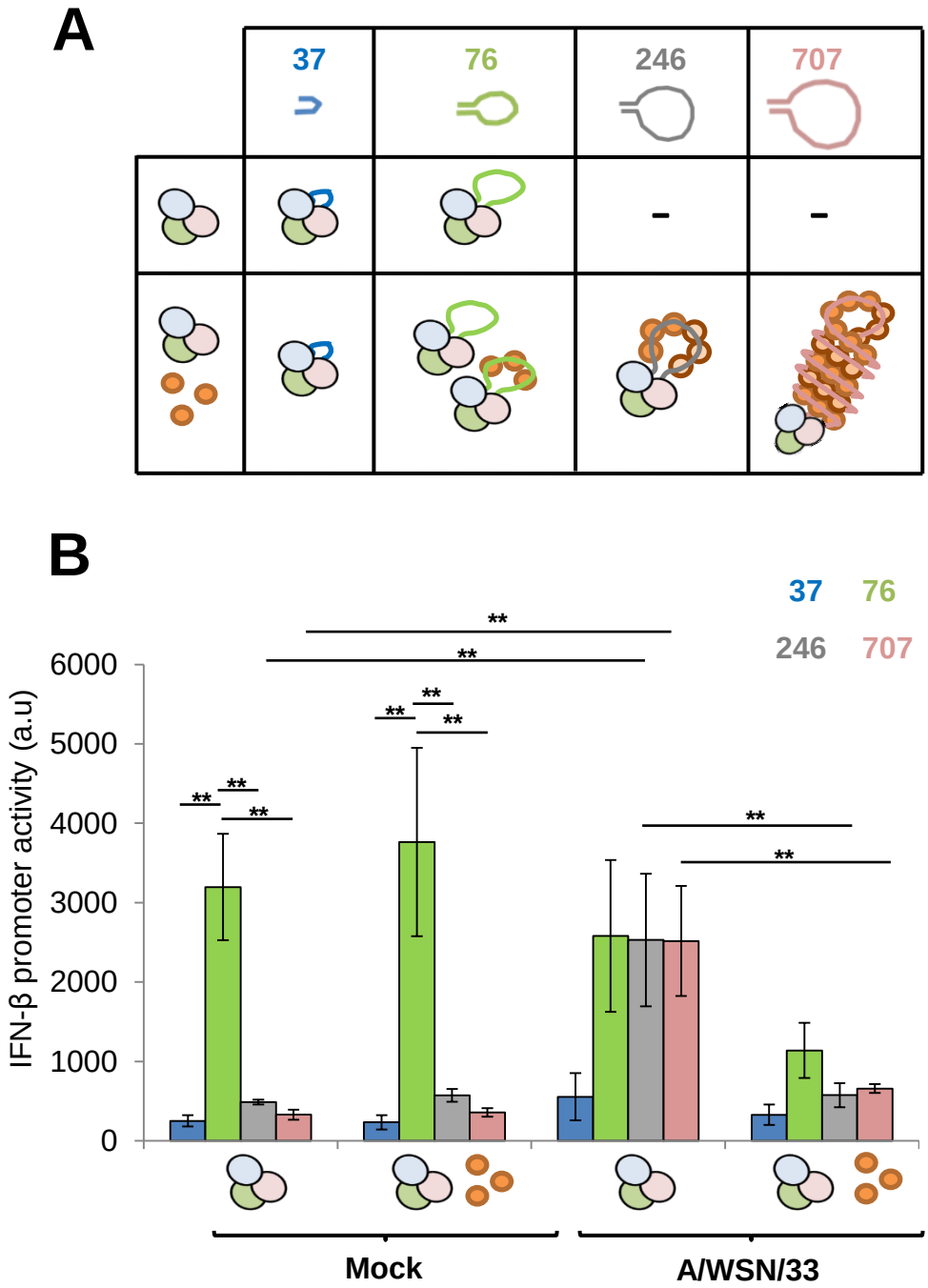


Figure 19 – Induction of interferon in RNP Reconstitutions – (A) Model of the types of RNPs generated in different RNP reconstitutions in the absence of infection. Different vRNA lengths, along with the presence or absence of NP and/or 3P are shown. A dash represents a lack of RNP accumulation due to the generated RNPs being inactive in terms of replication. (B) RNP reconstitutions were completed in cells co-transfected with the IFN-luciferase (luc) and β-galactosidase (β-gal) reporter plasmids. For the RNP reconstitutions, the cells were transfected with different combinations of the expression plasmids for vRNAs (of either 37, 76, 246 or 707 nt length), 3P and NP. 15 h post transfection, cells were either mock infected or infected with A/WSN/33 (MOI=5). 27 h post transfection, cells were lysed and the luc and β-gal activities were measured. The β-gal normalised luciferase signal is shown (IFN-β promoter activity). a.u = arbitrary units Bars represent standard deviations of average data from three independent experiments. Asterisks indicate a significant difference between the marked samples (one sample t-test; **, $p < 0.01$).

levels [37]. This leads to the conclusion that the 246 and 707 nt long RNAs, in conditions where 3P and NP are available to allow RNA encapsidation, are poor inducers of IFN (Fig. 19 B). Thus, the inability of 37, 246 and 707 nt long RNAs to induce IFN might be explained by effective 'shielding' of the RNAs by 3P and/or NP (Fig. 19 A and B). Finally, and most surprisingly, the transfection of cells with either *76 nt / 3P / NP* or *76 nt / 3P /* - induced significantly higher levels of IFN than equivalent RNP reconstitutions containing the 37, 246 or 707 nt long RNAs (Fig. 19 B – Mock). As stated above, in the*76 nt / 3P /* - condition, 'NP-less' RNPs are predicted to form whilst, in the *76 nt / 3P / NP* condition, it is unclear whether NP-less or NP-coated RNPs would form (Fig. 19 A). If 76 nt long RNPs are found in an NP-less state in both of these conditions, the ability of 76 nt long RNPs to induce IFN could give support to the initial hypothesis in one of two ways: i) Given that 3P interacts with the vRNA-associated NP, the 3P:NP interaction may be important in maintaining 3P's association to vRNA or cRNA. Thus, in NP-less RNPs, the association between 3P and RNA might be weaker and thus 3P might be more prone to dissociate from the RNA, allowing the RNA to be detected by receptors such as RIG-I. ii) The presence of a naked, NP-less RNA loop might allow access of a host factor to the RNA-loop of RNPs. This factor might then drive the dissociation of 3P, allowing the RNA to be detected by receptors such as RIG-I. It is of note that this host factor may be RIG-I itself.

The RNP reconstitutions described above lacked a number of viral proteins which are found in influenza A virus infected cells, such as NEP and M1. Thus, it was tested whether the same results would be observed if the full complement of viral proteins were present in the cell alongside the RNPs. To achieve a full complement of viral proteins, the RNP reconstitutions were performed as before, however, at 15 h post transfection, cells were infected with A/WSN/33 at an MOI of 5 (Fig. 19 B – A/WSN/33). The cells were lysed 27 h post transfection and IFN (using the luciferase system) was measured as before. When cells that

had been transfected with *37 nt / 3P / NP, 37 nt / 3P / -* or *246 nt / 3P / NP* were infected, there was no significant change in IFN levels relative to the equivalent mock infected sample (Fig. 22 B – A/WSN/33). Similarly, when cells that had been transfected with *76 nt / 3P / -* were infected, there was no significant difference in IFN induction relative to the equivalent mock infected sample (Fig. 22 B – A/WSN/33). However, when cells that had been transfected with *76 nt / 3P / NP* were infected, there was a significant, approximate 3-fold decrease in IFN levels relative to the equivalent mock infected sample (Fig. 19 B – A/WSN/33). This decrease in the IFN levels is likely due to the expression of NS1 in the infected cells, which would dampen the IFN response. Interestingly, however, when cells that had been transfected with *246 nt / 3P / -* or *707 nt / 3P / -* were infected, there was a significant, approximate 5-fold, increase in IFN levels relative to the equivalent mock infected samples (Fig. 19 B – A/WSN/33). This is despite the fact that, in these cells, prior to infection, there was likely non-detectable levels of 246 nt long and 707 nt long RNA. The difference between the effect of infection on cells transfected with i) *246 nt / 3P / NP* or *707 nt / 3P / NP* and ii) *246 nt / 3P / -* or *707 nt / 3P / -* indicates that the ability of 246 nt and 707 nt long RNAs to drive IFN induction within an infected cell may be affected by the levels of NP within said cell. Thus, in infected cells containing internal deletion RNAs of comparable sizes to DI-RNAs, the levels of NP may be important in determining IFN induction.

Overall, the data shown here indicates that both RNA length and the level of NP are key in the induction of IFN.

5.2.2 Intermediate Length RNPs are Potent Interferon Inducers

The results from Figure 19 were further investigated; initially focussing on those from the mock infected cells (Fig. 19 B - Mock). The simplest explanation of the data above was that 76 nt long RNPs were far more active in replication than their long (246 or 707 nt) or short (37 nt) counterparts and thus replicated far higher levels of RNA. If this was the case, then the strong induction of IFN in cells transfected with either *76 nt / 3P / NP* or *76 nt / 3P / -* could be explained by the higher level of viral RNA generated.

In order to test this, the experimental set-up from above (Fig. 19 B - Mock) was repeated, however, as well as measuring IFN induction, the RNA from the samples was extracted in order to quantify the levels of vRNA. This experiment was performed with a larger number of vRNA expression plasmids in order to express a wider range of vRNA lengths and thus gain more information on the importance of vRNA length in IFN induction. Secondly, in order to discount the possibility that the above results were caused by differences between the stocks of vRNA expression plasmids, such as mutations in the plasmid backbone, a separately prepared set of vRNA expression plasmids was used. Finally, a condition was added where cells were transfected with a control plasmid (3A) instead of the vRNA, NP and 3P expression plasmids. The luciferase signal from this condition was taken as 'background' and subtracted from that of the other samples.

When the RNP reconstitutions were performed, it was found that IFN was not induced significantly above background levels when no 3P expression plasmid was transfected (e.g. 76 nt / - / *NP* or *76 nt* / - / -) (Fig. 20 A). This is likely because, in these conditions, due to the absence of 3P, no active RNPs would form which would replicate the plasmid-expressed vRNAs. This result, however, is important as it demonstrates that the vRNA expressed from

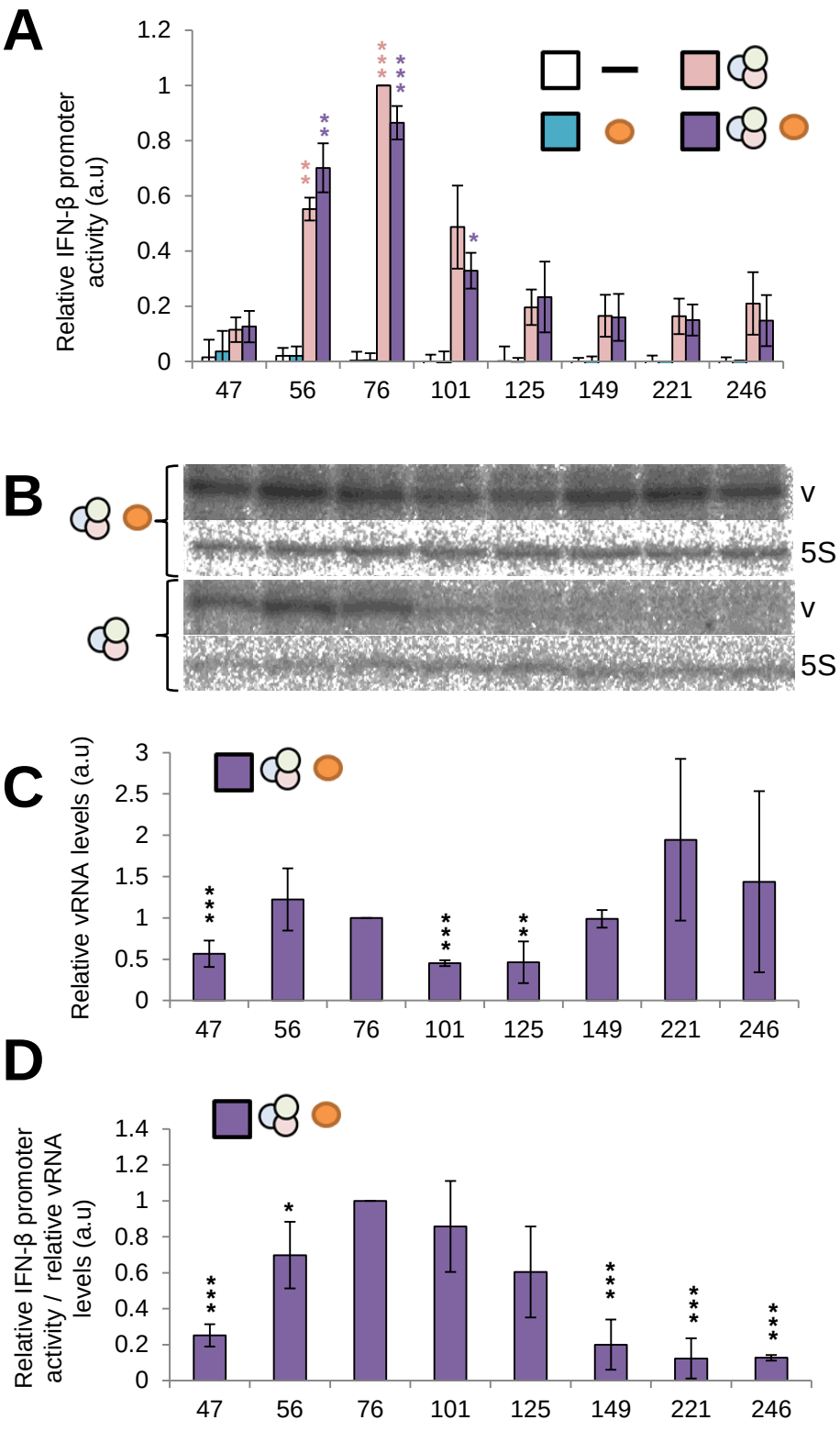

A
Relative IFN-β promoter activity (a.u)
1.2
1
0.8
0.6
0.4
0.2
0
47
56
76
101
125
149
221
246
B
v
5S
v
5S
C
Relative vRNA levels (a.u)
3
2.5
2
1.5
1
0.5
0
47
56
76
101
125
149
221
246
D
Relative IFN-β promoter activity / relative vRNA levels (a.u)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
47
56
76
101
125
149
221
246

Figure 20 – Intermediate length RNPs are strong interferon inducers – (A) RNP reconstitutions were performed in cells co-transfected with the IFN-luciferase (luc) and β-galactosidase (β-gal) reporter plasmids. For the RNP reconstitutions, the cells were transfected with different combinations of the expression plasmids for vRNAs (of varying lengths), 3P and NP. 27 h post transfection, cells were lysed and the luc and β-gal activities were measured. The background was subtracted using a control plasmid (3A) condition. The β-gal normalised luc signal is shown (relative IFN-β promoter activity). a.u = arbitrary units. Bars represent standard deviations of average data from three independent experiments. Pink asterisks indicate a significant difference between the marked sample and the *246/3P/-* sample. Purple asterisks indicate a significant difference between the marked sample and the *246/3P/NP* sample. (one sample t-test; *, $p < 0.05$. **, $p < 0.05$. ***, $p < 0.001$). **(B-C)** RNA was extracted from the samples from Panel A and the levels of vRNA and 5S rRNA were quantified by primer extension. (B) A representative result of the vRNA (v) and 5S rRNA (5S) signals from the *vRNA/3P/-* and *vRNA/3P/NP* samples are shown. (C) The 5S rRNA and vRNA signals from the *vRNA/3P/NP* samples were analysed by densitometry. The 5S rRNA-normalised vRNA levels (relative vRNA levels) from three independent repeats are shown. Bars represent standard deviations of average data from three independent experiments. Asterisks indicate a significant difference between the marked sample and the 76/3P/NP sample (one sample t-test; *, $p < 0.05$. **, $p < 0.05$. ***, $p < 0.001$). (D) The *vRNA/3P/NP* IFN data from Panel A was normalised by the corresponding vRNA levels shown in Panel C. Bars represent standard deviations of average data from three independent experiments. Asterisks indicate a significant difference between the marked sample and the 76/N/3P sample (one sample t-test; *, $p < 0.05$. **, $p < 0.05$. ***, $p < 0.001$).

the vRNA expression plasmid does not directly trigger the induction of IFN. However, when cells were transfected with the vRNA, NP and 3P expression plasmids, similar results were obtained as in Figure 19. That is, firstly, the transfection of cells with *246 nt / 3P / NP* did not induce significantly higher levels of IFN than the transfection of cells with *246 nt / 3P / -* (Fig. 20 A). The same observation was made with the 125, 149, 201 and 225 nt long RNAs (Fig. 20 A). Secondly, the transfection of cells with *56 nt / 3P / NP, 76 nt / 3P / NP, 56 nt / 3P / -* or *76 nt / 3P / -* induced significantly higher levels of IFN that equivalent RNP reconstitutions completed with the 47, 125, 149, 221 and 246 nt long RNAs (Fig. 20 A). The RNA from the 3P-containing samples was then extracted and the levels of vRNA and 5S rRNA were analysed by primer extension. Figure 20 B shows a representative repeat of this analysis. Firstly, as previously documented [37], RNPs containing RNAs of less than around 100 nt were active in the absence of NP, whilst those containing longer RNAs were not (Fig. 20 B). This is demonstrated by the lack of detectable levels of vRNA in cells transfected with *246 (as well as 125, 149 and 221) nt / 3P / -* (Fig. 20 B). Densitometry analysis was then performed on the primer extensions from the three repeats to calculate the 5S rRNA-normalised vRNA level in the different conditions. This was only completed on the RNP reconstitutions in which the NP expression plasmid had been transfected. The data demonstrated that there were no significant differences between the levels of vRNA in the cells transfected with i) *56 nt / 3P / NP* or *76 nt / 3P / NP* and those transfected with ii) *149 nt / 3P / NP*, 221 *nt / 3P / NP* or *246 nt / 3P / NP* (Fig. 20 C). Interestingly, there did appear to be significantly lower vRNA levels in the cells transfected with *47 / 3P / NP, 101 / 3P / NP* or *125 nt* / 3P / NP relative to those transfected with *76 / n / 3P,* however, the cause of this is unknown (Fig. 20 B and C).

This data, together, allowed a number of conclusions to be drawn:

- As stated above, the cells transfected with, for example, *246 nt / 3P / NP* containing higher levels of vRNA that those transfected with *246 nt / 3P /* -, however, did not induced higher levels of IFN (Fig. 20 A and B). This supports the hypothesis that RNAs over 149 nt long, when replicated in conditions where the RNA can be packaged into RNPs (i.e. in the presence of both NP and 3P), are poor inducers of IFN. This could be explained by 3P and NP 'shielding' the viral RNA from detection (**section 1.6.1.4**).

- The ability of cells transfected with *56 nt / 3P / NP* or *76 nt / 3P / NP* to induce higher levels of IFN than those transfected with *149 nt / 3P / NP*, *221 nt / 3P / NP* or 246 *nt / 3P / NP* (Fig. 20 A) was not explained by differences in the levels of vRNA generated between the conditions, as hypothesised in **section 5.2.1**. In order to calculate the relative IFN inducing activity of each RNA (or RNP) type relative to its levels, the IFN data from Figure 20 A was normalised by the corresponding vRNA/5S rRNA levels from Figure 20 C. The resultant vRNA-normalised IFN data demonstrates that, relative to their levels, 56 and 76 nt long RNA/RNPs are stronger inducers of IFN than their shorter (47 nt long) and longer (149, 221 and 246 nt long) counterparts (Fig. 20 D).

In this thesis, the generated RNA/RNPs will from now on be divided into three categories; i) 'short' RNAs / RNPs are those which can replicate without NP but do not induce a strong IFN response (e.g. 37 nt and 47 nt long RNAs), ii) 'intermediate' length RNAs / RNPs are those which can replicate without NP and induce a strong IFN response (e.g. 56 nt and 76 nt long RNAs), iii) 'long' RNAs / RNPs are those which cannot replicate without NP and, under conditions of high NP and 3P, do not induce a strong IFN response (e.g. 125, 149, 221 and 246 nt long RNAs).

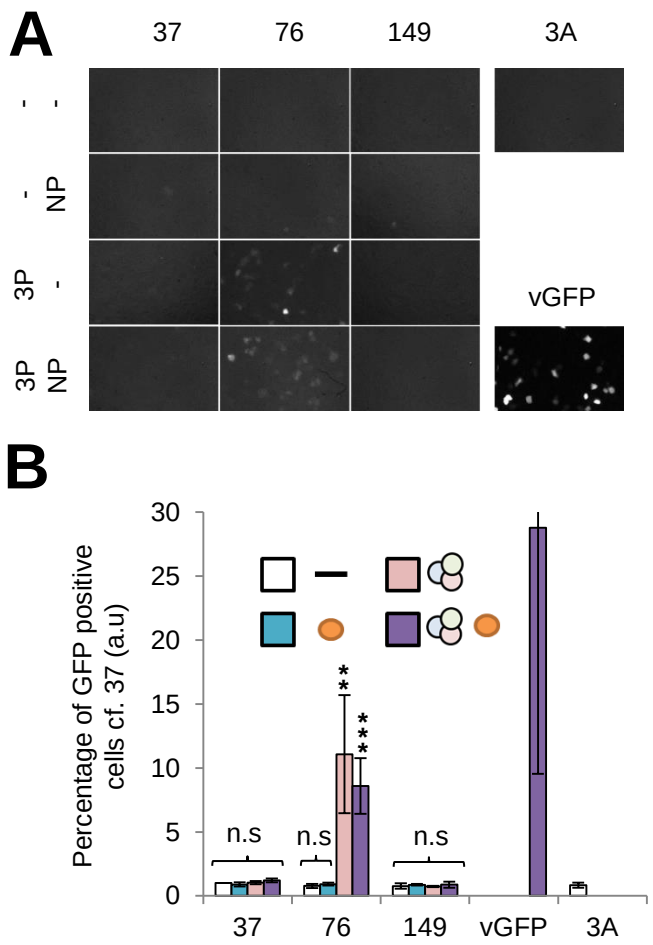


Figure 21 – Cells containing intermediate length RNPs induce GFP expression in an IFN-GFP cell line – RNP reconstitutions were performed in the IFN-GFP cell line, which encodes GFP under the control of the IFN promoter. For the RNP reconstitutions, cells were transfected with different combinations of the expression plasmids for vRNA (either 37, 76 or 149 nt long), 3P and NP. A positive control for the successful generation of RNPs was included using an expression plasmid for vGFP; a vRNA which, when transcribed by 3P, generates a GFP-encoding mRNA. Transfection of a control plasmid (3A) was included as a negative control. Cells were fixed 48 h post transfection. (A) Imaging of cells by fluorescence microscopy. Representative images are shown. The exposure for all conditions was identical apart from vGFP, for which the exposure was decreased due to the strength of the GFP signal. (B) Cells from 'A' were analysed by flow cytometry for GFP expression. The percentage of GFP positive cells relative to the '37 -' condition is shown. a.u. = arbitrary units. Bars represent standard deviations of average data from three independent experiments. Asterisks indicate a significant difference between the marked sample and the 3A sample (one sample t-test; *, $p < 0.05$. **, $p < 0.05$. ***, $p < 0.001$. n.s, $p > 0.05$).

The experimental set up from above was next repeated in a cell line which expresses GFP under the control of the IFN promoter. However, there were two differences in the transfection: i) Only three vRNA lengths were used; 37 nt (a representative 'short' RNA), 76 nt (a representative 'intermediate length' RNA) and 149 nt (a representative 'long' RNA). ii) A control condition was added in which an RNP reconstitution was performed using the pPolI GFP plasmid. The pPolI GFP plasmid expresses a vRNA (vGFP), which, when transcribed by 3P, leads to the expression of GFP. Therefore, this condition was used as a control for successful transfection and RNP reconstitution.

Microscopy analysis of the transfected cells supported the conclusions from above (Fig. 21 A). In short, the transfection of cells with a control plasmid (3A) led to only low levels of GFP positive cells. However, the transfection of cells with *vGFP / 3P / NP* led to high levels of GFP-positive cells, as expected. When the expression plasmids for 3P were not included in the transfection, the levels of GFP positive cells were comparable to that of the control plasmid (3A) transfected cells (Fig. 21 A). Furthermore, equivalently low levels of GFP positive cells were observed when cells were transfected with either *37 nt / 3P / NP*, *37 nt / 3P / -*, *149 nt / 3P / NP* or *149 nt / 3P / -* (Fig. 21 A). On the other hand, high numbers of GFP positive cells were visible when cells were transfected with either *76 nt / 3P / NP* or *76 nt / 3P / -* (Fig. 21 A). Analysis of the cells by flow cytometry confirmed that only transfection of cells with *76 / 3P / NP* or *76 / 3P / -* (and the vGFP condition) led to significantly higher levels of GFP positive cells than the control transfected cells (Fig. 21 B). This result supports the conclusion that intermediate length RNPs, which can replicate in the absence of NP (and thus without NP encapsidation), are strong inducers of IFN.

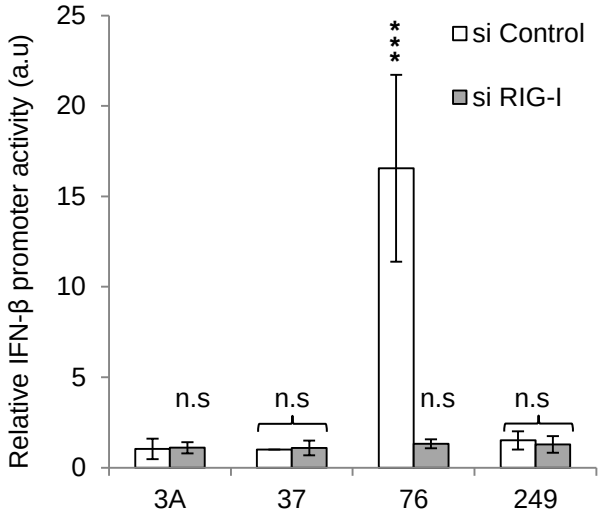


Figure 22 – RIG-I is essential for interferon induction in cells containing intermediate length RNPs – Cells were co-transfected with either a control- or RIG-I-specific siRNA, along with the IFN-luciferase (luc) and β-galactosidase (β-gal) reporter plasmids. 24 h post transfection, the cells were transfected with different combinations of the expression plasmids for vRNA (of varying lengths), 3P and NP. Transfection of empty vector (3A) was used as a negative control. 24 h post RNP reconstitution, cells were lysed and the luc and β-gal activities were measured. The β-gal normalised luc signal (relative IFN-β promoter activity) is shown relative to the 37/siControl sample. a.u. = arbitrary units. Bars represent standard deviations of average data from three independent experiments. Asterisks indicate a significant difference between the marked sample and the 3A/siControl sample (one sample t-test; ***, $p < 0.001$. n.s, $p > 0.05$).

