

DISSECTING MECHANISMS OF INTERSEGMENT EPISTASIS AND ITS ROLE IN
INFLUENZA A VIRUS BIOLOGY

BY

MEGHAN DIEFENBACHER

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Microbiology
in the Graduate College of the
University of Illinois Urbana-Champaign, 2022

Urbana, Illinois

Doctoral Committee:

Associate Professor Christopher Brooke, Chair
Professor Rachel Whitaker
Associate Professor Thomas Kehl-Fie
Professor Steven Blanke

ABSTRACT

Influenza A virus (IAV) is a significant respiratory pathogen that causes hundreds of thousands of deaths worldwide each year. The IAV genome is divided into eight negative-sense, single-stranded RNA segments. Genetic interactions within and between the eight segments and their encoded proteins, otherwise known as intersegment epistasis, play an important role in modulating IAV replication and evolution. However, the current understanding of the extent and nature of these epistatic interactions as well as their mechanisms remains incomplete. Here, I aim to investigate mechanisms by which intersegment epistasis modulates IAV replication and the genomic structure of IAV populations.

The eight gene segments are replicated and transcribed in the context of viral ribonucleoprotein complexes (vRNPs) in which each segment is encapsidated along its length by the viral nucleoprotein (NP) and is associated with a single copy of the viral RNA dependent-RNA polymerase (RdRp). Each segment exhibits its own unique level and temporal pattern of expression, however, the exact mechanisms underlying the independent regulation of individual gene segment expression are not well understood. I previously demonstrated that a single substitution in NP (F346S) selectively modulates neuraminidase (NA) gene segment expression while leaving other segments largely unaffected. The specificity of this effect for the NA segment is viral strain specific and depends on the UTR sequence of the NA segment. While the NP:F346S substitution did not significantly affect the ability of NP to bind to and oligomerize along the viral RNAs, it specifically decreased the ability of NP to promote NA segment vRNA

synthesis. In addition to NP residue F346, I identified two other adjacent aromatic residues in NP (Y385 & F479) capable of selectively regulating NA gene segment expression, suggesting a larger role for this domain in gene-segment specific regulation. These findings demonstrate a novel role for interactions between the gene segment UTRs and NP in regulating IAV gene segment expression.

RNA structures are proposed to play an important role in regulating IAV gene segment expression and packaging; however, their study has previously been limited to *in vitro* or *in silico* approaches that predicted structures present in naked IAV RNAs which were likely not representative of the structural conformations that would be present in the RNAs in the context of vRNPs in infected cells and virions. Using the selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) protocol, I was able to generate a comprehensive map of the RNA structural profile of the eight gene segments in the context of vRNPs isolated from virions purified from the A/Puerto Rico/8/1934 (PR8) and PR8 NP:F346S viruses. Each segment exhibited its own unique RNA structural profile. Initially, I used this data to determine whether the NP:F346S substitution affects the NP binding distribution along the NA segment and leads to the differential formation of RNA structures important for NP-dependent regulation. However, the RNA structural profiles of the NA segment were highly similar in the presence of NP:WT and NP:F346S, and introducing silent substitutions to disrupt predicted RNA structures in specific regions of the segment had no effect on the NP:F346S phenotype. Using this data set, I aim to

perform future experiments to probe the roles that the observed RNA structures in the eight segments and interactions between them may play in regulating gene segment expression, packaging, and the production of aberrant RNA replication species (DIs, mvRNAs etc...).

Only a small fraction of particles in IAV populations are fully infectious with the remaining portion consisting primarily of semi-infectious particles (SIPs) which fail to express one or more viral proteins. As SIPs lack one or more segments, they require co-infection with other IAV particles in order to obtain the complete complement of segments required for a productive infection. IAV strains can vary significantly in terms of the fraction of their population that exists as SIPs, suggesting that there is a genetic determinant underlying SIP production. However, the exact identity of the gene segment(s) that determine the SIP production phenotype of different IAV strains remains unknown. Using high SIP producing (A/Fort Monmouth/1/1947 (H1N1)) and low SIP producing (A/California/04/2009 (H1N1)) strains of IAV, I generated reassortant viruses containing a distinct set of gene segments from each parental strain and quantified the ratio of total particles to infectious particles for each reassortant using the hemagglutination assay and plaque assay respectively. I found that the SIP production phenotype was linked to the genetic background of the viral replicase complex (vRNP). These findings are suggestive of a potential role for gene segment replication efficiency and/or packaging in determining the fraction of IAV particles that are fully or semi-infectious.

My finding that the NP:F346S substitution can selectively regulate the expression of individual gene segments via interactions with the UTRs demonstrates a novel mechanism of intersegment gene expression regulation in IAV by which a segment's RNA sequence can influence its inherit susceptibility to regulation by the same viral protein. In addition, my finding that specific RNA structures within the NA segment influence its expression levels enhances our understanding of the role that RNA structures play in IAV biology. Finally, my finding that the vRNP complex genotype influences an IAV strain's SIP production phenotype suggests that the baseline efficiency of gene segment replication regulates SIP formation frequency by either modulating the efficiency by which IAV can successfully complete all steps of its replication cycle and evade the host immune response, or modulating the fraction of gene segments that get incorporated into viral particles. Altogether, these studies have important consequences for further developing our understanding of the factors that drive the replication and evolution of IAV.

To my family, for their love and support throughout my PhD journey.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Chris Brooke, my committee members, Dr. Rachel Whitaker and Dr. Thomas Kehl-Fie, and Diane Tsevelekos for taking the time to review and provide feedback on this thesis.

Additionally, I would like to thank Sonya Kumar Bharathkar for assistance with mammalian protein expression in HEK Expi-293-F cells, and Timothy J.C. Tan, Dr. Nicholas C. Wu, Sarah Leonard, and Dr. Beth Stadtmueller for assistance with and helpful discussions regarding NP-RNA binding experiments. I would also like to thank Dr. Andrew Mehle for helpful discussions regarding NP co-immunoprecipitation experiments. Finally, I would like to thank Dr. David L.V. Bauer for assistance with performing the SHAPE-MaP experiment and with data analysis.

This work has been generously supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health under awards K22AI116588 and R01AI139246, the Roy J. Carver Charitable Trust under award 17-4905, the Francis Crick Institute (which receives its core funding from Cancer Research UK) (FC011104), the UK Medical Research Council (FC011104), the Wellcome Trust (FC011104), and startup funds from the University of Illinois.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION	1
CHAPTER 2: INTERACTIONS BETWEEN INFLUENZA A VIRUS NUCLEOPROTEIN AND GENE SEGMENT UTRS FACILITATE SELECTIVE MODULATION OF VIRAL GENE EXPRESSION	32
CHAPTER 3: EVALUATING THE ROLE OF RNA STRUCTURES AND SEQUENCE MOTIFS IN IAV GENE SEGMENT EXPRESSION.....	86
CHAPTER 4: IDENTIFICATION OF THE GENETIC DETERMINANTS OF SEMI-INFECTIONOUS PARTICLE PRODUCTION IN INFLUENZA A VIRUS	106
CHAPTER 5: CONCLUSIONS AND DISCUSSION	118
REFERENCES.....	123

CHAPTER 1: INTRODUCTION

1.1 Influenza A Virus Significance and Overview

Significance to public health

Influenza A virus (IAV) is a respiratory pathogen that remains a significant threat to human health, causing hundreds of thousands of deaths worldwide each year (Iuliano et al., 2018). IAV infections can range from mild, upper respiratory tract infections with symptoms such as fever, sore throat, runny nose, cough, headache, muscle pain, and fatigue, to severe, lower respiratory tract infections with patients developing severe or lethal pneumonia as a result of the infection or indirectly through bacterial secondary infections (Krammer et al., 2018). Vaccines against IAV are available and have been shown to reduce the public health burden of seasonal Influenza by decreasing hospitalizations and deaths (CDC, 2020). However, the vaccines have relatively low efficacy that varies between ~10-60% each year ("Past Seasons Vaccine Effectiveness Estimates | CDC," 2022), and need to be continually updated in order to keep pace with the changing antigenicity of IAV that results from the continuous accumulation of antigenic escape substitutions in the viral population (antigenic drift). Antiviral drugs like neuraminidase inhibitors (NAIs) (Oseltamivir, Zanamivir, Peramivir etc...) and viral RNA-dependent RNA polymerase inhibitors (Baloxavir) are available for patients to take to reduce symptoms and disease severity, however, they must be applied early (<48hrs) during the course of infection, and IAV can develop resistance to these drugs (Lampejo, 2020; Świerczyńska et al., 2022). Compounding on the public health burden caused by seasonal IAV infections is

the possibility of pandemics caused by IAV strains that emerge from zoonotic reservoirs and are antigenically distinct from the seasonal strains (antigenic shift), allowing them to rapidly spread through the human population due to lack of pre-existing immunity, resulting in substantial increases in morbidity and mortality (Krammer et al., 2018). Over the past one hundred years there have been four major pandemics in 1918, 1957, 1968, and 2009 (Krammer et al., 2018). The deadliest pandemic, known as the “Spanish Flu,” occurred in 1918 and caused over 50 million deaths worldwide (Johnson and Mueller, 2002). As the pandemic IAVs are antigenically distinct from the seasonal IAVs, the seasonal vaccines do not protect against the pandemic viruses, leaving the antiviral drugs as the first line of defense against IAV pandemics (Krammer et al., 2018). Due to the substantial public health burden caused by seasonal epidemics of IAV and the potential for sporadic and deadly IAV pandemics, as well as the lack of highly effective antiviral drugs and vaccines, further research into the factors influencing IAV replication and evolution is needed to develop more effective therapeutics against IAV infections.

IAV particle and genome structure

IAV is a member of the family *Orthomyxoviridae* and has a genome divided into eight, negative-sense, single-stranded RNA segments that each encode one or more viral proteins (Fodor and Velthuis, 2019). The gene segments encode more than one protein through various mechanisms such as splicing (M and NS), alternative initiation, or ribosomal frameshifting (Dou et al., 2018). Each segment consists of an open reading frame (ORF) flanked by untranslated regions (UTRs)

consisting of segment specific ORF proximal regions and two highly conserved and partially complementary ends which interact with one another to form the viral promoter (Bae et al., 2001; Desselberger et al., 1980; Fodor and Velthuis, 2019). The UTRs and the ends of the ORFs form the packaging signals required for incorporation of the complete set of eight segments into virions (Fujii et al., 2005; Goto et al., 2013). Within the virion, each gene segment is maintained as a helical viral ribonucleoprotein complex (vRNP) in which the NP-encapsidated segment forms a filament that is closed by a loop at one end and bound by the viral RNA-dependent RNA polymerase (RdRp) at the opposite end (Arranz et al., 2012). The eight vRNPs are incorporated into viral particles in a 7+1 organization in which the remaining seven segments are arranged around a single segment (Noda et al., 2006). The viral membrane is supported underneath by the matrix (M1) protein, and the viral surface glycoproteins hemagglutinin (HA) and neuraminidase (NA), and the M2 ion channel protein are found embedded in the viral membrane (Krammer et al., 2018). HA promotes virion attachment and entry, while NA promotes virion release and escape from the host mucus layer (Byrd-Leotis et al., 2017). The ratio of HA to NA in virions varies between different IAV strains, but typically HA is present at higher abundances (~300-400 HA trimers compared to ~30-50 NA tetramers) (Harris et al., 2006). The distribution of HA and NA across the virion surface is uneven with NA clustered in particular locations (Vahey and Fletcher, 2019). This is thought to allow the virions to penetrate the host mucus layer while still being able to attach to host cells (Vahey and Fletcher, 2019). IAV virions are around 120nm in diameter and

exhibit either a spherical or filamentous morphology (Chu et al., 1949; Harris et al., 2006). The morphology of IAV virions varies depending upon the IAV strain and its passage history (lab adapted/passaged v. clinical isolate) (Choppin et al., 1960; Chu et al., 1949).

1.2 IAV Replication Cycle

Receptor specificity

The receptors for IAV are sialic acids (SAs) which are a class of carbohydrates characterized by a nine-carbon backbone that are present as the terminal carbohydrate moiety on host glycoproteins and glycolipids (Karakus et al., 2020). SAs are linked to the penultimate galactose residue via two different types of linkages: α 2-3 (*trans* conformation) and α 2-6 (*cis* conformation) (Shi et al., 2014). These different linkages cause SAs to adopt distinct structural conformations which necessitate changes in the IAV HA receptor binding pocket to alter its specificity (Shi et al., 2014). The varying receptor specificity of different IAV HAs for α 2-3-linked v. α 2-6-linked SAs is driven by the types and distributions of the receptors in their host species. For instance, avian IAVs preferentially bind to α 2,3-linked SAs while human IAVs bind to α 2,6-linked SAs which are more abundant in the areas where IAV replicates in those species- the epithelial cells of the intestine and upper respiratory tract (URT) respectively (Long et al., 2019). Previous pandemic IAV strains acquired mutations that altered the receptor specificity of HA from α 2,3-linked SAs to α 2,6-linked SAs (E190D/G225D for H1N1, and Q226L/G228S for H2N2 and H3N2) (Matrosovich et al., 2000; Pappas et al., 2010; Tumpey et al., 2007), and highly pathogenic avian IAV

strains, such as H5N1, require mutations altering HA receptor specificity to $\alpha 2,6$ -linked SAs in order to efficiently transmit between mammalian hosts (Herfst et al., 2012). Consequently, HA receptor specificity is hypothesized to be one of the host restriction factors that zoonotic IAV strains must overcome in order to efficiently replicate and transmit within the human population.

Receptor binding and cell entry

The IAV surface glycoprotein HA is necessary for both viral attachment and entry into host cells (Wu and Wilson, 2019). On viral particles, HA exists as a trimer and each individual monomer can bind the SA receptor (Wu and Wilson, 2019). Each HA monomer exhibits a very weak binding affinity ($\sim 0.1\text{mM}$) to SAs, so multiple interactions between HA trimers in the virion and SAs are required to increase the overall avidity and allow for sufficient attachment to host cells (Böttcher-Friebertshäuser et al., 2014). Once HA binds to the SA receptors, the virion is brought into the cells by receptor-mediated endocytosis (Dou et al., 2018). As the endosome acidifies, HA undergoes a drastic conformational rearrangement in which its hydrophobic fusion peptide inserts into the endosomal membrane and brings the viral and endosomal membranes closer together until they fuse and release the viral gene segments (in the form of vRNPs) into the cytoplasm (Dou et al., 2018; Wu and Wilson, 2019). The pH of fusion varies for different IAV strains (pH of ~ 5 -6) with human IAV HAs preferring more acidic conditions (DuBois et al., 2011; Galloway et al., 2013). Therefore, the pH of fusion of an HA is considered to be one of the important factors for determining the host range of different IAV strains. For HA to be fusion

competent, the original form of HA after translation (HA0) must be cleaved by host cell proteases, such as TMPRSS2 and HAT, into two domains (HA1 and HA2) which are connected by a single disulfide linkage in order to expose the fusion peptide (Böttcher-Friebertshäuser et al., 2014; Wu and Wilson, 2019). The major cleavage site for IAV HAs is between a conserved arginine-glycine bond, however, some highly pathogenic avian strains have a polybasic cleavage site (RKRR) which can be cleaved by furin proteases which are ubiquitously expressed allowing those strains to replicate in a wider range of tissues and cause more severe disease (Böttcher-Friebertshäuser et al., 2014). Therefore, the HA cleavage sequence of an IAV is considered a virulence marker (Böttcher-Friebertshäuser et al., 2014). Accompanying the conformational change in HA, protons enter the virion through the M2 ion channel, resulting in the acidification of the virion interior (Dou et al., 2018). This is necessary to dissociate the vRNPs from the virion matrix layer so that they are released into the host cytoplasm after the fusion between the viral and endosomal membranes (Dou et al., 2018).

Nuclear import

Once the vRNPs are released into the cytoplasm, they are trafficked to the nucleus via the host importin α/β nuclear import machinery (Eisfeld et al., 2015). While all of the protein components of the vRNP complex possess nuclear localization signals, NP, possessing a surface exposed, non-conventional NLS at its amino terminus (Wang et al., 1997), is considered to be the main driver of vRNP import as ~500 copies of each monomer are present in each vRNP complex as opposed to a single copy of each protein (PB2/PB1/PA) in the viral

RdRp (Tome-Amat et al., 2019). Whether the vRNPs remain in a complex or are transported individually to the nucleus has not been definitively determined, however, a study using smFISH to quantify the co-localization of individual vRNPs at early timepoints during infection found that the vRNPs entered the nucleus as a complex and that there was little dissociation of individual vRNPs prior to nuclear import (Chou et al., 2013).

Gene segment transcription and replication

Once the vRNPs enter the nucleus, the gene segments undergo primary transcription during which the negative sense viral genomic RNAs (vRNAs) present in the imported vRNP complexes are transcribed into positive sense mRNAs (Fodor and Velthuis, 2019). Transcription of the vRNAs occurs *in cis* i.e. catalyzed by the viral RdRp present in the vRNP complex (Fodor and Velthuis, 2019). The viral RdRp associates with the phosphorylated C-terminal domain of actively transcribing host RNA polymerase II (Martínez-Alonso et al., 2016), and through a process called cap-snatching, binds to the 5' cap of host mRNAs via its PB2 subunit and then cleaves 10-14 nucleotides downstream of the cap via its PA subunit (Fodor and Velthuis, 2019). The capped host RNA is then used as a primer for the synthesis of the viral mRNA which is catalyzed by the PB1 subunit of the RdRp (Fodor and Velthuis, 2019). Transcription terminates via the stuttering of the RdRp at the poly U tract at the 5' end of the vRNA, leading to the addition of a poly A tail to the viral mRNA (Fodor and Velthuis, 2019; Poon et al., 1999). The viral mRNAs are then exported from the nucleus and translated into proteins (Eisfeld et al., 2015). The mechanism that results in the RdRp switching

its function from a transcriptase to a replicase is unknown, as *in vitro* both mRNA and cRNA can be generated from vRNP complexes (Vreede and Brownlee, 2007). However, it is hypothesized that the translation and re-import of the RdRp proteins (PB2, PB1, and PA) and NP into the nucleus stabilizes the nascent cRNA so that it can direct replication of the vRNA template (Vreede et al., 2004).

The first stage of gene segment replication involves the synthesis of the positive sense complementary intermediate (cRNA) from the vRNA (Fodor and Velthuis, 2019). cRNA synthesis is catalyzed by the resident (vRNP-associated) RdRp and is primer-independent and initiated at the 3' end of the vRNA (Deng et al., 2006). The PB1 subunit catalyzes the elongation of the cRNA (Fodor and Velthuis, 2019). As the cRNA exits from the polymerase during elongation, it is encapsidated by the viral RdRp and NP to form cRNP complexes (Fodor and Velthuis, 2019). The second stage of gene segment replication involves the synthesis of vRNA using the cRNA as a template. In contrast to cRNA synthesis, vRNA synthesis is initiated internally at the 4th and 5th nucleotides of the 3' end of the cRNA (UC) resulting in the synthesis of an AG dinucleotide (Deng et al., 2006). The RdRp then backtracks to the 3' end of the cRNA and uses the AG nucleotide as a primer to begin the elongation of the vRNA (Deng et al., 2006). Similar to cRNAs, vRNAs are encapsidated by the RdRp and NP as they exit from the RdRp to form vRNPs (Fodor and Velthuis, 2019). Unlike cRNA synthesis which only requires the resident RdRp, vRNA synthesis requires dimerization between a regulatory polymerase and the resident polymerase (Fan et al., 2019; York et al., 2013). The dimerization of the polymerase has been

shown to promote backtracking and internal initiation by the RdRp, allowing for vRNA synthesis to occur (Fan et al., 2019). In addition to RdRp dimerization, the viral non-structural protein 2 (NS2) is also hypothesized to play a role in promoting viral gene segment replication. NS2 promotes cRNA synthesis and the synthesis of small viral RNAs (svRNAs), which are 22-27 nucleotide long RNAs corresponding to the 5' end of the vRNA and generated from the cRNA template, that are hypothesized promote replication in a segment-specific manner by binding to RdRps and converting them into replicases (Perez et al., 2010, 2012).