5.2.3 A RIG-I Ligand is Generated in Intermediate Length RNP Reconstitutions

Above, it was demonstrated that RNP reconstitutions completed with intermediate length RNAs lead to a strong induction of IFN. In order to determine by which mechanism IFN was being induced by these RNP reconstitutions, the experimental setup from above was repeated, however, cells were either transfected 24 h pre-RNP reconstitution with an anti-RIG-I siRNA or a control siRNA. It is important to note that the successful knockdown of RIG-I was not demonstrated due to the inability of the available RIG-I-specific antibodies to detect endogenous RIG-I (data not shown). When RNP reconstitutions were performed with control siRNA transfected cells, the same trend as above was found; that is the transfection of cells with *76 nt / 3P / NP* led to a significantly higher induction of IFN than the transfection of cells with *37 nt / 3P / NP* or *149 nt / 3P / NP* (Fig. 22). This response, however, was abolished when RIG-I was knocked down (Fig. 22). This indicated that the ligand which was triggering the IFN response in cells containing intermediate length RNPs (whether it be the intermediate length vRNA, intermediate length cRNA or an aberrant product of intermediate length RNA replication) was being detected by RIG-I. This is of interest as, as stated in **section 1.6.1.1**, in influenza A virus infections of lung epithelial cells, the IFN response is primarily triggered by the RIG-I receptor [98]. Therefore, it was possible that the ligand generated in intermediate length RNP reconstitutions might be responsible for IFN induction in influenza A virus infected cells.

5.2.4 The Presence of Limiting NP in RNP Reconstitutions Promotes the Induction of IFN

The data above demonstrates that some feature of intermediate length RNPs causes them to be stronger inducers than long RNPs. As stated above, the most obvious differences between intermediate length and full length RNPs are: i) the length of the RNA they contain and ii) their ability to replicate RNA in the absence of NP. Thus, it was possible that either i) the formation of intermediate length RNAs and/or ii) the formation of 'NP-less' RNPs might lead to the induction of IFN expression.

In order to separate these two possibilities, conditions were tested which would be expected to lead to the formation of 'NP-deficient' long RNPs (e.g. 246 nt long RNPs). Previously, the term 'NP-less RNP' was used to refer to RNPs with completely unencapsidated 'RNA-loops' (e.g. Fig, 23 A – B1). The term 'NP-deficient RNPs', however, is used to refer to both 'NP-less RNPs' and those RNPs which have only partially NP encapsidated RNA-loops (e.g. Fig. 23 A – B1 and B4).

In order to attempt to generate 'NP-deficient RNPs' containing RNAs longer than 100 nt, RNP reconstitutions were performed, however, varying amounts of NP expression plasmid were transfected. It was hypothesised that with low amount of NP expression plasmids, the level of NP would become 'limiting'. In this context, 'limiting NP' refers to there not being enough NP present to completely encapsidate all the 3P-bound RNA. Below, it is hypothesised as to whether appreciable levels of 'NP-deficient' RNPs would be generated in different conditions. Those points which are speculations are highlighted in italics:

- ***76 nt / 3P / -* transfected cells** (Fig. 23 A – B1)**:**
  - Robust vRNA replication, no NP for encapsidation
  - *High levels of 'NP-deficient' RNPs*

- ***76 nt / 3P / NP* transfected cells** (Fig. 23 A – B1)**:**
  - Robust vRNA replication, high levels of NP for encapsidation
  - *It is possible, however, that NP is not loaded onto these RNPs*
  - *High levels of 'NP-deficient' RNPs*
- ***246 nt / 3P / -* transfected cells** (Fig. 23 A – B2)**:**
  - Undetectable vRNA replication, no NP for encapsidation
  - *Low levels of 'NP-deficient RNPs'*
- ***246 nt / 3P / NP* transfected cells** (Fig. 23 A – B3)**:**
  - Robust vRNA replication, high levels of NP for encapsidation
  - *Low levels of 'NP-deficient RNPs'*
- ***246 nt / 3P / NP* transfected cells, NP limiting** (Fig. 23 A – B4)**:**
  - Fairly robust vRNA replication as some active RNPs formed
  - Low levels of NP for encapsidation, however, high levels of 3P
  - *Some progeny vRNAs and cRNAs are bound by 3P but not by NP*
  - *Moderate levels of 'NP-deficient' RNPs.*

The above conditions were tested in RNP reconstitutions, maintaining constant levels of the vRNA and 3P expression plasmids, whilst serially diluting the levels of the NP expression plasmid. The amount (nanograms) of each expression plasmid transfected in each condition is denoted alongside the transfection conditions. 24 h post transfection, the cells were lysed and the levels of NP quantified by Western Blot. It was found that a dilution series of NP was successfully achieved (Fig. 23 B - Top). The RNA was then extracted from the samples, the levels of vRNA and 5S rRNA were measured by primer extension, and a 5S rRNA normalised vRNA level was calculated. Similar to that observed in **section 5.2.2**, it was found that comparable levels of vRNA were generated when cells were transfected with *250 ng 76 nt / 250 ng 3P / 250 ng NP* as when cells were transfected with *250 ng 246 nt / 250 ng 3P / 250 ng NP* (Fig. 23 B – Top). When cells were transfected with 250 ng *246 nt / 250 ng 3P / 0 ng NP*, no detectable vRNA was observed (Fig. 23 B - Top). Furthermore, when cells were transfected with decreasing levels of NP expression plasmid (*e.g. 250 ng 246 / 250 ng 3P /*

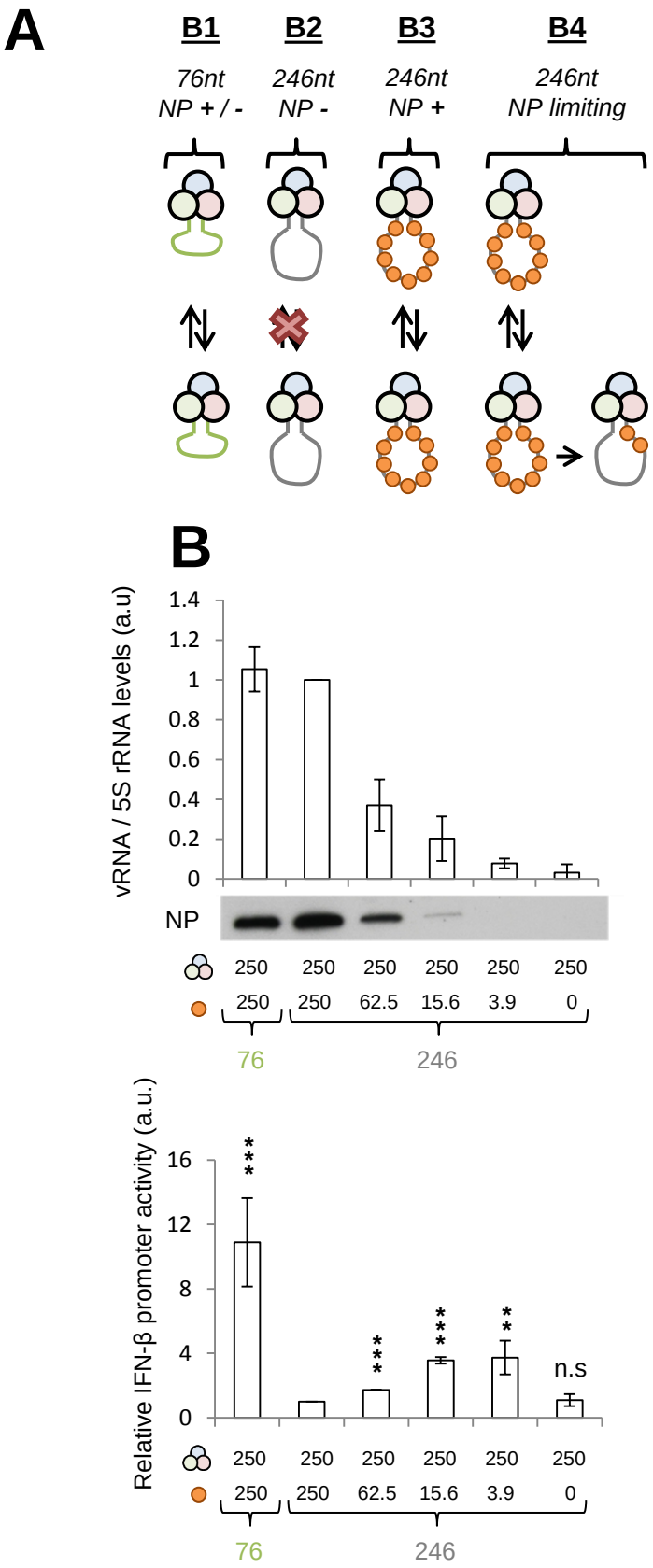

A
B1
B2
B3
B4
76nt
NP + / -
246nt
NP -
246nt
NP +
246nt
NP limiting
B
vRNA / 5S rRNA levels (a.u)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
NP
250 250 250 250 250 250
250 250 62.5 15.6 3.9 0
76
246
Relative IFN-β promoter activity (a.u.)
16
12
8
4
0
***
***
***
**
n.s
250 250 250 250 250 250
250 250 62.5 15.6 3.9 0
76
246

Figure 23 – Conditions which would likely promote the formation of NP-deficient RNPs lead to interferon induction - (A) Model of the RNP species formed under different conditions, B1 – Intermediate length (76 nt) RNPs can replicate viral RNA independently of the presence of NP. B2 – Under conditions of no NP, long (246 nt) RNPs are unable to replicate viral RNA. B3 – Under conditions of high NP, long (246 nt) RNPs are able to replicate viral RNA and progeny viral RNAs are likely efficiently encapsidated. B4 – Under conditions of limiting NP, some fully encapsidated, active RNPs are generated, however, not all progeny viral RNAs are able to be fully encapsidated by NP. (B) RNP reconstitutions were performed in cells co-transfected with the IFN-luciferase (luc) and β-galactosidase (β-gal) reporter plasmids. For the RNP reconstitutions, cells were transfected with a constant amount (250 ng) of the 3P and (either 76 nt or 246 nt long) vRNA expression plasmids. The amount of NP expression plasmid transfected, however, was varied. The amount (ng) of each plasmid transfected (except the vRNA expression plasmid) is shown. Top – 27 h post transfection, the cells were lysed. A representative western blot to analyse the levels of NP in the samples shown. The RNA was extracted from the samples and the levels of vRNA and 5S rRNA were quantified by primer extension. The 5S rRNA-normalised vRNA level is shown. Bars represent the variance of the average data from two independent experiments. Bottom - The luc and β-gal activities of each sample were measured. The background was subtracted using a control plasmid (3A) condition. The β-gal normalised luc signal is shown. Bars represent standard deviations of average data from three independent experiments. Asterisks indicate a significant difference between the marked sample and the 246/250ng NP sample (one sample t-test; ***, $p < 0.001$. **, $p < 0.01$. n.s, $p > 0.05$).

62.5, 15.6 or 3.9 ng NP), lower levels of vRNA were produced (Fig. 23 B – Top). This is as expected, as the limiting NP likely causes less NP-encapsidated active RNPs to be formed, meaning there is less efficient replication of the vRNA. When the corresponding IFN induction was measured for these samples, it was found that, as in **section 5.2.2**, cells transfected with 250 ng *76 nt / 250 ng 3P / 250 ng NP* induced IFN strongly (Fig. 23 B - Bottom). Furthermore, the transfection of cells with either *250 ng 246 nt / 250 ng 3P / 250 ng NP* or *250 ng 246 nt / 250 ng 3P / 0 ng NP* induced low levels of IFN (Fig. 23 B - Bottom). Interestingly, however, despite the lower levels of vRNA generated, the transfection of cells with either *250 ng 246 nt / 250 ng 3P / 15.6 ng NP* or *250 ng 246 nt / 250 ng 3P / 3.9 ng NP* induced IFN significantly more than the transfection of cells with *250 ng 246 nt / 250 ng 3P / 250 ng NP* (Fig. 23 B - Bottom). This indicates that, during the replication of long RNPs, conditions of limiting NP can lead to an IFN response. Furthermore, IFN induction occurred most strongly in the conditions in which NP-deficient RNPs had been predicted to form (Fig. 23 A and B). This fits the hypothesis that NP-deficient RNPs are able to induce IFN expression, although other possibilities are discussed in the discussion (Fig. 23 A and B).

.

5.2.5 vRNA Length, Not Sequence, Defines Interferon Induction

The next focus was on determining what feature of NP-deficient RNPs leads to IFN induction. One of the documented roles of NP is to block the formation of secondary RNA structures in RNPs [62]. Thus, it was hypothesised that a specific sequence in the loop region of NP-deficient 76 nt and 246 nt long RNPs (used in **sections 5.2.1, 5.2.2 and 5.2.3**) was forming a secondary structure which was important in driving the induction of IFN (Fig. 24 A). Importantly, given that the 76 nt and 246 nt long RNAs used above are both derived by

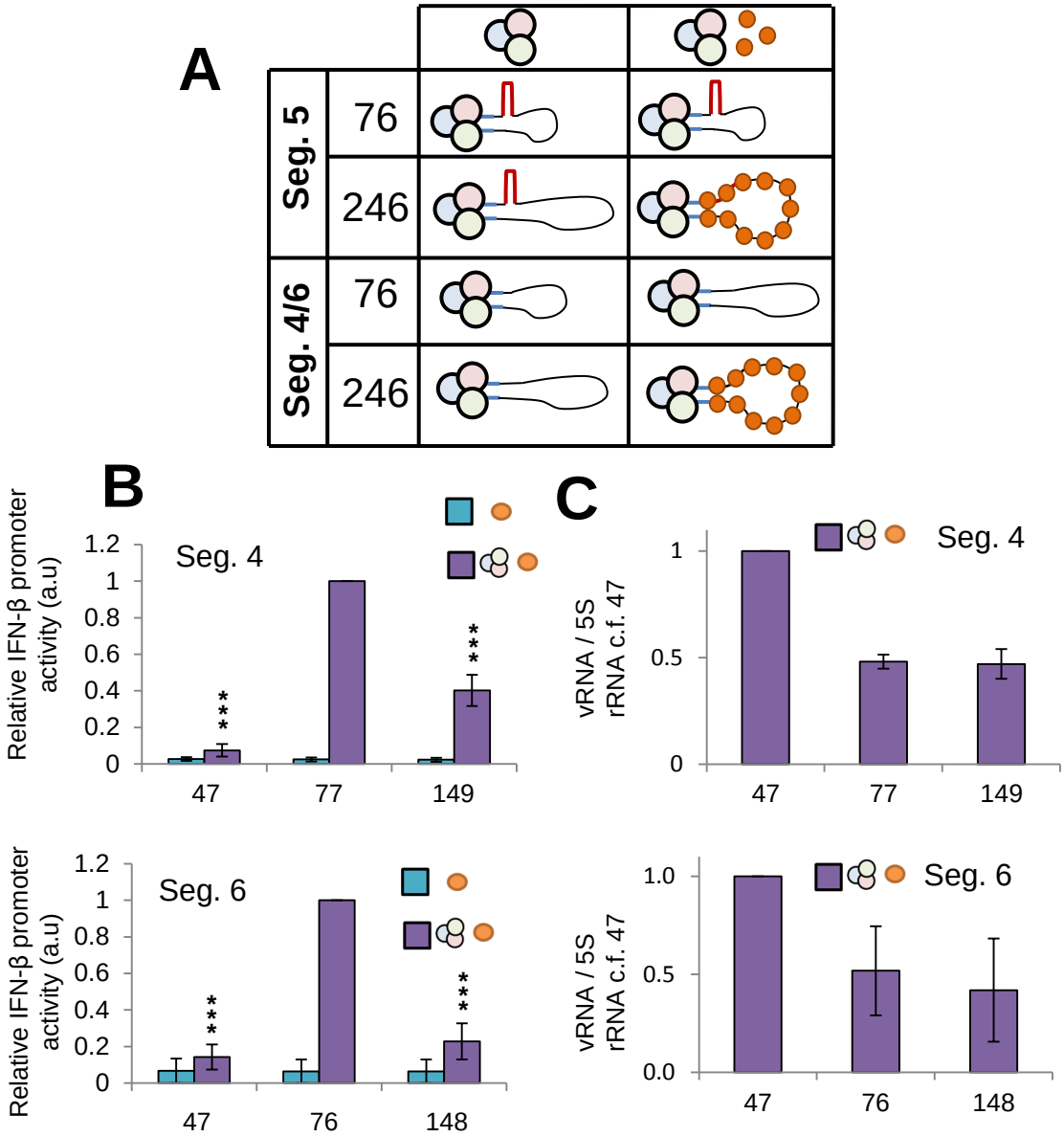


Figure 24 – The internal 'loop' sequence of the segment 5 derived vRNAs is not important for the induction of interferon by intermediate length RNPs. (A) Model of the possible structures of different RNP species formed from segment 4, 5 or 6 derived vRNAs when bound by 3P, or 3P and NP. In this model, the RNA-loop region (the region normally bound by NP in an RNP) of 76 nt long RNPs are modelled to not bind NP. The vRNA expression plasmids used in Figures 22 – 26 are all derived from the segment 5 vRNA sequence. Thus, all will contain certain conserved sequences within the RNA loop region. These sequences may be able to form into secondary structures in 'NP deficient RNPs'. Short vRNAs derived from different vRNA segments would have identical 5' and 3' promoter termini (blue), however, would have different internal RNA sequences and thus would form distinct RNA secondary structures. (B) RNP reconstitutions were performed in cells co-transfected with the IFN-luciferase (luc) and β-galactosidase (β-gal) reporter plasmids. For the RNP reconstitutions, cells were transfected with different combinations of the expression plasmids for NP, 3P and for a vRNA derived from the segment 4 (HA) or segment 6 (NA) vRNA. 27 h post transfection, the cells were lysed and the luc and β-gal activities were measured. The background was subtracted using a control plasmid (3A) condition. The β-gal normalised luc signal (relative IFN-β promoter activity) is shown. Data is normalised against the 76/77 nt NP/3P condition. a.u. = arbitrary units. Bars represent standard deviations of average data from three independent experiments Asterisks indicate a significant difference between the marked condition and the 76/77 nt NP/3P condition (one sample t-test; ***, $p < 0.001$). (C) RNA was extracted from the RNP reconstitution samples. The level of vRNA and 5S rRNA in each condition was quantified by primer extension. Densitometry of the vRNA and 5S rRNA signal was used to generate an average 5S rRNA-normalised vRNA level for each condition. Bars represent standard deviations of average data from three independent experiments

deletions from the same vRNA (the segment 5 vRNA), both RNAs would contain identical sequences in the loop region (though the 246 nt long RNA contains RNA sequence which the 76 nt long RNA does not). Thus, in an NP-deficient state, both the 76 nt and 246 nt RNPs might form the same IFN-inducing secondary structure (Fig. 24 A). To test this hypothesis, RNP reconstitutions were performed using vRNA expression plasmids derived from two different genome segments; segment 4 (s4) and segment 6 (s6). All of the deletion vRNAs, whether from segment 4, 5 or 6, have the same conserved termini, where 3P binds, however, have different internal sequences in the 'loop region' and thus would not form the same secondary structures (Fig. 24 A). RNP reconstitutions performed with short and long, segment 4 derived vRNAs (*s4 47 nt / 3P / NP* or *s4 149 nt / 3P / NP*) induced significantly less IFN than those performed with an intermediate length, segment 4 derived vRNA (*s4 77 nt / 3P / NP*) (Fig. 24 B – Top). The same result was found with segment 6 derived vRNAs, with RNP reconstitutions performed using short and long, segment 6 derived vRNAs (*s6 47 nt / 3P / NP* or *s6 148 nt / 3P / NP*) inducing significantly less IFN than those using an intermediate length, segment 6 derived vRNA (*s6 76 nt / 3P / NP*) (Fig. 24 B – Bottom). Furthermore, as with the segment 5 derived vRNAs, i) there were no significant differences between the levels of vRNA generated in cells transfected with *s4 77 nt / 3P / NP* and in those transfected with *s4 149 nt / 3P / NP* and ii) there were no significant differences between the levels of vRNA generated in cells transfected with *s6 76 nt / 3P / NP* and in those transfected with *s6 148 nt / 3P / NP* (Fig. 24 C – analysis by AJ Te Velthuis). Therefore, as in Figure 20 C, it appears that segment 4 and segment 6 derived intermediate length RNAs, in conditions where 3P is available to bind the RNAs, are strong IFN inducers relative to their levels (Fig. 24 B and C). This demonstrates that it is not the internal sequence of the intermediate length RNAs, but specifically their length, which is important in their ability to induce IFN.

5.2.6 Intermediate Length RNP Reconstitutions Generate a Strongly Inducing RNA Species

It was next investigated whether the difference in the ability of short, intermediate and long RNPs to induce IFN might be caused by differential shielding (**section 1.5.1.2**) of the vRNAs and cRNAs by 3P and NP. It was hypothesised that NP-deficient RNPs might be stronger inducers of IFN because they are less stable, more likely to experience 3P dissociation and thus more likely to expose the RNA they contain to RIG-I (in a similar manner to that shown in **Fig. 18 B**). To test this hypothesis, cells were transfected with either *37 nt / NP / 3P*, *76 nt / NP / 3P, 149 nt / NP / 3P* or with a control plasmid (3A). At 24 h post transfection, cells were lysed and the RNA was extracted. The process of RNA extraction would remove the viral proteins and thus remove any effect of protein shielding. Then, the extracted RNAs were transfected into cells along with the reporter plasmids to test the IFN-inducing activity of the RNAs. If shielding was the cause of the differential IFN induction, then, when the viral proteins were removed, all RNA species should induce comparably. This hypothesis is displayed graphically in Figure 25 A. The RNA extracted from the *76 nt / 3P / NP*-transfected cells, however, induced IFN significantly more than the RNA extracted from *37 nt / 3P / NP* or *149 / 3P / NP*-transfected cells (Fig. 25 B). This indicated that the shielding of vRNA and cRNA by the binding of 3P and NP did not explain the differential abilities of 'intermediate length' and 'short or long' RNAs to induce IFN. Therefore, it was hypothesised that there is an IFN inducing RNA species generated in cells containing i) intermediate length RNPs (*56 or 76 nt*) or ii) long RNPs (*e.g. 246 nt*) under conditions of limiting NP.

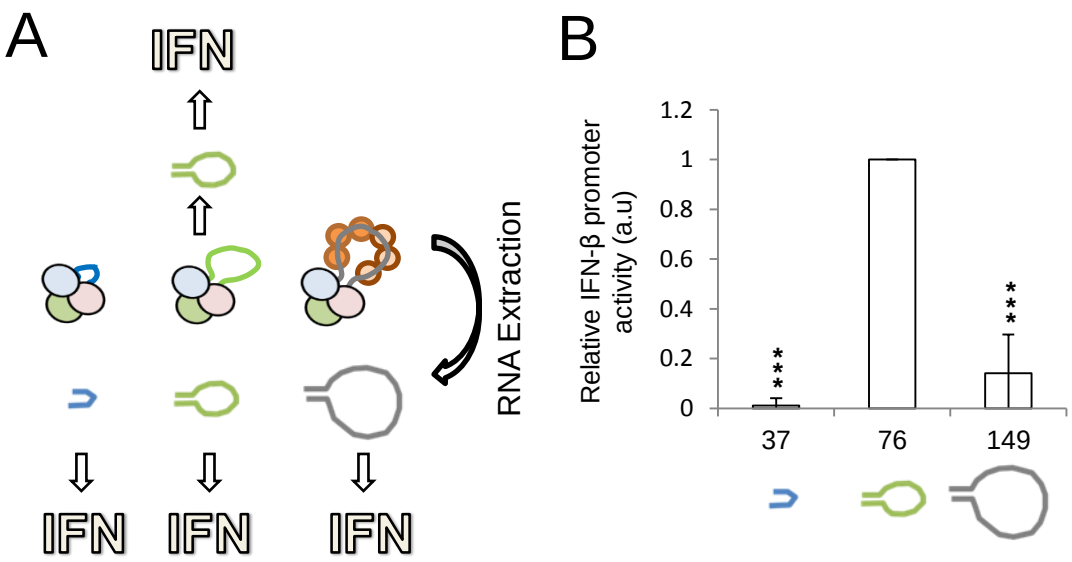
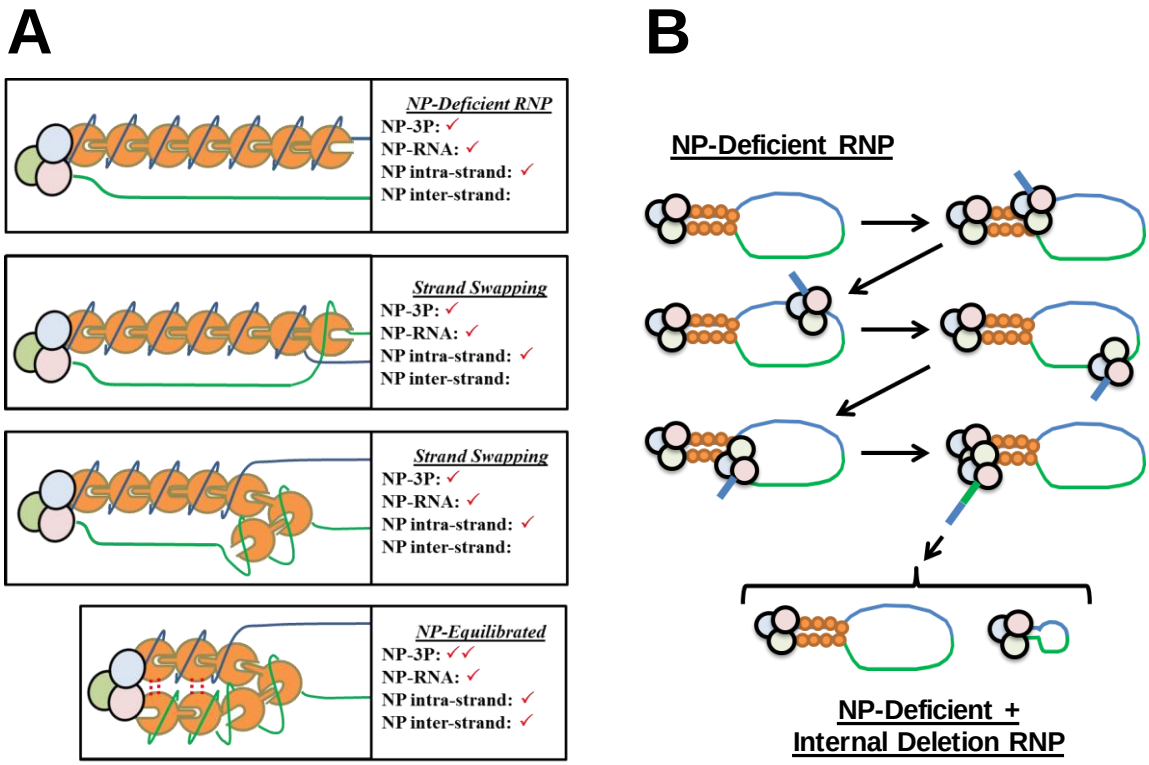
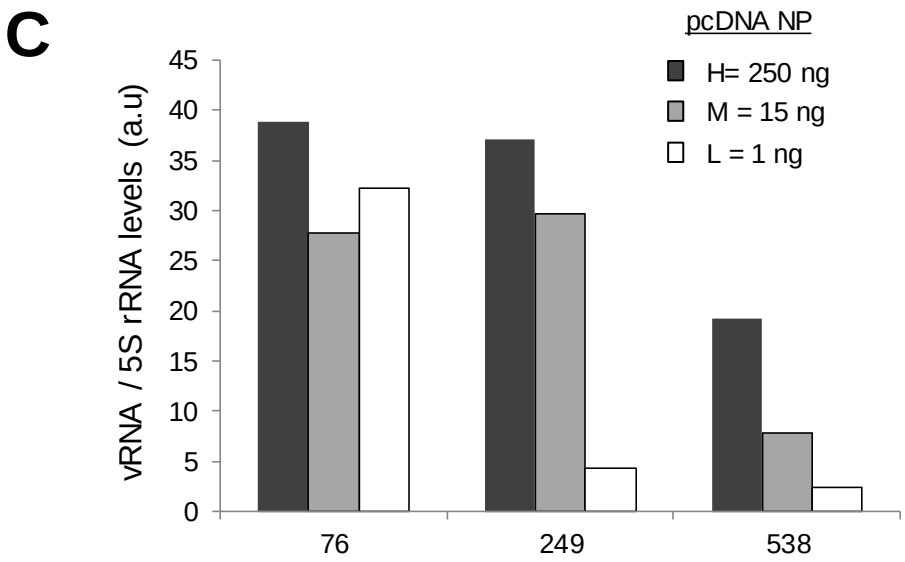


Figure 25 – Intermediate length RNPs do not induce interferon strongly due to decreased 'shielding' – (A) Model of how shielding could explain the differential ability of short, intermediate and long RNPs to induce IFN. The shielding of the vRNA or cRNA by 3P binding inhibits RIG-I binding to the RNA. The presence of a naked, NP-less RNA in the intermediate length RNPs (green) leads to a decreased stability of the RNP. Thus, 3P is more likely to dissociate, allowing RIG-I to bind the free RNA. If the vRNA and cRNAs were extracted from the RNPs, the effects of shielding would be removed. Thus, if these RNAs were transfected into cells, they should induce comparably. (B) Cells were transfected with the expression plasmids for vRNA (37, 76 or 149 nt long, segment 5 derived), 3P and NP. The transfection of a control plasmid (3A) was used as a negative control. 27 h post transfection, cells were lysed in trizol and the total RNA extracted. An equal amount of total RNA was then transfected into cells alongside the IFN-luciferase (luc) and β-galactosidase (β-gal) reporter plasmids. 24 h post transfection, cells were lysed and the reporter activities measured. The background was subtracted using the control plasmid (3A) condition. The β-gal normalised luc signal (relative IFN-β promoter activity) is shown. a.u. = arbitrary units. Bars represent standard deviations of average data from three independent experiments. Asterisks indicate a significant difference between the marked condition and the 76 nt condition (one sample t-test; ***, $p < 0.001$).