Temporal regulation of the expression of IAV RNA species (mRNA, cRNA, vRNA)

The three viral RNAs generated from each individual gene segment (mRNA, cRNA, vRNA) exhibit distinct temporal patterns of expression over the course of infection (Hatada et al., 1989; Phan et al., 2021; Shapiro et al., 1987). Viral mRNA begins to accumulate 1-2 hours post infection (hpi), peaks around 3-5hpi, and then decays over the remaining course of the infection (Hatada et al., 1989; Phan et al., 2021; Shapiro et al., 1987). The initial burst of transcription results from primary transcription by the imported vRNP complexes, while the second phase results from secondary transcription from newly replicated vRNP complexes whose replication was stimulated by the newly translated RdRp and NP proteins directed by the primary viral mRNA transcripts (Hatada et al., 1989; Shapiro et al., 1987). Protein synthesis is highly correlated with viral mRNA synthesis during the early phase of infection (1-4hrs), however, it becomes uncoupled during the later timepoints of infection with viral protein levels

continuing to increase while viral mRNA levels begin to decay (Shapiro et al., 1987). One possible explanation for the discrepancy between viral mRNA and protein synthesis during later timepoints of infection is that translation is directed by pre-existing instead of newly transcribed viral mRNA (Shapiro et al., 1987). In contrast to mRNA, vRNA synthesis begins to rapidly increase 2-4hpi and then continues to increase at a slower rate over the remaining course of the infection (Hatada et al., 1989; Phan et al., 2021; Shapiro et al., 1987). cRNA also begins to accumulate between 2-4hpi, but at a lower level compared to the vRNA, reaching a peak around 3-4hpi, and then stabilizing over the remaining course of infection (Hatada et al., 1989; Phan et al., 2021; Shapiro et al., 1987). The initial delay in vRNA and cRNA synthesis is thought to result from the requirement of newly translated RdRp and NP from viral mRNAs generated via primary transcription to stabilize the cRNA replicative intermediate (Hatada et al., 1989; Vreede et al., 2004). The mechanisms behind the differential patterns of accumulation of the viral mRNA, vRNA, and cRNA during later timepoints of infection remain unknown.

Heterogeneity in the temporal patterns and overall levels of expression of individual IAV gene segments

Further increasing the complexity involved in piecing together the mechanism(s) behind the coordinated regulation of IAV gene segment expression is the observation that in addition to the differences in the temporal expression patterns of the viral mRNA, vRNA, and cRNA, each individual gene segment exhibits its own unique level and temporal pattern of expression (Phan et al., 2021; Shapiro

et al., 1987). The eight gene segments are grouped into two different classes based on the temporal expression patterns of their encoded protein(s)- early (PB2, PB1, PA, NP, NS) and late (HA, NA, M1, M2) (Hatada et al., 1989; Shapiro et al., 1987). One hypothesis for the observed differences in the temporal expression patterns is the variation in the efficiency of nuclear export for each viral mRNA. The type of nuclear export machinery required for efficient nuclear export varies for each viral mRNA (Read and Digard, 2010). Several late viral transcripts (HA, M1, M2) utilize the cellular NXF1/TAP pathway for nuclear export, and require the NS1 protein to be synthesized beforehand to act as an adaptor for the recruitment of the cellular export machinery (Pereira et al., 2017; Read and Digard, 2010). A second hypothesis is that specific sequences within the viral mRNAs determine their translation efficiency. In the IAV strain A/WSN/1933 (H1N1) (WSN), the PB1 and NA mRNAs are predicted to have a suboptimal Kozak sequence next to their initiation codon and, therefore, lower levels of translation (Maeda et al., 2004; Wang et al., 2015). However, introducing mutations to optimize the Kozak sequence had conflicting results based on whether a native or reporter gene segment was used and the timepoint examined (Maeda et al., 2004; Wang et al., 2015), suggesting that additional sequence features of the mRNA and/or other viral or host factors are involved in modulating translation efficiency.

In addition to differences in the temporal expression patterns of their mRNAs and proteins, each gene segment exhibits different overall levels of expression and varies in the ratios of the different viral RNA species (mRNA v. vRNA v. cRNA)

produced (Phan et al., 2021; Shapiro et al., 1987). The segment UTRs play a major role in regulating the expression from individual gene segments (Lee et al., 2003; Lee and Seong, 1998; Ma et al., 2013; J. Wang, et al., 2016; Widjaja et al., 2012; Xiao et al., 2021.; Zheng et al., 1996). During infection, each gene segment must compete for limiting amounts of the viral RdRp for replication, therefore, segments whose UTRs more efficiently recruit the viral RdRp will be preferentially synthesized (Widjaja et al., 2012). Using a reporter system, Widjaja et al. found that the UTR sequences of the gene segments influenced their competitiveness for access to the viral RdRp under conditions where the RdRp was limiting (Widjaja et al., 2012). In addition, they found that introducing mutations in the UTR to increase the degree of base-pairing within the viral promoter (formed via base pairing interactions between the 3' 1-12 and 5' 1-13 nucleotides of the vRNA) made the segments more competitive (Widjaja et al., 2012). Other studies have also demonstrated the importance of the viral promoter structure in regulating viral gene segment expression (Lee et al., 2003; Lee and Seong, 1998). The promoter region is highly conserved in different IAV strains with the exception of the identity of the nucleotide at position 4 of the 3' UTR of the vRNA which varies (U or C) for different segments and between different IAV strains (Desselberger et al., 1980; Wang and Lee, 2009). The identity of the nucleotide at position 4 determines whether or not a base-pairing interaction occurs within the promoter between position 4 of the 3' UTR and position 4 of the 5' UTR, with a U but not C being able to basepair with the A at position 4 of the 5' UTR (Lee et al., 2003). The basepairing interaction between

positions 4 of the 3' and 5' UTRs influences the structure the promoter adopts and, therefore, its interactions with the RdRp (Lee et al., 2003). For instance, a C at position 4 biases the RdRp towards replication while a U biases the RdRp towards transcription (Lee et al., 2003; Lee and Seong, 1998). In addition to the conserved promoter region, the gene segment specific portions of the UTR play an important role in regulating the expression of their associated segment (Ma et al., 2013; J. Wang et al., 2016; Xiao et al., 2021; Zheng et al., 1996). A large number of gene segments exhibit basepairing between a majority of the nucleotides within the promoter proximal region (nucleotides 13-15 of the 3' UTR and 14-16 of the 5' UTR) (J. Wang et al., 2016). Basepairing interactions within this region, especially between position 13 of the 3' UTR and position 14 of the 5' UTR, have been demonstrated to be important for regulating gene segment synthesis (J. Wang et al., 2016). The identity of the remaining sequence present in the UTRs between the promoter proximal region and the ORF play an important role in regulating expression in a gene segment specific manner, but the mechanisms remain unknown (Ma et al., 2013; Xiao et al., 2021; Zheng et al., 1996). One of the studies examining a series of deletions within ORF proximal region of the 3' UTR of the HA segment found that there was a defect in HA vRNA synthesis during viral infection (Xiao et al., 2021). This suggests that the segment-specific UTR sequence is optimized to allow its associated segment to effectively compete with the remaining segments for expression and fine tune its expression level relative to the other segments to promote optimal viral replication.

The potential roles that RNA sequence motifs and secondary structures within the ORFs of the eight gene segments play in regulating gene segment expression have remained largely unexplored. A few studies predicted the formation of RNA secondary structures within the NP and M segments via chemical probing or computational methods whose disruption resulted in mild replication defects (Kobayashi et al., 2016; Soszynska-Jozwiak et al., 2015). Another study by Jiang et al. identified RNA structures located in the M and NS segment mRNAs that played a role in regulating the splicing of those segments (T. Jiang et al., 2016). Simon et al. performed DMS-MaPseq to probe the role that RNA structures present in the IAV mRNAs in the context of infected cells play in regulating IAV replication (Simon et al., 2019). Although they found that there were fewer RNA structures in the IAV mRNAs than were predicted *in vitro*, there were still RNA structures present, and disrupting these structures significantly attenuated viral replication (Simon et al., 2019).

Although many different mechanisms have been proposed to try and piece together the overall regulatory scheme for the temporal regulation of gene segment expression as well as the relative levels of expression of each segment, so far none have been able to comprehensively account for the observed patterns of viral gene segment expression.

Nuclear export and trafficking of vRNPs to the cell membrane

After vRNPs are replicated, they need to be exported from the nucleus and trafficked to the cell membrane for virion assembly (Dou et al., 2018). The host

CRM1 export machinery is responsible for exporting the vRNPs from the nucleus, and the viral NS2 protein acts as the adaptor between the CRM1 export machinery and the vRNP (Huang et al., 2013; Neumann et al., 2000; Paterson and Fodor, 2012). The viral NS2 protein contains an N-terminal nuclear export signal which interacts with CRM1 and a C-terminal domain that interacts with an N-terminal domain in M1 (Brunotte et al., 2014; Huang et al., 2013; Neumann et al., 2000; Paterson and Fodor, 2012; Shimizu et al., 2011). The M1 protein is recruited to vRNP complexes through interactions between its C-terminal domain and NP (Baudin et al., 2001). The CRM1-NS2-M1-NP interaction is called the “Daisy Chain” complex and is hypothesized to be the primary mechanism driving export of vRNPs from the nucleus (Brunotte et al., 2014; Eisfeld et al., 2015; Paterson and Fodor, 2012). In addition to NS2, NP also contains several putative nuclear export signals which allow for it to exit the nucleus independently, suggesting that NP may play a role in the nuclear export of vRNPs (Elton et al., 2001; Yu et al., 2012).

Once vRNPs exit the nucleus, they associate with Rab11+ recycling endosomes through the PB2 protein (Amorim et al., 2011; de Castro Martin et al., 2017; Eisfeld et al., 2011; Momose et al., 2011). The Rab11+ endosomes contain clusters of vRNPs that merge together as they traffic to the cell membrane, bringing the vRNPs in close proximity so that they can interact with one another via RNA-RNA interactions to form a complex of vRNPs consisting of the complete set of gene segments that will be packaged into virions (Lakdawala et al., 2014). The packaging signals for the eight gene segments have been

mapped to the UTRs and terminal regions of each gene segment (Fujii et al., 2005, 2003; Goto et al., 2013; Hutchinson et al., 2008, 2009; Liang et al., 2005; Marsh et al., 2007). Evidence from studies demonstrating that specific vRNAs have the capacity to selectively interact with one another *in vitro*, and data showing the presence of fibrous structures between the vRNPs in the 7+1 organization present in virions, suggested that the selective packaging of the gene segments is mediated by RNA-RNA interactions (Fournier et al., 2012; Gavazzi et al., 2013; Noda et al., 2012). Recent studies looking at the distribution of NP along the vRNA in vRNPs found that NP was unevenly distributed along the viral RNA and that regions unbound by NP had the potential to form RNA structures which played roles in gene segment packaging (Dadonaite et al., 2019; Williams et al., 2018). In addition to RNA-RNA interactions, other studies have proposed a role for specific residues in NP, engaging in either NP-RNA or NP-NP interactions, in facilitating the selective packaging of the eight gene segments into budding virions (Bolte et al., 2019; Moreira et al., 2016).

Virion budding and release

To assemble virions, the vRNPs are transported to the cell membrane where they associate with sections of the cell membrane containing HA, NA, and M2 (Dou et al., 2018). The virion membrane proteins HA, NA, and M2 are synthesized by endoplasmic reticulum (ER)-associated ribosomes and trafficked through the ER to lipid rich rafts in the cell membrane (Dou et al., 2018; Leser and Lamb, 2005; Takeda et al., 2003; Veit and Thaa, 2011). HA is targeted to these lipid rafts by a fatty acid modification of the C-terminal histidine (Wagner et

al., 2005), NA is targeted to these lipid rafts by the C-terminus of its transmembrane domain (Barman et al., 2004), and M2 associates with the boundaries of the lipid rafts through an unknown mechanism (Schroeder et al., 2005). The vRNPs are targeted to these lipid rafts via the interaction of M1 with the cytoplasmic tails of HA and NA (Ali et al., 2000). The budding of IAV virions from the cell membrane is stimulated by several different mechanisms. The first involves the role of viral proteins that induce membrane curvature. Crowding of the HA and NA proteins in lipid rafts as well as the association of curved viral proteins like M1 with the membrane induce membrane curvature (Chen et al., 2007; Chlanda et al., 2015; Dahmani et al., 2019; Takeda et al., 2003). Second, the viral M2 protein associates with the neck of the budding virions to induce membrane scission (Rossman et al., 2010). Once the virion buds from the cell membrane, it still remains tethered to the cell via interactions with virion-associated HA and cell SAs (Dou et al., 2018). The virion-associated NA is required to cleave the host SAs which allows the virion to be released from the host cell (Dou et al., 2018; McAuley et al., 2019).

1.3 Roles of Semi-Infectious Particles in IAV Populations

Genomic diversity present in IAV populations

Each step of the viral replication cycle is not 100% efficient, meaning that at the end of the infection process, a significant fraction of released virions are not fully replication competent (Diefenbacher et al., 2018). Historically, IAV has been shown to have a high ratio of non-infectious to fully infectious particles with fully infectious particles only accounting for ~1-50% of total particles depending on the

IAV strain examined (Brooke et al., 2013; Dou et al., 2017; Sun et al., 2020). The remaining non-infectious particles in the viral population consist of defective-interfering particles (DIs) and semi-infectious particles (SIPs) (Alnaji and Brooke, 2020; Brooke, 2014, 2017; Diefenbacher et al., 2018). SIPs are defined as particles that fail to express one or more gene segments in infected cells, while DIs are a distinct subset of particles characterized by large internal deletions (Alnaji and Brooke, 2020; Brooke et al., 2013; Brooke, 2014, 2017; Diefenbacher et al., 2018). The amount of SIPs produced by an IAV strain remains constant regardless of the host used to generate the virus (cell lines, eggs, mammalian animal models etc...) suggesting that SIP formation is determined by a viral genetic component that potentiates a replication defect during particular step(s) of the viral replication cycle (Brooke et al., 2013; Diefenbacher et al., 2018). However, the predominant stages of the viral replication cycle that drive SIP formation have not been definitively determined.

Efficiency of different stages of the IAV replication cycle

One stage that potentially poses a significant bottleneck for a productive infection is the nuclear import of vRNPs. This encompasses both transit across the cytoplasm from the point of fusion to the nuclear pore, as well as successful translocation through the nuclear pore complex. The success rates of these individual steps, and the specific factors that may influence them, remain poorly defined. Numerous studies have indicated that cytosolic trafficking of IAV vRNPs and the nuclear import of model cargo molecules are far from perfect (Babcock et al., 2004; Lowe et al., 2010; Schelker et al., 2016). Further, interferon inducible

antiviral effectors such as Mx1 and PLSCR1 may actively interfere with vRNP import (Götz et al., 2016; Luo et al., 2018). In their study examining the co-trafficking of individual IAV vRNPs, Chou et al. used smFISH to quantify the degree of co-localization between vRNPs during early infection (Chou et al., 2013). By comparing segment co-localization within the cytosol following inhibition of membrane fusion or nuclear export, they concluded that the vRNPs enter the nucleus as a complex and that dissociation of individual vRNPs prior to nuclear entry was minimal (Chou et al., 2013). This suggests that degradation or loss of vRNPs prior to nuclear entry is a minor contributor to the SIP phenotype.

Incomplete viral gene expression patterns may also arise through the failure to initiate vRNA transcription and replication in the nucleus or produce a functional gene product. In their single cell RNA sequencing data, Sun et al. found that IAV infected cells fail to express one of the individual gene segments 5-35% of the time under low MOI conditions (Sun et al., 2020). The high mutation rate of influenza viruses (predicted to be as high as 1.8×10^{-4} , or ~2–3 mutations per replicated genome), is likely to contribute to viral gene expression failure (Pauly et al., 2017). For instance, an incoming vRNA may carry a mutation in a critical cis-acting regulatory sequence that prevents replication and transcription of the gene segment or mRNA. Additionally, in cases where SIP production is examined via detection of viral proteins, non-synonymous mutations in coding sequences may result in the production of truncated or misfolded proteins that will not be detected by antibody staining. However, measurements of IAV mutant frequencies are generally much lower than the above rate estimate, around

1.5×10^{-5} to 2×10^{-6} mutations per nucleotide per infectious cycle, not high enough to explain observed SIP frequencies (Bloom, 2014; Nobusawa and Sato, 2006; Parvin et al., 1986; Suárez-López and Ortín, 1994; Suárez et al., 1992).

Ensuring the packaging of complete set of eight individual gene segments poses another potentially significant bottleneck during the IAV replication cycle. EM-based examination of a collection of IAV and IBV strains revealed that up to 20% of virions contained fewer than eight RNA segments, suggesting a greater degree of flexibility in packaging fidelity than was previously thought (Nakatsu et al., 2016). This number is likely an underestimation, given the observation from the same group that non-viral RNAs can be packaged into the 7 +1 array within virions in some cases (Noda et al., 2018).

Altogether, the small degree of flexibility in IAV packaging fidelity likely accounts for a portion of the SIP population, with the remaining population most likely generated through a combination of failures in gene segment trafficking or expression. These post-entry failure mechanisms may be probabilistic rather than deterministic in many cases. Thus, a perfectly intact virion containing eight gene segments of wild type sequence can end up being a SIP due to stochastic events following entry.

Consequences of SIPs for viral replication and evolution

One of the major processes driving the evolution of IAV populations is reassortment (Taubenberger and Kash, 2010). Reassortment is the genetic exchange of individual gene segments between two different IAV strains infecting

the same cell (either at the same time [co-infection] or sequentially [superinfection] (Sun and Brooke, 2018)), resulting in the generation of progeny with a distinct combination of segments derived from each parental strain (McDonald et al., 2016). The benefits of reassortment include the ability to more rapidly purge deleterious mutations from the population, and the potential to generate a virus with a novel genotype that enhances its replication and transmissibility or changes its antigenicity (McDonald et al., 2016). For instance, reassortment between zoonotic IAV strains and human IAV strains has led to the generation of pandemic IAVs strains which were antigenically distinct from the previously circulating human strains and, therefore, could spread rapidly within the human population that lacked pre-existing immunity (Taubenberger and Kash, 2010). Because SIPs require complementation to replicate, viral populations with more SIPs will be more dependent upon co-infection to propagate. A process that influences the co-infection rate is superinfection exclusion in which the virus currently replicating in the cell inhibits the subsequent infection of another virus (Sun and Brooke, 2018). Therefore, higher rates of superinfection exclusion would likely result in less co-infection and, therefore, reassortment. Sun et al. used UV irradiation to increase the SIP content of an IAV strain and found that populations with more SIPs exhibited less superinfection exclusion, which would suggest that populations with more SIPs would undergo higher rates of reassortment (Sun and Brooke, 2018). Using UV irradiation to artificially increase the SIP content within a viral population, Fonville et al. directly tested this hypothesis and found that as the proportion of SIPs

within a population increased, the proportion of productively infected cells that were co-infected also increased, as did the overall reassortment rate (Fonville et al., 2015). These studies suggest that IAV strains with distinct SIP production phenotypes may differ in their reassortment potential and, therefore, their pandemic risk.

As IAV populations with high amounts of SIPs contain a large proportion of virions that lack one or more segments, another potential consequence of SIPs for IAV replication and evolution is altering the effective gene dosage delivered to infected cells. In other viral systems, modulating the effective viral gene dose has important consequences for the expression of specific viral genes and overall levels of replication (Elde et al., 2012; Sicard et al., 2013). The importance of viral gene dosage effects for IAV replication has also been demonstrated. For instance, Brooke et al. found that artificially decreasing the amount of the NA segment present in a viral population led to a decrease in the overall expression level of the NA segment in infected cells (Brooke et al., 2014). Additionally, Phipps et al. & Martin et al. demonstrated that at sub-saturating MOIs, increasing the MOI, and, therefore, the overall gene dosage of each segment, increased IAV replication (Martin et al., 2020; Phipps et al., 2020). This effect was cell type dependent suggesting that different hosts, potentially hosts that a particular IAV strain is not sufficiently adapted to, require a higher gene dosage in order for IAV to successfully establish a productive infection (Martin et al., 2020; Phipps et al., 2020).

In addition to effects derived from decreasing the overall gene dosage, the lack of particular viral genes themselves in IAV populations with high amounts of SIPs has the potential to influence the host response to infection. In their single cell RNA sequencing data, Sun et al. found that infected cells lacking the expression of the PA and NS gene segments exhibited the greatest changes in the overall transcriptional response to IAV infection (Sun et al., 2020). Additionally, they found that for the A/Perth/16/2009 (H1N1) IAV strain, infected cells lacking the expression of the NS segment exhibited increased expression of host interferon (IFN) and interferon-stimulated genes (ISGs) (Sun et al., 2020). Therefore, the presence and/or absence of specific viral genes has the potential to affect the overall infection outcome.

1.4 Role of Epistasis in IAV Evolution

Defining epistasis

Epistasis is defined as genetic interactions between individual mutations and/or genes in a genome, and it plays an important role in determining the evolutionary paths available to an organism (Lyons and Lauring, 2018). In most cases epistasis is believed to constrain evolution as deleterious mutations often require several additional permissive mutations to arise by chance before they can be fixed in the population (Lyons and Lauring, 2018). And once those mutations are fixed, the organism can no longer revert back to its original genetic state as the reversion of those mutations would lead to a decrease in fitness (Lyons and Lauring, 2018). Additionally, the greater the degree of epistasis, the higher likelihood that a mutation would have pleiotropic, and potentially deleterious,

effects, requiring additional mutations to restore fitness (Lyons and Luring, 2018). However, epistasis also has the potential to increase evolutionary flexibility by allowing a wider range of mutations to compensate for the deleterious effects of others (Lyons and Luring, 2018). Epistasis is predicted to be highly prevalent in IAV and an important factor in driving its evolution (Lyons and Luring, 2018; Nshogozabahizi et al., 2017).

Epistasis within individual IAV proteins

Epistasis plays a significant role in the evolution of many individual IAV proteins. HA, one of the IAV surface glycoproteins that binds to SA receptors on host cells, is the one of the major antigenic determinants of IAV (Wu and Wilson, 2019). The antigenic evolution of HA is punctuated in nature with the HA protein accumulating mutations but remaining in the same antigenic cluster for several years before acquiring a mutation that changes its antigenicity and moves it into a novel antigenic cluster (Koelle et al., 2006). The observations that HA requires specific mutations to accumulate before a mutation causing antigenic drift can emerge, that the genetic background of the HA influences its mutational tolerance as well as the types of mutations accumulated, and that once the antigenic change occurs HA can no longer revert back to its original state substantiate the important role that epistasis plays in the evolution of HA (Lee et al., 2018; Wu et al., 2017, 2018). Additionally, the mutations that allow for antigenic escape often have pleiotropic and potentially deleterious effects on HA protein function which further constrains its evolution via epistasis. For instance, mutations that alter the antigenicity of HA can alter the receptor binding avidity of

HA, and require mutations in HA that either directly modulate its receptor binding avidity or indirectly modulate it via altering the number of glycosylation sites on HA to restore viral fitness (Hensley et al., 2009; Kosik et al., 2018; Myers et al., 2013; Prachanronarong et al., 2019).

The remaining surface glycoprotein NA, whose function is to cleave SA receptors from host cells to allow for viral release (McAuley et al., 2019), also exhibits many epistatic signatures in its evolution. NA inhibitors (NAIs) are one of the few classes of antiviral drugs used to treat IAV infections (Świerczyńska et al., 2022). Mutations in NA that prevent binding of NAIs to the NA active site are often deleterious because they decrease NA activity and, therefore, require compensatory mutations to restore NA activity (L. Jiang et al., 2016). In 2007-2008 an oseltamivir resistance mutation (NA H275Y) rose to high frequencies in circulating IAV strains, however, by itself the mutation was deleterious, and it was later shown that it was able to emerge as a result of permissive mutations (NA V234M/R222Q) that emerged shortly before (Abed et al., 2011; Bloom et al., 2010; Duan et al., 2014). Additionally, the fitness effects of NAI resistance mutations vary depending upon the viral genetic background (L. Jiang et al., 2016).

The evolution of the NP protein is also influenced by epistasis. For instance, Gong et al. examined the amino acid changes that emerged in the NP of IAV strains of the H3N2 subtype between 1968 and 2007 and found that maintaining the overall stability of the NP protein was the primary driver of its evolution and

that mutations that altered immune epitopes in NP but decreased its stability were accompanied by stabilizing mutations (Gong et al., 2013). In addition, Göte et al. found that mutations that allowed NP to escape restriction by the host innate immune effector MxA compromised its ability to promote vRNP nuclear import (Götz et al., 2016).

Epistasis between viral gene segments and proteins (intersegment epistasis)

Epistasis within individual viral proteins clearly plays an important role in IAV evolution, however, epistasis also is an important factor in determining the degree of parallel evolution between different IAV genes and/or proteins. For instance, epistasis plays an important role during IAV reassortment. Patterns of reassortment between different IAV strains are not random with specific segments co-segregating with one another in different genetic backgrounds (Chen et al., 2008; Lubeck et al., 1979; Rabadan et al., 2008). Many reassortants are less fit than their parental strains and increases in mutation frequencies are often observed after reassortment events, suggesting that the reassorted gene segments are initially incompatible with one another (Neverov et al., 2014). This incompatibility likely arises from mismatches in RNA-RNA, RNA-protein, or protein-protein interactions (White and Lowen, 2018).

An example of an RNA-RNA mismatch would be a reassortment event that brings together non-compatible packaging signals from different IAV strains and, therefore, impairs gene segment packaging (White and Lowen, 2018). Dadonaite et al. used a technique called sequencing of psoralen-cross-linked, ligated and

selected hybrids (SPLASH) to probe RNA-RNA interactions between vRNPs in virions and identified a RNA-RNA interaction between the PB1 and NA segments that drove their co-segregation in different H3N2 reassortants (Dadonaite et al., 2019). RNA-protein interactions have also been shown to modulate gene segment packaging. For instance, individual mutations in NP resulted in packaging defects for specific sets of segments (Bolte et al., 2019; Moreira et al., 2016).

Incompatibilities between different viral proteins also likely play a role in reassortment outcomes. A well-established example is the need to maintain the functional balance between the receptor binding avidity of HA and the receptor cleavage activity NA (Casalegno et al., 2014; Gaymard et al., 2016; Wagner et al., 2002). For instance, in cases where either HA or NA activity was disrupted by mutations arising from antigenic escape or NAI resistance respectively, compensatory mutations in HA and/or NA were required to restore their functional balance and viral fitness (Cohen-Daniel et al., 2009; Das et al., 2011; Ginting et al., 2012; Hensley et al., 2011). Additionally, Ward et al. found that the evolutionary rate of HA varied depending upon the genetic background of the NA segment, suggesting the presence of variation in the functional compatibility between the activities of HA and the NAs from different IAV strains (Ward et al., 2013). As the PB2, PB1, and PA proteins are required to form a functional viral RdRp (Fodor and Velthuis, 2019), they also have been shown to interact epistatically. Chen et al. found that a reassortant that contained the PB2 segment from the A/WSN/1933 strain and the remaining seven segments from the

A/Puerto Rico/8/1934 strain was attenuated even though the strains exhibited significant genetic similarity (Chen et al., 2019). Compensatory mutations in regions of the PB2, PB1, and PA proteins involved in the formation of the RdRp dimerization domain, which plays an important role for promoting vRNA replication, were found to restore viral fitness (Chen et al., 2019).

Aside from influencing viral fitness by determining the compatibility between different IAV RNA segments and proteins, epistatic interactions between gene segments also can regulate viral gene expression in a segment-specific manner. Although the mechanisms remain unknown, several studies demonstrated that the genetic origin of the M, PB2, and PB1 segments determined overall NA protein expression levels and incorporation into virions (Campbell et al., 2014; Moulès et al., 2011). Additionally, we found that a single amino acid substitution in NP (F346S), that conferred enhanced replication and transmissibility of the mouse-adapted A/Puerto Rico/8/1934 (H1N1) (PR8) strain in guinea pigs, resulted in a selective decrease in the expression levels and packaging frequencies of the NA gene segment (Brooke et al., 2014; Ince et al., 2013). Even though the mechanisms remain to be determined, intersegment epistasis is clearly highly prevalent in IAV populations and plays an important role in IAV replication and evolution.

1.5 Aims of Thesis

Genetic interactions within and between segments play important roles in modulating IAV replication and evolutionary potential, however, the molecular

mechanisms regulating those interactions have remained underexplored and incompletely defined. In this thesis, I discuss work that I have done to more precisely uncover specific mechanisms by which intersegment epistasis modulates IAV gene segment expression and the genomic structure of IAV populations.

In the following two chapters of this thesis, I discuss work that I have done to improve our understanding of the potential mechanisms behind how genetic interactions within and between individual gene segments (intra/intersegment epistasis) can regulate the expression patterns of individual gene segments (Fig. 1.1). In chapter 2, I highlight our study utilizing a NP mutant (F346S) that exhibited a defect in expression of a single segment (NA) to probe the mechanism by which intersegment epistasis can regulate the expression of individual gene segments. Building upon the work described in chapter 2, in chapter 3, I discuss work that aims to identify specific RNA motifs and structures within individual segments and interactions between them that are important for modulating various steps of the IAV replication cycle. Together, these two chapters highlight the importance of intra/intersegment epistasis in regulating individual gene segment expression patterns.

In chapter 4 of this thesis, I explore the role of intersegment epistasis in regulating the genomic structure of IAV populations (Fig. 1.1). Using two IAV strains that exhibit distinct SIP production phenotypes, I generated reassortants between the two strains to identify the genetic determinants driving SIP

formation. I aim to further define how those genetic determinants mechanistically regulate SIP production. Altogether, these studies provide novel insights into the mechanisms of intersegment epistasis in IAV, and provide a framework to better understand and predict how these epistatic interactions influence IAV replication and evolution.

1.6 Figure

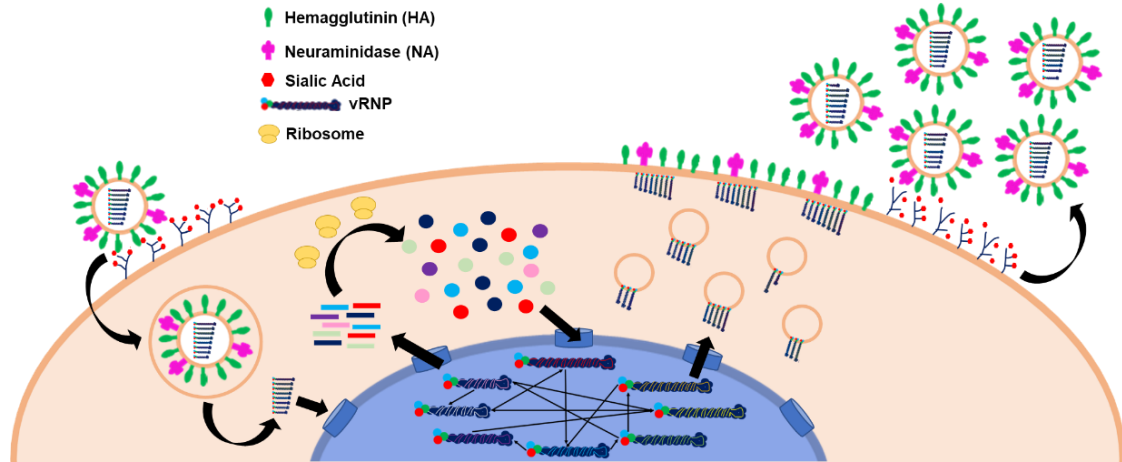


Fig. 1.1. Evaluating the role of intersegment epistasis during the IAV replication cycle and its consequences for the genomic structure of IAV populations. IAV virions bind to the host cell sialic acid-containing receptors via HA and are internalized via receptor mediated endocytosis. The eight vRNPs are then released into the cytoplasm and transported into the nucleus where they undergo a round of primary transcription to generate mRNAs necessary for the synthesis of IAV proteins. The IAV proteins required for the initiation of gene segment replication are imported into the nucleus, and gene segment replication commences. During the gene segment transcription and replication processes, epistatic interactions between the segments (indicated by black arrows) modulate their expression kinetics via unknown mechanisms. The vRNPs are then exported from the nucleus and virion assembly occurs. Virions are then released via the cleavage of host cell sialic acid residues by NA. Epistatic interactions between segments that modulate the efficiency of each step of the IAV replication cycle ultimately influence the production of semi-infectious particles that lack one or more segments or contain a defective segment (truncated vRNP).

CHAPTER 2: INTERACTIONS BETWEEN INFLUENZA A VIRUS NUCLEOPROTEIN AND GENE SEGMENT UTRS FACILITATE SELECTIVE MODULATION OF VIRAL GENE EXPRESSION

2.1 Introduction

Influenza A virus (IAV) is a major respiratory pathogen that causes seasonal epidemics and occasional pandemics that result in substantial morbidity and mortality (Krammer et al., 2018). The genome of IAV is divided into eight negative-sense, single-stranded RNA segments that encode one or more viral proteins (Fodor and Velthuis, 2019). These negative-sense genomic RNAs (viral RNAs [vRNAs]) are used as templates to synthesize both the mRNA needed for protein synthesis and the positive-sense replicative intermediates (cRNAs) for genome replication (Fodor and Velthuis, 2019). The individual IAV gene segments vary in both overall expression levels and timing, but the specific factors that govern this variation are poorly understood (Phan et al., 2021; Shapiro et al., 1987).

Each IAV gene segment consists of one or more open reading frames (ORFs) flanked by untranslated regions (UTRs) (Fodor and Velthuis, 2019). The UTRs consist of both segment-specific sequences and highly conserved sequences at the 3' and 5' termini which interact with one another to form the viral promoter (Fodor and Velthuis, 2019). Previous studies have established roles for segment-specific sequences within the UTRs in modulating gene expression in a segment-specific manner (Ma et al., 2013; Maeda et al., 2004; Wang et al., 2015; J. Wang et al., 2016; Xiao et al., 2021; Zheng et al., 1996). In virions and within infected cells, the gene segments are maintained as viral ribonucleoprotein complexes

(vRNPs) in which the viral RNA is bound along its length by nucleoprotein (NP) and is associated with the viral RNA-dependent RNA polymerase (RdRp) (Fodor and Velthuis, 2019).

NP is a highly conserved (Babar and Zaidi, 2015) and multifunctional protein. To perform its integral role in vRNP formation, NP has two major known activities: RNA binding and oligomerization. NP binds RNA nonspecifically through a positively charged groove located between its head and body domains (Ng et al., 2008; Ye et al., 2006). Oligomerization of individual NP protomers occurs through the insertion of a C-terminal tail loop into the receptor groove of the neighboring protomer (Chan et al., 2010; Turrell et al., 2013; Ye et al., 2006). As a key component of the vRNP complex, NP plays an essential role in vRNA replication and mRNA transcription. NP is hypothesized to act as an elongation factor for the viral polymerase, as only short transcripts (<100 nucleotides [nt]) can be generated in its absence or in the presence of binding/oligomerization-deficient NP mutants (Turrell et al., 2013). NP facilitates the import and export of vRNPs from the nucleus (Eisfeld et al., 2015; Tome-Amat et al., 2019; Wang et al., 1997). Finally, NP is critically involved in the selective packaging of the viral gene segments—both directly through specific amino acid residues (Bolte et al., 2019; Moreira et al., 2016) and indirectly through determining the accessibility of RNA structures important for packaging (Dadonaite et al., 2019; Williams et al., 2018).

We previously identified an NP substitution (F346S) that was sufficient to significantly enhance the replication and transmissibility of the A/Puerto

Rico/8/1934 (PR8) strain of IAV in guinea pigs while selectively decreasing the expression of the neuraminidase (NA) gene segment (Brooke et al., 2014; Ince et al., 2013). This finding suggested (i) that NP plays an unappreciated role in selectively regulating the expression of individual viral genes and (ii) that this mode of gene regulation may be involved in modulating transmission potential. Given that gene segment replication and transcription occur in the context of the vRNP, and that the vRNPs of all eight gene segments are thought to largely be structurally and functionally equivalent, it is not clear how substitutions in NP could result in selective modulation of NA segment expression.

Here, we dissect the mechanism by which specific residues in NP selectively modulate NA segment expression. In addition, we pinpoint the specific determinants within the NA genomic RNA that are required for susceptibility to selective regulation by NP. Altogether, these results illuminate a new mode of selective gene regulation by influenza viruses that may play an important role in host adaptation and transmission.

2.2 Results

NP:F346S suppresses NA segment replication but not mRNA transcription

We examined the effects of NP:F346S on NA vRNA abundance over the course of a single PR8 replication cycle. Similar to our previous findings, NP:F346S reduced NA expression nearly 20-fold by 12 h post-infection (hpi) (Brooke et al., 2014), while leaving HA expression largely unaffected (Fig. 2.1A). We previously showed that NP:F346S also decreased NA mRNA abundance, raising the

possibility that this substitution directly affected all NA segment-derived RNA species (Brooke et al., 2014). To determine whether the effects of NP:F346S are specific for vRNA, we compared levels of primary NA mRNA synthesis between wild-type NP (NP:WT) and NP:F346S in the presence of 100 µg/mL of cycloheximide. Cycloheximide blocks translation of the viral replicase machinery needed for vRNA synthesis, thus only allowing primary transcription of viral mRNAs from incoming vRNPs (Vreede et al., 2004). We observed no differences in the NA mRNA levels between the WT and NP:F346S in the presence of cycloheximide, indicating that NP:F346S has no effects on primary mRNA transcription (Fig. 2.1B).

To determine whether the effect of NP:F346S on NA vRNA abundance is due to reduced synthesis, as opposed to a decrease in stability, we performed a 4-thiouridine (4SU) pulse assay on infected cells. During the pulse, 4SU is incorporated only into newly synthesized RNAs which then can be biotinylated and selectively purified (Rutkowski and Dölken, 2017), allowing us to specifically examine the effect of NP:F346S on the *de novo* synthesis rate of NA vRNA. We pulsed infected cells with 4SU for 1 h and measured the amount of vRNA synthesized during the pulse by performing reverse transcription-quantitative PCR (RT-qPCR) on 4SU-labeled RNAs using a universal, vRNA-specific primer for the RT reaction followed by segment-specific primers for the qPCR (Fig. 2.1C). The levels of the 4SU-containing newly synthesized NA vRNA were over 10-fold lower during NP:F346S infection than with NP:WT (Fig. 2.1C). Altogether,

our data indicate that NP:F346S specifically affects the synthesis of new NA vRNA molecules during infection.

The effect of NP:F346S on NA segment expression is strain specific

Given that NP is thought to play the same role in the replication of all viral gene segments, how can substitutions in NP selectively reduce synthesis of the NA RNA while leaving the other segments largely unaffected (Brooke et al., 2014)? We hypothesized that this specificity must depend upon unique motifs present with the NA segment. To test this hypothesis, we introduced the NP:F346S substitution into a divergent IAV strain of the H3N2 subtype, A/Udorn/307/72 (Udorn), and examined whether it reduced Udorn NA segment expression similar to what was observed with PR8.