5.2.7 Limiting NP Promotes the Accumulation of Internal Deletion RNAs

In order to determine how limiting NP might lead to the induction of IFN, the structure of NP-deficient RNPs was first considered. It was hypothesised that, due to the proposed mechanism of NP assembly onto RNPs (**section 1.3.5.5**) [37], NP-deficient RNPs would initially have NP loaded onto only one vRNA terminus (Fig. 26 A). The NP molecules in such an RNP, however, would likely not be able to form all types of NP-interactions that are normally present in an RNP (**described in section 1.3.3.2**). Specifically, the NP within NP-deficient RNPs would be able to form the NP:3P, the NP intra-strand (NP:NP; tail loop, insertion groove mediated) and the NP:RNA interactions, however, given it is only localised on one of the RNA termini, it would not be able to form the NP inter-strand interaction (and potentially not a second NP:3P interaction at the 5′ proximal region of the RNA) (Fig. 26 A – Top). However, if NP was able to undergo 'strand switching' and equilibrate between the 5′ and 3′ termini of the RNA, then all of the interactions might be able to form and the resultant RNP might be more stable (Fig. 26 A). When such a hypothetical RNP undergoes replication (demonstrated in Fig. 26 B with a trans-acting polymerase), 3P might only be able to replicate the NP coated RNA termini, but not the internal, NP-less RNA (based on the requirement of NP for elongation by 3P [37]). Interestingly, if NP was equilibrated between the two RNA termini, this would generate an internal deletion RNA with equal lengths of the 5′ and 3′ RNA termini replicated (Fig. 26 B); an effect that has been observed with DI RNAs [147]. Therefore, this appeared to be an attractive model for the formation of DI RNAs.


A
NP-Deficient RNP
NP-3P: ✓
NP-RNA: ✓
NP intra-strand: ✓
NP inter-strand:
Strand Swapping
NP-3P: ✓
NP-RNA: ✓
NP intra-strand: ✓
NP inter-strand:
Strand Swapping
NP-3P: ✓
NP-RNA: ✓
NP intra-strand: ✓
NP inter-strand:
NP-Equilibrated
NP-3P: ✓✓
NP-RNA: ✓
NP intra-strand: ✓
NP inter-strand: ✓
B
NP-Deficient RNP
NP-Deficient +
Internal Deletion RNP


C
pcDNA NP
H= 250 ng
M = 15 ng
L = 1 ng
vRNA / 5S rRNA levels (a.u)
45
40
35
30
25
20
15
10
5
0
76
249
538

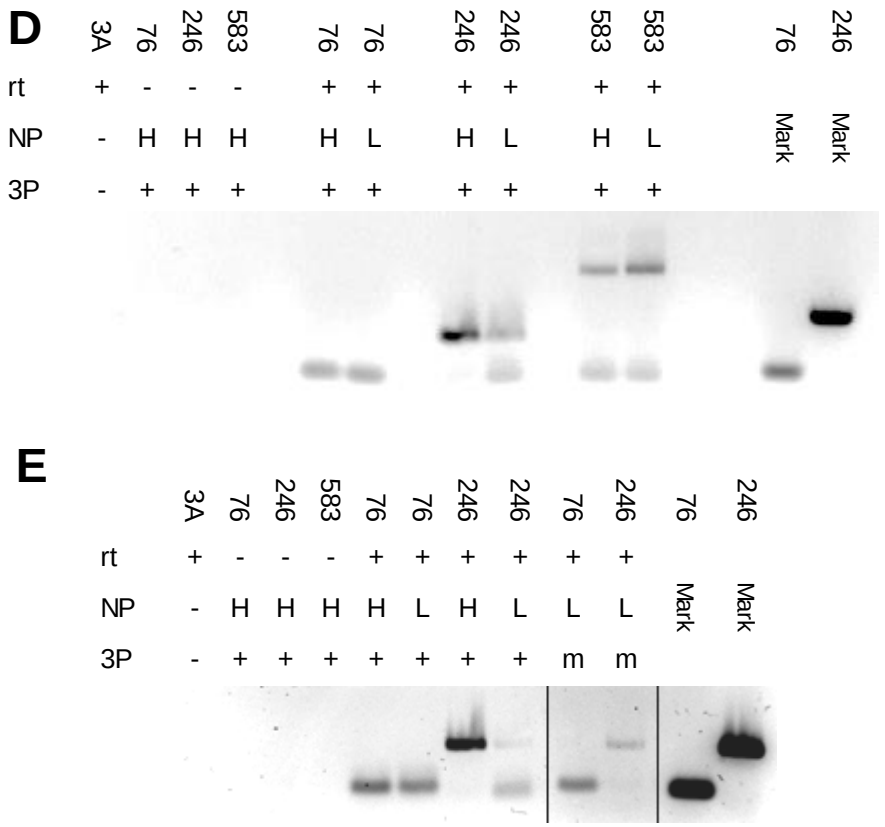


Figure 26 – Conditions of low NP drive the accumulation of internal deletion RNAs – (A) Model of the NP interactions which may form in 'NP deficient' RNPs. Panel 1 – Given NP loads linearly onto vRNAs and cRNAs, an NP deficient RNP likely only has NP loaded initially onto one of the vRNA or cRNA terminini. This would mean that NP would not be able to form the inter-strand interaction. Panel 2 / 3 – In such an RNP, NP may be able to undergo 'strand switching' and equilibrate between the two termini. Panel 4 – This would allow the formation of the inter-strand interaction whilst maintaining the other interactions (B) Model of NP-deficient RNP replication. An NP-deficient RNP from Panel A is shown. Replication is shown via a 'trans-acting polymerase'. The trans-acting polymerase replicates the NP-encapsidated 3′-proximal region of the template vRNA or cRNA. The polymerase is unable to copy the NP-deficient regions of vRNA or cRNA. The polymerase either slides along the template (or potentially dissociates and reassociates) until it reaches the NP-encapsidated 5′-proximal region of the template. The polymerase, upon interaction with NP, is able to re-initiate and replicate the 5'-proximal region of the template. This leads to the generation of an internal deletion RNA with similar lengths of 5'- and 3'-proximal RNA replicated. (C) RNP reconstitutions were completed with a constant amount of vRNA expression plasmid (either 76 nt, 246 nt or 583 nt) and 3P expression plasmids. The level of NP expression plasmid, however, was serially diluted (250 ng, 15 ng or 1 ng). 24 h post transfection, the RNA was extracted from the samples and the vRNA and 5S rRNA levels were quantified by primer extension. The 5S rRNA normalised vRNA levels are shown. a.u = arbitrary units. (D) The *vRNA / 250 ng NP* (H = high NP) and *vRNA / 1 ng NP* (L = low NP) RNA samples from Panel C were used in a reverse transcription reaction with a 3' vRNA primer. This was completed + / - reverse transcriptase. A PCR reaction was then completed on the products using 5' and 3' vRNA primers to amplify any vRNAs which contain both of the vRNA termini. PCR amplification of the vRNA sequences from the vRNA expression plasmids (76 and 246 nt) were used to generate size markers. (E) The transfection from Panel D were repeated, however using only the 76 nt and 246 nt long vRNA expression plasmids. Alongside this, RNP reconstitutions were performed using the *76 nt / 1 ng NP* and *246 nt / 1 ng* conditions, however, cells were transfected with the expression plasmids for an inactive polymerase (m) instead of a wt polymerase. The cells were lysed, the RNA extracted and rt PCR was performed on the extracted RNA as per Panel D.

Therefore, it was tested whether limiting NP (and potentially thus the generation of NP-deficient RNPs) might cause the accumulation of internal deletion RNAs. To test this, cells were transfected with the expression plasmids for a vRNA (either a 76 nt, 246 nt or 583 nt long segment 5 derived RNA), 3P and NP. The amount of NP expression plasmid within the transfections, however, was varied between 250 ng (high), 15 ng (medium) and 1 ng (low), whilst the vRNA and 3P expression plasmids were maintained at 250 ng. Below, for simplicity, only the vRNA length and amount of NP expression plasmid will be stated (e.g. *76 nt / 250 ng* NP).

It was hypothesised that, under conditions of limiting NP, the 246 nt and 583 nt long vRNAs would be more poorly replicated. Thus, in order to determine in which conditions NP might be limiting, the level of vRNA in each of the conditions at 24 h post transfection was measured by primer extension. It was found that comparable levels of vRNA were generated in cells transfected with *76 nt / 250 ng NP*, *76 nt / 15 ng NP* and *76 nt / 1 ng NP* (Fig. 26 C). This is likely due to the fact that 76 nt RNAs do not require NP for replication. Furthermore, comparable levels of vRNA were generated in cells transfected with *246 nt /250 ng NP* and *246 nt / 15 ng NP*, however, lower levels were generated in cells transfected with *246 nt / 1 ng NP* (Fig. 26 C). This could be explained by the levels of NP being limiting only in the *246 nt / 1 ng NP* condition. Finally, lower levels of vRNA were generated in cells transfected with *583 nt / 250 ng NP*, *583 nt / 15 ng NP* and *583 nt / 1 ng NP* (Fig. 26 C). This implies that, with the 583 nt long RNA, which requires more NP per vRNA molecule for complete encapsidation, NP might be limiting even in cells transfected with *583 nt / 250 ng NP* (Fig. 26 C).

In order to determine whether internal deletion RNAs were present in the conditions which had been hypothesised to contain 'limiting NP', the RNAs from the 'high NP' and 'low NP' conditions were used in a reverse transcription (rt) PCR (rt PCR) reaction using a primer

against the terminal 3′ sequence of the segment 5 vRNA. The rt reaction was performed in the presence (+ rt) or absence (- rt) of the reverse transcriptase. A PCR reaction was then performed on the resultant cDNAs using primers against terminal 3′ and 5′ sequence of the segment 5 vRNA. These primers would, therefore, amplify any vRNA species which contains both of the vRNA termini (full length and internal deletion RNAs), but, would not amplify those which contain just the 5′ or 3′ terminus (stalled vRNAs or vRNA degradation products). PCRs performed using the expression plasmids for the 76 nt and 246 nt long vRNA were used to generate size markers. The PCR products were run on an agarose gel. The 76 nt and 246 nt long 'markers' show the size of the plasmid expressed vRNAs (Fig. 26 D). In the minus rt samples, no bands were seen, however, in all of the plus rt samples, full length vRNAs of the relevant sizes were found (Fig. 26 D). Interestingly, however, a second lower molecular weight band was found in the *246 nt / 1 ng NP*, *583 nt / 250 ng NP* and *583 nt / 1ng NP* conditions (Fig. 26 D); all of which had been hypothesised to have 'limiting NP' (Fig. 26 C). This low molecular weight band, on the other hand, was not present in the *246 nt / 250 ng NP* sample (Fig. 26 D), which had been hypothesised not to have limiting levels of NP (Fig. 26 C). Overall, this supports the hypothesis that under limiting NP conditions, internal deletion RNAs accumulate in cells.

It was next tested whether the activity of 3P was required for the generation of the internal deletion RNAs seen in Figure 26 D. Thus, the transfections from Figure 26 C were repeated, however, a set of conditions was added where a replication inactive 3P (m) was expressed in cells instead of the wt 3P. This mutant 3P was hypothesised to be able to shelter the vRNA generated from the vRNA expression plasmid, however, not to be able to generate product vRNAs. With active 3P, a low molecular weight band was once again seen when cells were transfected with *246 nt / 1 ng NP* (Fig. 26 E). However, only full length RNAs were present when inactive 3P was expressed under conditions of low NP. This demonstrates that the

vRNA expression plasmid is not the direct source of the short RNAs, and thus that the internal deletion RNAs are indeed generated by 3P.

Overall, the data within the section supports the hypothesis that conditions of limiting NP favour the generation (or accumulation) of internal deletion RNAs.

5.2.8 Disruption of the Ratio of 3P and NP Leads to the Accumulation of Intermediate Length RNA and Interferon Induction in Influenza Virus Infected Cells

It was next investigated as to whether limiting NP can also lead to the generation of internal deletion RNAs during influenza A virus infections. To test this, two different techniques were used to attempt to disrupt the ratio of 3P to NP in an infected cell. The first consisted of transfecting cells with either an NP-specific RNA (siNP) or a control siRNA (siCo), and infecting the cells at 24 h post transfection. The second consisted of transfecting cells with expression plasmid for the subunits of 3P, and infecting the cells at 24 h post transfection. The cells were lysed at 12 h post infection, total RNA was extracted and rt PCR was completed using primers specific to the 3' and 5' termini of either the segment 1 (PB2) or segment 3 (PA) vRNA. PCRs amplification of the coding region from the expression plasmids for the 76 nt long vRNA, 246 nt long vRNA, segment 1 vRNA and segment 3 vRNA were used to generate markers. It was found that no PCR products were generated using the RNA from uninfected cells, or if a reverse transcriptase was not added to the rt reaction (Fig. 27 A). On the other hand, full length vRNA products were generated from RNAs extracted from infected cells when an rt reaction was performed. The infection of siNP-transfected cells did not lead to the clear generation of short products (Fig. 27 A). It is important to note, however, that it was not clear as to whether the siRNA caused an efficient

depletion of NP in infected cells. On the other hand, the pre-expression of 3P (whether in the presence or absence of siNP) before infection led to the generation of a product of comparable size to the 76 nt marker. This indicates that high levels of 3P in an infected cell promote the formation of internal deletion RNAs.

As stated above, it was hypothesised that the overexpression of 3P might promote the formation of short RNAs by disrupting the ratio of 3P to NP. In order to test this hypothesis, cells were transfected with the expression plasmids for either 3P, or 3P and NP. At 24 h post transfection, cells were infected and, at 12 h post infection, the cells were lysed, total RNA extracted and analysed by rt PCR, as above. It was found that the infection of cells expressing 3P led to the formation of the internal deletion RNA species, however, the infection of cells expressing 3P and NP did not (Fig. 27 B). Once again the internal deletion RNA species were of a comparable size to the 76 nt marker. This indicates that the disruption of the ratio of 3P and NP in an infected cell leads to the generation of detectable levels of internal deletion RNAs which are of comparable size to intermediate length RNA, which were shown above to trigger the expression of IFN.

Given that the disruption of the ratio of 3P to NP in infected cells appeared to promote the accumulation of RNAs of comparable size to intermediate length RNAs, it was hypothesised that it might also lead to the induction of IFN expression. To determine whether this was the case, the transfections from above were repeated including the IFN-luc and β-gal reporter plasmids. The cells were infected 24 h post transfection and then lysed 12 h post infection, before the reporter activities were measured. It was found that the transfection of cells with the expression plasmids for 3P or 3P and NP only lead to a modest induction of IFN expression above background (Fig. 27 C). The infection of cells transfected with control plasmid (3A) also did not lead to the induction of IFN above background levels. However,

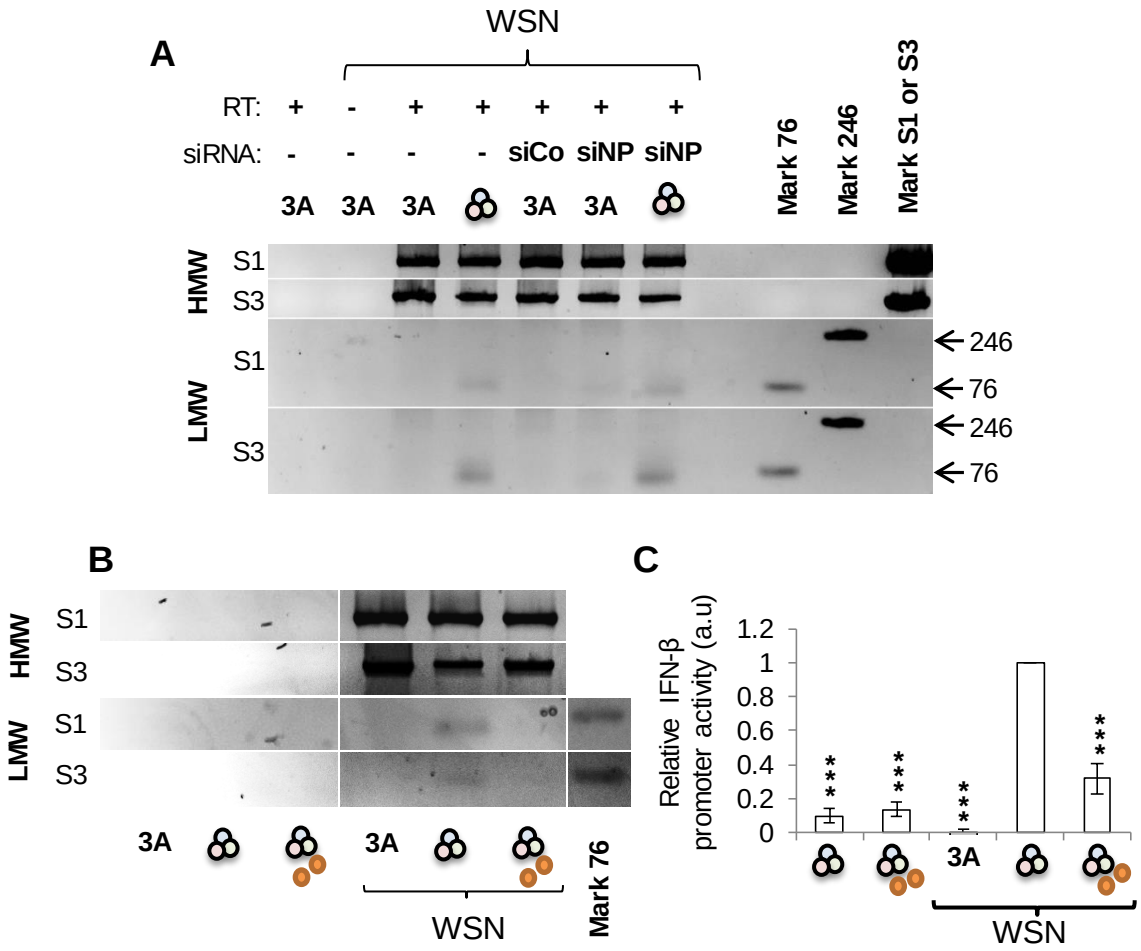


Figure 27 – Influenza viruses generate intermediate length RNAs – (A) HEK 293T cells were transfected with different combinations of empty vector (3A), expression plasmids for 3P (denoted by triplet of circles), an siRNA against NP (siNP) or a control siRNA (siCo). 24 h post transfection, cells were infected with A/WSN/33 (MOI=5). 12 h post infection, cells were lysed and the total RNA extracted. Reverse transcription was performed using a primer against either the 3'-termini of either the segment 1 (S1) or segment 3 (S3) vRNA. A PCR reaction was then performed using primers against the 3'- and 5'-termini of either the segment 1 or segment 3 vRNA. The cDNA products were analysed by electrophoresis. The observed high molecular weight (HMW) and low molecular weight (LMW) products are shown. Markers of 76 nt, 246 nt, or of the length of segment 1 or 3 vRNAs were used. n=1 (B) HEK 293T cells were transfected with different combinations of 3A, an expression plasmid for NP (denoted by an orange circle) or expression plasmids for 3P (denoted by a triplet of circles). 24 h post transfections, cells were infected, lysed and analysed as in Panel A. Representative image of three independent repeats. (C) Cells were transfected and infected as in panel C, however, the IFN-luciferase (IFN-luc) and β-galactosidase (β-gal) were also included. 12 h post infection, cells were lysed and the reporter activities measured. The background was subtracted using the mock infected, control plasmid (3A) condition. The β-gal normalised luc signal (relative IFN-β promoter activity) is shown. a.u. = arbitrary units. Bars represent standard deviations of average data from three independent experiments. Asterisks indicate a significant difference between the marked condition and the WSN-infected, 3P transfected (3P/WSN) condition (one sample t-test; ***, $p < 0.001$).

.

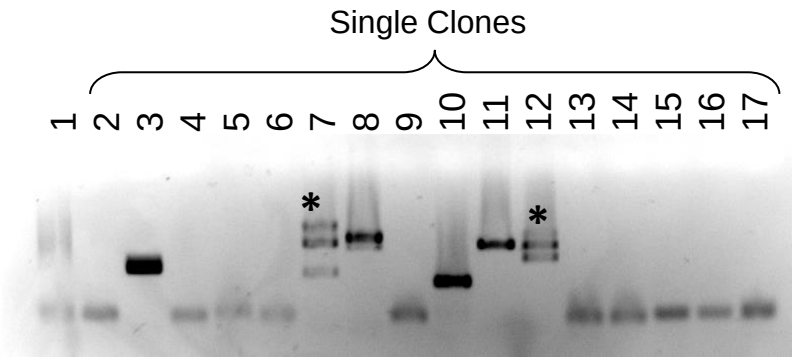
the infection of cells transfected with 3P expression plasmids led to a robust induction of IFN expression above background. Interestingly, this induction of IFN was significantly reduced if cells were co-transfected with the expression plasmid for NP (Fig. 27 C).

Overall, the data above support the conclusions that conditions which disrupt the ratio of 3P and NP and promote the accumulation of internal deletion RNAs of comparable size to intermediate length RNAs in infected cells, also trigger the induction of IFN expression. Interestingly, this result could potentially explain the observation from Figure 19 that infection of cells transfected with 3P expression plasmid, but not 3P and NP expression plasmids, led to the induction of IFN expression. Furthermore, given that *Laske et al.* have hypothesised that the presence of DI RNAs within infected cells leads to the levels of free NP becoming limiting, it is possible that DI viruses trigger the induction of IFN expression by disrupting the ratio of 3P to NP within infected cells [239].

5.2.9 Influenza A Virus Generates A Range of Intermediate Length RNA Species

In **section 5.2.8**, it was found that the disruption of the ratio of 3P to NP in infected cells promotes the accumulation of internal deletion RNAs of comparable size to intermediate length RNAs. In order to investigate whether the disruption of the ratio of 3P to NP within influenza A virus infected cells indeed leads to the generation of intermediate length (56-76 nt long) internal deletions RNAs (which were found in Figure 20 to induce IFN strongly), experiments were performed to characterise the internal deletion RNA species observed in **section 5.2.8**. Cells expressing exogenous 3P were infected with WSN and the presence of internal deletion RNAs was investigated by rt PCR using primers specific to the 3' and 5' termini of the segment 1 (PB2) vRNA. As before, the generation of cDNA products of comparable size to the intermediate length RNAs were observed (Fig. 28 A, Lane 1). The

A



B

Length (nt)	51	52	55	58	60	64	293	479
vRNA 5' end length (nt)	29	26	30	35	35	25	177	279

Figure 28 – Influenza viruses generate intermediate length RNAs between the lengths of 56 and 101 nt – (A) HEK 293T cells were transfected with expression plasmids for 3P. 24 h post transfection, the cells were infected with A/WSN/33 (MOI = 5). Total RNA was extracted and reverse transcription was performed using primers against 3'-termini of the segment 1 (PB2) vRNA. A PCR reaction was then performed using primers against the 3'- and 5'-termini of the segment 1 vRNA. The cDNA products were analysed by electrophoresis (Lane 1). The cDNAs were ligated into the TOPO-TA cloning plasmid, the resultant plasmids transfected into DH5α cells and the cells plated on ampicillin containing plates. Single clones were selected and used in PCR reactions performed using primers against the 3'- and 5'-termini of the segment 1 vRNA. The resultant DNA products were analysed by electrophoresis (lanes 2 -17). Lanes which contain multiple products are marked with an asterisks. (B) A number of clones which produced only a single PCR product in Panel A were analysed by sanger sequencing. The resultant sequences were aligned to the full length segment 1 vRNA sequence and the length of the internal deletion products was calculated. Furthermore, the position at which the junction occurs was calculated.

resultant cDNA products were ligated into the TOPO® TA vector, transformed into DH5α chemically-competent cells and single colonies were selected. PCR was performed on the lysates of the selected clones using primers specific to the 3' and 5' termini of the segment 1 (PB2) vRNA. The selected clones yielded a range of different lengths cDNA products (Fig. 28 A, Lanes 2 – 17). The plasmid DNA from a selection of the clones which yielded only a single cDNA product (colonies which yielded multiple products are marked with an asterisk) was submitted for sequencing. It was found that the cDNA products were derived from internal deletion RNAs derived from the segment 1 genomic RNA. These internal deletion products were of lengths between 51 and 479 nt (Fig. 28 B). As previously observed for DI-RNAs, the majority of the cDNAs were made up of comparable lengths of the 3' and 5' sequence of the segment 1 vRNA (Fig 28. B). Interestingly, a number of the detected products were between the lengths of 56 and 76 nt (Fig. 28 B). Overall, this data demonstrates that influenza A virus generates internal deletion RNAs of lengths between 56 and 76 nt, which have been shown above to be strong inducers of IFN expression.

5.2.10 3P of the 1918 Virus Strain Generates High Levels of Intermediate Length RNAs

The activation of RIG-I does not only lead to the induction of the expression of IFN. It has also been found to directly lead to the activation of the inflammasome and the generation of pro-inflammatory cytokines, both in an IFN-dependent and IFN-independent manner [240, 241]. It was thus likely that the generation of intermediate length vRNAs would lead to the generation of pro-inflammatory cytokines. Importantly, the high pathogenicity of the 1918 pandemic influenza virus, HPAIV H5N1 viruses and HPAIV H7N9 viruses in humans is associated with the release of high levels of pro-inflammatory cytokines; termed a cytokine storm. It was thus hypothesised that, in infected human cells, the 1918-virus and HPAIV H5N1 and H7N9 viruses might generate high levels of intermediate length vRNAs.