NP residue F346 is highly conserved (Shu et al., 1993) and is identical in both the PR8 and Udorn IAV strains. In the Udorn background, NP:F346S had no appreciable effect on NA segment expression, indicating that the effects of NP:F346S are virus strain dependent (Fig. 2.2A). To test whether this strain specificity arises from the differences in NA segment (versus NP or the viral polymerase complex), we generated chimeric viruses encoding the Udorn HA and NA segments along with the remaining six segments from PR8. Introduction of the NP:F346S substitution into the PR8:Udorn HA, NA virus similarly had no effect on NA expression, indicating that the effects of NP:F346S on NA expression depend upon the specific sequence of the NA segment (Fig. 2.2B).

Selective modulation of gene expression by NP:F346S depends upon segment UTR sequences

To pinpoint the specific motif(s) within the NA gene segment that confers susceptibility to selective modulation by NP:F346S, we first divided the NA segment into three broad functional regions: (i) the portion of the NA ORF that does not overlap known packaging signals, (ii) the portions of the NA ORF that do overlap known packaging signals, and (iii) the NA UTRs. We then tested each for their role in conferring sensitivity to the effects of NP:F346S.

To determine if any RNA sequence elements within the NA ORF (exclusive of the packaging signals as defined based on retention within defective interfering particles in a previous study (Alnaji et al., 2019)) were important for the effects of NP:F346S, we used the CodonShuffle package (Jorge et al., 2015) to introduce 227 silent substitutions within the region encompassing nucleotides 38 to 1319 of the NA segment while minimizing effects on codon frequency or dinucleotide content (Fig. 2.3A). In Fig. 2.3 and subsequent figures, we show the segments in the vRNA sense with the numbering corresponding to the first nucleotide position in the cRNA (i.e., numbered from the 3' end of the vRNA). The codon-shuffled NA segment exhibited a decrease in its expression level similar to that of WT NA in the presence of NP:F346S, indicating that the effects of NP:F346S on NA replication do not require motifs within the nonpackaging signal region of the NA ORF (Fig. 2.3B).

Not surprisingly, viruses in which the CodonShuffle program was used to mutagenize the regions of the NA ORF that overlap the packaging signals were

unable to be rescued after two attempts (data not shown). As an alternative approach to examine the roles of the NA segment packaging signal regions and UTRs in determining sensitivity to NP:F346S, we generated two sets of recombinant PR8 viruses in which we swapped terminal sequences between the HA segment (which is unaffected by NP:F346S) and the NA segment and paired them with either NP:WT or NP:F346S (Fig. 2.3C). One set of viruses contained chimeric HA-NA segments in which both the UTRs and packaging signals present within the terminal coding regions of the PR8 HA and NA segments were swapped (UTR+Pack swap) based on a previously described set of viable chimeric HA-NA segments (Zheng et al., 2020). The other set of viruses contained segments in which only the UTRs of the PR8 HA and NA segments were swapped (UTR swap). For the UTR swap viruses, a segment containing the HA ORF with the NA UTRs exhibited a severe packaging deficiency (Supplemental Fig. 2.1A), and the abundance of the segment in infected cells was below the limit of detection for the qPCR assay.

We infected Madin-Darby canine kidney (MDCK) cells with these recombinant viruses and quantified the effects of NP:F346S on intracellular HA and NA vRNA levels (Fig. 2.3D). Replacing the packaging signals and UTRs of the NA segment with those of the HA segment reduced the effect of NP:F346S on NA expression ~3-fold (Fig. 2.3D). Similarly, while WT HA expression was unaffected by NP:F346S, an HA segment containing the packaging signals and UTRs from the NA segment exhibited a >100-fold reduction in expression in the context of NP:F346S versus NP:WT (Fig. 2.3D). The observation that swapping the

packaging signals and UTRs of the PR8 HA and NA segments caused their phenotypes to reverse in the presence of NP:F346S (HA became sensitive and NA became more resistant) suggested that the packaging signals and UTRs may play a role in this NP-dependent regulation. However, looking at the ratio of the HA and NA segments with the swapped UTRs and packaging signals in the viral stocks, the decrease in their abundance in the presence of NP:F346S corresponds to the observed expression decrease, suggesting that the observed changes in gene expression largely stem from changes in gene segment packaging ratios—likely a result of removing the packaging signals from their native context (Supplemental Figs. 2.1A and 2.1B). This made it difficult to separate the effects of decreased packaging efficiency from regulation by NP:F346S. However, for viruses in which we only swapped the UTRs between PR8 HA and NA, we found that replacing the NA segment UTRs with those from HA completely eliminated the effect of NP:F346S on NA vRNA abundance, suggesting that the ability of NP:F346S to regulate NA segment expression largely depends on the UTR sequences (Fig. 2.3D). Altogether, while we have not completely ruled out the role of the packaging signals in influencing the effects of NP:F346S on segment expression, our observations (i) that saturation of a majority of the NA ORF with silent substitutions was insufficient to decrease sensitivity to the effects of NP:F346S and (ii) that replacing the NA segment UTRs with those of the PR8 HA segment was sufficient to render the segment resistant to the effects of NP:F346S suggest that the selective effects of

NP:F346S on vRNA synthesis primarily depend upon the segment UTR sequences.

Identification of specific nucleotides within the NA UTR that determine sensitivity to NP:F346S

We next sought to identify which specific elements within the NA UTR are required for susceptibility to modulation by NP:F346S. To do this, we took advantage of the high degree of similarity between the NA segment UTRs of PR8 (susceptible to the effects of NP:F346S) and Udorn (resistant to the effects of NP:F346S) (Figs. 2.2 and 2.4A). The 3' UTRs of the PR8 and Udorn NA segments differ in (i) the identity of the nucleotides at positions 4 and 13 (CC/UU for PR8/Udorn, respectively) and (ii) the sequence directly upstream of the initiating Met codon of the NA ORF ([AAAUUU/ACUUC] for PR8/Udorn, respectively) (Fig. 2.4A). In the 5' NA segment UTR, Udorn has a 9-bp insertion relative to the PR8 sequence plus a few additional nucleotide substitutions (Fig. 2.4A).

We first confirmed that the difference in susceptibility of the PR8 and Udorn NA segments to the effects of NP:F346S is associated with the UTR sequences. We generated a virus in which the UTRs of the PR8 NA segment were replaced with those from the NA segment of Udorn (PR8:NA^{Udorn UTR}) (Figs. 2.4A and 2.5). The effects of NP:F346S on PR8:NA^{Udorn UTR} were reduced compared to WT PR8 NA, again indicating that the segment UTR sequences play a significant role in determining the segment specificity of the effects of NP:F346S on gene expression (Fig. 2.4B). We also attempted to generate viruses in which the UTRs

of the Udorn NA segment were replaced with those from PR8; however, we were unable to rescue a virus with this chimeric Udorn/PR8 NA segment and NP:F346S. By replacing the internal gene segments of Udorn with those of PR8, we were able to rescue viruses containing a segment with the Udorn NA ORF and PR8 UTRs and NP:WT/F346S (PR8:Udorn HA,NA^{PR8 UTR}). The viruses were still highly attenuated, reaching titers of only 10⁴ to 10⁵ infectious particles/mL. Replacing the Udorn NA UTRs with those of PR8 made it susceptible to the effects of NP:F346S, although to a lesser degree than PR8 NA (~30% versus 5 to 10% of NP:WT, respectively), further substantiating the role of the NA UTRs in regulation by NP:F346S (Fig. 2.4C). Additionally, while the Udorn NA segment paired with a PR8 backbone exhibited no apparent defects in genome packaging, the Udorn:NA^{PR8 UTR} segment in a PR8 backbone did exhibit decreased packaging efficiency in the presence of NP:F346S, suggesting that some of the observed decrease in expression levels within infected cells might be due to decreases in delivered NA gene dose due to decreased packaging efficiency of the Udorn:NA^{PR8 UTR} segment (Supplemental Fig. 2.1C).

We next generated a panel of recombinant viruses with chimeric PR8-Udorn NA UTR sequences (Fig. 2.5). We first examined both the polymorphisms at positions 4 and 13 of the promoter and the promoter-proximal region in the 3' UTR. Introducing individual C4U or C13U substitutions into the PR8 NA UTR slightly reduced the extent to which NP:F346S decreased NA vRNA levels (Fig. 2.6A). Introducing C4U and C13U together decreased the relative effect of NP:F346S on NA vRNA levels roughly 6-fold (Fig. 2.6A). Retaining the C's at

positions 4 and 13 while swapping the ORF proximal region of the PR8 NA 3' UTR with that of Udorn (PR8:NA^{Udorn 3' ORF Proximal UTR}) slightly reduced the extent to which NP:F346S decreased NA vRNA levels, to an extent comparable with those obtained with the individual C4U or C13U mutations (Fig. 2.6A). Replacing the entire PR8 NA 3' UTR with that of Udorn (PR8:NA^{Udorn 3' UTR}) (has U4/13 and Udorn ORF proximal UTR sequence) decreased the relative effect of NP:F346S to approximately the same level as with the dual C4/13U mutations (Fig. 2.6A). Taken together, these data demonstrate that sensitivity to the effects of NP:F346S is largely determined by the identity of nucleotides 4 and 13 in the 3' UTR; however, additional motifs within the segment also likely contribute.

We next examined the roles of the ORF proximal regions of the NA UTRs. As described above, just replacing the 3' PR8 NA ORF proximal region with that of Udorn (PR8:NA^{Udorn 3' ORF Proximal UTR}) did not have much of an effect on susceptibility to NP:F346S (Fig. 2.6A). Replacing the 5' UTR of the PR8 NA with that of Udorn (PR8:NA^{Udorn 5' UTR}) enhanced susceptibility to NP:F346S ~4-fold (Fig. 2.6A), likely due to the fact that there was an additional packaging defect for the segment in the presence of NP:F346S (Supplemental Fig. 2.1D).

Interestingly, replacing both the ORF proximal regions of the PR8 NA with those from Udorn (PR8:NA^{Udorn ORF Proximal UTR}) (has 4/13C and Udorn NA ORF proximal sequences) made the segment resistant to NP:F346S, actually increasing expression ~5- to 6-fold relative to that of NP:WT (Fig. 2.6B).

The only differences between the PR8:NA^{Udorn UTR} segment, which was partially resistant to the effects of NP:F346S (Fig. 2.4B), and PR8:NA^{Udorn ORF Proximal UTR}, which was completely resistant, were the identities of nucleotides 4 and 13 of the 3' UTR (Fig. 2.5), so we next asked whether these nucleotides were responsible for the resistance phenotype observed for PR8:NA^{Udorn ORF Proximal UTR}. Mutating the C at position 4 or 13 of the PR8:NA^{Udorn ORF Proximal UTR} segment to U (PR8:NA^{C4U/Udorn ORF Proximal UTR}/PR8:NA^{C13U/Udorn ORF Proximal UTR}) restored the susceptibility of the segment to NP-dependent regulation, again emphasizing the importance of positions 4 and 13 of the PR8 NA 3' UTR for determining the effects of NP:F346S (Fig. 2.6B). Interestingly, we also found that replacing the ORF proximal regions of the Udorn NA UTR with those of PR8 (PR8:Udorn HA,NA^{PR8 ORF Proximal UTR}) (has 4/13U and PR8 NA ORF proximal sequences) increased the expression level of the segment to near WT Udorn NA segment expression levels (~80 to 90%) in the presence of NP:F346S (Fig. 2.6C). In conclusion, susceptibility of a gene segment to the effects of NP:F346S depends upon a specific combination of nucleotide identities at positions 4 and 13 of the 3' UTR and ORF proximal sequences (4/13C and Udorn NA ORF proximal UTRs or 4/13U and PR8 NA ORF proximal UTRs).

As several of the PR8-Udorn NA UTR mutant viruses were highly attenuated, we wanted to determine whether there were any compensatory mutations that may have emerged that could potentially confound our results. We performed next-generation sequencing on these viruses and found that the only virus with any mutations over ~30% in the population was PR8:NA^{Udorn UTR} NP:F346S, which

had a fixed nonsynonymous substitution in PB2 (E191G). As the PB2 E191 residue is not a part of any major functional domain of PB2 (Fodor and Velthuis, 2019; L. Wang et al., 2016), and its function has not been previously explored, its effect on PB2 function and how that could potentially affect NA expression independently or dependently of the context of NP:F346S are unclear.

We hypothesized that the effects of different UTR mutations on sensitivity of NA expression levels to NP:F346S could arise from two distinct mechanisms: (i) abrogating the selective effect of NP:F346S on NA vRNA synthesis, restoring NA vRNA levels to those observed in the context of NP:WT, and (ii) reducing NA levels in the context of NP:WT, bringing them closer to what is observed with NP:F346S. To distinguish between these possibilities, we compared the expression of the different NA UTR mutant constructs to NP vRNA levels (which are unaffected by the NP:F346S substitution) (Figs. 2.6D and 2.6E). We observed that the expression of all the PR8/Udorn NA UTR chimeric constructs except for PR8:NA^{C4U} was reduced compared to that for WT NA in the context of NP:WT, and in some cases was lower than the level observed for the WT NA segment in the presence of NP:F346S (Figs. 2.6D and 2.6E). Thus, none of the UTR mutants tested mitigated the effects of NP:F346S by simply restoring NA levels to those observed in the context of NP:WT. The second possibility, that these mutations appeared to reduce sensitivity to NP:F346S because they reduced NA levels in the context of NP:WT to levels associated with NP:F346S, was also not supported by these data. For instance, the PR8:NA^{Udorn ORF Proximal UTR} segment exhibited increased expression in the presence of NP:F346S, while

the PR8:NA^{Udom 5' UTR} segment exhibited an even more substantial decrease in the presence of NP:F346S than the WT NA (Fig. 2.6D). Altogether, our data suggest that it is impossible to cleanly separate the effects of mutations in the UTR on susceptibility to the effects of NP:F346S from their broader effects on baseline expression levels in the context of NP:WT.

NP:F346S has minimal effects on NP RNA binding or oligomerization

While the UTR sequences of the NA segment are clearly involved in determining the segment specificity of the effects of NP:F346S on gene expression, the specific mechanisms involved are not obvious. NP:F346S is not located in any previously described functional domains; thus, it was not immediately apparent how the NP:F346S substitution might alter NP function (Albo et al., 1995; Chan et al., 2010; Elton et al., 1999b; Kobayashi et al., 1994; Mondal et al., 2015; Turrell et al., 2015; Ye et al., 2006).

NP has two well-described biochemical activities that are required for the synthesis of full-length viral RNA transcripts: RNA binding and oligomerization (Turrell et al., 2013; Vreede et al., 2011; Ye et al., 2006). To determine whether F346S affects the RNA-binding activity of NP, we compared the *in vitro* RNA binding affinities of purified C-terminally His-tagged versions of the NP:WT and NP:F346S proteins using biolayer interferometry (BLI) (Fig. 2.7A). NP:F346S was associated with a slightly higher K_d (dissociation constant) than NP:WT (3.216 nM versus 2.477 nM) (Fig. 2.7A). However, as the K_Ds previously measured for known RNA binding- or oligomerization-deficient NP mutants

ranged between ~100 nM to 1 μ M, compared to a range of ~10 to 40 nM for the WT NP protein (Chan et al., 2010; Chenavas et al., 2013; Ng et al., 2008; Bogdan Tarus et al., 2012; B. Tarus et al., 2012), the changes that we measured in this study are likely not functionally significant. Although these data suggest that the F346S substitution has minimal effects on the RNA binding affinity of NP, our *in vitro* assay may have failed to fully recapitulate conditions as they occur during infection.

We next evaluated whether NP:F346S affects the oligomerization of NP monomers. We overexpressed both His-tagged and enhanced green fluorescent protein (eGFP)-tagged versions of either NP:WT or NP:F346S in 293T cells and quantified the amount of eGFP-NP that coimmunoprecipitated with His-NP by Western blotting. As a positive control, we assessed the effect of the oligomerization-deficient R416A mutant (Chan et al., 2010; Elton et al., 1999; Tarus et al., 2012; Ye et al., 2006) in our assay and measured a substantial reduction in pulldown efficiency (Fig. 2.7B). In contrast, we did not observe any effect of F346S on the coimmunoprecipitation efficiencies of His-NP and eGFP-NP, suggesting that NP:F346S does not significantly affect the ability of NP to oligomerize, at least under *in vitro* overexpression conditions (Fig. 2.7B).

Additionally, we evaluated whether NP:F346S decreases the NP content of vRNPs. We overexpressed PB2, PB1, HA-tagged PA, and NP (WT/F346S/R416A) and a NA vRNA template in 293T cells to generate vRNPs and visualized the amount of NP that coimmunoprecipitated with HA-tagged PA

via Western blotting. Mirroring our data looking at NP monomer association, there was a substantial reduction in pulldown efficiency for the oligomerization-deficient NP:R416A mutant but no difference in the pulldown efficiencies between NP:WT and F346S (Fig. 2.7C). These data suggest that the NP content of vRNPs is not affected by NP:F346S.

Altogether, these data suggest that F346S has minimal effects on the RNA-binding and oligomerization activities of NP, at least in *in vitro* binding assays. If true in the context of infection, it would suggest that the effects of this substitution on NA segment replication occur through some other, uncharacterized feature of NP biology.

A cluster of aromatic residues within NP governs NA segment replication

Finally, we examined the effects of alternative substitutions at the NP:F346 locus on NA expression. Introducing a silent substitution (T1082C) into the NP:F346 codon had no effects on NA protein expression levels, suggesting that the effects of NP:F346S require the amino acid substitution (Fig. 2.8A). Interestingly, we observed that a wide array of amino acid substitutions at position 346 (representative of amino acids with distinct functional groups) resulted in a selective decrease in NA expression, suggesting that a phenylalanine is required at NP position 346 for maximal NA expression (Fig. 2.8A). To better understand the need for a phenylalanine residue at this position, we examined the surrounding protein structure. We noticed two additional aromatic residues (Y385 and F479) directly adjacent to F346 that could potentially interact via π - π

stacking interactions (Fig. 2.8B). Mutation of either Y385 or F479 to alanine resulted in a selective decrease in NA abundance, though not quite as pronounced as that observed for NP:F346S (Fig. 2.8C). These data suggest that maximal expression of the NA gene segment depends upon a cluster of aromatic residues (F346, Y385, and F479) in NP.

2.3 Discussion

Our results describe a surprising role for NP in the selective regulation of NA segment synthesis during IAV infection. We found that substitutions at NP:F346 can specifically decrease the rate of NA vRNA synthesis while leaving the other gene segments largely unaffected. The specificity of this effect largely depends upon specific sequence motifs within the NA segment UTRs, demonstrating how interactions between NP and the individual gene segment UTRs can selectively modulate gene segment replication and expression.

Our results raise several additional questions, one of which concerns the role of the F346-Y385-F479 motif in NP function. The F346/Y385/F479 residues are highly conserved among IAV NP genes (Shu et al., 1993), indicating the importance of this motif for viral fitness in humans. NP promotes vRNA replication by stabilizing the positive-sense cRNA replicative intermediate and by acting as an elongation factor (Turrell et al., 2013; Vreede et al., 2004). Previous studies have demonstrated that these functions require both the RNA-binding and oligomerization activities of NP (Turrell et al., 2013; Vreede et al., 2011). Surprisingly, we found that NP:F346S does not appreciably affect the RNA

binding or oligomerization activities of NP in vitro (Fig. 2.7); however, more in-depth studies examining NP assembly and recruitment to vRNPs in the context of infection would aid in further substantiating whether NP:F346S affects these activities. If NP-RNA binding and oligomerization are not affected by NP:F346S, this raises the question of how substitutions at NP:F346 can modulate vRNA replication kinetics. One possibility is that NP positions F346-Y385-F479 govern interactions with other viral and/or cellular proteins involved in IAV gene segment replication (Lin et al., 2017; Marklund et al., 2012; Mayer et al., 2007; Minakuchi et al., 2017; Momose et al., 2001; Mondal et al., 2015; Morris et al., 2020; Newcomb et al., 2009; Robb et al., 2011; Turrell et al., 2015). Our results in Fig. 2.7C demonstrating that NP:F346S does not affect the association between PA and NP suggest that NP-polymerase complex interactions are not affected; however, we have not directly tested the effects of NP:F346S on interactions between NP and PB2/PB1 individually. We performed a preliminary mass spectrometry experiment to identify cellular and/or viral proteins that differentially associate with NP:WT and NP:F346S and found no clear hits (data not shown). Another intriguing possibility is that substitutions at NP positions F346-Y385-F479 affect the types of specific viral RNA species that are produced, such as small viral RNAs (svRNAs), which can modulate replication in a segment-specific manner (Perez et al., 2010, 2012). Finally, based on the NP structure, π - π stacking interactions between these residues may stabilize the structure of the loop regions where Y385 and F479 are located (Fig. 2.8); thus, these residues

may play a role in maintaining the structural integrity and conformational stability of NP.

The effects of the NP:F346S substitution clearly depend upon specific sequence motifs in the NA segment UTRs; however, this relationship is complicated. We identified two regions in the NA UTR that were important for determining both susceptibility to NP-dependent regulation and baseline NA segment expression levels in the context of WT NP: nucleotides 4 and 13 in the promoter/extended duplex region, and the 3' and 5' ORF proximal regions. The NA segment that demonstrated the lowest sensitivity to the effects of NP:F346S was one that contained a combination of UTR features from PR8 and Udorn: C's at positions 4 and 13 (as in WT PR8) and the Udorn-derived 3' and 5' ORF proximal sequences (Figs. 2.5 and 2.6B). Interestingly, this specific NA segment also exhibited a >100-fold reduction in baseline NA expression in the absence of NP:F346S (Fig. 2.6D). Disrupting this pairing by mutating one of the C's at position 4 or 13 or replacing one of the Udorn-derived ORF proximal regions with that from PR8 restored the effects of NP:F346S on NA synthesis (Figs. 2.5, 2.6A, and 2.6B). All NA constructs that we tested that were less sensitive to the effects of NP:F346S, with the exception of PR8:NA^{C4U}, also exhibited significantly lower baseline levels of NA expression under WT NP conditions, indicating that sensitivity to the effects of NP:F346S cannot be uncoupled from baseline NA expression levels.

For all eight genome segments, gene expression is influenced by the structure of the viral promoter that is formed by base-pairing interactions between the 3' and

5' UTRs (Lee et al., 2003). Base-pairing between positions 4 of the 3' and 5' UTRs and between positions 13 of the 3' UTR and 14 of the 5' UTR influences the promoter structure and stability and has been shown to be important for regulating gene segment replication and transcription (Lee et al., 2003; Lee and Seong, 1998; J. Wang et al., 2016; Wang and Lee, 2009; Widjaja et al., 2012). For the NA segment, a C at position 4 in the 3' UTR (which is unable to base-pair with the A at position 4 of the 5' UTR) promotes genome replication, while a U at this position favors mRNA transcription (Lee and Seong, 1998). Although it is intriguing to speculate that changing the identity of position 4 to a U would alter polymerase interactions with the promoter of NA to favor transcription and, therefore, decrease the effect of NP:F346S, which primarily affects vRNA synthesis (Fig. 2.1B and C), this is not supported by our data. In the context of NP:WT, the C4U substitution did not decrease NA vRNA abundance (Fig. 2.6D) as would be expected if the U at position 4 favors transcription at the cost of replication. Additionally, the NA:C4U mutation resulted in only a modest increase in expression (~2-fold) in the presence of NP:F346S (Figs. 2.6A and 2.6D). In the NA 3' UTR, it was a combination of the NA C4/13U mutations that accounted for the majority of the increase in resistance to NP:F346S, suggesting that the pairing of these two nucleotides is important (Fig. 2.6A).

As positions 4 and 13 of the 3' UTR are both involved in base-pairing interactions in the traditional panhandle structure of the IAV promoter (Ferhadian et al., 2018), they could affect the stability of the promoter and thus its interactions with the RNA polymerase. Although the ORF proximal region of the PR8 NA 3' UTR is

highly conserved compared to other N1 NA segments, the PR8 NA 3' UTR is one of the only N1 genes to have C's at both positions 4 and 13 (Zhao et al., 2014). Further, NA is the only PR8 gene segment to have C's at both of these positions. Having C's at both positions 4 and 13 of the NA 3' UTR decreases the number of predicted base-pairing interactions within the traditional panhandle structure of the promoter and would therefore be expected to decrease its stability. This unique feature may make the NA segment of PR8 uniquely dependent on WT NP to facilitate the stable interaction between the promoter and the viral replicase and, therefore, promote maximal vRNA synthesis.

The ORF proximal regions of the UTRs also influence gene segment expression; however, the exact mechanism(s) remains unclear (Ma et al., 2013; Maeda et al., 2004; Wang et al., 2015; Xiao et al., 2021; Zheng et al., 1996). Consistent with this, we observed that the ORF proximal regions of the NA UTRs play important roles both in regulating the baseline expression level of the NA segment and in determining sensitivity to NP:F346S. Relative to Udorn and the remaining seven gene segments of PR8, the PR8 NA segment UTRs harbor a unique extended stretch of base-pairing from 3' nt 14/5' nt 15 to 3' nt 17/5' nt 18 located within the poly U stretch of the 5' UTR. Given the hypothesized role for NP in promoter escape and elongation (Turrell et al., 2013), the NA segment may be particularly dependent on WT NP (and thus sensitive to NP:F346S) to allow the polymerase to bypass this extended base-pairing region during elongation of the nascent vRNA. Another potential explanation of our observation that specific pairings between nucleotides 4 and 13 of the NA 3' UTR and the NA UTR ORF proximal

regions were associated with the greatest resistance to the effects of NP:F346S is that these regions interact to form an RNA secondary structure (either within a single UTR and potentially the proximal ORF sequence or between the 3' and 5' UTRs) that may compete with the promoter structure for polymerase binding or stall the polymerase at the beginning of the template. WT NP could be required to melt this structure to allow replication to move beyond the initiation step. Substitutions at positions F346-Y385-F479 may compromise the ability of NP to facilitate bypassing of this RNA structure, resulting in the decreased expression of the segment. The *in silico* structural prediction algorithm M-fold does not predict that any of the NA mutations tested would result in major structural differences in the 3' UTR (where segment replication initiates) (data not shown). This does not preclude these mutations having significant effects on the UTR structure in the dynamic context of active RNPs during infection. Altogether, our data suggest that the unique sequence and presumed structure of the PR8 NA segment UTRs confer elevated sensitivity to perturbations of NP function and thus likely explain the segment specificity of the effects of NP:F346S.

Through mutations in the UTR sequences of individual segments, IAV can more finely coordinate the expression of the eight individual gene segments without altering the protein coding capacity of the segments or polymerase activity. Similar to our findings, a recent study demonstrated that the 3' UTR of the HA segment played a role in regulating HA expression in a segment-specific manner and that this regulation was observed only when the HA segment had to compete with the remaining seven segments for replication/transcription (Xiao et al.,

2021). An additional study using a reporter system also found that the UTR sequences of the segments affected the ability of the segments to compete with one another for access to the viral polymerase (Widjaja et al., 2012). Altogether, these studies highlight the importance of the individual segment UTR sequences in maintaining the optimal balance in expression of the eight gene segments during infection. Our results further demonstrate how slight perturbations in polymerase or NP function can affect the expression of specific segments to a greater degree than others.

What are the implications of NA (and potentially other viral gene segments) being uniquely sensitive to individual substitutions in NP? HA and NA facilitate viral particle attachment and release, respectively, and balancing these activities is necessary for maintaining viral fitness (Benton et al., 2015; Byrd-Leotis et al., 2017; Casalegno et al., 2014; Gaymard et al., 2016; Guo et al., 2018; Lai et al., 2019; McAuley et al., 2019; Wagner et al., 2002; Xu et al., 2012; Yen et al., 2011). HA and NA evolve at higher rates than the rest of the IAV genome due to immune selection (Bhatt et al., 2011). Immune escape substitutions within HA and/or NA often alter glycoprotein function and require compensatory mutations to restore their functional balance and viral fitness (Dai et al., 2021.; Das et al., 2011, 2013; Doud et al., 2017; Hensley et al., 2009, 2011, Kosik et al., 2018, 2019; Prachanronarong et al., 2019). If substitutions in NP can tune NA expression, it expands the number of available genetic pathways for maintaining HA/NA functional balance. The functional link between the NP and NA segments that we establish here also has important consequences for reassortment, as the

need to maintain compatible NP and NA genotypes may constrain the repertoire of viable reassortant progeny when heterologous viral strains mix. Finally, variation in NP requirements between segments could influence patterns of expression kinetics, as the concentration of NP within the cell is dynamic over time. Altogether, our results highlight the potential of genome segmentation to facilitate dynamic changes in gene expression patterns through mechanisms that are not readily available to non-segmented viruses. This regulatory agility may help promote viral adaptation in response to changing host environments.

In summary, we identified a novel mechanism through which interactions between NP and other gene segment UTRs facilitate selective regulation of viral gene expression. Our data reveal a new functional domain in NP and suggest a broader role for NP in selective regulation of individual viral gene segments. These findings demonstrate how the expression of individual gene segments can be modulated to maximize viral fitness under different host conditions.

2.4 Materials and Methods

Plasmids

The A/Puerto Rico/8/34 and A/Udorn/72 reverse genetics plasmids were gifts from Adolfo Garcia-Sastre (Icahn School of Medicine at Mt. Sinai) and Kanta Subbarao (Doherty Institute), respectively. The pCI vector was provided by Joanna Shisler (National Science Foundation [NSF]). The lentivirus generation plasmids pHAGE2-EF1aInt WSN HA W, HDM Hgpm2, HDM tatlb, pRC CMV Rev1b, and HDM VSV-G were provided by Jesse Bloom (Fred Hutch Cancer

Research Center). The peGFP-C1 plasmid for generating C-terminally eGFP-tagged proteins was provided by Andrew Mehle (University of Wisconsin—Madison).

NP mutant, PR8 HA/NA chimera, and PR8/Udorn NA chimera inserts were generated via site-directed mutagenesis and overlap extension PCR. For the chimeric Udorn-PR8 NA inserts with the Udorn NA ORF, the A763C silent mutation was introduced into the Udorn NA sequence to disrupt an internal *Lgul* restriction site. *Lgul* restriction sites were added to both ends of the inserts for cloning into the pDZ vector. To clone IAV ORFs into the pCI mammalian expression vector, inserts were generated by PCR with primers that bound to the terminal regions of the PR8 NA/HA/NP/PB1 ORFs. *EcoRI/SalI* or *MluI/KpnI* restriction sites were added to the 5'/3' ends of the PR8 NA/HA/NP and PR8 PB1 inserts, respectively, and internal *EcoRI/SalI* restriction sites in HA were removed via site-directed mutagenesis. For epitope- and eGFP-tagged NP expression vectors, mutations in PR8 NP were introduced via site-directed mutagenesis. C-terminal 6× His tags were introduced by performing PCR with primers designed to add a C-terminal 6× His tag before the stop codon of the PR8 NP ORF. A C-terminal HA tag was added to PR8 PA by performing PCR with primers designed to add a C-terminal HA tag before the stop codon of the PR8 PA ORF. *EcoRI/SalI* restriction sites were added to the 5'/3' ends of the inserts, respectively. For cloning into the peGFP-C1 plasmid, *BspEI/KpnI* restriction sites were introduced to the 5'/3' ends, respectively, of the PR8 NP (WT/F346S/R416A) ORFs. For the lentiviral expression vector (pHAGE-EF1aInt),

inserts were generated by PCR with primers designed to bind to the 5'- and 3'-terminal regions of the PR8 HA ORF and introduce *Bam*HI/*Not*I restriction sites to the 5'/3' ends, respectively.

The inserts were then digested with the following restriction enzymes: for pDZ, *L*gull (Thermo Fisher; FD1934); for pCI, *E*coRI/*S*alI (Thermo Fisher; FD0274/FD0644) and *M*luI/*K*pnl (Thermo Fisher; FD0564/FD0524) for the PR8 HA/NP/NP (WT/F346S/R416A)-C-His tag/NA/PA-HA tag ORF and PR8 PB1 ORF inserts, respectively; for peGFP-C1, *B*spEI/*K*pnl (Thermo Fisher; FD0534/FD0524); and for pHAGE-EF1aInt, *B*amHI/*Not*I (Thermo Fisher; FD0054/FD0593). The inserts were then ligated into the desired vectors (New England BioLabs [NEB]; T4 DNA ligase; M0202S), and transformed into chemically competent *Escherichia coli* cells (NEB; Turbo C2984H) via the heat shock method. Insert sequences were confirmed via Sanger sequencing.

Cells

Madin-Darby canine kidney (MDCK), MDCK-SIAT1 cells, and 293T cells were obtained from Jonathan Yewdell (NIH), Jesse Bloom (Fred Hutchinson Cancer Research Center), and Joanna Shisler (NSF), respectively, and were maintained in Gibco's minimal essential medium (MEM) with GlutaMax (Life Technologies) supplemented with 8.3% fetal bovine serum (Seradigm) (MEM+FBS) and incubated at 37°C with 5% CO₂. MDCK and MDCK-SIAT1 cells were used for virus propagation and infection experiments due to their ability to support robust replication of IAV and their tolerance to the trypsin necessary to convert the HA

protein into its fusion-competent state. 293T cells were used for transfections to generate viruses via the reverse genetics system or to express viral proteins due to their high transfection efficiency.

MDCK cells expressing the PR8 HA protein were generated via lentiviral transduction. 293T cells were transfected with 250 ng each of the plasmids required for lentivirus generation (HDM Hgpm2, HDM tatlb, pRC CMV Rev1b, and HDM VSV-G) and 1 µg of the transfer vector pHAGE PR8 HA (generated as described above). One day post-transfection, the transfection medium was replaced with 2 mL of MEM+FBS. The next day, the lentiviral supernatant was collected and 1 mL was used to infect three wells of MDCK cells plated in a 6-well plate at 10% confluency. Two days post-transduction, the MDCK cells were harvested and combined and surface stained with an anti-HA antibody (H36-26 AF488), and positive cells were sorted out via fluorescence-activated cell sorting (FACS). MDCK-HA cells were used to propagate viruses that had defects in HA protein expression and replicated poorly in the MDCK and MDCK-SIAT1 cell lines.

Viruses

Recombinant A/Puerto Rico/8/1934 (H1N1) (rPR8) and A/Udorn/72 (H3N2) (Udorn) viruses were generated using 8- and 12-plasmid reverse genetics systems, respectively. The rPR8 clones differ from the published sequence (GenBank accession no. AF389115 to AF389122) at two positions: PB1 A549C (K175N) and HA A651C (I207L) (numbering from initiating Met). Viruses

containing point mutations in the NP or NA segments or the PR8 HA/NA or PR8/Udorn NA chimeric segments were generated by rescuing the viruses using plasmids containing those specific segments (generated as described above). The Udorn HA segment-encoding plasmids used were found to have the following mutations: A81G (N18D), C129T (H34Y), G1103T (silent), T1486A (F486Y), and A1614G (N529D) relative to the Udorn HA reference sequence (GenBank accession no. AX350190). The Udorn NA segment-encoding plasmids used were found to have the following mutations: G992A (A325P) and C1216G (D399E) relative to the Udorn NA reference sequence (GenBank accession no. CY009638).

Viruses were rescued by transfecting 293T cells with 500 ng each of the relevant reverse genetics plasmids using JetPrime (Polyplus) according to the manufacturer's instructions. For the PR8 HA/NA chimeric and PR8/Udorn NA chimeric PR8 NA ORF-containing viruses with NP:F346S, the cells were also transfected with 500 ng of the pCI PR8 NP, and the pCI PR8 HA/NA plasmids or pCI PR8 NA plasmid, respectively, to promote viral growth via expression of the native viral proteins. At 18 to 24 h post-transfection, the medium was replaced with viral growth medium (MEM, 1 mM HEPES, 1 µg/mL of tosylsulfonyl phenylalanyl chloromethyl ketone [TPCK] trypsin [Worthington Biochemical Corporation, Lakewood, NJ], 50 µg/mL of gentamicin) containing 2×10^5 MDCK cells. For the viruses with the chimeric PR8 HA/NA segments, the viral growth medium was modified by adding 2×10^5 PR8 HA+ MDCK cells instead of MDCK cells. Transfection supernatants were collected 24 h after medium change.

To generate the seed stocks of the PR8 NP:WT/F346S, Udorn NP:WT/F346S, PR8 Udorn HA/NA NP:WT/F346S, PR8 NP point mutant viruses, PR8 NA codon shuffle NP:WT/F346S viruses, PR8:NA^{C13U} NP:WT/F346S, PR8:NA^{C4/13U} NP:WT/F346S, PR8:NA^{Udorn UTR} NP:WT/F346S, and PR8:Udorn HA,NA^{PR8 ORF Proximal UTR} NP:WT/F346S viruses, transfection supernatants were plaqued, and a single plaque was used to infect a single well of MDCK cells in a 6-well plate. Viral growth was performed in viral growth medium (MEM, 1 mM HEPES, 1 µg/mL of TPCK trypsin, 50 µg/mL of gentamicin). Seed stocks were harvested and clarified (14,000 rpm, 15 min, 4°C) between 24 and 72 h post-infection. Only seed stocks were generated for the PR8 NA codon shuffle NP:WT/F346S, PR8:NA^{C13U} NP:WT/F346S, PR8:NA^{C4/13U} NP:WT/F346S, PR8:NA^{Udorn UTR} NP:WT/F346S, and PR8:Udorn HA,NA^{PR8 ORF Proximal UTR} NP:WT/F346S viruses. MDCK cells in a T75 or T175 flask were then infected with the seed stocks at a multiplicity of infection (MOI) of 0.001 or 0.01 50% tissue culture infective dose (TCID₅₀)/cell, respectively, in viral growth medium, and the working stocks were harvested 24 to 72 h post-infection and clarified (3,500 rpm, 15 min, 4°C).

To generate the seed stocks of PR8 HA/NA chimeric viruses, PR8 HA+ MDCK cells in a single well of a 6-well plate were infected with 1 mL of transfection supernatant for 1 h at 37°C with rocking, and then the transfection supernatant was removed and replaced with 3 mL of viral growth medium with 0.5 µg/mL of TPCK trypsin and left to incubate for up to 48 h. The infection supernatants were harvested and clarified (14,000 rpm, 15 min, 4°C), and 1 mL of the infection supernatant was used to perform the next passage of the viruses in the PR8 HA+

MDCK cells. The passaging continued until cytopathic effect was observed. This occurred within the first two passages for all the viruses and between 24 and 48 h post-infection. The PR8 HA/NA UTR swap NP:WT virus had a mutation in the PR8 HA ORF NA UTR segment (C33A→L5I). The seed stocks of the PR8:NA^{C4U} NP:WT/F346S, PR8:NA^{Udorn ORF proximal UTR} NP:WT/F346S, PR8:NA^{C4U/Udorn ORF proximal UTR} NP:WT/F346S, PR8:NA^{C13U/Udorn ORF Proximal UTR} NP:WT/F346S, PR8:NA^{Udorn 3' ORF Proximal UTR} NP:WT/F346S, PR8:NA^{Udorn 3' UTR} NP:WT/F346S, PR8:NA^{Udorn 5' UTR} NP:WT/F346S, and PR8:Udorn HA,NA^{PR8 UTR} NP:WT/F346S viruses were generated as described above in MDCK cells with 1 µg/mL of TPCK trypsin used in the viral growth medium.