In order to test whether 3P of the 1918 pandemic influenza virus strain replicates high levels of intermediate length vRNAs, RNP reconstitutions were performed as above using the 246 nt vRNA expression plasmid with varying levels of NP (A/WSN/33) expression plasmid, however, using either the expression plasmids for the WSN or 1918 3P subunits. 24 h post transfection, the cells were lysed, the total RNA extracted and rt PCR was performed using primers against the 3' and 5'-termini of the segment 5 vRNA. The resultant cDNA products were analysed by electrophoresis as above. As before, WSN 3P only generated detectable levels of intermediate length RNAs at low levels of NP expression plasmid. On the other hand, the 1918 3P generated intermediate length RNAs at higher levels than the WSN 3P in all tested conditions (Fig. 29 – performed by Arend Te Velthuis).

Overall, this supports the hypothesis that the 3P of the 1918 pandemic influenza virus strain generates higher levels of intermediate length RNAs than non-avian derived 3Ps. This could potentially be a cause of the strong inflammatory response induced by the 1918 pandemic virus strain.

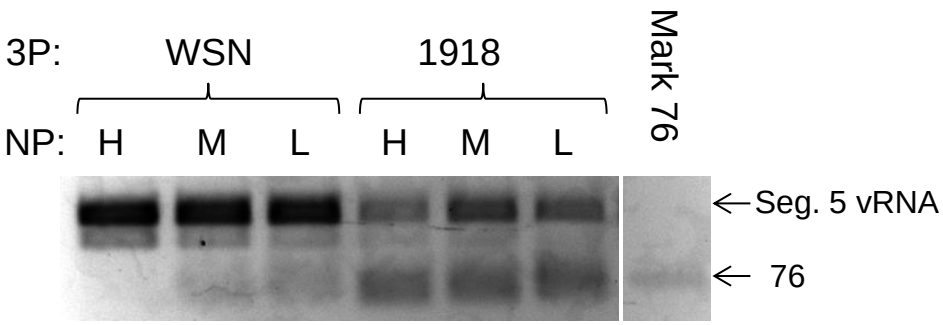


Figure 29 – The polymerase of the 1918 pandemic virus strain generates high levels of intermediate length RNAs – (A) HEK 293T cells were transfected with expression plasmids (250ng) for the 246 nt long vRNA and the subunits of 3P (either of the A/WSN/33 or 1918 pandemic virus strain). The cells were also transfected with either 250 ng (H), 15 ng (M) or 1 ng (L) of expression plasmid for NP (of the A/WSN/33 strain). 24 h post transfection, cells were lysed and total RNA was extracted. Reverse transcription was performed using primers against 3'-termini of the segment 5 (NP) vRNA. A PCR reaction was then performed using primers against the 3'- and 5'-termini of the segment 5 vRNA. The cDNA products were analysed by electrophoresis. A 76 nt long marker was included. Analysis performed in collaboration with Arend Te Velthuis.

5.3 Part II – Final Discussion

5.3.1 Limiting NP Promotes the Accumulation of Internal Deletion RNAs and Leads to the Induction of IFN Expression

Influenza viruses are able to induce IFN expression in infected cells. The production of IFN leads to the expression of a number of ISGsd, many of which act to limit the replication and spread of influenza virus. The manner by which influenza viruses trigger the induction of IFN expression is poorly understood. Previous work has demonstrated that the PRR RIG-I is key in the induction of IFN expression in response to influenza viruses [98]. Furthermore, the 5'-ppp of the influenza viral RNAs and the viral RNA termini have been shown to be key for IFN induction [114]. Given that RIG-I principally recognises regions of double stranded RNA that contain a 5'-ppp [100, 242], it is likely that during infection the 5'-ppp-containing terminus of the influenza virus viral RNAs forms into a double stranded structure. However, it is, unclear as to i) which specific double stranded RNA structure is involved in IFN induction and ii) how this structure is formed during infection.

The induction of IFN in influenza virus infected cells has been shown to be heterogeneous, with only a small proportion of infected cells robustly expressing IFN [143]. This highlights the possibility that IFN induction is associated with erroneous infections. In support of this, DI viruses, which contain aberrantly replicated genomic RNAs, promote the induction of IFN expression [138, 145, 153]. The aberrant RNAs contained by DI viruses have internal deletions and are larger than 200 nt long [147]. Interestingly, passaging of influenza viruses through IFN deficient cell lines, which likely allows the accumulation of IFN-inducing species, leads to the accumulation of multiple DI RNA species [146]. Together this indicates either that i) DI RNAs are themselves a potent IFN inducing RNA or ii) the presence of DI RNAs within infected cells promotes the generation of a potent IFN inducing RNA.

The idea that DI RNAs are potent IFN inducing RNAs is supported by work which showed that, in influenza virus infected cells, RIG-I binds to high levels of internal deletion RNAs of comparable length to those found in influenza DI viruses (greater than 200 nt long) [18, 147, 156]. However, it is not clear why DI RNAs would be more potent IFN inducers than full length genomic viral RNAs. Given the 'junction sites' in DI RNAs have been found to be normally at least 100 nt downstream and upstream of the 3′ and 5′ viral RNA termini, respectively, DI RNAs and full length viral RNAs have identical viral RNA termini [147]. Therefore, as RIG-I likely recognises the 5'ppp termini of viral RNAs [97, 100, 114], there is no apparent reason as to why DI RNAs would be stronger RIG-I ligands than full length viral RNAs.

As stated above, it is also possible that DI viruses lead to the induction of IFN expression by disrupting the normal life cycle of a co-infecting wild type 'helper virus' and thus leading to the formation of a strongly IFN inducing viral RNA species. In fact, it is the ability of DI viruses to disrupt the replication of a wild type 'helper virus' which gives 'defective ***interfering***' viruses their name [18]. Interestingly, *Laske et al.* have hypothesised that the presence of DI RNAs within an infected cell leads the levels of free NP to become limiting before that of free 3P [239]. This led to the hypothesis that DI RNAs might act to disrupt the levels of viral proteins in infected cells and thus lead to the generation of an IFN inducing species.

In Chapter 5, it was found that viral RNAs species with internal deletions accumulate to detectable levels under conditions where free NP likely becomes limiting before free 3P. This was found to occur both in RNP reconstitution assays and in virus infected cells. These observations are substantiated by work which showed that when an influenza virus infection is reconstituted in cells by exogenously expressing all of the viral genomic RNAs, except that which encodes NP, a number of short viral RNA species are observed [71].This led to the hypothesis that limiting NP promotes the accumulation of detectable levels of internal

deletion RNAs. The observation that limiting NP promotes the accumulation of internal deletion RNAs could be explained by two, non-mutually exclusive, models: i) Under conditions of limiting NP, but non-limiting 3P, viral RNAs are bound by 3P and only by a limited number of NP molecules, resulting in a structure which is denoted here as an 'NP-deficient RNP'. Given that NP is important in the process of elongation during viral RNA synthesis, NP-deficient RNPs are only able to replicate viral RNA that is bound by NP. If the NP on NP-deficient RNPs is preferentially localised near the RNA termini and not the central region of the RNA, replication of this RNA might lead to the generation of internal deletion RNAs (as illustrated in Fig. 26). ii) Given that shorter viral RNAs require less NP than full length viral RNAs to be encapsidated and efficiently replicated, internal deletion RNAs are likely more efficiently replicated than full length viral RNAs under conditions of limiting NP. Furthermore, the replication of RNAs shorter than 100 nt long, which do not require NP for replication, would likely be unaffected by conditions of limiting NP. This would give a replicative advantage to internal deletion RNAs relative to full length viral RNAs. If the assumption is made that internal deletion RNAs are generated stochastically at low levels, they would only accumulate to detectable levels under conditions of limiting NP, where they have a replicative advantage over the highly abundant full length viral RNAs (Fig. 30).

Interestingly, it was also found in Chapter 5 that a number of the formed internal deletion RNAs were shorter than the DI RNAs observed in DI viruses. Specifically, internal deletion RNAs were detected of length of approximately 50 – 65 nt long. Interestingly, IFN was found to be expressed in conditions under which these internal deletion RNAs formed. Indeed, it was found that the replication of internal deletion viral RNAs of lengths between 56 and 76 nt by 3P leads to the generation of a species which was a strong IFN inducer. This occurred independently of the internal sequence of the 56 – 76nt long RNA and was mediated by the RIG-I receptor. This leads to the conclusion that the generation of internal deletion viral

RNAs of approximately 56 – 76 nt in length (termed in this thesis as 'intermediate' length viral RNAs) within infected cells causes the activation of RIG-I and the induction of IFN expression. Therefore, overall, it appears that conditions of limiting NP in influenza virus infected cells favour the accumulation of intermediate length viral RNAs, which may be involved in the induction of IFN expression (Fig. 30).

In summary, the data above supports the hypothesis that, during influenza virus infection, the depletion of NP in infected cells might lead to the generation of internal deletion RNAs of different sizes (Fig. 30 – 1). Given that the internal deletion RNAs likely require less NP for encapsidation than full length viral RNAs, these internal deletion RNAs might replicate more efficiently than full length viral RNAs under the conditions of limiting NP (Fig 30 -2). Thus, conditions of limiting NP might promote the accumulation of high levels of internal deletion RNAs. A number of these internal deletion RNAs would likely have defective interfering activity and thus might act to further deplete the levels of free NP (Fig. 30 – 3) [239]. Interestingly, this would further promote the formation and accumulation of internal deletion RNAs (Fig. 30 – 2 and 4). These conditions would also favour the replication of intermediate length RNAs, which do not require NP for replication. The accumulation of high levels of intermediate length RNA might then lead to the induction of IFN expression (Fig. 30 – 5). Interestingly, the intermediate length vRNAs would likely lack the packaging sequences required for their incorporation into virions, and so only longer internal deletion RNAs would likely be packaged into virions (Fig. 30 – 6). The fact that intermediate length RNAs can likely not undergo cell to cell transmission may explain i) why they have not been observed in DI viruses and ii) why high levels of intermediate length viral RNAs have not been observed in influenza virus infected cells. However, as hypothesised by *Laske et al.*, the infection of a cell with a DI virus, along with a wild type 'helper' virus to provide the full complement of viral proteins, would introduce a DI RNA into the infected cell and thus might

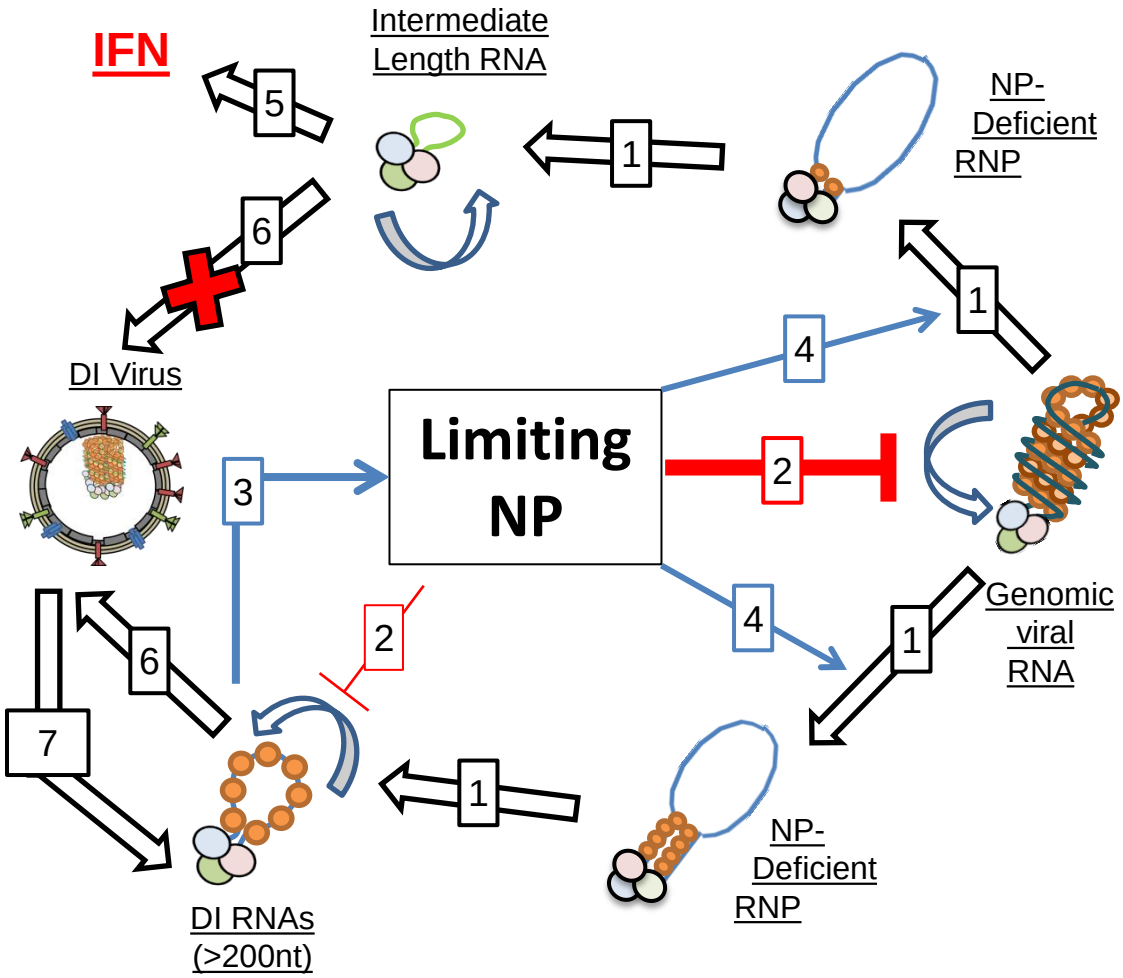


Figure 30 – Model of how limiting NP might promote the accumulation of intermediate length RNPs – (1) During virus replication, if NP becomes limiting RNPs can be formed which are deficient in NP. These RNPs are termed here as 'NP deficient RNPs'. Due to their deficiency in NP, during viral RNA replication, these RNPs generate internal deletion RNA products. These RNAs can be of a range of sizes, and can be bound by 3P and/or NP. (2) The conditions of limiting NP inhibit the replication of longer viral RNA species more than that of short viral species. (3) A number of the internal deletion RNAs will have comparable size to DI RNAs. The presence of DI RNAs within the cell, according to *Laske et al.*, leads to further depletion of free NP. (4) The increasingly limiting NP would favour the production of new internal deletion RNAs and the replication of internal deletion RNAs in a positive feedback manner. (5) Intermediate length RNPs, which are generated and accumulate under the conditions of limiting NP lead to the induction of IFN via an unknown mechanism. (6) The internal deletion RNAs which contain the requisite packaging signals are incorporated into viruses. Due to their size, however, intermediate length RNAs lack these signals and thus are not incorporated into forming virions. (7) However, when a cell is infected with a DI-RNA containing virus (DI virus) along with a wild type helper virus, the DI-RNA leads to the levels of NP to become limiting at earlier time points than in wild type virus-only infected cells. This promotes the formation of intermediate length RNAs *de novo* in cells, leading to the induction of IFN expression.

cause the levels of NP to once again become limiting, promoting the *de novo* formation of intermediate length RNAs (Fig. 30 - 7).

To further investigate this model, it would be of interest to determine whether higher levels of intermediate lengths RNAs are generated in infected cells which are co-infected with a DI virus than in those cells which are not. Furthermore, in order to determine whether DI viruses induce IFN expression by leading to limiting levels of NP, it would be interesting to determine whether the over-expression of NP in cells inhibits the ability of DI viruses to induce IFN expression.

5.3.2 Mechanisms Which May Lead to Limiting NP or NP-Deficient RNPs

As stated above, it is possible that presence of DI RNAs within influenza virus infected cells may lead to the depletion of free-NP and the accumulation of DI RNAs. Interestingly, a number of viral proteins have also been found to be involved in the generation of DI RNAs. Specifically, certain mutations in PA and NEP have been correlated with the accumulation of DI RNAs [151, 152]. One of the documented roles of NEP is to regulate the switch between transcription and replication, and thus likely control the level of viral mRNAs which are generated before viral genomic replication initiates [71, 73]. Furthermore, the mutation in PA has been found to decrease the transcriptional activity of 3P and increase the replicative activity of 3P [243]. It is thus possible that both of these mutations disrupt the equilibrium between viral transcription, which is required for viral protein synthesis, and viral genomic RNA replication. This might limit the levels to which the viral proteins NP and 3P accumulate before viral genomic RNA replication initiates. Under these conditions, given that more molecules of NP than 3P are required for the encapsidation of viral RNAs, the levels of free

NP might become quickly limiting. This might thus promote the accumulation of internal deletion RNAs in the manner detailed above.

Furthermore, certain mutations within the insertion groove of NP have been documented which lead to dominant negative NP mutants, which are able to bind into the growing chain of NP on a viral RNA, however, act as chain terminators [37]. Interestingly, preliminary data has demonstrated that the presence of such an NP mutant in RNP reconstitutions promotes the induction of IFN (data not shown). It is possible that this occurs due to this NP mutant promoting the formation of NP-deficient RNPs.

Finally, the ISGs Mx1 and TRIM22 have been shown to disrupt the 3P:NP interaction and target NP for degradation, respectively. It is thus possible that these ISGs may also promote the accumulation of NP-deficient RNPs [119, 122] and it would be of interest to test their effect on IFN induction and the formation of intermediate length RNAs.

Overall, therefore, it is possible that conditions of limiting NP and/or the formation of NP-deficient RNPs can be promoted in influenza virus infected cells by mutations which disrupt the normal process of viral protein expression or by mutations or factors which disrupt the correct formation of RNPs.

5.3.3 Potential Mechanisms by Which Intermediate Length RNAs Induce the Expression of IFN

In this thesis, due to time constraints, it has not been possible to unequivocally determine the mechanisms by intermediate length RNAs induce IFN expression. Interestingly, however, in the results described in section **5.2.2**, the viral RNA species which was found to be the strongest inducer of IFN (76 nt long RNAs) is of comparable length to the longest viral RNA species that can be efficiently replicated in the absence of NP[37]. Thus, it is possible that the

replication of non-NP bound regions of vRNA by 3P might lead to the formation of a strong IFN inducing species. Given that NP appears to be important in the process of elongation, it is likely that the replication of non-NP bound viral RNA leads to the generation of high levels of abortive replication products (Fig. 31). It is possible, therefore, that these high levels of abortive viral RNA products are able to anneal to the non-NP bound regions of other viral RNAs (abortive cRNAs to vRNAs and abortive vRNAs to cRNAs) to form double stranded RNAs (Fig. 31). Unlike the influenza virus panhandle structure, which is a poor IFN inducer due to the presence of mismatches [142], the resultant cRNA/vRNA hybrids would be a perfectly double stranded viral RNA species. This species would therefore likely be a potent activator of RIG-I and thus lead to high levels of IFN expression. It is thus possible that the reason why *Baum et al*. found that high levels of internal deletion RNAs bind to RIG-I is due to the fact that the presence of high levels of internal deletion RNAs within a cell promotes the formation of high levels of abortive RNA, which then anneal to the viral RNAs within the cell and thus allow their detection by RIG-I [156].

In order to further investigate this model, it would be of great interest to firstly generate highly pure preparations of 76 nt long RNAs to determine whether these RNAs are intrinsically potent inducers of IFN expression, or whether it is their replication which generates the potent inducer. Furthermore, it would be of interest to determine whether i) RNPs containing 76 nt long RNAs generate high levels of abortive viral RNAs and ii) whether high levels of short, abortive viral RNAs species co-purify with RIG-I from DI virus infected cells. This could be tested by Northern blot or radioactive labelling of the RNAs from RNP reconstitutions or from RIG-I affinity purifications.

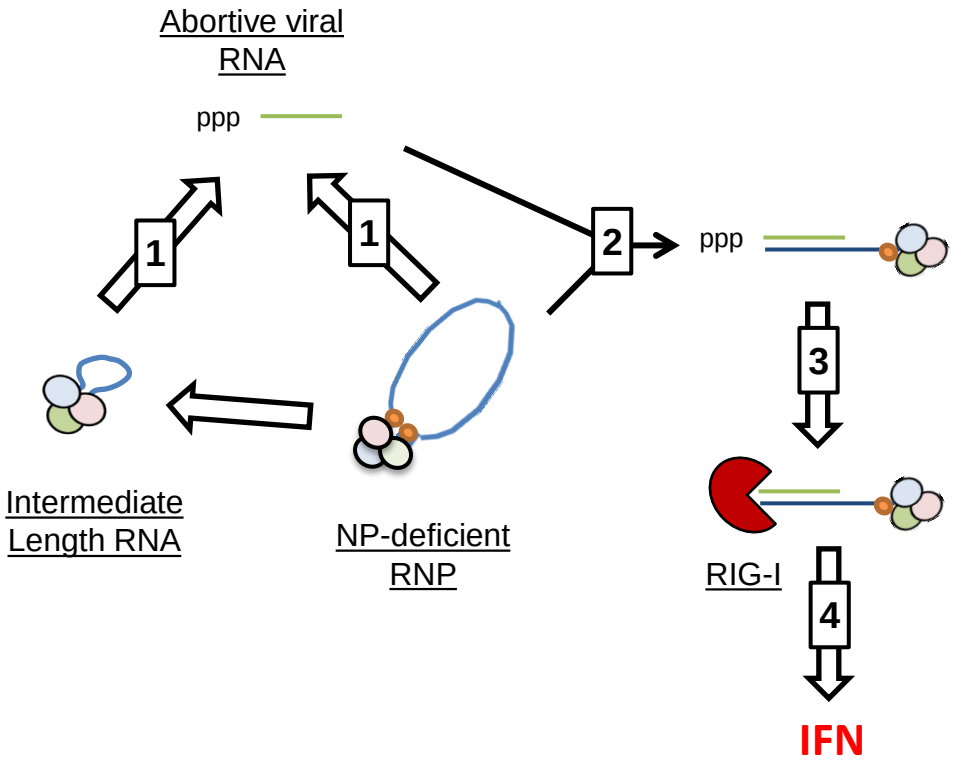


Figure 31 – Model of how intermediate length RNAs might lead to the induction of IFN expression – (1) Due to the absence of NP, which is important in the process of elongation during RNA replication, NP-deficient RNPs (whether they contain intermediate length or long viral RNAs) generate high levels of abortive viral RNA replication products. (2) These replication products can associate to the complementary RNA in NP-deficient RNPs due to the absence of NP. (3) Unlike the influenza virus panhandle structure, which contains RNA mismatches, the resultant RNA:RNA hybrid is a perfect double stranded RNA. Due to this, this RNA is efficiently recognised by RIG-I and leads to the activation of RIG-I. (4) The activation of RIG-I leads to the induction of IFN expression.

5.3.4 Avian Viruses and Intermediate Length RNAs

The activation of RIG-I does not only lead to the induction of IFN expression, but also promotes the activity of the inflammasome, resulting in the production of pro-inflammatory cytokines [240, 241]. The infection of humans with the 1918 pandemic virus strain, HPAIV H5N1 viruses and HPAIV H7N9 viruses leads to an over activation of the innate immune response and a subsequent release of high levels of pro-inflammatory cytokines. This 'cytokine storm' leads to pulmonary oedema which causes death in a high proportion of cases. Interestingly, in this thesis, it was found that the replication of viral RNAs by the 3P of the 1918 pandemic virus strain leads to the accumulation of high levels of intermediate length RNAs. Furthermore, this accumulation occurred independently of the level of NP. This would imply that in cells containing the 3P of the 1918 virus, intermediate length viral RNAs might accumulate to high levels, leading to robust activation of RIG-I and the subsequent activation of the inflammasome, and thus the generation of pro-inflammatory cytokines. Therefore, this supports the conclusion that the 1918 pandemic virus may induce a cytokine storm due to generating high levels of immunostimulatory intermediate length viral RNAs.

In this thesis, due to time restraints, work was not completed to determine whether H5N1 and H7N9 viruses also generate high levels of intermediate length RNAs in human cells. It is of interest to note, however, that it has been found that in human cells, avian-adapted 3Ps, such as those of H5N1 and H7N9 viruses, replicate full length viral RNAs inefficiently [193], however, are able to efficiently replicate intermediate length vRNAs [244]. Therefore, it is possible that, in human cells, avian adapted 3Ps favour the accumulation of intermediate length viral RNAs (Fig. 32). Indeed, investigation of clinical samples from a H7N9 infected individual has shown that both avian-adapted and human-adapted 3P was present in the virus population [245]. Thus, in such infections, it is possible that the avian adapted 3P within the

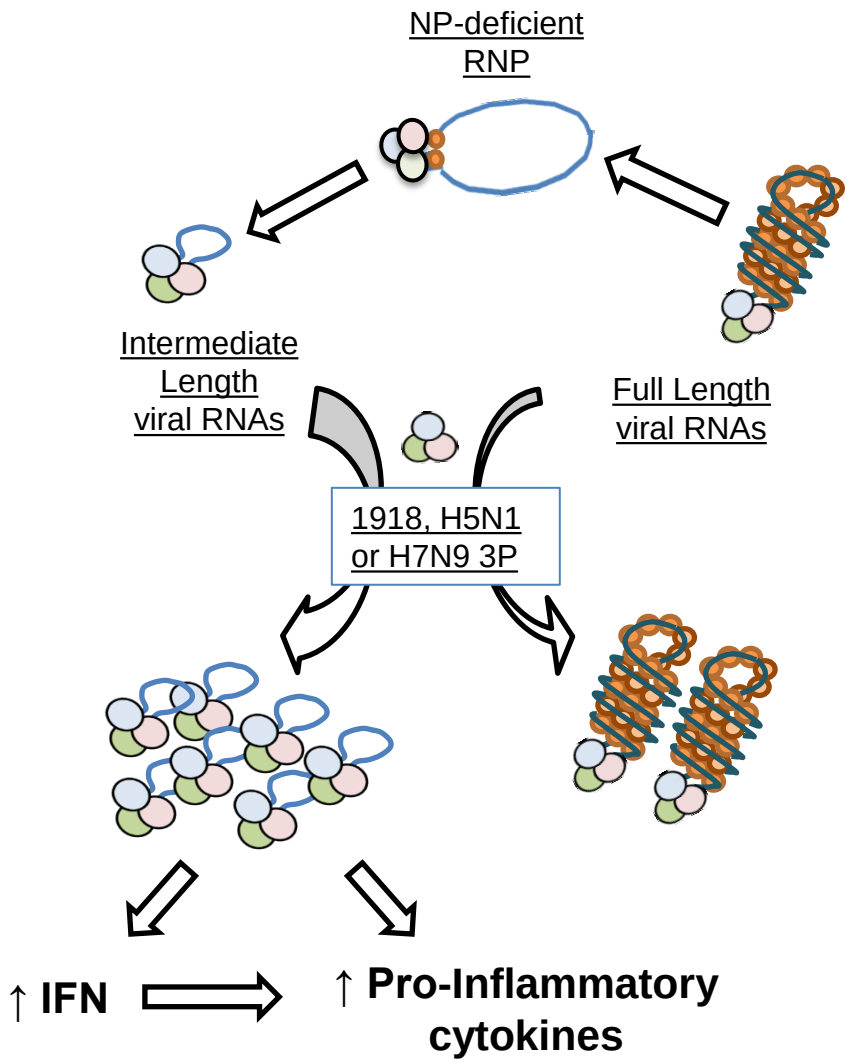


Figure 32 – Model of how avian adapted influenza virus polymerases might promote the accumulation of intermediate length RNAs – During the infection of a human cell with the 1918 pandemic influenza virus, an H5N1 influenza virus strains or an H7N9 influenza virus strain intermediate length RNAs are generated from full length viral genomic RNAs. The 3P of the 1918 virus strain in human cells has been found to replicate high levels of intermediate length RNAs. Furthermore, the avian-adapted 3Ps of H5N1 and H7N9 influenza viruses in human cells are poorly able to replicate full length viral RNAs but are able to efficiently replicate intermediate length RNAs [245]. Thus, the infection of human cells with the 1918 pandemic virus, HPAIV H5N1 viruses or H7N9 viruses leads to the accumulation of intermediate length vRNAs within cells. These intermediate length RNAs lead to the activation of RIG-I and induce the production of high levels of IFN. The activation of RIG-I also leads to the activation of the inflammasome, both directly and via the induction of IFN expression. The activation of the inflammasome leads to the high levels of pro-inflammatory cytokines termed the 'cytokine storm'.

infected cells promotes the accumulation of intermediate length vRNAs, leading to robust activation of RIG-I, subsequent activation of the inflammasome, and thus the generation of pro-inflammatory cytokines (Fig. 32). In support of this model, recent work found that RNPs containing an avian adapted 3P led to the activation of RIG-I more potently than RNPs containing a human adapted 3P [238]. Thus, as above, it would also be of interest to determine whether human infections with HPAIV H5N1 and H7N9 virus strains lead to the accumulation of high levels of intermediate length viral RNAs.