Virus titers were determined by TCID₅₀ assay on MDCK cells, or by determining the fraction of viral particles expressing NP (NP-expressing unit [NPEU]) via flow cytometry on infected MDCK cells using the anti-NP AF647 (HB65) antibody.

Next-generation sequencing of viruses

Viral RNA was extracted from 140 µL of the viral infection supernatant using the QIAamp kit (Qiagen) and eluted in 60 µL of nuclease-free water (Ambion).

Contaminating DNA was removed using the Qiagen RNase-free DNase set, and then the RNA was cleaned using the RNeasy kit (Qiagen) and eluted in 30 µL of nuclease-free water (Ambion). cDNA was synthesized using the Superscript III reverse transcriptase kit (Thermo Fisher) as follows: 1 µL of 2 µM MBTUni-12 primer (5'-ACGCGTGATCAGCRAAAGCAGG-3'), 1 µL of 10 mM deoxynucleoside triphosphate (dNTP) mix (NEB; N0447S), and 8 µL of nuclease-

free water (Ambion) were added to 3 μ L of RNA and then incubated at 65°C for 5 min and then 4°C for 2 min. Four microliters of 5 $\times$ first-strand cDNA synthesis buffer, 1 μ L of 0.1 M dithiothreitol (DTT), 1 μ L of SUPERase-In RNase inhibitor (Invitrogen; AM2696), and 1 μ L of Superscript III reverse transcriptase (Invitrogen; 18080-044) were added to the reaction, and then the reaction mixture was incubated at 45°C for 50 min. The cDNA was then stored at -20°C. The PCR to simultaneously amplify all eight gene segments was performed using Phusion polymerase (NEB; M0530L) as follows: 2.5 μ L of 10 μ M MBTUni-12 primer, 2.5 μ L of 10 μ M MBTUni-13 primer (5'-ACGCGTGATCAGTAGAAACAAGG-3'), 10 μ L 5 $\times$ HF Phusion buffer, 1 μ L of 10 mM dNTP mix (NEB; N0447S), 0.5 μ L of Phusion polymerase (NEB; M0530L), and 28.5 μ L of nuclease-free water (Ambion) were added to 5 μ L of cDNA. The cycling conditions for the PCR were as follows: 98°C for 30 s, 25 cycles of 98°C for 10 s/57°C for 30 s/72°C for 1 min 30 s, 72°C for 5 min, and a 4°C hold. The PCR products were cleaned using the Invitrogen PureLink PCR purification kit (Invitrogen; K310002) using the buffer for the <300-bp cutoff and eluted in 30 μ L of nuclease-free water (Ambion). The PCR products were then subjected to next-generation sequencing using the Illumina NovaSeq or MiSeq platform.

Time course infections

MDCK cells were infected with the PR8 NP:WT/F346S or Udorn NP:WT/F346S viruses at an MOI of 0.1 NPEU/cell in a 24-well plate for 1 h at 37°C. At 1 h post-infection, the virus supernatant was replaced with 0.5 mL of MEM+FBS. At 3 h post-infection, the MEM+FBS was replaced with 0.5 mL of NH₄Cl medium (MEM,

50 mM HEPES buffer, 20 mM NH₄Cl [pH 7.2]) to prevent viral spread (Sun et al., 2020; Sun and Brooke, 2018). The cell monolayers were harvested 4, 8, and 12 h post-infection, and cellular RNA was extracted using the RNeasy kit (Qiagen). Reverse transcription was performed using the Verso cDNA synthesis kit (Thermo Fisher). The reaction mixtures were set up as follows: 4 µL of RNA, 4 µL of 5× cDNA synthesis buffer, 2 µL of dNTP mix, 1 µL of 10 µM PR8 RT_4A primer (5'-AGCAAAAGCAGG-3'), 1 µL of RT enhancer, 1 µL of Verso enzyme mix, 7 µL of nuclease-free water; mixtures were incubated at 45°C for 50 min and 95°C for 2 min and held at 4°C. cDNA was stored at -20°C. Quantitative real-time PCR on cDNA was carried out using Power SYBR green PCR master mix (MM) (Thermo Fisher) on a QuantStudio 3 thermal cycler (Thermo Fisher). qPCR primer sets used are listed in Table 2.1. Reactions were set up as follows: 2 µL of cDNA, 10 µL of 2× Power SYBR green MM, 1 µL of 10 µM forward primer, 1 µL of 10 µM reverse primer, and 6 µL of nuclease-free water (Ambion). The cycling conditions were as follows: 50°C for 2 min, 95°C for 10 min, and then 95°C for 15 s, followed by 60°C for 1 min repeated 40 times.

Analysis of primary viral transcription in infected cells

MDCK-SIAT1 cells were infected with the PR8 NP:WT/F346S viruses at an MOI of 5 TCID₅₀s/cell in a 6-well plate in the presence of 100 µg/mL of cycloheximide (Sigma-Aldrich). Infected cells were harvested at 2 and 6 h post-infection, and cellular RNA was extracted using the RNeasy kit (Qiagen). Reverse transcription was performed using the Superscript III reverse transcriptase kit (Thermo Fisher). The reactions were set up as follows: 4 µL of RNA, 0.5 µL of 100 µM

oligo(dT)₂₀ primer (IDT), 1 µL of 10 mM dNTP, and 6.5 µL of nuclease-free water (Ambion) were incubated at 65°C for 5 min and then 4°C for 1 min. Four microliters of 5× first-strand RNA buffer, 1 µL of 0.1 M DTT, 2 µL of Superscript III RNase inhibitor (Thermo Fisher), and 1 µL of Superscript III reverse transcriptase (Thermo Fisher) were added to the previous reaction and incubated at 50°C for 60 min and 70°C for 15 min and held at 4°C. cDNA was stored at -20°C. Quantitative real-time PCR on cDNA was carried out using Power SYBR green PCR master mix (Thermo Fisher) on a QuantStudio 3 thermal cycler (Thermo Fisher) as described in “Time course infections” above, with the exception of 1.5 µL of cDNA being used as a template.

4SU RNA pulse

MDCK cells at 60 to 70% confluency were infected with the PR8 NP:WT/F346S viruses at an MOI of 5 TCID₅₀/cell and incubated at 37°C for 1 h. The virus was aspirated and then 3 mL of MEM+FBS was added to each well. At 7 hpi, the MEM+FBS was replaced with 1 mL of fresh MEM+FBS containing 500 µM 4-thiouridine (4SU; Tri-Link Biotechnologies; N-1025). The cells were kept in the dark during the labeling process to prevent cross-linking of 4SU to cellular proteins. At 1 h after labeling, the cells were harvested and cellular RNA was extracted using the RNeasy kit (Qiagen).

The cellular RNA was then biotinylated by performing a reaction with EZ-Link HPDP-biotin (Thermo Scientific). Reactions conditions were as follows: 10 µg of RNA and biotin-HPDP (0.2-µg/µL final concentration) were added to biotinylation

buffer (10 mM Tris-HCl [pH 7.5], 1 mM EDTA), resulting in a total reaction volume of 250 μ L, and then incubated at room temperature with end-over-end rotation for 2 h protected from light. RNA was then extracted using the chloroform-isoamyl alcohol procedure performed as follows. A total of 400 μ L of chloroform-isoamyl alcohol (49:1 ratio) was added to each reaction, mixed, and then added to a 2-mL Quanta Bio 5PRIME Phase Lock Heavy Tube. The phases were separated by centrifugation (full speed, 5 min, 4°C), and the aqueous layer (top) was transferred to a new 1.5-mL microcentrifuge tube. The previous steps were then repeated one additional time. RNA was precipitated as follows. One volume of isopropanol, 1/10 volume of 5 M NaCl, and 1 μ L of 15- μ g/mL GlycoBlue coprecipitant (Invitrogen) were added to each sample and mixed. The samples were then frozen at -70°C overnight. On the next day, the samples were thawed and RNA was pelleted via centrifugation (full speed, 20 min, 4°C). The pellet was then washed twice with 400 μ L of 80% ethanol (full speed, 5 min, 4°C). The pellet was then air dried at room temperature for 5 min and resuspended in 20 μ L of nuclease-free water (Ambion).

Biotinylated RNAs were then selectively purified using the μ MACS streptavidin kit (Miltenyi Biotec) as follows: A total of 15 μ L of RNA was added to 85 μ L of nuclease-free water (Ambion) and denatured by incubation at 65°C for 10 min, followed by cooling on ice for 5 min. A total of 100 μ L of Miltenyi streptavidin beads was added to each reaction and incubated at room temperature for 15 min with end-over-end rotation. Meanwhile, the μ MACS columns were equilibrated by adding 100 μ L of the equilibration buffer for nucleic acid applications and then

washed with 1 mL of wash buffer (100 mM Tris-HCl [pH 7.5], 10 mM EDTA, 1 M NaCl). The biotinylated RNA-streptavidin bead solution was then added to the columns, and then the columns were washed with 0.9 mL of 65°C wash buffer three times, followed by 0.9 mL of room temperature wash buffer three times. The biotinylated RNA was then eluted with 150 µL of 0.1 M DTT. The RNA was stored at -70°C.

Reverse transcription was performed using the Superscript III reverse transcriptase kit (Thermo Fisher). The reactions were set up as follows: 2 µL of RNA, 0.5 µL of 10 µM PR8 RT_4A primer (5'-AGCAAAAGCAGG-3'), 1 µL of 10 mM dNTP, and 8.5 µL of nuclease-free water (Ambion) were incubated at 65°C for 5 min and then 4°C for 1 min. Four microliters of 5× first-strand RNA buffer, 1 µL of 0.1 M DTT, 2 µL of Superscript III RNase inhibitor (Thermo Fisher), and 1 µL of Superscript III reverse transcriptase (Thermo Fisher) were added to the previous reaction, and the reaction mixtures were incubated at 50°C for 60 min and 70°C for 15 min and held at 4°C. cDNA was stored at -20°C.

Quantitative real-time PCR on cDNA was carried out using Power SYBR green PCR master mix (Thermo Fisher) on a QuantStudio 3 thermal cycler (Thermo Fisher) as described in “Time course infections” above, with the exceptions of using 1.5 µL of cDNA and the following cycling conditions: 95°C for 10 min and then 95°C for 15 s, followed by 54/57°C for PR8 NA/HA vRNA, respectively, for 1 min, repeated 40 times.

Quantification of single-replication-cycle viral gene expression levels

MDCK cells were infected with viruses at an MOI of 0.1 NPEU/cell (0.03 NPEU/cell for PR8:Udorn HA,NA^{PR8} UTR NP:WT/F346S and PR8:Udorn HA,NA^{PR8} ORF Proximal UTR NP:WT/F346S viruses due to their low titer) or an MOI of 0.1 TCID₅₀/cell (PR8 NP:WT versus F346S versus Y385A versus F479A) in a 24-well plate. At 1 hpi, the viral supernatant was removed and replaced with TCID₅₀ medium (MEM, 1 mM HEPES, 0.5 or 1 µg/mL of TPCK trypsin, 50 µg/mL of gentamicin). Cells were harvested at 8 hpi, and cellular RNA was extracted using the RNeasy kit (Qiagen). For viruses derived from direct passage of the transfection supernatant in MDCK cells or a single infection with a plaque supernatant, the cellular RNA was treated with RNase-free DNase I (Qiagen) and cleaned using the RNeasy kit (Qiagen). RT-qPCR was carried out as described in “Time course infections” above.

Flow cytometry to detect viral protein expression in singly infected cells

MDCK cells were infected at an MOI of 0.03 TCID₅₀/cell in a 6-well plate, or 1.5×10^6 MDCK cells per well were infected with 10^{-1} to 10^{-4} dilutions of the viral stock in a 6-well plate (for NPEU determination). At 1 hpi, the virus was replaced with 3 mL of MEM+FBS. At 3 hpi, the MEM+FBS was replaced with 3 mL of NH₄Cl medium (MEM, 50 mM HEPES buffer, 20 mM NH₄Cl [pH 7.2]) to prevent secondary infection (Sun et al., 2020; Sun and Brooke, 2018). At 16 hpi, the cells were harvested, fixed, and permeabilized with foxP3 fix/perm buffer (eBioscience). Cells were subsequently stained with one or multiple of the following antibodies: PR8 NA: rabbit anti-NA (08-0096-03 EXSANG 3/9/09)

followed by donkey anti-rabbit phycoerythrin (PE) (711-116-152, lot 121465), PR8 NP: HB65 AF647 or PacB, Udorn NA: goat anti-N2 NA primary followed by donkey anti-goat (705-116-147) PE secondary, and Udorn HA:H14A2 AF647. The cells were run on a BD LSR II flow cytometer and analyzed using FlowJo version 10.1 (Tree Star, Inc.). Viral protein expression levels were determined from the geometric mean fluorescence intensity (GMFI) of the fluorophore associated with each protein. NPEU titers were calculated by dividing the number of infected cells (percent NP positive $\times$ total number of cells) by the dilution factor and the volume of the inoculum.

PR8 NP:WT/F346S-C-His protein purification

A total of 100 μ g of the pCI PR8 NP:WT-C-His or pCI PR8 NP:F346S-C-His plasmid was transiently transfected into 100 mL cultures of HEK Expi-293-F cells with ExpiFectamine according to the company protocol (Thermo Fisher). The transfected cells were pelleted (1,000 rpm, 5 min, 4°C) and resuspended in 3 mL of equilibration buffer (phosphate-buffered saline [PBS], 10 mM imidazole, 1 $\times$ cOmplete EDTA-free protease inhibitor cocktail [Sigma-Aldrich]). Cells were kept frozen at -70°C until the next step of the purification procedure was performed. The cells were thawed and then lysed via the freeze-thaw method: incubation in a dry ice-ethanol bath followed by a 42°C water bath, repeated twice. The chromosomal DNA was then sheared by passing the lysate through an 18-gauge needle 4 times. The cellular debris was pelleted by centrifugation (3,000 $\times$ g, 15 min, 4°C), and the clarified lysates were transferred to new tubes. The

clarified lysates were then treated with RNase A (50 µg/mL final concentration) for 2 h at room temperature.

His-tagged proteins were then selectively purified using the HisPur nickel-nitrilotriacetic acid (Ni-NTA) spin purification kit, 0.2 mL (Thermo Fisher). To improve the purity of the eluted protein fractions and increase the protein yield, the eluate fractions for the PR8 NP:WT/F346S C-His proteins were combined, the buffer for the eluate fractions was exchanged to equilibration buffer using Pierce protein concentrators (Polyethersulfone [PES] 10-kDa molecular weight cutoff [MWCO]; 0.5 mL; Thermo Fisher Scientific), and the HisPur Ni-NTA column purification was repeated once more. The eluate fractions for each protein were then combined, added to a Slide-A-Lyzer dialysis cassette (extra strength; 10-kDa MWCO; 0.5 to 3 mL capacity; Thermo Scientific), and dialyzed in PBS at 4°C overnight. The dialyzed protein samples were concentrated using Pierce concentrators (PES; 10 kDa MWCO; 0.5 mL; Thermo Scientific), and the protein concentrations were determined using the Pierce Coomassie Plus Bradford assay kit (Thermo Scientific).

BLI

Kinetics of binding of PR8 NP:WT/F346S C-His proteins to biotinylated single-stranded RNA (ssRNA) was determined by biolayer interferometry (BLI) using an Octet Red96e instrument (FortéBio) at room temperature. Anti-streptavidin biosensors (FortéBio) were washed in 180 µL of 1× kinetics buffer (0.002% [vol/vol] Tween 20 in 1× PBS [pH 7.4]) for 30 min at room temperature. The

experiment consisted of five steps: (i) baseline, 60 s with 1× kinetics buffer; (ii) loading, 300 s with 5' biotinylated 24-nt ssRNA (5'-UUUGUUACACACACACACGCUGUG-3'; IDT) at 1 μM; (iii) baseline, 60 s with 1× kinetics buffer; (iv) association, 60 s with 50, 100, 200, and 400 nM PR8 NP:WT/F346S C-His protein; and (v) dissociation: 300 s with 1× kinetics buffer. Octet Data Acquisition (version 11.1; FortéBio) software was used to obtain biolayer interferometry data. To calculate the dissociation constant (K_D) via curve fitting, a 1:1 binding model was used. Octet Data Analysis (version 11.1; FortéBio) software was used to analyze binding kinetics. Negative controls were set up with 1× kinetics buffer replacing 1 μM biotinylated ssRNA in the loading step.

Transfection protocols for eGFP/His-tagged NP protein coimmunoprecipitation and HA-tagged vRNP complex pulldown experiments

For the eGFP/His-tagged NP protein coimmunoprecipitation experiment, 293T cells in a 10 cm dish were transfected with 5 μg each of one of the following pairs of plasmids using JetPrime (Polyplus) according to the manufacturer's instructions: pCI PR8 NP:WT-C-His and peGFP-PR8 NP:WT, pCI PR8 NP:F346S-C-His and peGFP-PR8 NP:F346S, or pCI PR8 NP:R416A-C-His and peGFP-PR8 NP:R416A. For the HA-tagged vRNP complex pulldown experiment, 293T cells in a 10 cm dish were transfected with 2 μg each of plasmids pCI PR8 PB2, pCI PR8 PB1, pCI PR8 PA-HA tag, pCI PR8 NP:WT/F346S/R416A, and pHH21 PR8 NA using JetPrime (Polyplus) according to the manufacturer's instructions.

Coimmunoprecipitation

At 24 h post-transfection, the cells were lysed in MOPS coimmunoprecipitation lysis buffer (20 mM 3-morpholinopropane-1-sulfonic acid [MOPS; pH 7.5], 150 mM NaCl, 0.5% Igepal CA-630, 1× cOmplete EDTA-free protease inhibitor cocktail [Sigma-Aldrich]) and clarified via centrifugation (20,000 × g, 15 min, 4°C). Mouse anti-His (HIS.H8) antibody (Invitrogen; for eGFP/His-tagged NP protein coimmunoprecipitation) or mouse anti-HA tag (2-2.2.14) antibody (Invitrogen; for the HA-tagged vRNP complex pulldown experiment) was added to the clarified lysates to a final dilution of 1:100, and the lysates were incubated with the antibody with end-over-end rotation overnight at 4°C. Antigen-antibody complexes were selectively purified using Pierce protein A agarose (Thermo Fisher Scientific) according to the manufacturer's instructions.

Western blotting

For the eGFP/His-tagged NP protein coimmunoprecipitation experiment, Western blotting was performed on the immunoprecipitated protein samples using the mouse anti-eGFP (F56-6A.1.2.3) primary antibody (Invitrogen; 1:1,000) followed by rat anti-mouse horseradish peroxidase (HRP)-conjugated (187.1) secondary antibody (BD Biosciences; 1:500) to detect coimmunoprecipitated PR8 NP:WT/F346S/R416A eGFP-tagged proteins, or mouse anti-His (HIS.H8) primary antibody (Invitrogen; 1:1,000) followed by the rat anti-mouse HRP-conjugated (187.1) secondary antibody (BD Biosciences; 1:500) to assess the pulldown efficiency of the PR8 NP:WT/F346S/R416A-C-His proteins. A Western blot was also performed on the cell lysates using the rabbit anti-NP primary

antibody (GeneTex; GTX125989; 1:1,000) followed by the goat anti-rabbit HRP-conjugated (G-21234) secondary antibody (Invitrogen; 1:10,000) to detect the expression efficiencies of the His-tagged and eGFP-tagged PR8 NPs in transfected cells.

For the HA-tagged vRNP complex pulldown experiment, Western blotting was performed on the immunoprecipitated protein samples using mouse anti-HA tag (2-2.2.14) primary antibody (Invitrogen; 1:1,000) followed by the rat anti-mouse HRP-conjugated (187.1) secondary antibody (BD Biosciences; 1:500) to assess pulldown efficiency of the HA-tagged PA, or rabbit anti-NP polyclonal antibody (GeneTex; GTX125989; 1:1,000) followed by goat anti-rabbit HRP-conjugated (G-21234) secondary antibody (Invitrogen; 1:10,000) to detect coimmunoprecipitated NP.

Proteins were visualized using the SuperSignal Pico West Plus chemiluminescent substrate (Thermo Fisher Scientific) and imaged using the iBright CL1000 imaging system (Invitrogen).

Quantifying gene segment ratios in viral stocks

140 μ L of the viral stock was treated with 0.25 μ g of RNaseA for 30min at 37°C. The viral RNA was then extracted using the QIAamp Viral RNA Extraction Kit (Qiagen) and eluted in 60 μ L of nuclease-free water (Ambion). RNA was treated with RNase-free DNaseI (Qiagen) and cleaned using the RNeasy Kit (Qiagen). RT-qPCR was carried out as described in the “Time-course infections” section.

2.5 Figures and Table

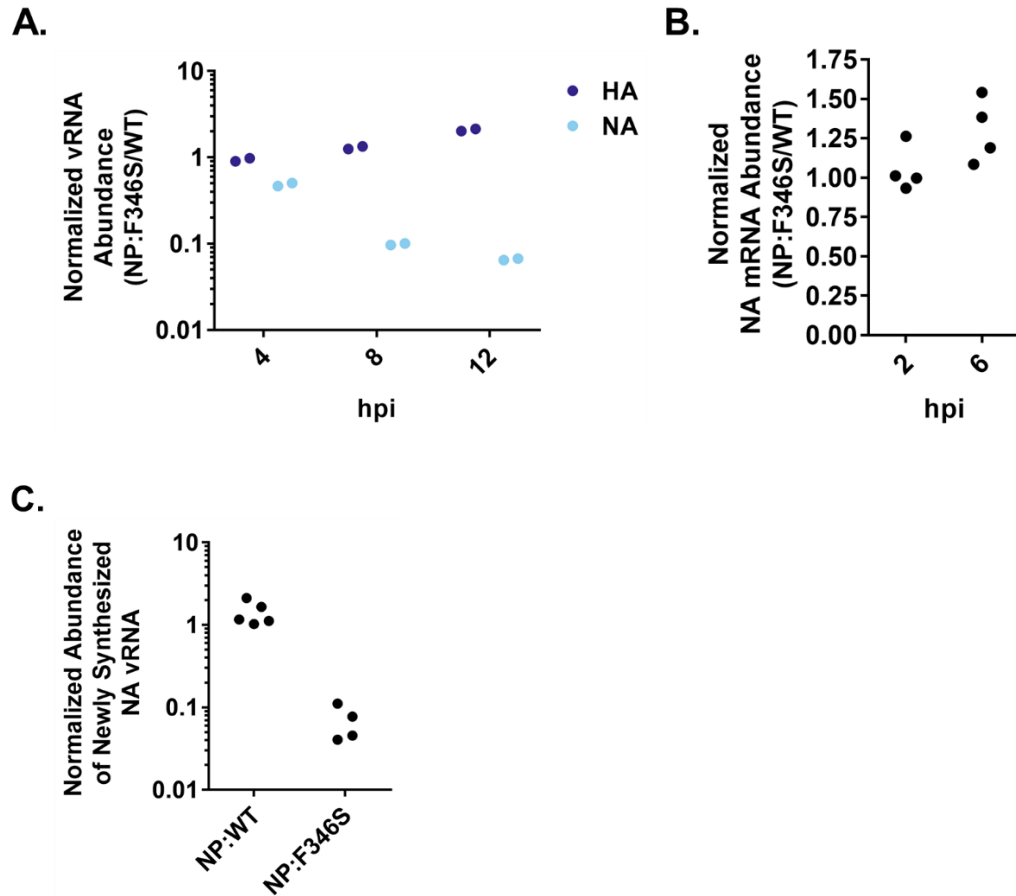


Fig. 2.1. NP:F346S affects NA vRNA replication but not primary transcription. (A) NP-normalized NA and HA vRNA abundances as determined by RT-qPCR on total cellular RNA in MDCK cells infected with the PR8 NP:F346S virus (MOI = 0.1 NPEU/cell) relative to PR8 NP:WT at the indicated hours post-infection (hpi). Secondary infection was blocked via the addition of 20 mM NH_4Cl 3 hpi (Sun et al., 2020; Sun and Brooke, 2018). The data shown are for individual cell culture well replicates representative of the data obtained through two similar experiments. (B) NP-normalized NA mRNA abundance as determined by RT-qPCR on total cellular RNA in PR8 NP:F346S-infected MDCK-SIAT1 cells relative to PR8 NP:WT at the indicated hours post-infection. Infections were initiated at a MOI of 5 TCID₅₀/cell in the presence of 100 $\mu\text{g}/\text{mL}$ of cycloheximide. Data points indicate individual cell culture well replicates pooled from two independent experiments. (C) NP-normalized abundances of newly synthesized NA vRNA in PR8 NP:F346S- and PR8 NP:WT-infected cells, as measured by 4-thiouridine (4SU) pulse-labeling. MDCK cells were infected with PR8-NP:WT or PR8-NP:F346S at an MOI of 5 TCID₅₀/cell for 7 h, followed by a 1 h pulse with 500 μM 4SU. Cellular RNA was then harvested and the abundances of 4SU-labeled viral RNAs were determined by RT-qPCR. Data points represent individual cell culture well replicates pooled from two independent experiments.

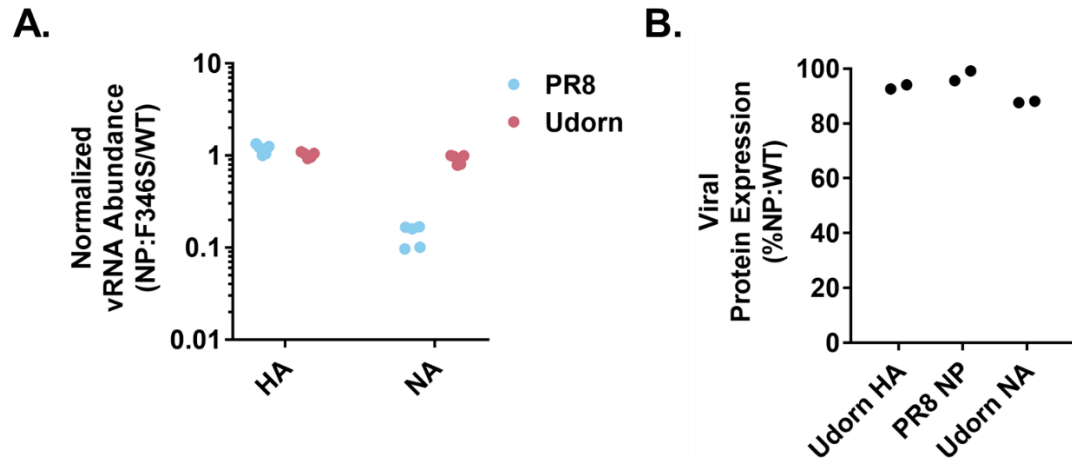
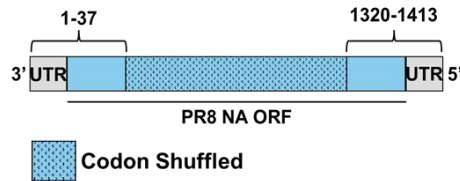
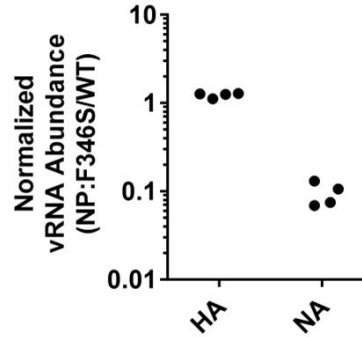


Fig. 2.2. Susceptibility to the effects of NP:F346S is NA segment genotype specific. (A) NP-normalized vRNA abundances as determined by RT-qPCR on total cellular RNA in PR8 NP:F346S- or Udorn NP:F346S-infected MDCK cells (MOI = 0.1 NPEU/cell, 8 hpi) expressed as fraction of PR8 NP:WT or Udorn NP:WT, respectively. Secondary infection was blocked via the addition of 20 mM NH_4Cl 3 hpi (Sun et al., 2020; Sun and Brooke, 2018). The data shown are for individual cell culture well replicates pooled from two independent experiments. (B) Viral protein expression levels as determined by geometric mean fluorescence intensity (GMFI) measured via flow cytometry in rPR8 Udorn HA/NA NP:F346S-infected MDCK cells (MOI = 0.03 $\text{TCID}_{50}/\text{cell}$; 16 hpi) expressed as a percentage of rPR8 Udorn HA/NA NP:WT. Secondary infection was blocked via the addition of 20 mM NH_4Cl 3 hpi (Sun et al., 2020; Sun and Brooke, 2018). The data shown are for individual cell culture well replicates representative of the data obtained through two similar experiments.

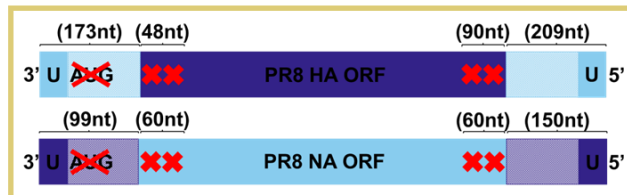
A. PR8 NA Codon Shuffle



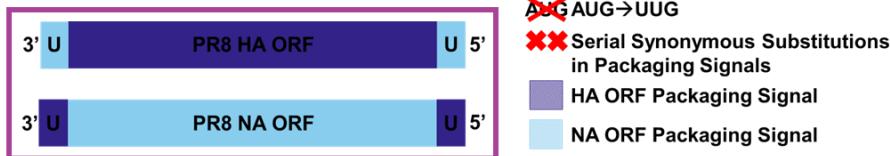
B.



C. PR8 HA/NA UTR+Pack Swap



PR8 HA/NA UTR Swap



D.

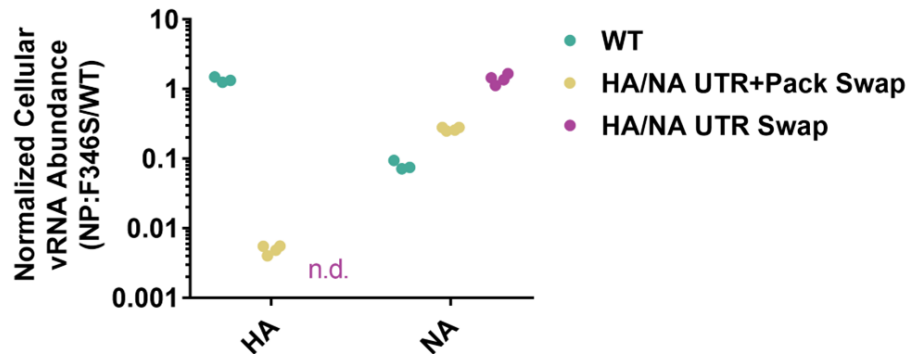


Fig. 2.3. Susceptibility to NP-dependent regulation maps to the UTRs of the NA segment. (A) Schematic depiction of the codon shuffled PR8 NA construct. (B) NP-normalized abundances of HA and NA vRNA following infection of MDCK cells with the PR8 NA Codon Shuffle NP:F346S virus (MOI = 0.1 NPEU/cell; 8 hpi) as determined by RT-qPCR on total cellular RNA relative to PR8 NA Codon Shuffle NP:WT. (C) Schematic depictions of the PR8 HA/NA UTR+Pack swap and PR8 HA/NA UTR swap gene segments. In the segments in which the PR8 HA and NA UTRs and packaging signals were swapped, the start codon of the newly appended packaging signal for each segment was mutated to prevent the

Fig. 2.3. (continued)

expression of any protein encoded by the packaging signal sequence, and the packaging signals within the native ORFs were disrupted via the addition of silent substitutions to prevent duplication of the packaging signals. (D) NP-normalized abundances of the HA ORF-containing or NA ORF-containing segments from the PR8 NP:F346S, PR8 HA/NA UTR+Pack swap NP:F346S, and PR8 HA/NA UTR swap NP:F346S viruses in infected MDCK cells (MOI = 0.1 NPEU/cell; 8 hpi) as determined by RT-qPCR on total cellular RNA, expressed as a fraction of PR8 NP:WT, PR8 HA/NA UTR+Pack swap NP:WT, and PR8 HA/NA UTR swap NP:WT, respectively. “n.d.” indicates that the segment was below the limit of detection for the assay. In panels B and D, each data point represents a cell culture well replicate pooled from two independent experiments.

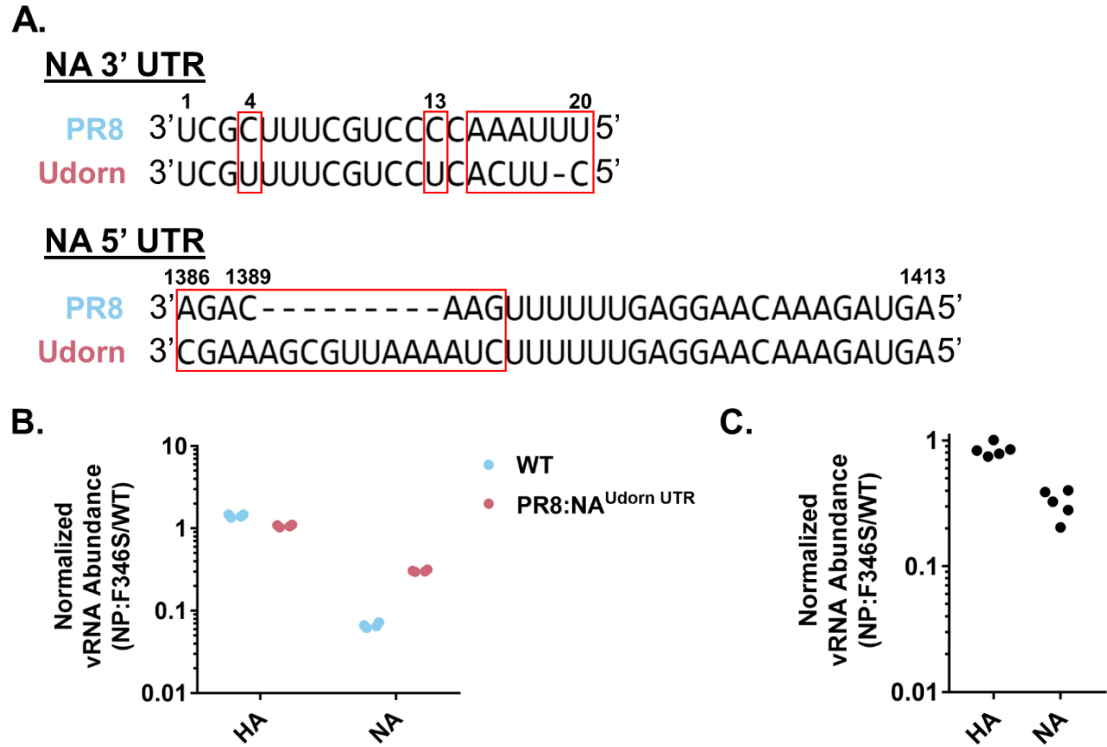


Fig. 2.4. The UTRs of the Udorn NA segment confer resistance to regulation by NP:F346S. (A) Alignment of the PR8 NA and Udorn NA 3' and 5' UTRs using the M-Coffee alignment algorithm on the T-Coffee web server. Regions of interest are boxed in red. PR8 NA nucleotide numbering is shown. (B) NP-normalized abundances of the HA and NA segments as determined by RT-qPCR on total cellular RNA in MDCK cells infected with the PR8 NP:F346S and PR8:NA^{Udorn UTR} NP:F346S (MOI = 0.1 NPEU/cell; 8 hpi) viruses as determined by RT-qPCR relative to PR8 NP:WT and PR8:NA^{Udorn UTR} NP:WT, respectively. (C) NP-normalized abundances of the HA and NA segments as determined by RT-qPCR on total cellular RNA in MDCK cells infected with PR8:Udorn HA,NA^{PR8 UTR} NP:F346S (MOI = 0.03 NPEU/cell, 8 hpi) virus relative to PR8:Udorn HA,NA^{PR8 UTR} NP:WT. In panels B and C, each data point represents a cell culture well replicate pooled from two independent experiments.

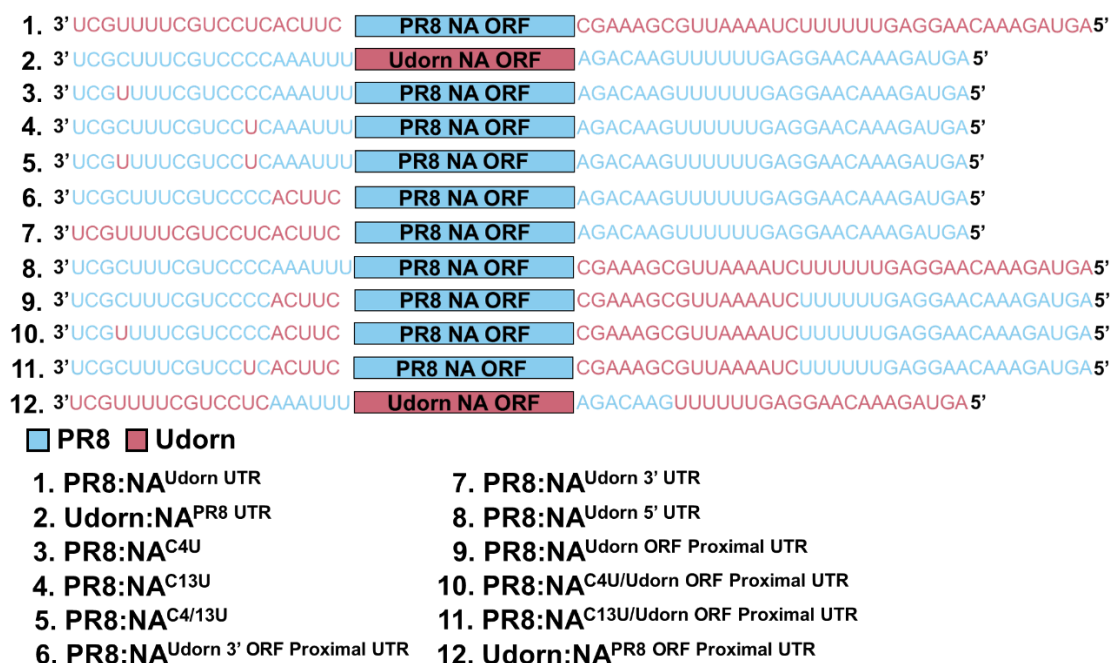


Fig. 2.5. UTR sequences of the PR8-Udorn NA UTR chimeric constructs. The sequences derived from PR8 and Udorn NA are colored blue and pink, respectively. The sequences are shown in negative sense, 3'→5'.