5.3.5 Intermediate Length RNA Generation as a Therapeutic Target

The generation of intermediate length viral RNAs might represent a process which could be targeted by therapeutics. Interestingly, the infection of mice with DI viruses has already been shown to give protection from a subsequent influenza virus infection in an IFN-dependent manner. Firstly, therefore, it is possible that during the infection of humans with seasonal H1N1 or H3N2 influenza viruses, the promotion of the formation of intermediate length viral RNAs might act to restrict the severity of the viral infection by promoting the induction of IFN and pro-inflammatory cytokine expression. Interestingly, a drug has been discovered (termed nucleozin) which promotes the aggregation of NP and thus could be potentially used to make NP limiting within infected cells [246]. It would thus be of interest to determine whether nucleozin promotes the formation of intermediate length viral RNAs and the induction of IFN expression. Secondly, it is possible that drugs which inhibit the formation of intermediate length viral RNAs might ameliorate the inflammatory pathology of HPAIV H5N1 and H7N9 viruses.

Chapter 6

Appendix

6.1 Reagent Tables

6.1.1 Oligonucleotide Table

	Description	Sequence
A	N9T-PB2-GFP	GAATAAAAGAACTAAGGACTCTAATGTCGCAGTCTCGCACTCGC
B	N9T-PB2-GFP	GCGAGTGCGAGACTGCGACATTAGAGTCCTTAGTTCTTTATTC
C	PCR APEX	ATATGCGGCCGCCATGGGAAAGTCTTACCCACT
D	PCR APEX	ATATTCTAGATCATTAGGCATCAGCAAA
E	siRNA TUFM 1	AGGUGGAGGCCCAGGUUUAUU
F	siRNA TUFM 2	GGUCAUGCAGAUUAUGUUAUU
G	siRNA TUFM 3	GCAUUGAGAUGUUCCACAAUU
H	siRNA GFP	UUUCGCGUAUGGUCUUCAAUU
I	siRNA RIG-I	CAGAAGAUCUUGAGGAUAAUU
J	siRNA NP	GGAUCUUAUUUCUUCGGAGTT
K	GC- Primer Ext	AGCAAAAGCAGGGTAGACTAGT
L	GC-2 Primer Ext	AAAGAAAAATACCCTT
M	5S rRNA Primer Ext	TCCCAGGCGGTCTCTCCCATCC
N	dT20	TTTTTTTTTTTTTTTTTTTT
O	MAVS RT-qPCR	CTGCCTCACAGCTAGTGACC
P	MAVS RT-qPCR	CCGGCGCTGGAGATTATTG
Q	β-ACTIN qPCR	CTCCTTAATGTCACGCACGAT
R	β-ACTIN qPCR	CATGTACGTTGCTATGCAGGC
S	Segment 5 RT PCR (Rt and PCR)	AGCAAAAGCAGGGTAGACTAGT
T	Segment 5 RT PCR (PCR only)	AGTAGAAACAAGGGTATTTTTCT
U	Segment 1 RT PCR (Rt and PCR)	AGCGAAAGCAGGTCAATTATAT
V	Segment 1 RT PCR (PCR only)	AGTAGAAACAAGGTCGTTTTTAAAC
W	Segment 3 RT PCR (Rt and PCR)	AGCGAAAGCAGGTACTGATTC
X	Segment 3 RT PCR (PCR only)	AGTAGAAACAAGGTACTTTTTTGGAC

6.1.2 Antibody Table

Antibody	Function
Polyclonal Rabbit-α-PB2[175]	IF, WB
Polyclonal Rabbit-α-NP (PR8)	IF, WB
Cy2-conjugated Donkey-α-Rabbit	IF
Polyclonal Goat- α-p32 (sc10258)	WB
Polyclonal Rabbit-α-VDAC (10527)	WB
Polyclonal Goat- α-Importin α1 (ab6039)	WB
Monoclonal Mouse-α-Flag (F3165)	WB

6.2 Supplementary Figures

Supplementary Figure 1 – Analysis the Levels of NP in N9-WSN and D9-WSN Infected SILAC-labelled Cells

	N9-WSN	D9-WSN
Repeat 1	6108	4018
Repeat 2	11202	12142
Repeat 3	1.52	15

Supplementary Figure 2 - Data Sets N9 Repeat 1, 2 and 3

	Relative protein level log2(Inf/Uninf)						Relative protein level log2(Inf/Uninf)				
Accession	R1	R2	R3	Average	P-value	Accession	R1	R2	R3	Average	P-value
O00154	0.98	2.16	0.84	0.407	5.5E-01	P36542	1.09	1.05	0.98	0.060	2.7E-01
O00299	0.96	0.59	0.53	-0.531	9.1E-02	P36543	0.89	0.67	0.77	-0.370	3.5E-02
O00571	0.98	1.02	1.32	0.146	3.8E-01	P36957	0.90	0.96	1.16	0.010	9.9E-01
O14548	1.49	1.22	1.04	0.322	1.1E-01	P38117	0.90	0.83	1.11	-0.076	5.2E-01
O14949	0.65	0.91	1.64	0.094	9.7E-01	P38646	0.91	0.94	1.10	-0.023	7.6E-01
O43169	0.89	1.18	1.47	0.239	3.8E-01	P40925	0.81	0.66	1.36	-0.086	6.4E-01
O43837	0.05	0.67	0.83	-0.951	2.6E-01	P40926	0.91	0.96	1.07	-0.026	6.8E-01
O75438	1.32	1.06	1.05	0.193	1.6E-01	P40939	0.85	1.17	1.05	0.032	8.9E-01
O94826	0.67	1.00	0.97	-0.183	3.3E-01	P45880	1.05	1.05	1.07	0.080	4.7E-04
O95573	1.01	1.44	0.43	-0.062	6.8E-01	P47985	1.34	0.95	1.07	0.162	3.6E-01
O95831	0.99	0.99	1.12	0.047	4.8E-01	P48047	1.60	1.16	1.02	0.335	1.9E-01
O95994	0.87	1.54	0.81	0.098	9.1E-01	P49327	0.88	0.86	0.75	-0.277	1.8E-02
O96000	1.02	1.27	1.10	0.175	1.4E-01	P49411	1.02	0.96	1.06	0.018	7.1E-01
O96008	0.90	1.08	1.10	0.041	7.2E-01	P51659	0.90	1.14	0.99	0.013	9.6E-01
P00167	1.00	1.10	1.04	0.063	2.0E-01	P52815	0.84	1.12	1.22	0.082	7.1E-01
P00338	1.14	0.72	0.86	-0.139	4.4E-01	P53396	1.10	0.66	0.82	-0.223	3.0E-01
P00367	1.11	1.08	0.97	0.073	3.2E-01	P55084	0.91	1.04	1.00	-0.029	6.2E-01
P00390	0.93	0.57	0.68	-0.464	7.9E-02	P55957	0.96	0.32	0.81	-0.520	2.5E-01
P00403	1.10	1.03	2.37	0.584	2.9E-01	P56134	1.10	1.03	1.13	0.120	4.6E-02
P00505	0.88	0.95	1.10	-0.035	6.8E-01	P56385	0.80	0.95	0.95	-0.153	1.4E-01
P02786	1.18	1.02	0.64	-0.083	6.5E-01	P56556	1.09	0.97	0.72	-0.110	5.0E-01
P04040	1.16	0.75	8.36	1.775	4.2E-01	P62258	0.99	0.76	0.20	-0.623	2.7E-01
P04179	0.91	1.14	0.86	-0.043	6.8E-01	P62263	1.04	0.95	0.84	-0.085	3.8E-01
P06576	1.04	1.07	1.06	0.075	4.6E-03	P62906	1.10	0.94	0.95	-0.006	9.0E-01
P07195	1.19	0.73	0.57	-0.267	3.4E-01	P63104	1.05	0.81	0.83	-0.152	2.5E-01
P07339	1.08	0.79	0.65	-0.249	2.5E-01	P63167	0.10	0.84	0.68	-0.883	2.3E-01
P07900	1.15	0.96	0.70	-0.096	5.8E-01	P63244	0.76	0.86	0.99	-0.201	1.2E-01
P07954	0.65	0.96	1.12	-0.137	5.0E-01	P99999	0.96	1.36	0.93	0.119	6.0E-01
P08107	1.20	0.83	0.71	-0.130	5.0E-01	Q00325	1.09	1.15	1.29	0.230	3.4E-02
P08238	1.13	1.04	0.85	0.008	9.9E-01	Q00610	0.94	0.89	0.80	-0.189	5.2E-02
P09211	1.04	0.70	0.48	-0.436	1.9E-01	Q04837	0.95	1.07	1.06	0.040	5.4E-01
P09622	0.86	1.04	1.10	-0.004	9.2E-01	Q06830	0.88	0.88	0.68	-0.295	7.1E-02
P09669	0.94	1.12	0.53	-0.211	4.4E-01	Q07021	0.94	1.13	1.23	0.135	3.4E-01
P09972	1.47	0.87	0.74	0.042	9.4E-01	Q13011	0.85	0.87	1.04	-0.120	2.5E-01
P10620	1.26	1.67	2.24	0.785	3.5E-02	Q14318	0.50	0.84	3.32	0.636	8.6E-01
P10809	1.01	1.03	1.07	0.052	1.2E-01	Q16555	1.01	0.67	0.34	-0.572	2.0E-01
P11177	0.51	1.12	0.96	-0.216	4.5E-01	Q16698	0.68	1.12	0.87	-0.163	4.0E-01
P11586	1.16	0.70	0.62	-0.272	3.0E-01	Q16719	1.09	0.86	0.78	-0.142	3.4E-01
P12236	0.98	1.78	1.12	0.371	2.8E-01	Q16762	0.62	1.41	0.60	-0.187	4.9E-01
P12931	0.84	0.97	1.20	0.006	9.5E-01	Q16891	1.09	1.11	1.03	0.109	2.5E-02
P13073	0.95	1.06	1.04	0.025	6.4E-01	Q56VL3	0.76	1.02	0.96	-0.129	3.3E-01
P13804	0.93	0.86	1.04	-0.082	3.2E-01	Q5RI15	0.58	1.91	3.32	0.954	4.5E-01
P14618	0.99	0.77	0.61	-0.340	1.5E-01	Q6NZI2	1.17	0.74	0.59	-0.263	3.3E-01
P15531	0.47	0.82	0.64	-0.641	4.3E-02	Q6YN16	1.16	1.58	0.91	0.284	3.4E-01
P15924	1.24	0.87	12.50	2.284	3.6E-01	Q7Z434	0.89	0.92	1.10	-0.048	6.0E-01
P15954	1.04	1.08	1.01	0.064	8.4E-02	Q96CW1	1.29	0.84	0.41	-0.237	4.7E-01

P19404	1.00	1.06	1.00	0.024	4.4E-01	**Q96EL3**	0.13	1.28	1.17	-0.220	5.1E-01
P21796	1.19	1.01	1.02	0.103	2.7E-01	**Q96HS1**	1.11	0.85	0.97	-0.035	7.2E-01
P21980	1.25	0.66	0.58	-0.270	3.6E-01	**Q96I99**	1.24	0.74	0.98	-0.018	8.3E-01
P22392	0.79	0.84	0.64	-0.397	2.7E-02	**Q96IX5**	0.87	1.01	0.80	-0.163	1.6E-01
P22695	1.00	0.98	1.01	-0.003	7.7E-01	**Q99623**	1.03	1.16	1.11	0.136	4.6E-02
P24752	0.90	0.97	1.02	-0.049	3.8E-01	**Q99714**	1.02	0.86	1.05	-0.033	6.9E-01
P25705	1.02	1.06	1.01	0.044	1.3E-01	**Q99798**	2.76	1.21	0.97	0.721	2.8E-01
P28331	0.88	1.15	1.18	0.096	5.8E-01	**Q99832**	1.30	1.09	1.13	0.231	4.4E-02
P30048	0.88	0.94	1.07	-0.054	5.2E-01	**Q9BZF9**	0.90	0.88	0.56	-0.364	1.5E-01
P30084	0.60	0.97	1.02	-0.212	3.7E-01	**Q9C002**	1.53	1.12	0.93	0.255	3.5E-01
P31040	1.16	0.99	1.02	0.081	3.4E-01	**Q9GZT3**	1.18	0.98	1.05	0.094	3.2E-01
P31930	0.99	1.03	1.03	0.029	2.4E-01	**Q9H0U4**	1.40	0.97	1.00	0.169	4.3E-01
P34897	0.93	1.22	1.21	0.161	3.2E-01	**Q9NS69**	0.80	1.14	1.95	0.373	5.1E-01
P34932	1.03	0.94	0.67	-0.180	3.4E-01	**Q9NVI7**	1.06	1.03	1.14	0.109	6.3E-02
P35232	1.01	1.13	1.07	0.094	1.1E-01	**Q9NX40**	0.89	1.08	0.95	-0.039	6.2E-01
P35613	1.08	0.83	0.40	-0.373	3.2E-01	**Q9NX63**	0.86	1.08	1.02	-0.023	7.8E-01
P35998	1.47	0.81	0.67	-0.023	7.7E-01	**Q9P0J0**	1.10	0.92	0.96	-0.007	8.9E-01
P36507	1.07	0.81	0.92	-0.102	4.0E-01	**Q9UI09**	0.83	1.00	0.97	-0.098	2.7E-01
Q9Y277	1.14	0.99	1.09	0.102	1.8E-01	**Q9UJZ1**	0.85	1.07	1.12	0.019	9.5E-01
Q9Y6N5	0.63	1.10	0.31	-0.556	2.4E-01						

Supplementary Figure 3 - Data Sets D9 Repeat 1, 2 and 3

	Relative protein level log2(Inf/Uninf)						Relative protein level log2(Inf/Uninf)				
Accession	R1	R2	R3	Average	P-value	Accession	R1	R2	R3	Average	P-value
O00154	0.80	0.90	1.06	-0.119	6.8E-01	P35998	0.89	1.06	0.81	-0.124	4.6E-01
O00299	0.87	0.44	0.60	-0.654	6.4E-01	P36507	1.15	0.63	1.06	-0.085	7.2E-01
O00571	0.71	0.77	1.54	0.011	7.5E-01	P36542	1.09	1.02	1.09	0.091	2.9E-01
O14548	1.18	1.23	1.00	0.185	1.0E-01	P36543	0.86	0.57	0.78	-0.440	8.7E-01
O14949	1.12	1.01	1.47	0.260	3.0E-01	P36957	0.94	0.96	1.08	-0.011	4.7E-01
O43169	0.79	1.02	1.29	0.049	5.4E-01	P38117	0.98	0.97	1.14	0.042	4.1E-01
O43837	0.04	1.59	1.08	-0.143	6.3E-01	P38646	0.92	0.98	1.21	0.054	4.5E-01
O75438	1.02	0.98	1.21	0.097	3.8E-01	P40925	0.81	0.56	1.78	0.070	8.1E-01
O75964	1.57	1.30	1.20	0.439	3.7E-02	P40926	1.06	1.08	1.14	0.129	2.7E-01
O94826	0.60	1.01	0.97	-0.218	8.7E-01	P40939	1.01	1.03	1.06	0.048	3.4E-01
O95573	0.87	0.92	0.62	-0.317	7.3E-01	P45880	1.02	0.96	1.19	0.079	3.9E-01
O95831	0.89	1.02	1.08	-0.003	4.6E-01	P47985	1.31	0.90	1.46	0.291	3.0E-01
O95994	0.86	1.20	0.83	-0.054	4.0E-01	P48047	0.97	1.03	1.03	0.017	3.7E-01
O96000	1.05	1.23	1.09	0.164	1.9E-01	P49327	0.86	0.62	0.94	-0.309	9.8E-01
O96008	0.91	0.95	1.20	0.026	4.9E-01	P49411	0.97	1.42	1.12	0.227	1.9E-01
P00167	1.27	0.88	0.78	-0.036	3.3E-01	P51659	1.14	0.93	1.18	0.114	3.4E-01
P00338	1.07	0.46	0.96	-0.266	9.8E-01	P52815	0.81	0.79	1.10	-0.151	7.7E-01
P00367	1.20	1.18	1.16	0.237	1.4E-01	P53396	1.05	0.41	1.16	-0.197	9.8E-01
P00390	1.23	0.63	0.84	-0.152	7.4E-01	P55084	0.94	1.03	1.06	0.012	4.1E-01
P00403	1.25	0.97	1.49	0.307	2.7E-01	P55957	1.01	0.36	0.71	-0.531	7.1E-01
P00505	1.00	1.05	1.37	0.188	3.4E-01	P56385	1.13	1.16	1.12	0.188	1.8E-01
P02786	0.87	0.53	0.77	-0.470	8.2E-01	P56556	1.07	0.93	0.73	-0.135	3.8E-01
P04040	1.41	2.63	0.97	0.738	2.2E-02	P61604	1.03	1.01	1.10	0.067	3.4E-01
P04179	0.78	0.97	0.78	-0.246	7.3E-01	P62258	0.91	0.64	1.12	-0.167	8.4E-01
P06576	1.04	0.99	1.07	0.048	3.5E-01	P62263	0.89	1.16	0.92	-0.013	3.7E-01
P07195	1.07	0.47	0.59	-0.496	8.0E-01	P62906	1.25	0.80	0.86	-0.045	4.5E-01
P07339	0.52	1.71	0.62	-0.073	7.9E-01	P63104	0.92	0.56	0.83	-0.379	9.3E-01
P07900	1.00	0.65	0.91	-0.225	8.3E-01	P63167	0.08	0.65	0.72	-1.048	4.1E-01
P07954	0.51	1.07	1.02	-0.204	9.3E-01	P63244	0.83	1.97	0.78	0.255	2.7E-01
P08107	0.92	0.66	0.81	-0.325	9.4E-01	P99999	0.96	1.16	1.20	0.145	3.0E-01
P08238	0.97	0.67	0.96	-0.205	8.1E-01	Q00325	1.16	1.33	0.70	0.090	3.4E-02
P09211	1.11	0.53	0.52	-0.478	8.5E-01	Q00610	0.90	0.65	0.95	-0.263	9.1E-01
P09622	0.96	1.10	1.10	0.074	3.4E-01	Q03135	1.17	0.53	0.82	-0.250	9.3E-01
P09669	1.80	0.99	0.75	0.237	1.2E-01	Q04837	0.82	1.15	1.09	0.027	4.6E-01
P09972	1.10	0.56	0.77	-0.303	9.6E-01	Q06830	0.93	0.75	0.84	-0.251	7.9E-01
P10620	1.25	1.46	1.43	0.464	8.5E-02	Q07021	1.02	1.65	1.19	0.364	1.3E-01
P10809	1.00	1.04	1.16	0.090	3.5E-01	Q13011	0.99	0.96	1.08	0.012	4.2E-01
P11177	0.97	1.97	0.94	0.372	1.4E-01	Q14318	1.12	0.80	3.07	0.734	4.1E-01
P11586	1.18	1.07	0.65	-0.047	1.2E-01	Q16555	0.84	0.41	0.66	-0.651	6.3E-01
P12236	0.88	1.70	0.91	0.218	2.2E-01	Q16698	0.87	1.29	0.93	0.040	3.3E-01
P12931	0.88	0.72	0.25	-0.698	5.5E-01	Q16719	0.91	0.41	0.31	-0.886	4.8E-01
P13073	1.00	1.15	1.04	0.092	2.6E-01	Q16762	0.46	1.17	0.65	-0.400	8.8E-01
P14618	0.92	0.50	0.81	-0.431	8.5E-01	Q16891	1.04	0.95	0.99	-0.008	3.9E-01
P15531	0.75	0.60	0.93	-0.401	8.9E-01	Q56VL3	1.33	0.88	0.99	0.091	3.0E-01
P15924	0.44	2.80	1.59	0.687	5.2E-01	Q5RI15	1.01	1.70	1.93	0.629	1.8E-01
P15954	1.13	1.08	1.00	0.100	2.1E-01	Q6NZI2	0.93	0.48	0.76	-0.467	8.1E-01

P19404	0.93	1.38	0.93	0.114	2.2E-01	**Q6YN16**	1.58	1.33	1.99	0.707	1.1E-01
P21796	1.24	0.90	1.18	0.143	3.2E-01	**Q7Z434**	0.81	0.87	0.21	-0.662	5.8E-01
P21980	1.62	0.38	0.85	-0.069	9.4E-01	**Q96CW1**	0.87	0.63	0.68	-0.464	8.8E-01
P22392	0.94	0.68	0.85	-0.285	8.8E-01	**Q96EL3**	0.11	1.44	1.13	-0.167	7.2E-01
P22695	1.06	0.96	1.09	0.051	3.6E-01	**Q96HS1**	0.95	0.80	1.18	-0.034	6.1E-01
P24752	0.84	1.04	1.02	-0.049	5.1E-01	**Q96IX5**	0.97	0.95	0.88	-0.099	4.6E-01
P25705	1.00	1.22	1.08	0.139	2.3E-01	**Q99623**	0.92	1.04	1.02	-0.012	4.3E-01
P28331	1.01	1.29	0.76	0.030	1.5E-01	**Q99714**	0.83	2.29	0.99	0.452	2.3E-01
P30044	1.02	0.91	0.80	-0.137	4.7E-01	**Q99798**	0.89	1.03	1.13	0.022	4.6E-01
P30048	0.95	1.00	1.09	0.019	4.2E-01	**Q99832**	1.13	0.80	1.20	0.055	4.8E-01
P30084	0.58	1.13	1.01	-0.143	8.1E-01	**Q9BPW8**	0.83	0.83	0.92	-0.218	7.6E-01
P31040	0.82	1.00	0.85	-0.164	6.1E-01	**Q9BZF9**	0.72	0.69	0.57	-0.605	7.0E-01
P31930	1.04	0.99	0.89	-0.040	3.5E-01	**Q9C002**	1.06	1.18	0.93	0.080	1.8E-01
P34897	0.87	1.01	1.26	0.066	4.8E-01	**Q9H0U4**	0.96	0.69	1.20	-0.077	7.2E-01
P34932	1.46	0.59	0.85	-0.049	6.9E-01	**Q9NS69**	0.84	1.10	1.20	0.064	4.6E-01
P35232	1.01	1.08	1.13	0.102	3.0E-01	**Q9NVI7**	1.13	0.98	1.02	0.060	2.9E-01
P35613	0.90	0.60	0.65	-0.481	8.5E-01	**Q9NX40**	0.29	1.12	1.05	-0.289	8.3E-01
Q9UJZ1	0.97	1.18	1.17	0.145	2.8E-01	**Q9NX63**	0.85	0.96	1.10	-0.042	5.5E-01
Q9Y277	0.95	0.88	1.18	0.005	5.2E-01	**Q9P0J0**	1.04	0.92	0.90	-0.069	4.2E-01
Q9Y2R0	1.10	0.89	1.37	0.162	4.0E-01	**Q9UI09**	1.12	1.21	0.96	0.131	1.4E-01
Q9Y6N5	0.99	1.23	0.88	0.048	2.2E-01						

Supplementary Figure 4 - Data Sets N9 Repeat 1 and 2, and D9 Repeat 1 and 2

	Relative protein level log2(Inf/Uninf)							Relative protein level log2(Inf/Uninf)					
Accession	R1N	R2N	R1D	R2D	Average	**P-value**	Accession	R1N	R2N	R1D	R2D	Average	**P-value**
O00154	1.11	-0.15	-0.25	0.09	0.201	5.4E-01	**P48047**	0.21	0.05	0.03	0.04	0.083	9.9E-02
O00217	1.00	1.02	0.13	0.10	0.563	7.3E-02	**P49327**	-0.23	-0.68	-0.42	-0.10	-0.357	3.2E-02
O00299	-0.75	-1.20	-0.93	-0.73	-0.901	1.7E-04	**P49411**	-0.06	0.50	0.08	0.17	0.171	2.0E-01
O00330	-0.31	1.23	-0.03	0.00	0.222	5.4E-01	**P50213**	-0.32	0.52	0.06	-0.03	0.058	7.5E-01
O00483	-0.08	-0.21	-0.07	-0.25	-0.155	1.6E-02	**P51659**	0.19	-0.11	-0.01	0.24	0.078	3.8E-01
O00571	0.02	-0.37	0.40	0.62	0.170	4.7E-01	**P51970**	0.34	0.33	-0.01	0.02	0.172	1.3E-01
O14548	0.29	0.30	0.06	0.00	0.162	8.4E-02	**P52815**	0.16	-0.35	0.28	0.14	0.060	6.8E-01
O14949	-0.14	0.01	0.72	0.55	0.284	2.2E-01	**P53396**	-0.61	-1.30	-0.29	0.21	-0.497	1.7E-01
O14957	-0.93	-0.72	0.32	0.11	-0.305	3.6E-01	**P53597**	0.10	0.23	0.72	0.04	0.274	1.3E-01
O43169	0.24	0.03	0.56	0.37	0.299	3.6E-02	**P54819**	-0.12	0.20	0.85	0.02	0.237	3.1E-01
O43181	0.05	0.15	0.04	-0.01	0.057	1.3E-01	**P55084**	0.05	0.04	-0.01	0.08	0.043	6.3E-02
O43402	1.09	-0.05	-0.07	0.31	0.320	2.8E-01	**P55957**	-1.63	-1.49	-0.31	-0.49	-0.981	2.7E-02
O43678	0.61	0.62	0.11	-0.18	0.288	2.0E-01	**P56134**	0.04	0.47	0.18	0.68	0.343	5.5E-02
O43819	0.28	0.30	0.30	1.06	0.487	4.3E-02	**P56378**	0.35	0.67	0.19	0.35	0.390	8.3E-03
O43837	-0.57	0.67	-0.28	0.11	-0.015	9.6E-01	**P56381**	0.13	-0.21	0.09	0.02	0.008	9.2E-01
O43920	0.20	0.07	0.08	-0.01	0.083	1.0E-01	**P56385**	-0.07	0.22	-0.08	0.17	0.059	4.7E-01
O75251	-0.38	-1.08	-0.20	-0.03	-0.422	1.2E-01	**P56556**	-0.05	-0.10	-0.47	-0.46	-0.270	5.5E-02
O75438	0.08	-0.03	0.07	0.28	0.101	1.7E-01	**P60900**	-1.14	-0.31	-0.45	-0.33	-0.557	3.0E-02
O75608	0.35	-0.29	-0.55	0.76	0.067	8.3E-01	**P61604**	-0.04	0.02	0.04	0.14	0.041	3.0E-01
O75947	0.18	0.11	0.00	0.18	0.119	3.3E-02	**P62244**	-0.27	-0.15	-0.16	-0.12	-0.174	2.0E-03
O75964	0.32	0.38	0.19	0.26	0.285	4.3E-04	**P62258**	-0.40	-0.64	-2.31	0.17	-0.797	1.9E-01
O76031	-0.55	2.15	0.14	-0.40	0.335	6.1E-01	**P62263**	-0.07	0.22	-0.26	-0.12	-0.059	5.8E-01
O94760	0.28		-1.13	-0.78	-0.542	1.9E-01	**P62906**	-0.08	-0.32	-0.08	-0.22	-0.174	2.4E-02
O94826	0.00	0.01	-0.04	-0.05	-0.020	2.0E-01	**P63104**	-0.30	-0.84	-0.27	-0.27	-0.419	2.5E-02
O95139		2.08	-0.02	0.36	0.807	2.0E-01	**P63167**	-0.24	-0.62	-0.56	-0.48	-0.474	1.1E-03
O95168	-0.13	0.85	-0.80	0.11	0.004	9.9E-01	**P63244**	-0.22	0.98	-0.02	-0.36	0.096	7.6E-01
O95571		1.54	0.38	-0.27	0.551	2.7E-01	**P82650**	-0.42	-0.19	0.46	0.22	0.017	9.4E-01
O95573	0.52	-0.12	-1.23	-0.68	-0.377	3.5E-01	**P82663**	0.89	1.13	0.23	0.52	0.693	1.3E-02
O95831	-0.02	0.03	0.16	0.11	0.071	1.1E-01	**P82673**	0.49	0.89	0.89	0.11	0.596	1.9E-02
O95994	0.62	0.26	-0.31	-0.26	0.079	7.4E-01	**P82675**	0.11	-0.19	-0.45	-0.50	-0.260	1.1E-01
O96000	0.34	0.30	0.14	0.12	0.224	6.7E-03	**P82932**	-0.72	-0.19	-0.50	-0.27	-0.421	1.3E-02
O96008	0.11	-0.07	0.14	0.26	0.111	1.5E-01	**P99999**	0.44	0.21	-0.10	0.26	0.204	1.2E-01
P00167	0.13	-0.19	0.06	-0.36	-0.088	4.6E-01	**Q00059**	-0.19	-0.06	0.06	0.31	0.030	7.9E-01
P00338	-0.48	-1.12	-0.21	-0.06	-0.466	9.3E-02	**Q00325**	0.20	0.42	0.36	-0.52	0.115	6.1E-01
P00367	0.11	0.24	-0.05	0.21	0.127	1.0E-01	**Q00610**	-0.16	-0.62	-0.33	-0.07	-0.295	5.0E-02
P00390	-0.82	-0.67	-0.56	-0.25	-0.577	3.0E-03	**Q02127**	-0.51	-0.19	0.57	3.02	0.724	4.0E-01
P00403	0.04	-0.04	1.24	0.57	0.454	1.8E-01	**Q03135**	-0.33	-0.92		-0.28	-0.509	3.1E-02
P00505	-0.07	0.07	0.13	0.46	0.147	2.4E-01	**Q04837**	0.10	0.20	0.08	0.13	0.128	2.0E-03
P02786	0.03	-0.92	-0.66	-0.38	-0.482	5.5E-02	**Q06830**	-0.18	-0.41	-0.55	-0.26	-0.350	5.8E-03
P04040	-0.42	1.39	3.06	-0.05	0.998	2.5E-01	**Q07021**	0.17	0.72	0.30	0.25	0.363	2.6E-02
P04083	-0.28	-0.72	-0.73	-0.43	-0.540	2.8E-03	**Q12906**	1.54	2.40	0.02	-0.43	0.882	2.3E-01
P04156	-0.67	-1.48		-1.08	-1.077	2.7E-03	**Q13011**	-0.20	-0.06	0.06	0.11	-0.024	7.5E-01
P04179	0.19	-0.05	-0.21	-0.36	-0.109	3.9E-01	**Q13162**	0.17	-3.20	0.55	0.61	-0.467	6.3E-01
P06576	0.10	-0.01	0.08	0.09	0.065	3.8E-02	**Q13405**	-0.31	0.44	0.00	0.18	0.077	6.4E-01
P07195	-0.45	-1.08	-0.80	-0.77	-0.777	9.0E-04	**Q14197**	-0.41	0.19	-0.33	0.76	0.054	8.5E-01
P07339	-0.33	0.77	-0.61	-0.69	-0.216	5.5E-01	**Q14318**	-0.26	-0.32	1.73	1.62	0.693	2.7E-01

P07602	-0.16	-0.57	-0.92	-0.87	-0.631	1.1E-02	Q16540	1.42	0.11	0.37	0.43	0.582	9.0E-02
P07858	-0.48	-0.15	-0.77	-0.82	-0.555	1.1E-02	Q16555	-0.57	-1.28	-1.57	-0.59	-1.003	7.1E-03
P07900	-0.06	-0.61	-0.51	-0.13	-0.330	5.2E-02	Q16698	0.17	0.37	-0.20	-0.11	0.058	6.7E-01
P07919	0.08	0.11	0.00	0.09	0.070	2.6E-02	Q16718	0.22	0.41	-0.35	-0.27	0.003	9.9E-01
P07954	-0.06	0.09	0.16	0.03	0.056	2.9E-01	Q16719	-0.23	-1.29	-0.37	-1.70	-0.897	4.6E-02
P08107	-0.27	-0.59	-0.50	-0.30	-0.415	1.7E-03	Q16762	0.49	0.22	-0.73	-0.63	-0.160	6.2E-01
P08238	0.05	-0.57	-0.23	-0.06	-0.203	1.8E-01	Q16836	0.19	0.14	-0.20	-0.23	-0.024	8.4E-01
P08559	0.08	0.92	0.15	0.14	0.321	1.6E-01	Q16891	0.15	-0.07	0.05	-0.01	0.028	5.6E-01
P09211	-0.51	-0.93	-1.06	-0.94	-0.859	3.7E-04	Q53H12	0.15	-0.18	0.20	0.54	0.179	2.7E-01
P09622	0.05	0.14	0.13	0.14	0.115	1.6E-03	Q53H82	-1.56	-1.20	-0.64	-0.20	-0.901	2.4E-02
P09669	0.16	-0.01	-0.91	-0.42	-0.295	2.6E-01	Q56VL3	0.02	-0.19	-0.05	-0.01	-0.058	2.5E-01
P09972	-0.20	-0.84	-0.43	-0.37	-0.458	1.5E-02	Q5RI15	0.94	0.76	1.73	0.95	1.095	2.3E-03
P10515	-0.29	1.04	0.42	0.15	0.329	2.8E-01	Q5T1M5	0.02		-0.91	-1.81	-0.898	9.8E-02
P10599	-0.69	0.39	-0.81	-0.42	-0.383	2.0E-01	Q5T653	3.57		-0.12	0.12	1.188	2.9E-01
P10606	0.41	0.05	0.31	-0.10	0.169	2.0E-01	Q5TGZ0	0.23	-0.14	-0.25	0.23	0.018	8.9E-01
P10620	0.74	0.55	1.16	0.51	0.739	2.6E-03	Q5XKP0	0.15	0.28	0.09	0.34	0.217	9.4E-03
P10809	0.04	0.05	0.10	0.21	0.101	4.2E-02	Q6NZI2	-0.43	-1.06	-0.77	-0.40	-0.666	5.5E-03
P11177	0.16	0.98	-0.06	-0.09	0.249	3.6E-01	Q6YN16	0.66	0.41	-0.13	0.99	0.483	8.9E-02
P11310	1.51	0.43	-0.14	-0.21	0.398	3.6E-01	Q7Z434	-0.12	-0.20	0.13	-2.23	-0.605	3.1E-01
P11586	-0.51	0.10	-0.68	-0.61	-0.425	5.6E-02	Q8N183	0.38	0.05	0.52	0.41	0.341	1.5E-02
P12236	0.83	0.77	0.16	-0.14	0.406	1.4E-01	Q8N5N7	0.07	0.39	0.19	0.14	0.198	3.0E-02
P12931	-0.05	-0.47	0.26	-2.00	-0.563	3.0E-01	Q8N983	0.13	0.01	0.02	0.15	0.076	8.3E-02
P13073	0.08	0.20	0.06	0.06	0.101	2.5E-02	Q8WUK0	0.40	2.04	0.37	0.83	0.910	5.9E-02
P13804	-0.21	0.01	0.05	0.15	0.001	9.9E-01	Q96A26	-0.02	0.26	0.43	0.04	0.177	1.4E-01
P14406	0.21	0.10	-0.08	-0.21	0.006	9.5E-01	Q96A35	0.33	0.23	0.59	-0.06	0.274	8.5E-02
P14618	-0.37	-1.00	-0.72	-0.30	-0.600	1.0E-02	Q96AG4	0.22	-0.10	-0.06	0.66	0.180	3.4E-01
P14854	0.31	0.86	-0.09	-0.07	0.253	3.0E-01	Q96CW1	-0.24	-0.67	-1.30	-0.57	-0.694	2.0E-02
P14927	0.00	-0.43	0.07	0.27	-0.022	8.8E-01	Q96E11	-0.75	-0.06	0.00	-0.06	-0.218	2.7E-01
P15531	-0.29	-0.74	-0.65	-0.11	-0.448	2.3E-02	Q96EL2	-0.78		0.78	0.23	0.077	8.5E-01
P15924	-0.21	1.48	3.64	0.67	1.398	1.4E-01	Q96EL3	0.35	0.52	0.22	0.17	0.318	6.6E-03
P15954	0.12	0.11	0.02	0.01	0.064	7.2E-02	Q96EY1	-0.50	-0.16	-0.02	1.17	0.122	7.5E-01
P17931	-0.37	-1.00	-0.90	-0.66	-0.732	2.1E-03	Q96EY8	-0.31	-0.02	-0.47	-0.02	-0.205	1.1E-01
P18859	0.32	0.31	0.20	0.18	0.253	4.4E-04	Q96HS1	-0.24	-0.33	-0.05	0.24	-0.094	4.8E-01
P19404	0.08	0.46	-0.01	-0.10	0.109	4.1E-01	Q96I99	-0.43	-0.18	-0.03	0.38	-0.065	7.2E-01
P20674	0.17	-0.01	-0.08	0.00	0.021	7.1E-01	Q96IX5	0.01	-0.07	-0.33	-0.18	-0.142	1.0E-01
P21796	0.01	-0.16	0.03	0.24	0.030	7.2E-01	Q99497	-0.43	-0.65	-0.75	-0.39	-0.557	6.5E-04
P21912	-0.08	0.07	0.08	0.18	0.063	2.9E-01	Q99623	0.21	0.05	0.14	0.03	0.110	3.9E-02
P21980	-0.61	-1.38	-0.78	-0.23	-0.748	2.0E-02	Q99714	-0.22	1.19	0.07	-0.02	0.256	4.5E-01
P22307	-0.47	0.63	-0.07	-0.06	0.008	9.7E-01	Q99798	0.28	0.04	-0.04	0.18	0.114	1.6E-01
P22392	-0.25	-0.56	-0.64	-0.24	-0.422	6.6E-03	Q99832	0.12	-0.33	0.18	0.26	0.057	6.8E-01
P22570	-1.31	0.88	0.25	0.22	0.012	9.8E-01	Q9BPW8	-0.63	-0.26	0.41	-0.13	-0.153	5.0E-01
P22695	-0.02	-0.06	0.01	0.12	0.011	7.8E-01	Q9BTZ2	0.15		-0.04	0.02	0.045	3.9E-01
P23919	-0.49	-0.36	-0.27	-0.21	-0.335	1.6E-03	Q9BXK5	0.76	1.39	-0.01	-0.24	0.473	2.5E-01
P24752	-0.04	0.05	0.03	0.03	0.020	3.6E-01	Q9BYD1	1.13	0.61	-0.27	-0.45	0.255	5.2E-01
P25325	-0.25	0.07	-0.53	-0.87	-0.395	9.9E-02	Q9BYD3	-0.63	0.20	-1.69	-0.57	-0.669	1.4E-01
P25705	0.09	0.28	0.01	0.12	0.125	7.0E-02	Q9BYN8	0.88	1.39	0.31	0.49	0.766	1.8E-02
P28331	0.20	0.37	0.24	-0.40	0.104	5.6E-01	Q9BZF9	-0.19	-0.54	-0.85	-0.82	-0.600	7.8E-03
P29353	-0.16	-0.41	-0.60	-0.23	-0.351	1.2E-02	Q9C002	0.16	0.24	-0.11	-0.10	0.047	6.2E-01
P30042	1.11	1.10	-0.18	-0.13	0.476	2.4E-01	Q9GZT3	-0.04	-0.08	0.07	0.02	-0.005	8.9E-01
P30044	-0.17	-0.14	-0.40	-0.33	-0.259	5.9E-03	Q9H0U4	-0.04	-0.54	0.00	0.27	-0.079	6.6E-01

P30048	-0.08	0.00	0.09	0.13	0.034	5.0E-01
P30049	0.09	0.28	0.05	0.14	0.141	2.8E-02
P30084	-0.05	0.17	0.03	0.01	0.044	3.8E-01
P30086	-1.83	-0.64	-0.97	-0.27	-0.927	3.2E-02
P30405	-0.13	-0.04	-0.01	-0.17	-0.087	6.3E-02
P31040	-0.02	0.00	0.03	-0.23	-0.052	4.1E-01
P31930	0.05	-0.02	0.05	-0.17	-0.021	6.9E-01
P31937	-0.11	-0.25	-0.59	0.00	-0.239	1.1E-01
P34897	0.28	0.02	0.28	0.34	0.229	1.9E-02
P34932	-0.08	-0.76	-0.57	-0.23	-0.414	3.8E-02
P35232	0.17	0.11	0.09	0.17	0.138	5.2E-04
P35579	-0.53	-1.14	0.58	0.28	-0.204	6.2E-01
P35613	-0.26	-0.73	-1.31	-0.63	-0.734	1.5E-02
P35998	-0.30	0.08	-0.58	-0.31	-0.277	8.8E-02
P36507	-0.31	-0.68	-0.13	0.08	-0.258	1.6E-01
P36542	0.08	0.02	-0.03	0.12	0.048	1.9E-01
P36543	-0.59	-0.80	-0.38	-0.35	-0.531	2.2E-03
P36957	-0.06	-0.06	0.22	0.11	0.051	4.9E-01
P38117	-0.26	-0.04	0.15	0.18	0.008	9.4E-01
P38646	-0.08	-0.02	0.14	0.28	0.076	3.8E-01
P39748	0.17	-0.50	0.13		-0.069	7.2E-01
P40925	-0.60	-0.84	0.44	0.84	-0.042	9.2E-01
P40926	-0.06	0.11	0.10	0.19	0.085	1.5E-01
P40939	0.22	0.05	0.07	0.08	0.105	3.8E-02
P42765	0.10	0.30	0.08	-0.01	0.116	1.2E-01
P43897	-0.02	0.42	0.29	0.18	0.219	5.4E-02
P45880	0.07	-0.06	0.10	0.25	0.090	2.0E-01
P47985	-0.07	-0.15	0.10	0.55	0.105	5.2E-01
Q9Y6M9	0.11	0.44	0.12	0.87	0.385	7.6E-02
Q9Y6N5	0.14	0.30	-1.70	-0.18	-0.360	4.6E-01
Q9H2U2	0.01	0.07	0.05	-0.05	0.020	5.0E-01
Q9H2W6	-0.08	-0.18	-0.37	0.44	-0.047	8.0E-01
Q9HAV7	0.82	0.39	0.39	0.43	0.507	2.8E-03
Q9NQR4	-0.66	-2.20	-0.14	0.01	-0.749	1.9E-01
Q9NRX2	0.11	0.22	0.09	0.10	0.131	4.5E-03
Q9NS69	0.19	0.14	0.96	0.26	0.389	9.1E-02
Q9NVI7	0.05	-0.03	0.19	0.02	0.056	2.8E-01
Q9NX40	0.11	0.16	-0.07	0.07	0.067	2.2E-01
Q9NX63	0.11	-0.06	0.03	0.14	0.053	2.8E-01
Q9NZ45	-0.04	0.11	0.04	0.09	0.053	1.6E-01
Q9P015	-0.07	0.05	-0.29	0.11	-0.051	5.8E-01
Q9P0J0	-0.13	-0.11	-0.05	-0.16	-0.111	2.2E-03
Q9P0K7	-0.74	-3.08	-1.05	-0.69	-1.393	5.0E-02
Q9UBC2	-0.89	-0.41	-0.91	-0.52	-0.682	1.8E-03
Q9UI09	0.00	0.27	-0.04	-0.06	0.042	6.1E-01
Q9UIJ7	-1.02		0.06	-0.03	-0.328	3.1E-01
Q9UJZ1	0.09	0.24	0.17	0.22	0.182	1.5E-03
Q9ULD2	-0.31	-0.74	-0.54	-0.50	-0.521	1.0E-03
Q9Y237	-0.26	-0.63	0.15	-0.28	-0.254	1.6E-01
Q9Y277	-0.02	-0.18	0.12	0.24	0.041	6.7E-01
Q9Y2Q9	0.13	-0.07	0.15	0.13	0.084	1.6E-01
Q9Y2R0	0.10	-0.17	0.10	0.46	0.122	3.8E-01
Q9Y2R9	-0.65	-0.44	0.49	0.30	-0.077	7.9E-01
Q9Y2S7	-0.12	0.40	0.31	0.15	0.183	1.6E-01
Q9Y3B7	0.03	0.21	0.35	0.06	0.162	7.1E-02
Q9Y3D9	-0.44	0.08	0.40	0.20	0.060	7.5E-01
Q9Y3F4	-0.76	-0.26	-0.44	-0.59	-0.512	3.2E-03
Q9Y5J7	0.70	0.91	-0.01	0.31	0.478	5.7E-02
Q9Y5L4	0.84	1.62	0.01	0.06	0.632	1.5E-01

Supplementary Figure 5 - Mitochondrial PB2 StrepTactin Purifications: Relative Protein Fold Changes

Accession	Relative Protein Level (Strep-PB2/WSN) R1	R2	R3	Average	Accession	Relative Protein Level (Strep-PB2/WSN) R1	R2	R3	Average
A3EZ79			0.10	**0.10**	P42704	10.000			**10.000**
A3EZ82	0.10		0.10	**0.10**	P43243	10.000			**10.000**
B9A064	2.34	0.10		**1.22**	P45880	1.319			**1.319**
C9JLW8			4.02	**4.02**	P46776	1.231			**1.231**
HA	2.28	1.31	8.87	**4.15**	P46783	0.890			**0.890**
M1	2.53	0.55	11.51	**4.87**	P46940	0.878			**0.878**
M2	2.21		10.00	**6.10**	P47914	10.000			**10.000**
NP	1179	508	1332	**1006.56**	P47929	0.229	1.984	0.027	**0.747**
NS1	8.09	0.53	2.74	**3.79**	P48047	0.501	0.456	3.841	**1.600**
O00148	10.00			**10.00**	P48163			0.100	**0.100**
O00159		1.93	0.94	**1.43**	P48444	10.000			**10.000**
O00425	10.00		10.00	**10.00**	P48643	10.000			**10.000**
O00505	10.00			**10.00**	P49207			10.000	**10.000**
O00571	10.00			**10.00**	P49327	1.911	1.840	0.131	**1.294**
O14556			0.10	**0.10**	P49368	10.000			**10.000**
O14818			0.10	**0.10**	P49411	5.423		12.019	**8.721**
O14950		0.79	2.58	**1.68**	P49748			1.121	**1.121**
O14980	5.16			**5.16**	P49862	0.100			**0.100**
O15235			7.62	**7.62**	P50395			0.100	**0.100**
O43707	10.00		0.10	**5.05**	P50990	1.010		0.100	**0.555**
O43795	1.96	4.64	2.73	**3.11**	P51148	2.026		2.086	**2.056**
O43852	6.34		31.37	**18.86**	P51149	0.739			**0.739**
O60437			0.10	**0.10**	P51571	2.109		10.000	**6.055**
O60684	10.00		10.00	**10.00**	P51648			0.100	**0.100**
O60814	5.38	0.75		**3.07**	P51659			0.100	**0.100**
O75223	0.10		0.16	**0.13**	P51991	10.000			**10.000**
O75342	0.27		0.10	**0.18**	P52294	10.000		10.000	**10.000**
O75369	0.10		0.10	**0.10**	P52815			10.000	**10.000**
O75955		0.73		**0.73**	P52907			0.125	**0.125**
O95831			8.73	**8.73**	P54652	0.100			**0.100**
P00338			0.08	**0.08**	P55000	0.100			**0.100**
P00403	0.10			**0.10**	P55060	6.871	3.739	2.082	**4.231**
P00558			0.10	**0.10**	P55072			0.100	**0.100**
P01036	0.10			**0.10**	P55084			2.784	**2.784**
P01037	0.10			**0.10**	P55786			0.100	**0.100**
P01040	0.82			**0.82**	P58107			0.100	**0.100**
P01591	0.48			**0.48**	P60174	0.100		0.100	**0.100**
P01620	1.41	0.17		**0.79**	P60660	3.321		5.524	**4.422**
P01625	6.36	0.10		**3.23**	P60709	2.782	2.058		**2.420**
P01765		0.10		**0.10**	P60842	3.535		1.963	**2.749**
P01766		0.18		**0.18**	P60866	1.815		7.918	**4.866**
P01833	3.41			**3.41**	P60900			0.100	**0.100**
P01834	2.30	0.12		**1.21**	P61019			7.341	**7.341**
P01857	6.04	0.24	0.61	**2.30**	P61026	6.705	0.121	3.554	**3.460**
P01859	6.10	0.09		**3.09**	P61106	0.100	0.100	0.100	**0.100**

P01860	4.32	0.10		**2.21**	P61204	1.827		10.000	**5.914**
P01876	0.49		0.47	**0.48**	P61224	3.240			**3.240**
P02545	0.10		0.01	**0.05**	P61247	10.000			**10.000**
P02647	10.00			**10.00**	P61254			7.079	**7.079**
P02768	1.11	0.02	0.19	**0.44**	P61513			10.000	**10.000**
P02787	0.45			**0.45**	P61586			10.000	**10.000**
P02788	0.10		0.46	**0.28**	P61626	3.952	1.534	2.289	**2.592**
P04040	0.46		0.25	**0.36**	P61978	15.934		0.862	**8.398**
P04062			0.10	**0.10**	P61981	10.000		0.100	**5.050**
P04075			0.10	**0.10**	P62081	1.331			**1.331**
P04083	1.29		0.51	**0.90**	P62158	0.100	1.809	1.913	**1.274**
P04406	0.18	0.72	0.23	**0.38**	P62241	1.575		6.569	**4.072**
P04792	0.67		0.17	**0.42**	P62244			10.000	**10.000**
P04843			0.10	**0.10**	P62258			0.100	**0.100**
P04908		1.71		**1.71**	P62263	5.304	0.296	3.641	**3.080**
P05089	0.93		0.10	**0.51**	P62266		10.000	2.174	**6.087**
P05090	0.10			**0.10**	P62269	0.556	0.602	3.129	**1.429**
P05109	0.96		5.54	**3.25**	P62277	10.000		10.000	**10.000**
P05141	1.44	0.71	1.34	**1.16**	P62280			5.585	**5.585**
P05165	2.16	0.48	3.83	**2.16**	P62424	2.240			**2.240**
P05166	0.99	0.92	2.18	**1.36**	P62701	2.663	3.496	24.575	**10.245**
P05387	0.10		0.54	**0.32**	P62736	10.000			**10.000**
P05388	1.39		10.00	**5.70**	P62750	3.085	0.217		**1.651**
P06396			0.10	**0.10**	P62753	10.000		19.120	**14.560**
P06576	1.76		0.52	**1.14**	P62805	0.792	0.580	0.457	**0.610**
P06702	1.01	1.81	0.84	**1.22**	P62820			1.063	**1.063**
P06733	0.30		0.10	**0.20**	P62826	0.773	0.100	0.546	**0.473**
P06744			0.10	**0.10**	P62829	2.871	41.167	4.201	**16.080**
P06748	109.27			**109.27**	P62851	6.601	4.286	0.100	**3.662**
P06753			0.10	**0.10**	P62857	4.355			**4.355**
P07237			0.10	**0.10**	P62861			5.913	**5.913**
P07339	1.25		0.10	**0.67**	P62899	7.232		10.000	**8.616**
P07355	0.49	0.30	0.88	**0.56**	P62913			2.337	**2.337**
P07384	0.10		0.10	**0.10**	P62917	0.100			**0.100**
P07437	8.45	2.90	10.00	**7.12**	P62937	0.100			**0.100**
P07476			0.10	**0.10**	P63104	1.298	1.081	0.441	**0.940**
P07858			10.00	**10.00**	P63173	4.155		2.528	**3.342**
P07900	4.48	1.15	2.78	**2.80**	P63244	10.000			**10.000**
P07910	10.00			**10.00**	P63261			0.946	**0.946**
P08107	5.03		0.32	**2.67**	P67809	10.000		38.861	**24.430**
P08195	10.00		5.89	**7.94**	P68032		1.188	0.884	**1.036**
P08238	3.00	2.33	4.49	**3.27**	P68104	1.998	1.835	3.239	**2.357**
P08670	5.49	2.05	2.01	**3.18**	P68363			0.215	**0.215**
P08758			0.10	**0.10**	P68366	1.267		2.459	**1.863**
P08865	1.71		0.57	**1.14**	P68371	10.000	10.000	7.759	**9.253**
P09211			0.10	**0.10**	P68431			5.784	**5.784**
P09228	0.10			**0.10**	P68871	4.033		0.100	**2.067**
P09496			0.10	**0.10**	P78371	4.335			**4.335**
P09525			0.10	**0.10**	P80748	0.100			**0.100**
P09651	10.00			**10.00**	P81605	0.196	1.376	0.971	**0.848**

P0C0S8	8.98			**8.98**	P83731	5.606			**5.606**
P0C7P4			2.52	**2.52**	P84098	4.860		7.377	**6.119**
P0CG05	1.18	0.15		**0.67**	PA	536	10.000	3620	**1388.533**
P0CG47	1.35	0.38	2.79	**1.51**	PB1	1817	10.000	6004	**2610.351**
P10599	0.93	0.10	0.10	**0.38**	PB2	2604	10.000	2861	**1824.977**
P10809	12.34	10.00	46.85	**23.06**	Q00059			9.251	**9.251**
P11021	2.08	0.62	1.68	**1.46**	Q00325			9.993	**9.993**
P11142	4.44	2.56	2.13	**3.04**	Q00610	4.686	10.000	0.951	**5.212**
P11166	2.64			**2.64**	Q00839	10.000			**10.000**
P11182			7.18	**7.18**	Q01105			3.284	**3.284**
P11279	0.61			**0.61**	Q01469	0.100	0.100	0.100	**0.100**
P11498	2.18	0.55	2.51	**1.75**	Q02218			3.826	**3.826**
P11940	3.40		8.95	**6.17**	Q02413	0.490	2.183	0.075	**0.916**
P12081			0.10	**0.10**	Q02878	1.538		0.100	**0.819**
P12236	3.95	1.03	5.95	**3.64**	Q04837			1.115	**1.115**
P12268	4.07	0.54	5.97	**3.53**	Q06830		0.708	1.741	**1.224**
P12273	2.80			**2.80**	Q07021	2.817	10.000	4.718	**5.845**
P13489	5.08		0.29	**2.68**	Q07065			0.100	**0.100**
P13639			0.10	**0.10**	Q08188	0.178	1.046	0.037	**0.420**
P13797			0.10	**0.10**	Q08211	10.000			**10.000**
P13987	2.07			**2.07**	Q08554	0.308	0.862	3.819	**1.663**
P14618	0.91		0.29	**0.60**	Q09666	0.100		0.330	**0.215**
P14735			0.10	**0.10**	Q12860	5.040			**5.040**
P14923	0.80	4.01	0.09	**1.63**	Q13011	0.577		9.753	**5.165**
P15311		0.64	0.10	**0.37**	Q13085	1.886	0.504	3.110	**1.833**
P15531			0.10	**0.10**	Q13151	10.000			**10.000**
P15559			10.00	**10.00**	Q13813			0.100	**0.100**
P15924	0.70	1.74	0.03	**0.82**	Q13835	0.548	1.776	0.012	**0.779**
P16989			10.00	**10.00**	Q13867			0.100	**0.100**
P17900	0.60			**0.60**	Q14134			0.100	**0.100**
P17931			0.10	**0.10**	Q14204	1.245	10.000	0.100	**3.782**
P18085	3.99			**3.99**	Q14254		0.895		**0.895**
P18206	0.21		0.10	**0.16**	Q14257			10.000	**10.000**
P18510	0.10		0.10	**0.10**	Q14574	0.347		0.100	**0.224**
P18621	10.00			**10.00**	Q14974		10.000		**10.000**
P19338	10.00			**10.00**	Q15149	0.100	1.684	0.116	**0.633**
P20073	4.49			**4.49**	Q15165	10.000			**10.000**
P20339			0.10	**0.10**	Q15293	2.300	0.749	7.441	**3.497**
P20930	0.24		0.00	**0.12**	Q15365	4.419	2.325	7.647	**4.797**
P21333	10.00		13.42	**11.71**	Q15366	1.674			**1.674**
P21796	0.92			**0.92**	Q15517	0.375		0.100	**0.237**
P22061			5.23	**5.23**	Q16610	0.405		0.100	**0.252**
P22626	10.00			**10.00**	Q27J81	2.402		10.000	**6.201**
P22735	0.09		0.10	**0.09**	Q3MHD2	1.297	2.144	0.684	**1.375**
P23396	2.49	0.28	2.81	**1.86**	Q3ZCQ8	0.605			**0.605**
P23526			0.10	**0.10**	Q562R1		0.100	0.403	**0.251**
P25311	0.12	0.30	0.10	**0.17**	Q5D862	0.258	0.101	0.004	**0.121**
P25705	1.99	1.37	1.89	**1.75**	Q5T749	3.263	1.073	0.725	**1.687**
P25786			0.10	**0.10**	Q5T9A4			10.000	**10.000**
P25788			0.10	**0.10**	Q6NZI2			29.586	**29.586**

P25789			0.10	**0.10**	Q6UWP8	0.100		0.100	**0.100**
P26038			0.10	**0.10**	Q6ZVX7	0.100		0.100	**0.100**
P26599	10.00			**10.00**	Q71U36	2.332			**2.332**
P26641			0.10	**0.10**	Q7L576	1.362		4.135	**2.748**
P27348			0.10	**0.10**	Q7Z406			0.136	**0.136**
P27482	9.21		0.10	**4.66**	Q7Z417	2.028	0.489	8.869	**3.795**
P27635			16.74	**16.74**	Q86W56	0.100			**0.100**
P28072			0.10	**0.10**	Q86YZ3	0.249	1.082	2.311	**1.214**
P28074			0.10	**0.10**	Q8IW75	0.100		0.100	**0.100**
P29401			0.10	**0.10**	Q8IZP0	2.137		5.388	**3.763**
P29508	0.10			**0.10**	Q8NC51	10.000			**10.000**
P29692			1.99	**1.99**	Q8TDL5	10.000			**10.000**
P30050	3.30			**3.30**	Q8WUM4	1.763			**1.763**
P30086			0.10	**0.10**	Q8WUW1			4.561	**4.561**
P31025	4.34		0.96	**2.65**	Q8WVV4	0.100		0.100	**0.100**
P31040			10.00	**10.00**	Q8WWM7	1.865	1.076	7.244	**3.395**
P31151	2.23			**2.23**	Q92817			0.100	**0.100**
P31943	10.00		0.10	**5.05**	Q92820	0.100			**0.100**
P31944	0.44	0.68	0.23	**0.45**	Q92928	0.100			**0.100**
P31947			0.10	**0.10**	Q92973	10.000			**10.000**
P31949			0.10	**0.10**	Q96DA0	1.885			**1.885**
P32119	0.61		0.23	**0.42**	Q96DR8	0.100			**0.100**
P32969	5.74			**5.74**	Q96HS1			0.100	**0.100**
P34931		10.00		**10.00**	Q96P63	0.885		0.100	**0.493**
P35579	2.28	2.31	0.99	**1.86**	Q96RQ3	3.412	0.492	4.794	**2.899**
P35754			0.10	**0.10**	Q99623	10.000			**10.000**
P36542	3.06		5.18	**4.12**	Q9HAV7	1.614	0.948	6.545	**3.036**
P36578	10.00			**10.00**	Q9HCC0	2.838	0.723	3.894	**2.485**
P36952			0.10	**0.10**	Q9HCY8			0.100	**0.100**
P36957	11.73	2.79	7.09	**7.20**	Q9NUL7			10.000	**10.000**
P38646	6.43	1.56	10.68	**6.22**	Q9NX20			0.863	**0.863**
P39019	2.80	0.27	2.95	**2.00**	Q9NZH8	0.100			**0.100**
P39023			5.25	**5.25**	Q9NZI8	10.000		10.000	**10.000**
P40227	10.00			**10.00**	Q9NZT1	1.062	0.200	0.013	**0.425**
P40925			0.10	**0.10**	Q9UGM3			2.137	**2.137**
P40939	4.71			**4.71**	Q9UI42	0.313		0.100	**0.206**
P42357	0.83		0.10	**0.46**	Q9UJS0	2.909		7.024	**4.967**
P42677			1.53	**1.53**	Q9Y2A7	0.942		3.301	**2.121**
					Q9Y3R4	0.100			**0.100**

6.3 Supplementary Figure Legends

Supplementary Figure 1 – Analysing the level of NP in the SILAC samples – The lowest SINQ value assigned to any protein within each of the SILAC data sets was used to set the background level (detection threshold). SINQ analysis was used to calculate the level of NP in each of the samples. The NP level in each sample was then divided by the background level to calculate the level of NP above background for the three N9-WSN data sets and the three D9-WSN data sets.

Supplementary Figure 2 – Analysis of the SILAC N9-WSN data sets from repeat 1, 2 and 3 - The relative protein levels on infection with N9-WSN from repeats 1, 2 and 3 are shown, as well as the average relative protein levels. Only mitochondrial annotated proteins which were detected and quantified in all the repeats (n=3) are displayed.

Supplementary Figure 3 – Analysis of the SILAC D9-WSN data sets from repeat 1, 2 and 3 - The relative protein levels on infection with D9-WSN from repeats 1, 2 and 3 are shown, as well as the average relative protein levels. Only mitochondrial annotated proteins which were detected and quantified in all the repeats (n=3) are displayed.

Supplementary Figure 4 – Analysis of the SILAC N9-WSN data sets from repeat 1 and 2, and the D9-WSN data sets from repeat 1 and 2 - The relative protein levels on infection with N9-WSN from repeats 1 and 2, and with D9-WSN from repeats 1 and 2 are shown. The average relative protein levels from the four data sets is also shown. Only mitochondrial annotated proteins which were detected and quantified in all the repeats (n=3) are displayed.

Supplementary Figure 5 – SINQ analysis of the proteins which copurify with Strep-PB2 or PB2 from isolated mitochondria – SINQ analysis was used to quantify the levels of mitochondrial proteins in the elution from the StrepTactin purifications of Strep-PB2 from isolated mitochondria. These were then normalised against the level of 'contaminant proteins' in the elution. A fold difference in the level of each protein between the PB2-Strep-WSN and WSN samples (relative protein level) was calculated for repeat 1, 2 and 3 (R1/2/3) and an average relative protein level was calculated. If a protein was only detected in the Strep-PB2-WSN sample then the relative protein level was arbitrarily set to 10. If a protein was only detected in the Strep-PB2-WSN sample, then the relative protein level was arbitrarily set as 0.1.

Declaration of Authorship

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

Joshua. C. D. Long

Chapter 7

References

1. Potter, C.W. A history of influenza. *J Appl Microbiol* **91**, 572-579 (2001).

2. WHO. H5N1 influenza.

3. Shaw ML, P.P., Wright PF, Neumann G, Kawaoka Y *Fields Virology Sixth Edition*, Sixth edn, vol. 1, 2013.

4. WHO. Influenza (Seasonal). 2014 [cited]Available from: http://www.who.int/mediacentre/factsheets/fs211/en

5. Rossman, J.S. & Lamb, R.A. Influenza virus assembly and budding. *Virology* **411**, 229-236 (2011).

6. Blumenkrantz, D., Roberts, K.L., Shelton, H., Lycett, S. & Barclay, W.S. The short stalk length of highly pathogenic avian influenza H5N1 virus neuraminidase limits transmission of pandemic H1N1 virus in ferrets. *Journal of virology* **87**, 10539-10551 (2013).

7. Wilson, I.A., Skehel, J.J. & Wiley, D.C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. *Nature* **289**, 366-373 (1981).

8. Sriwilaijaroen, N. & Suzuki, Y. Molecular basis of the structure and function of H1 hemagglutinin of influenza virus. *Proc Jpn Acad Ser B Phys Biol Sci* **88**, 226-249 (2012).

9. Noda, T. Native morphology of influenza virions. *Front Microbiol* **2**, 269 (2011).

10. Calder, L.J., Wasilewski, S., Berriman, J.A. & Rosenthal, P.B. Structural organization of a filamentous influenza A virus. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 10685-10690 (2010).

11. Chen, B.J., Leser, G.P., Jackson, D. & Lamb, R.A. The influenza virus M2 protein cytoplasmic tail interacts with the M1 protein and influences virus assembly at the site of virus budding. *Journal of virology* **82**, 10059-10070 (2008).

12. Ali, A., Avalos, R.T., Ponimaskin, E. & Nayak, D.P. Influenza virus assembly: effect of influenza virus glycoproteins on the membrane association of M1 protein. *Journal of virology* **74**, 8709-8719 (2000).

13. Noda, T. *et al*. Architecture of ribonucleoprotein complexes in influenza A virus particles. *Nature* **439**, 490-492 (2006).

14. Seladi-Schulman, J., Steel, J. & Lowen, A.C. Spherical influenza viruses have a fitness advantage in embryonated eggs, while filament-producing strains are selected in vivo. *Journal of virology* **87**, 13343-13353 (2013).

15. Hutchinson, E.C. *et al*. Conserved and host-specific features of influenza virion architecture. *Nature communications* **5**, 4816 (2014).

16. Richardson, J.C. & Akkina, R.K. NS2 protein of influenza virus is found in purified virus and phosphorylated in infected cells. *Arch Virol* **116**, 69-80 (1991).

17. Brooke, C.B., Ince, W.L., Wei, J., Bennink, J.R. & Yewdell, J.W. Influenza A virus nucleoprotein selectively decreases neuraminidase gene-segment packaging while enhancing viral fitness and transmissibility. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 16854-16859 (2014).

18. Dimmock, N.J. & Easton, A.J. Defective interfering influenza virus RNAs: time to reevaluate their clinical potential as broad-spectrum antivirals? *Journal of virology* **88**, 5217-5227 (2014).

19. Bottcher, E. *et al*. Proteolytic activation of influenza viruses by serine proteases TMPRSS2 and HAT from human airway epithelium. *Journal of virology* **80**, 9896-9898 (2006).

20. Kido, H. *et al*. Isolation and characterization of a novel trypsin-like protease found in rat bronchiolar epithelial Clara cells. A possible activator of the viral fusion glycoprotein. *The Journal of biological chemistry* **267**, 13573-13579 (1992).

21. Boycott, R., Klenk, H.D. & Ohuchi, M. Cell tropism of influenza virus mediated by hemagglutinin activation at the stage of virus entry. *Virology* **203**, 313-319 (1994).

22. Steinhauer, D.A. Role of hemagglutinin cleavage for the pathogenicity of influenza virus. *Virology* **258**, 1-20 (1999).

23. Schnell, J.R. & Chou, J.J. Structure and mechanism of the M2 proton channel of influenza A virus. *Nature* **451**, 591-595 (2008).

24. Ye, Z., Liu, T., Offringa, D.P., McInnis, J. & Levandowski, R.A. Association of influenza virus matrix protein with ribonucleoproteins. *Journal of virology* **73**, 7467-7473 (1999).

25. Bui, M., Whittaker, G. & Helenius, A. Effect of M1 protein and low pH on nuclear transport of influenza virus ribonucleoproteins. *Journal of virology* **70**, 8391-8401 (1996).

26. Hutchinson, E.C. & Fodor, E. Nuclear import of the influenza A virus transcriptional machinery. *Vaccine* **30**, 7353-7358 (2012).

27. Wu, W.W., Sun, Y.H. & Pante, N. Nuclear import of influenza A viral ribonucleoprotein complexes is mediated by two nuclear localization sequences on viral nucleoprotein. *Virol J* **4**, 49 (2007).

28. Desselberger, U., Racaniello, V.R., Zazra, J.J. & Palese, P. The 3' and 5'-terminal sequences of influenza A, B and C virus RNA segments are highly conserved and show partial inverted complementarity. *Gene* **8**, 315-328 (1980).

29. Cheong, H.K., Cheong, C. & Choi, B.S. Secondary structure of the panhandle RNA of influenza virus A studied by NMR spectroscopy. *Nucleic acids research* **24**, 4197-4201 (1996).

30. Bae, S.H. *et al*. Structural features of an influenza virus promoter and their implications for viral RNA synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 10602-10607 (2001).

31. Tomescu, A.I., Robb, N.C., Hengrung, N., Fodor, E. & Kapanidis, A.N. 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**, E3335-3342 (2014).

32. Pflug, A., Guilligay, D., Reich, S. & Cusack, S. Structure of influenza A polymerase bound to the viral RNA promoter. *Nature* **516**, 355-360 (2014).

33. Reich, S. *et al*. Structural insight into cap-snatching and RNA synthesis by influenza polymerase. *Nature* **516**, 361-366 (2014).

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

35. Poole, E., Elton, D., Medcalf, L. & Digard, P. Functional domains of the influenza A virus PB2 protein: identification of NP- and PB1-binding sites. *Virology* **321**, 120-133 (2004).

36. Biswas, S.K., Boutz, P.L. & Nayak, D.P. Influenza virus nucleoprotein interacts with influenza virus polymerase proteins. *Journal of virology* **72**, 5493-5501 (1998).

37. Turrell, L., Lyall, J.W., Tiley, L.S., Fodor, E. & Vreede, F.T. The role and assembly mechanism of nucleoprotein in influenza A virus ribonucleoprotein complexes. *Nature communications* **4**, 1591 (2013).

38. Ng, A.K. *et al*. Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *FASEB J* **22**, 3638-3647 (2008).

39. Moeller, A., Kirchdoerfer, R.N., Potter, C.S., Carragher, B. & Wilson, I.A. Organization of the influenza virus replication machinery. *Science* **338**, 1631-1634 (2012).

40. Arranz, R. *et al*. The structure of native influenza virion ribonucleoproteins. *Science* **338**, 1634-1637 (2012).

41. Duesberg, P.H. Distinct subunits of the ribonucleoprotein of influenza virus. *J Mol Biol* **42**, 485-499 (1969).

42. Hengrung, N. *et al*. Crystal structure of the RNA-dependent RNA polymerase from influenza C virus. *Nature* **527**, 114-117 (2015).

43. Thierry, E. *et al*. Influenza Polymerase Can Adopt an Alternative Configuration Involving a Radical Repacking of PB2 Domains. *Molecular cell* **61**, 125-137 (2016).

44. Fodor, E., Pritlove, D.C. & Brownlee, G.G. Characterization of the RNA-fork model of virion RNA in the initiation of transcription in influenza A virus. *Journal of virology* **69**, 4012-4019 (1995).

45. Koppstein, D., Ashour, J. & Bartel, D.P. Sequencing the cap-snatching repertoire of H1N1 influenza provides insight into the mechanism of viral transcription initiation. *Nucleic acids research* **43**, 5052-5064 (2015).

46. Engelhardt, O.G., Smith, M. & Fodor, E. Association of the influenza A virus RNA-dependent RNA polymerase with cellular RNA polymerase II. *Journal of virology* **79**, 5812-5818 (2005).

47. Martinez-Alonso, M., Hengrung, N. & Fodor, E. 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* (2016).

48. Dias, A. *et al*. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* **458**, 914-918 (2009).

49. Yuan, P. *et al*. Crystal structure of an avian influenza polymerase PA(N) reveals an endonuclease active site. *Nature* **458**, 909-913 (2009).

50. Rao, P., Yuan, W. & Krug, R.M. Crucial role of CA cleavage sites in the cap-snatching mechanism for initiating viral mRNA synthesis. *The EMBO journal* **22**, 1188-1198 (2003).

51. Poon, L.L., Pritlove, D.C., Fodor, E. & Brownlee, G.G. 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 (1999).

52. Bier, K., York, A. & Fodor, E. Cellular cap-binding proteins associate with influenza virus mRNAs. *The Journal of general virology* **92**, 1627-1634 (2011).

53. Yamayoshi, S., Watanabe, M., Goto, H. & Kawaoka, Y. Identification of a Novel Viral Protein Expressed from the PB2 Segment of Influenza A Virus. *Journal of virology* **90**, 444-456 (2016).

54. Vreede, F.T., Ng, A.K., Shaw, P.C. & Fodor, E. Stabilization of influenza virus replication intermediates is dependent on the RNA-binding but not the homo-oligomerization activity of the viral nucleoprotein. *Journal of virology* **85**, 12073-12078 (2011).

55. Vreede, F.T., Chan, A.Y., Sharps, J. & Fodor, E. Mechanisms and functional implications of the degradation of host RNA polymerase II in influenza virus infected cells. *Virology* **396**, 125-134 (2010).

56. York, A., Hengrung, N., Vreede, F.T., Huiskonen, J.T. & Fodor, E. 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**, E4238-4245 (2013).

57. Jorba, N., Coloma, R. & Ortin, J. Genetic trans-complementation establishes a new model for influenza virus RNA transcription and replication. *PLoS pathogens* **5**, e1000462 (2009).

58. Sugiyama, K., Kawaguchi, A., Okuwaki, M. & Nagata, K. pp32 and APRIL are host cell-derived regulators of influenza virus RNA synthesis from cRNA. *Elife* **4** (2015).

59. Te Velthuis, A.J., Robb, N.C., Kapanidis, A.N. & Fodor, E. The role of the priming loop in RNA synthesis. *Nat Microbiol* **1** (2016).

60. Deng, T., Vreede, F.T. & Brownlee, G.G. 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 (2006).

61. Resa-Infante, P., Recuero-Checa, M.A., Zamarreno, N., Llorca, O. & Ortin, J. Structural and functional characterization of an influenza virus RNA polymerase-genomic RNA complex. *Journal of virology* **84**, 10477-10487 (2010).

62. Baudin, F., Bach, C., Cusack, S. & Ruigrok, R.W. Structure of influenza virus RNP. I. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *The EMBO journal* **13**, 3158-3165 (1994).

63. Chan, W.H. *et al*. 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 (2010).

64. Hutchinson, E.C. *et al*. Mapping the phosphoproteome of influenza A and B viruses by mass spectrometry. *PLoS pathogens* **8**, e1002993 (2012).

65. Turrell, L., Hutchinson, E.C., Vreede, F.T. & Fodor, E. Regulation of influenza A virus nucleoprotein oligomerization by phosphorylation. *Journal of virology* **89**, 1452-1455 (2015).

66. Mondal, A., Potts, G.K., Dawson, A.R., Coon, J.J. & Mehle, A. Phosphorylation at the homotypic interface regulates nucleoprotein oligomerization and assembly of the influenza virus replication machinery. *PLoS pathogens* **11**, e1004826 (2015).

67. Kawaguchi, A., Momose, F. & Nagata, K. Replication-coupled and host factor-mediated encapsidation of the influenza virus genome by viral nucleoprotein. *Journal of virology* **85**, 6197-6204 (2011).

68. Naito, T. *et al*. An influenza virus replicon system in yeast identified Tat-SF1 as a stimulatory host factor for viral RNA synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 18235-18240 (2007).

69. Zhou, Z. *et al*. Fragile X mental retardation protein stimulates ribonucleoprotein assembly of influenza A virus. *Nature communications* **5**, 3259 (2014).

70. Perez, J.T. *et al*. A small-RNA enhancer of viral polymerase activity. *Journal of virology* **86**, 13475-13485 (2012).

71. Perez, J.T. *et al*. 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 (2010).

72. Vreede, F.T., Jung, T.E. & Brownlee, G.G. Model suggesting that replication of influenza virus is regulated by stabilization of replicative intermediates. *Journal of virology* **78**, 9568-9572 (2004).

73. Robb, N.C., Smith, M., Vreede, F.T. & Fodor, E. NS2/NEP protein regulates transcription and replication of the influenza virus RNA genome. *The Journal of general virology* **90**, 1398-1407 (2009).

74. Olson, A.C., Rosenblum, E. & Kuchta, R.D. Regulation of influenza RNA polymerase activity and the switch between replication and transcription by the concentrations of the vRNA 5' end, the cap source, and the polymerase. *Biochemistry* **49**, 10208-10215 (2010).

75. Paterson, D. & Fodor, E. Emerging roles for the influenza A virus nuclear export protein (NEP). *PLoS pathogens* **8**, e1003019 (2012).

76. Elton, D. *et al*. Interaction of the influenza virus nucleoprotein with the cellular CRM1-mediated nuclear export pathway. *Journal of virology* **75**, 408-419 (2001).

77. Wurzer, W.J. *et al*. Caspase 3 activation is essential for efficient influenza virus propagation. *The EMBO journal* **22**, 2717-2728 (2003).

78. Amorim, M.J. *et al*. A Rab11- and microtubule-dependent mechanism for cytoplasmic transport of influenza A virus viral RNA. *Journal of virology* **85**, 4143-4156 (2011).

79. Eisfeld, A.J., Kawakami, E., Watanabe, T., Neumann, G. & Kawaoka, Y. RAB11A is essential for transport of the influenza virus genome to the plasma membrane. *Journal of virology* **85**, 6117-6126 (2011).

80. Momose, F. *et al*. Apical transport of influenza A virus ribonucleoprotein requires Rab11-positive recycling endosome. *PloS one* **6**, e21123 (2011).

81. Hutchinson, E.C., von Kirchbach, J.C., Gog, J.R. & Digard, P. Genome packaging in influenza A virus. *The Journal of general virology* **91**, 313-328 (2010).

82. Gerber, M., Isel, C., Moules, V. & Marquet, R. Selective packaging of the influenza A genome and consequences for genetic reassortment. *Trends Microbiol* **22**, 446-455 (2014).

83. Chen, B.J., Leser, G.P., Morita, E. & Lamb, R.A. Influenza virus hemagglutinin and neuraminidase, but not the matrix protein, are required for assembly and budding of plasmid-derived virus-like particles. *Journal of virology* **81**, 7111-7123 (2007).

84. Takeda, M., Leser, G.P., Russell, C.J. & Lamb, R.A. Influenza virus hemagglutinin concentrates in lipid raft microdomains for efficient viral fusion. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 14610-14617 (2003).

85. Scheiffele, P., Roth, M.G. & Simons, K. Interaction of influenza virus haemagglutinin with sphingolipid-cholesterol membrane domains via its transmembrane domain. *The EMBO journal* **16**, 5501-5508 (1997).

86. Zhang, J., Leser, G.P., Pekosz, A. & Lamb, R.A. The cytoplasmic tails of the influenza virus spike glycoproteins are required for normal genome packaging. *Virology* **269**, 325-334 (2000).

87. Zhang, J., Pekosz, A. & Lamb, R.A. Influenza virus assembly and lipid raft microdomains: a role for the cytoplasmic tails of the spike glycoproteins. *Journal of virology* **74**, 4634-4644 (2000).

88. Jin, H., Leser, G.P., Zhang, J. & Lamb, R.A. Influenza virus hemagglutinin and neuraminidase cytoplasmic tails control particle shape. *The EMBO journal* **16**, 1236-1247 (1997).

89. Rossman, J.S., Jing, X., Leser, G.P. & Lamb, R.A. Influenza virus M2 protein mediates ESCRT-independent membrane scission. *Cell* **142**, 902-913 (2010).

90. Crotta, S. *et al*. Type I and type III interferons drive redundant amplification loops to induce a transcriptional signature in influenza-infected airway epithelia. *PLoS pathogens* **9**, e1003773 (2013).

91. Thompson, M.R., Kaminski, J.J., Kurt-Jones, E.A. & Fitzgerald, K.A. Pattern recognition receptors and the innate immune response to viral infection. *Viruses* **3**, 920-940 (2011).

92. Garcia-Sastre, A. Induction and evasion of type I interferon responses by influenza viruses. *Virus Res* **162**, 12-18 (2011).

93. Jensen, S. & Thomsen, A.R. Sensing of RNA viruses: a review of innate immune receptors involved in recognizing RNA virus invasion. *Journal of virology* **86**, 2900-2910 (2012).

94. Ishikawa, H. & Barber, G.N. STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature* **455**, 674-678 (2008).

95. Ma, Z. & Damania, B. The cGAS-STING Defense Pathway and Its Counteraction by Viruses. *Cell Host Microbe* **19**, 150-158 (2016).

96. Ishikawa, H., Ma, Z. & Barber, G.N. STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. *Nature* **461**, 788-792 (2009).

97. Goubau, D. *et al*. Antiviral immunity via RIG-I-mediated recognition of RNA bearing 5'-diphosphates. *Nature* **514**, 372-375 (2014).

98. Kato, H. *et al*. Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* **441**, 101-105 (2006).

99. Lee, M.K. *et al*. Structural features of influenza A virus panhandle RNA enabling the activation of RIG-I independently of 5'-triphosphate. *Nucleic acids research* (2016).

100. Kolakofsky, D., Kowalinski, E. & Cusack, S. A structure-based model of RIG-I activation. *RNA* **18**, 2118-2127 (2012).

101. Kowalinski, E. *et al*. Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA. *Cell* **147**, 423-435 (2011).

102. Anchisi, S., Guerra, J. & Garcin, D. RIG-I ATPase activity and discrimination of self-RNA versus non-self-RNA. *MBio* **6**, e02349 (2015).

103. Bamming, D. & Horvath, C.M. Regulation of signal transduction by enzymatically inactive antiviral RNA helicase proteins MDA5, RIG-I, and LGP2. *The Journal of biological chemistry* **284**, 9700-9712 (2009).

104. Gack, M.U. *et al*. TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. *Nature* **446**, 916-920 (2007).

105. Loo, Y.M. & Gale, M., Jr. Immune signaling by RIG-I-like receptors. *Immunity* **34**, 680-692 (2011).

106. Seth, R.B., Sun, L., Ea, C.K. & Chen, Z.J. Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. *Cell* **122**, 669-682 (2005).

107. Meylan, E. *et al*. Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. *Nature* **437**, 1167-1172 (2005).

108. Xu, L.G. *et al*. VISA is an adapter protein required for virus-triggered IFN-beta signaling. *Molecular cell* **19**, 727-740 (2005).

109. Kawai, T. *et al*. IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. *Nature immunology* **6**, 981-988 (2005).

110. Hou, F. *et al*. MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response. *Cell* **146**, 448-461 (2011).

111. Xu, H. *et al*. Structural basis for the prion-like MAVS filaments in antiviral innate immunity. *Elife* **3**, e01489 (2014).

112. Borden, E.C. *et al*. Interferons at age 50: past, current and future impact on biomedicine. *Nat Rev Drug Discov* **6**, 975-990 (2007).

113. Schneider, W.M., Chevillotte, M.D. & Rice, C.M. Interferon-stimulated genes: a complex web of host defenses. *Annu Rev Immunol* **32**, 513-545 (2014).

114. Rehwinkel, J. *et al*. RIG-I detects viral genomic RNA during negative-strand RNA virus infection. *Cell* **140**, 397-408 (2010).

115. Amini-Bavil-Olyaee, S. *et al*. The antiviral effector IFITM3 disrupts intracellular cholesterol homeostasis to block viral entry. *Cell Host Microbe* **13**, 452-464 (2013).

116. Desai, T.M. *et al*. IFITM3 restricts influenza A virus entry by blocking the formation of fusion pores following virus-endosome hemifusion. *PLoS pathogens* **10**, e1004048 (2014).

117. Haller, O., Frese, M. & Kochs, G. Mx proteins: mediators of innate resistance to RNA viruses. *Rev Sci Tech* **17**, 220-230 (1998).

118. Haller, O. & Kochs, G. Human MxA protein: an interferon-induced dynamin-like GTPase with broad antiviral activity. *J Interferon Cytokine Res* **31**, 79-87 (2011).

119. Verhelst, J., Parthoens, E., Schepens, B., Fiers, W. & Saelens, X. Interferon-inducible protein Mx1 inhibits influenza virus by interfering with functional viral ribonucleoprotein complex assembly. *Journal of virology* **86**, 13445-13455 (2012).

120. Turan, K. *et al*. Nuclear MxA proteins form a complex with influenza virus NP and inhibit the transcription of the engineered influenza virus genome. *Nucleic acids research* **32**, 643-652 (2004).

121. Xiao, H., Killip, M.J., Staeheli, P., Randall, R.E. & Jackson, D. The human interferon-induced MxA protein inhibits early stages of influenza A virus infection by retaining the incoming viral genome in the cytoplasm. *Journal of virology* **87**, 13053-13058 (2013).

122. Di Pietro, A. *et al*. TRIM22 inhibits influenza A virus infection by targeting the viral nucleoprotein for degradation. *Journal of virology* **87**, 4523-4533 (2013).

123. Wang, X., Hinson, E.R. & Cresswell, P. The interferon-inducible protein viperin inhibits influenza virus release by perturbing lipid rafts. *Cell Host Microbe* **2**, 96-105 (2007).

124. Garcia, M.A. *et al*. Impact of protein kinase PKR in cell biology: from antiviral to antiproliferative action. *Microbiol Mol Biol Rev* **70**, 1032-1060 (2006).

125. Tal, M.C. *et al*. Absence of autophagy results in reactive oxygen species-dependent amplification of RLR signaling. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 2770-2775 (2009).

126. Filomeni, G., De Zio, D. & Cecconi, F. Oxidative stress and autophagy: the clash between damage and metabolic needs. *Cell Death Differ* **22**, 377-388 (2015).

127. Tait, S.W. & Green, D.R. Mitochondria and cell death: outer membrane permeabilization and beyond. *Nat Rev Mol Cell Biol* **11**, 621-632 (2010).

128. Chen, H. & Chan, D.C. Physiological functions of mitochondrial fusion. *Ann N Y Acad Sci* **1201**, 21-25 (2010).

129. McBride, H.M., Neuspiel, M. & Wasiak, S. Mitochondria: more than just a powerhouse. *Curr Biol* **16**, R551-560 (2006).

130. Koshiba, T., Yasukawa, K., Yanagi, Y. & Kawabata, S. Mitochondrial membrane potential is required for MAVS-mediated antiviral signaling. *Science signaling* **4**, ra7 (2011).

131. Li, X.D., Sun, L., Seth, R.B., Pineda, G. & Chen, Z.J. Hepatitis C virus protease NS3/4A cleaves mitochondrial antiviral signaling protein off the mitochondria to evade innate immunity. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 17717-17722 (2005).

132. Lei, Y. *et al*. The mitochondrial proteins NLRX1 and TUFM form a complex that regulates type I interferon and autophagy. *Immunity* **36**, 933-946 (2012).

133. Moore, C.B. *et al*. NLRX1 is a regulator of mitochondrial antiviral immunity. *Nature* **451**, 573-577 (2008).

134. Castanier, C., Garcin, D., Vazquez, A. & Arnoult, D. Mitochondrial dynamics regulate the RIG-I-like receptor antiviral pathway. *EMBO Rep* **11**, 133-138 (2010).

135. Park, S. *et al*. The mitochondrial antiviral protein MAVS associates with NLRP3 and regulates its inflammasome activity. *Journal of immunology* **191**, 4358-4366 (2013).

136. Subramanian, N., Natarajan, K., Clatworthy, M.R., Wang, Z. & Germain, R.N. The adaptor MAVS promotes NLRP3 mitochondrial localization and inflammasome activation. *Cell* **153**, 348-361 (2013).

137. Zhou, R., Yazdi, A.S., Menu, P. & Tschopp, J. A role for mitochondria in NLRP3 inflammasome activation. *Nature* **469**, 221-225 (2011).

138. Killip, M.J., Fodor, E. & Randall, R.E. Influenza virus activation of the interferon system. *Virus Res* **209**, 11-22 (2015).

139. Weber, M. *et al*. Incoming RNA virus nucleocapsids containing a 5'-triphosphorylated genome activate RIG-I and antiviral signaling. *Cell Host Microbe* **13**, 336-346 (2013).

140. Killip, M.J., Smith, M., Jackson, D. & Randall, R.E. Activation of the interferon induction cascade by influenza a viruses requires viral RNA synthesis and nuclear export. *Journal of virology* **88**, 3942-3952 (2014).

141. Lahaye, X. *et al*. The capsids of HIV-1 and HIV-2 determine immune detection of the viral cDNA by the innate sensor cGAS in dendritic cells. *Immunity* **39**, 1132-1142 (2013).

142. Anchisi, S., Guerra, J., Mottet-Osman, G. & Garcin, D. Mismatches in the Influenza A Virus RNA Panhandle Prevent Retinoic Acid-Inducible Gene I (RIG-I) Sensing by Impairing RNA/RIG-I Complex Formation. *Journal of virology* **90**, 586-590 (2016).

143. Chen, S. *et al*. Heterocellular induction of interferon by negative-sense RNA viruses. *Virology* **407**, 247-255 (2010).

144. Killip, M.J. *et al*. Failure to activate the IFN-beta promoter by a paramyxovirus lacking an interferon antagonist. *Virology* **415**, 39-46 (2011).

145. Tapia, K. *et al*. Defective viral genomes arising in vivo provide critical danger signals for the triggering of lung antiviral immunity. *PLoS pathogens* **9**, e1003703 (2013).

146. Perez-Cidoncha, M. *et al*. An unbiased genetic screen reveals the polygenic nature of the influenza virus anti-interferon response. *Journal of virology* **88**, 4632-4646 (2014).

147. Saira, K. *et al*. Sequence analysis of in vivo defective interfering-like RNA of influenza A H1N1 pandemic virus. *Journal of virology* **87**, 8064-8074 (2013).

148. Jennings, P.A., Finch, J.T., Winter, G. & Robertson, J.S. Does the higher order structure of the influenza virus ribonucleoprotein guide sequence rearrangements in influenza viral RNA? *Cell* **34**, 619-627 (1983).

149. Marriott, A.C. & Dimmock, N.J. Defective interfering viruses and their potential as antiviral agents. *Rev Med Virol* **20**, 51-62 (2010).

150. Duhaut, S.D. & Dimmock, N.J. Heterologous protection of mice from a lethal human H1N1 influenza A virus infection by H3N8 equine defective interfering virus: comparison of defective RNA sequences isolated from the DI inoculum and mouse lung. *Virology* **248**, 241-253 (1998).

151. Odagiri, T. & Tobita, K. Mutation in NS2, a nonstructural protein of influenza A virus, extragenically causes aberrant replication and expression of the PA gene and leads to generation of defective interfering particles. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 5988-5992 (1990).

152. Fodor, E., Mingay, L.J., Crow, M., Deng, T. & Brownlee, G.G. A single amino acid mutation in the PA subunit of the influenza virus RNA polymerase promotes the generation of defective interfering RNAs. *Journal of virology* **77**, 5017-5020 (2003).

153. Frensing, T., Pflugmacher, A., Bachmann, M., Peschel, B. & Reichl, U. Impact of defective interfering particles on virus replication and antiviral host response in cell culture-based influenza vaccine production. *Appl Microbiol Biotechnol* **98**, 8999-9008 (2014).

154. Easton, A.J. *et al*. A novel broad-spectrum treatment for respiratory virus infections: influenza-based defective interfering virus provides protection against pneumovirus infection in vivo. *Vaccine* **29**, 2777-2784 (2011).

155. Scott, P.D., Meng, B., Marriott, A.C., Easton, A.J. & Dimmock, N.J. Defective interfering influenza A virus protects in vivo against disease caused by a heterologous influenza B virus. *The Journal of general virology* **92**, 2122-2132 (2011).

156. Baum, A., Sachidanandam, R. & Garcia-Sastre, A. Preference of RIG-I for short viral RNA molecules in infected cells revealed by next-generation sequencing. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 16303-16308 (2010).

157. Strahle, L., Garcin, D. & Kolakofsky, D. Sendai virus defective-interfering genomes and the activation of interferon-beta. *Virology* **351**, 101-111 (2006).

158. Killip, M.J. *et al*. Deep sequencing analysis of defective genomes of parainfluenza virus 5 and their role in interferon induction. *Journal of virology* **87**, 4798-4807 (2013).

159. Boergeling, Y. *et al*. Evidence for a Novel Mechanism of Influenza Virus-Induced Type I Interferon Expression by a Defective RNA-Encoded Protein. *PLoS pathogens* **11**, e1004924 (2015).

160. Hale, B.G., Randall, R.E., Ortin, J. & Jackson, D. The multifunctional NS1 protein of influenza A viruses. *The Journal of general virology* **89**, 2359-2376 (2008).

161. Gack, M.U. *et al*. Influenza A virus NS1 targets the ubiquitin ligase TRIM25 to evade recognition by the host viral RNA sensor RIG-I. *Cell Host Microbe* **5**, 439-449 (2009).

162. Nemeroff, M.E., Barabino, S.M., Li, Y., Keller, W. & Krug, R.M. Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. *Molecular cell* **1**, 991-1000 (1998).

163. Chen, Z., Li, Y. & Krug, R.M. Influenza A virus NS1 protein targets poly(A)-binding protein II of the cellular 3'-end processing machinery. *The EMBO journal* **18**, 2273-2283 (1999).

164. Chen, W. *et al*. A novel influenza A virus mitochondrial protein that induces cell death. *Nature medicine* **7**, 1306-1312 (2001).

165. Yamada, H., Chounan, R., Higashi, Y., Kurihara, N. & Kido, H. Mitochondrial targeting sequence of the influenza A virus PB1-F2 protein and its function in mitochondria. *FEBS letters* **578**, 331-336 (2004).

166. Dudek, S.E. *et al*. The influenza virus PB1-F2 protein has interferon antagonistic activity. *Biol Chem* **392**, 1135-1144 (2011).

167. Gibbs, J.S., Malide, D., Hornung, F., Bennink, J.R. & Yewdell, J.W. The influenza A virus PB1-F2 protein targets the inner mitochondrial membrane via a predicted basic amphipathic helix that disrupts mitochondrial function. *Journal of virology* **77**, 7214-7224 (2003).

168. Zamarin, D., Garcia-Sastre, A., Xiao, X., Wang, R. & Palese, P. Influenza virus PB1-F2 protein induces cell death through mitochondrial ANT3 and VDAC1. *PLoS pathogens* **1**, e4 (2005).

169. Chanturiya, A.N. *et al*. PB1-F2, an influenza A virus-encoded proapoptotic mitochondrial protein, creates variably sized pores in planar lipid membranes. *Journal of virology* **78**, 6304-6312 (2004).

170. Varga, Z.T., Grant, A., Manicassamy, B. & Palese, P. Influenza virus protein PB1-F2 inhibits the induction of type I interferon by binding to MAVS and decreasing mitochondrial membrane potential. *Journal of virology* **86**, 8359-8366 (2012).

171. Varga, Z.T. *et al*. The influenza virus protein PB1-F2 inhibits the induction of type I interferon at the level of the MAVS adaptor protein. *PLoS pathogens* **7**, e1002067 (2011).

172. Jagger, B.W. *et al*. An overlapping protein-coding region in influenza A virus segment 3 modulates the host response. *Science* **337**, 199-204 (2012).

173. Hu, J. *et al*. PA-X decreases the pathogenicity of highly pathogenic H5N1 influenza A virus in avian species by inhibiting virus replication and host response. *Journal of virology* **89**, 4126-4142 (2015).

174. Hayashi, T., MacDonald, L.A. & Takimoto, T. Influenza A Virus Protein PA-X Contributes to Viral Growth and Suppression of the Host Antiviral and Immune Responses. *Journal of virology* **89**, 6442-6452 (2015).

175. Carr, S.M., Carnero, E., Garcia-Sastre, A., Brownlee, G.G. & Fodor, E. Characterization of a mitochondrial-targeting signal in the PB2 protein of influenza viruses. *Virology* **344**, 492-508 (2006).

176. Graef, K.M. *et al*. The PB2 subunit of the influenza virus RNA polymerase affects virulence by interacting with the mitochondrial antiviral signaling protein and inhibiting expression of beta interferon. *Journal of virology* **84**, 8433-8445 (2010).

177. Iwai, A. *et al*. Influenza A virus polymerase inhibits type I interferon induction by binding to interferon beta promoter stimulator 1. *The Journal of biological chemistry* **285**, 32064-32074 (2010).

178. Shapira, S.D. *et al*. A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. *Cell* **139**, 1255-1267 (2009).

179. Nandi, S. *et al*. MAVS protein is attenuated by rotavirus nonstructural protein 1. *PloS one* **9**, e92126 (2014).

180. La Gruta, N.L. & Turner, S.J. T cell mediated immunity to influenza: mechanisms of viral control. *Trends Immunol* **35**, 396-402 (2014).

181. Skountzou, I. *et al*. Immunity to pre-1950 H1N1 influenza viruses confers cross-protection against the pandemic swine-origin 2009 A (H1N1) influenza virus. *Journal of immunology* **185**, 1642-1649 (2010).

182. Kreijtz, J.H., Fouchier, R.A. & Rimmelzwaan, G.F. Immune responses to influenza virus infection. *Virus Res* **162**, 19-30 (2011).

183. Ekiert, D.C. *et al*. A highly conserved neutralizing epitope on group 2 influenza A viruses. *Science* **333**, 843-850 (2011).

184. Worobey, M., Han, G.Z. & Rambaut, A. A synchronized global sweep of the internal genes of modern avian influenza virus. *Nature* **508**, 254-257 (2014).

185. Taubenberger, J.K., Reid, A.H., Krafft, A.E., Bijwaard, K.E. & Fanning, T.G. Initial genetic characterization of the 1918 "Spanish" influenza virus. *Science* **275**, 1793-1796 (1997).

186. Reid, A.H., Fanning, T.G., Hultin, J.V. & Taubenberger, J.K. Origin and evolution of the 1918 "Spanish" influenza virus hemagglutinin gene. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 1651-1656 (1999).

187. Tumpey, T.M. *et al*. Characterization of the reconstructed 1918 Spanish influenza pandemic virus. *Science* **310**, 77-80 (2005).

188. McAuley, J.L. *et al*. Expression of the 1918 influenza A virus PB1-F2 enhances the pathogenesis of viral and secondary bacterial pneumonia. *Cell Host Microbe* **2**, 240-249 (2007).

189. Reid, A.H., Fanning, T.G., Janczewski, T.A. & Taubenberger, J.K. Characterization of the 1918 "Spanish" influenza virus neuraminidase gene. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 6785-6790 (2000).

190. Gao, R. *et al*. Human infection with a novel avian-origin influenza A (H7N9) virus. *N Engl J Med* **368**, 1888-1897 (2013).

191. Herfst, S. *et al*. Airborne transmission of influenza A/H5N1 virus between ferrets. *Science* **336**, 1534-1541 (2012).

192. Imai, M. *et al*. Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. *Nature* **486**, 420-428 (2012).

193. Subbarao, E.K., London, W. & Murphy, B.R. A single amino acid in the PB2 gene of influenza A virus is a determinant of host range. *Journal of virology* **67**, 1761-1764 (1993).

194. Mehle, A. & Doudna, J.A. 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 (2009).

195. Bussey, K.A., Bousse, T.L., Desmet, E.A., Kim, B. & Takimoto, T. PB2 residue 271 plays a key role in enhanced polymerase activity of influenza A viruses in mammalian host cells. *Journal of virology* **84**, 4395-4406 (2010).

196. Labadie, K., Dos Santos Afonso, E., Rameix-Welti, M.A., van der Werf, S. & Naffakh, N. Host-range determinants on the PB2 protein of influenza A viruses control the interaction between the viral polymerase and nucleoprotein in human cells. *Virology* **362**, 271-282 (2007).

197. Ng, A.K. *et al*. Influenza polymerase activity correlates with the strength of interaction between nucleoprotein and PB2 through the host-specific residue K/E627. *PloS one* **7**, e36415 (2012).

198. Long, J.S. *et al*. Species difference in ANP32A underlies influenza A virus polymerase host restriction. *Nature* **529**, 101-104 (2016).

199. Moncorge, O., Mura, M. & Barclay, W.S. Evidence for avian and human host cell factors that affect the activity of influenza virus polymerase. *Journal of virology* **84**, 9978-9986 (2010).

200. Zell, R. *et al*. Prevalence of PB1-F2 of influenza A viruses. *The Journal of general virology* **88**, 536-546 (2007).

201. Krumbholz, A. *et al*. Current knowledge on PB1-F2 of influenza A viruses. *Medical microbiology and immunology* **200**, 69-75 (2011).

202. Yoshizumi, T. *et al*. Influenza A virus protein PB1-F2 translocates into mitochondria via Tom40 channels and impairs innate immunity. *Nature communications* **5**, 4713 (2014).

203. Miotto, O., Heiny, A., Tan, T.W., August, J.T. & Brusic, V. Identification of human-to-human transmissibility factors in PB2 proteins of influenza A by large-scale mutual information analysis. *BMC bioinformatics* **9 Suppl 1**, S18 (2008).

204. Fislova, T., Thomas, B., Graef, K.M. & Fodor, E. Association of the influenza virus RNA polymerase subunit PB2 with the host chaperonin CCT. *Journal of virology* **84**, 8691-8699 (2010).

205. Martell, J.D. *et al*. Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. *Nat Biotechnol* **30**, 1143-1148 (2012).

206. York, A., Hutchinson, E.C. & Fodor, E. Interactome analysis of the influenza A virus transcription/replication machinery identifies protein phosphatase 6 as a cellular factor required for efficient virus replication. *Journal of virology* **88**, 13284-13299 (2014).

207. Trudgian, D.C. *et al*. Comparative evaluation of label-free SINQ normalized spectral index quantitation in the central proteomics facilities pipeline. *Proteomics* **11**, 2790-2797 (2011).

208. Watanabe, T. *et al*. Influenza virus-host interactome screen as a platform for antiviral drug development. *Cell Host Microbe* **16**, 795-805 (2014).

209. Ge, Q. *et al*. RNA interference of influenza virus production by directly targeting mRNA for degradation and indirectly inhibiting all viral RNA transcription. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 2718-2723 (2003).

210. Fodor, E. & Smith, M. The PA subunit is required for efficient nuclear accumulation of the PB1 subunit of the influenza A virus RNA polymerase complex. *Journal of virology* **78**, 9144-9153 (2004).

211. Patel, D., Schultz, L.W. & Umland, T.C. Influenza A polymerase subunit PB2 possesses overlapping binding sites for polymerase subunit PB1 and human MAVS proteins. *Virus Res* **172**, 75-80 (2013).

212. von Heijne, G. Mitochondrial targeting sequences may form amphiphilic helices. *The EMBO journal* **5**, 1335-1342 (1986).

213. Neupert, W. & Herrmann, J.M. Translocation of proteins into mitochondria. *Annual review of biochemistry* **76**, 723-749 (2007).

214. Schulz, C., Schendzielorz, A. & Rehling, P. Unlocking the presequence import pathway. *Trends Cell Biol* **25**, 265-275 (2015).

215. Momose, F. *et al*. Identification of Hsp90 as a stimulatory host factor involved in influenza virus RNA synthesis. *J Biol Chem* **277**, 45306-45314 (2002).

216. Deng, T., Sharps, J., Fodor, E. & Brownlee, G.G. 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 (2005).

217. Manzoor, R. *et al*. Heat shock protein 70 modulates influenza A virus polymerase activity. *J Biol Chem* **289**, 7599-7614 (2014).

218. Westermann, B. Bioenergetic role of mitochondrial fusion and fission. *Biochimica et biophysica acta* **1817**, 1833-1838 (2012).

219. Yasukawa, K. *et al*. Mitofusin 2 inhibits mitochondrial antiviral signaling. *Science signaling* **2**, ra47 (2009).

220. Arnoult, D. Mitochondrial fragmentation in apoptosis. *Trends Cell Biol* **17**, 6-12 (2007).

221. Bhola, P.D. & Letai, A. Mitochondria-Judges and Executioners of Cell Death Sentences. *Molecular cell* **61**, 695-704 (2016).

222. Weinberg, S.E., Sena, L.A. & Chandel, N.S. Mitochondria in the regulation of innate and adaptive immunity. *Immunity* **42**, 406-417 (2015).

223. Anand, S.K. & Tikoo, S.K. Viruses as modulators of mitochondrial functions. *Advances in virology* **2013**, 738794 (2013).

224. Geng, X. *et al*. Hepatitis B virus X protein targets Bcl-2 proteins to increase intracellular calcium, required for virus replication and cell death induction. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 18471-18476 (2012).

225. Perez, D. & White, E. TNF-alpha signals apoptosis through a bid-dependent conformational change in Bax that is inhibited by E1B 19K. *Molecular cell* **6**, 53-63 (2000).

226. Kim, J.H. *et al*. Role of host-specific amino acids in the pathogenicity of avian H5N1 influenza viruses in mice. *The Journal of general virology* **91**, 1284-1289 (2010).

227. Wu, X. *et al*. Mitochondrial proteomic analysis of human host cells infected with H3N2 swine influenza virus. *J Proteomics* **91**, 136-150 (2013).

228. Munday, D.C., Howell, G., Barr, J.N. & Hiscox, J.A. Proteomic analysis of mitochondria in respiratory epithelial cells infected with human respiratory syncytial virus and functional implications for virus and cell biology. *J Pharm Pharmacol* **67**, 300-318 (2015).

229. Rebsamen, M. *et al*. NLRX1/NOD5 deficiency does not affect MAVS signalling. *Cell Death and Differentiation* **18**, 1387-1387 (2011).

230. Arnoult, D. *et al*. An N-terminal addressing sequence targets NLRX1 to the mitochondrial matrix. *J Cell Sci* **122**, 3161-3168 (2009).

231. Jaworska, J. *et al*. NLRX1 prevents mitochondrial induced apoptosis and enhances macrophage antiviral immunity by interacting with influenza virus PB1-F2 protein. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E2110-2119 (2014).

232. Cheng, M.Y., Hartl, F.U. & Horwich, A.L. The mitochondrial chaperonin hsp60 is required for its own assembly. *Nature* **348**, 455-458 (1990).

233. Del Arco, A., Agudo, M. & Satrustegui, J. Characterization of a second member of the subfamily of calcium-binding mitochondrial carriers expressed in human non-excitable tissues. *Biochem J* **345 Pt 3**, 725-732 (2000).

234. Akram, M. Citric acid cycle and role of its intermediates in metabolism. *Cell Biochem Biophys* **68**, 475-478 (2014).

235. Kohler, A. & Hurt, E. Exporting RNA from the nucleus to the cytoplasm. *Nat Rev Mol Cell Biol* **8**, 761-773 (2007).

236. Nagarajan, V.K., Jones, C.I., Newbury, S.F. & Green, P.J. XRN 5'-->3' exoribonucleases: structure, mechanisms and functions. *Biochimica et biophysica acta* **1829**, 590-603 (2013).

237. Makela, S.M. *et al*. RIG-I Signaling Is Essential for Influenza B Virus-Induced Rapid Interferon Gene Expression. *Journal of virology* **89**, 12014-12025 (2015).

238. Weber, M. *et al*. Influenza virus adaptation PB2-627K modulates nucleocapsid inhibition by the pathogen sensor RIG-I. *Cell Host Microbe* **17**, 309-319 (2015).

239. Laske, T., Heldt, F.S., Hoffmann, H., Frensing, T. & Reichl, U. Modeling the intracellular replication of influenza A virus in the presence of defective interfering RNAs. *Virus Res* **213**, 90-99 (2016).

240. Poeck, H. *et al*. Recognition of RNA virus by RIG-I results in activation of CARD9 and inflammasome signaling for interleukin 1 beta production. *Nature immunology* **11**, 63-69 (2010).

241. Pothlichet, J. *et al*. Type I IFN triggers RIG-I/TLR3/NLRP3-dependent inflammasome activation in influenza A virus infected cells. *PLoS pathogens* **9**, e1003256 (2013).

242. Hornung, V. *et al*. 5'-Triphosphate RNA is the ligand for RIG-I. *Science* **314**, 994-997 (2006).

243. Fodor, E. *et al*. 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 (2002).

244. Paterson, D., te Velthuis, A.J., Vreede, F.T. & Fodor, E. Host restriction of influenza virus polymerase activity by PB2 627E is diminished on short viral templates in a nucleoprotein-independent manner. *Journal of virology* **88**, 339-344 (2014).

245. Jonges, M. *et al*. Emergence of the virulence-associated PB2 E627K substitution in a fatal human case of highly pathogenic avian influenza virus A(H7N7) infection as determined by Illumina ultra-deep sequencing. *Journal of virology* **88**, 1694-1702 (2014).

246. Kao, R.Y. *et al*. Identification of influenza A nucleoprotein as an antiviral target. *Nat Biotechnol* **28**, 600-605 (2010).