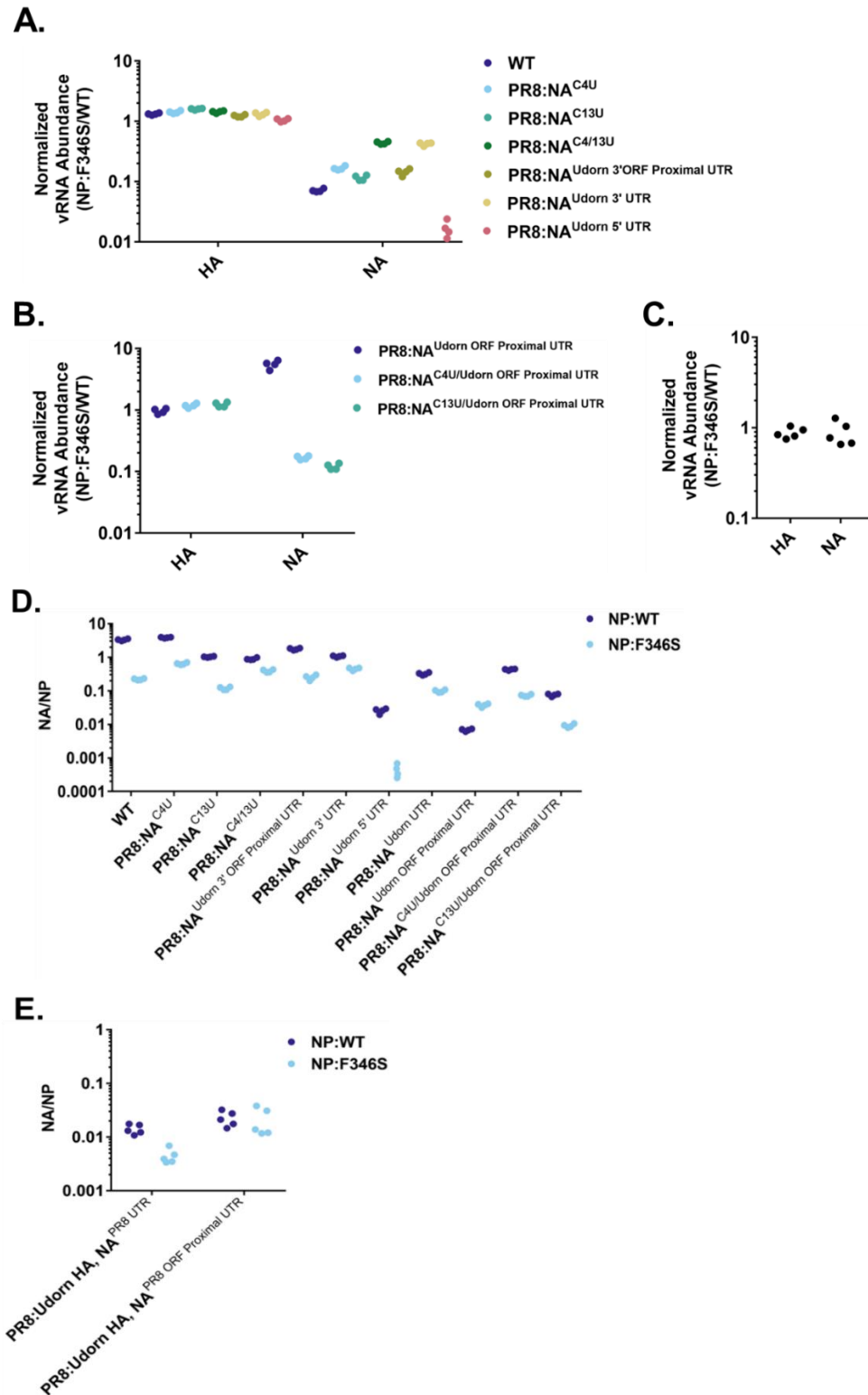


Fig. 2.6. The effect of mutations in the PR8 NA and Udorn NA UTRs on baseline expression levels and sensitivity to NP:F346S. (A) NP-normalized abundances of the HA and NA segments as determined by RT-qPCR on total cellular RNA in

Fig. 2.6. (continued)

MDCK cells infected with the indicated viruses encoding NP:F346S at a MOI of 0.1 NPEU/cell at 8 hpi, relative to the NP:WT-encoding versions of the same viruses. (B) NP-normalized abundances of the HA and NA segments as determined by RT-qPCR on total cellular RNA in MDCK cells infected with the indicated viruses encoding NP:F346S (MOI = 0.1 NPEU/cell; 8 hpi), relative to the NP:WT-encoding versions of the same viruses. (C) NP-normalized abundances of the HA and NA segments as determined by RT-qPCR on total cellular RNA in MDCK cells infected with the PR8:Udorn HA,NA^{PR8 ORF Proximal UTR} NP:F346S virus (MOI = 0.03 NPEU/cell; 8 hpi) relative to PR8:Udorn HA,NA^{PR8 ORF Proximal UTR} NP:WT. (D and E) Data from experiments shown in Fig. 4B and panels A and B, and Fig. 4C and panel C, respectively, showing the intracellular abundances of the indicated chimeric NA segment vRNAs normalized to NP vRNA levels (in the context of NP:WT or NP:F346S) in infected MDCK cells (MOI = 0.1 [D] or 0.03 [E] NPEU/cell 8 hpi) as determined by qRT-PCR on total cellular RNA. For all panels, the data represent individual cell culture well replicates pooled from two independent experiments.

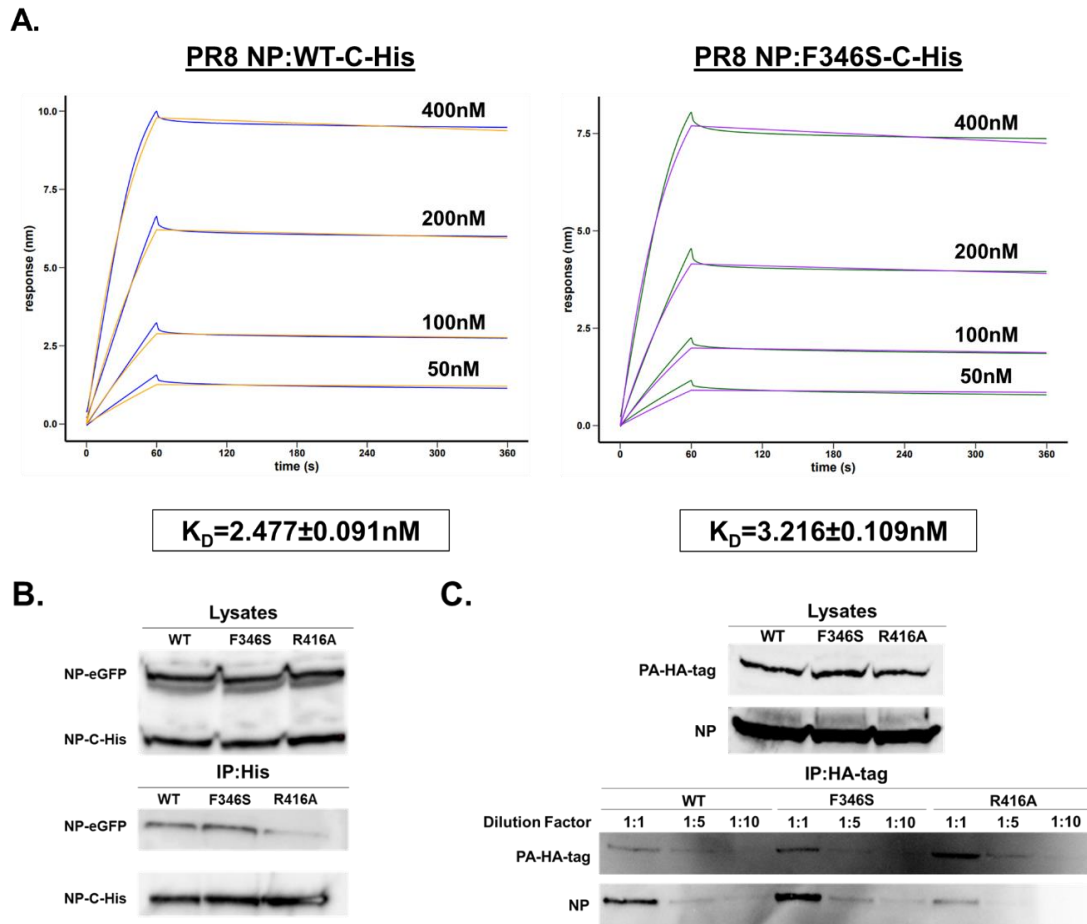


Fig. 2.7. NP:F346S does not affect NP RNA binding or oligomerization. (A) RNA binding kinetics of the PR8 NP:WT-C-His and PR8 NP:F346S-C-His proteins as determined by BLI. The raw data are colored blue/green and the fitted data are colored orange/purple for the NP:WT/F346S-C-His proteins. The data are representative of those from a single experiment. The dissociation constants (K_D) of the PR8 NP:WT-C-His and PR8 NP:F346S-C-His proteins are shown with the standard errors of the means (SEM). (B) Western blots for NP or eGFP/His-tagged NP performed on either the total cell lysate or His-tagged NP (WT/F346S/R416A) coimmunoprecipitated protein lysate fraction, respectively, from 293T cells transfected with expression vectors encoding the eGFP- and His-tagged versions of either the WT, F346S, or R416A NP 24 h post-transfection. (C) Western blots for HA-tagged PA or NP performed on either the whole-cell lysate or HA-tagged PA coimmunoprecipitated protein lysate fraction (1:1/1:5/1:10 diluted) from 293T cells transfected with expression vectors encoding the vRNP complex (PB2, PB1, PA-HA tag, and NP [WT, F346S, or R416A]) and a vRNA template (NA vRNA) 24 h post-transfection.

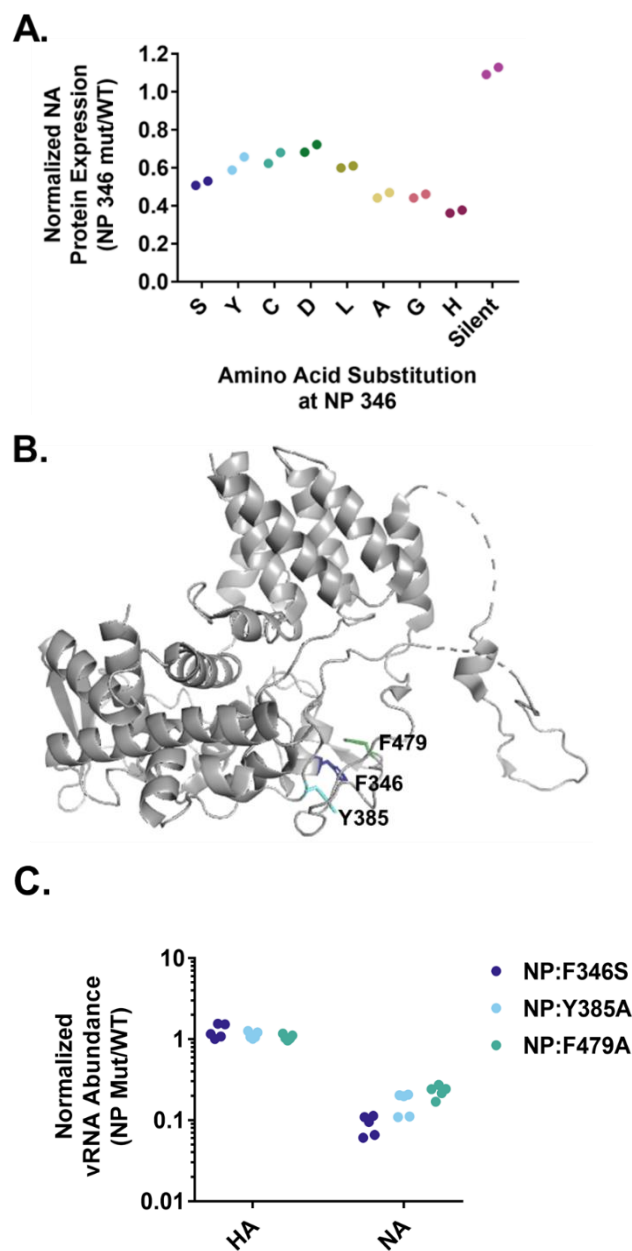
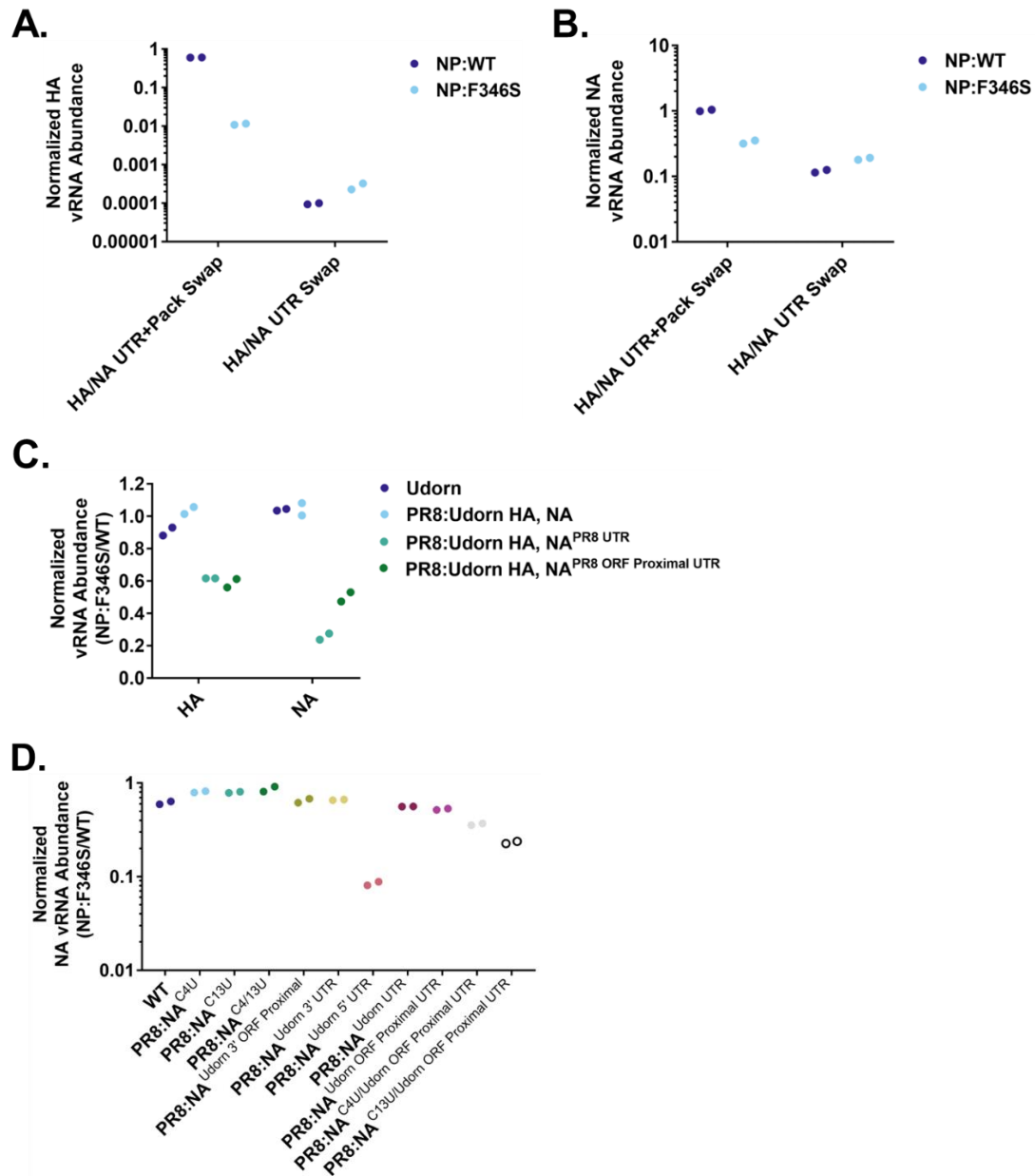


Fig. 2.8. A cluster of aromatic residues is involved in the regulation of NA gene segment expression. (A) NP-normalized NA protein expression levels as determined by geometric mean fluorescent intensity (GMFI) measured via flow cytometry in cells infected with the indicated PR8 NP 346 mutant viruses (MOI = 0.03 TCID₅₀/cell; 16 hpi) expressed as a fraction of PR8 NP:WT. Secondary infection was blocked via the addition of 20 mM NH₄Cl 3 hpi (Sun et al., 2020; Sun and Brooke, 2018). The data shown are for individual cell culture well replicates representative of the data obtained through two similar experiments. (B) Location of F346, Y385, and F479 in the NP protein visualized using the PyMol software (PDB code 2IQH). (C) NP-normalized viral RNA abundance as determined by RT-qPCR on total cellular RNA in PR8 NP:F346S-,

Fig. 2.8. (continued)

PR8 NP:Y385A-, and PR8 NP:F479A-infected MDCK cells (MOI = 0.1 TCID₅₀/cell; 8 hpi) relative to PR8 NP:WT. The data shown are for individual cell culture well replicates representative of the data from two independent experiments.



Supplemental Fig. 2.1. Quantification of gene segment ratios in viral RNA stocks. (A/B) Normalized HA ORF (A) or NA ORF (B) containing viral RNA abundance in the viral stocks of the PR8 HA/NA UTR+Pack Swap NP:WT/F346S and PR8 HA/NA UTR Swap NP:WT/F346S viruses. (C) Normalized viral RNA abundance in the viral stocks of the Udorn NP:F346S, PR8:Udorn HA, NA NP:F346S, PR8:Udorn HA, NA^{PR8 UTR} NP:F346S, and PR8:Udorn HA, NA^{PR8 ORF Proximal UTR} NP:F346S viruses relative to Udorn NP:WT, PR8:Udorn HA, NA NP:WT, PR8:Udorn HA, NA^{PR8 UTR} NP:WT, and PR8:Udorn HA, NA^{PR8 ORF Proximal UTR} NP:WT viruses respectively. D.) Normalized viral RNA abundance in the viral stocks of the indicated viruses with NP:F346S as determined by RT-qPCR normalized to the NP:WT versions of each virus. Each data point represents a qPCR technical replicate.

<u>Segment</u>	<u>Forward Primer (5'-->3')</u>	<u>Reverse Primer (5'-->3')</u>
PR8 HA	AAGGCAAACCTACTGG TCCTGTT	AATTGTTTCGCATGGTA GCCTATAC
PR8 NP	AGGCACCAAACGGTCT TACG	TTCCGACGGATGCTCT GATT
PR8 NA/PR8 NA ORF Containing UTR Mutants/PR8 NA ORF HA UTR	AAATCAGAAAATAACAA CCATTGGA	ATTCCCTATTTGCAATA TTAGGCT
Udorn HA	GACTATCATTGCTTTGA GC	CACTAGTGTTCCGTTT GGC
Udorn NP	CGGTCTTATGAACAGAT GG	TCGTCCAATTCCATCA ATC
Udorn NA/Udorn NA ORF containing UTR mutants	AACAATTGGCTCTGTCT CTC	GTCGCACTCATATTGC TTG
PR8 NA Codon Shuffle	CAAATGGGACCGTCAA AGACCGC	GATGGGGCTTCACCGA CTGG
PR8 HA ORF NA UTR+Pack	AAATCAGAAAATAACAA CCATTGGA	CAACAATACCAACAGA TTAGC
PR8 NA ORF UTR+Pack	AAGGCAAACCTACTGG TCCTGTT	ATCAGCCCTACCACGA GGC
PR8 HA ORF NA UTR	CAGGAGTGCCAAATTG AGGATGG	CCGGCAATGGCTCCAA ATAGACC

Table 2.1. qPCR primer sets used for the indicated segments

CHAPTER 3: EVALUATING THE ROLE OF RNA STRUCTURES AND SEQUENCE MOTIFS IN IAV GENE SEGMENT EXPRESSION

3.1 Introduction

RNA structures within the ORFs of the eight IAV RNA species play important roles in IAV biology. RNA structures within the IAV gene segments are usually more favored to form within the positive sense RNA species (mRNA/cRNA) relative to the negative sense RNA species (vRNA) (Moss et al., 2011). As a result, multiple functions have been attributed to RNA structures within the positive sense RNAs such as regulating gene segment splicing and expression levels (T. Jiang et al., 2016; Simon et al., 2019; Soszynska-Jozwiak et al., 2015). The main function attributed to RNA structures within the negative sense vRNA is coordinating the intersegment interactions responsible for the selective packaging of a complete set of eight unique gene segments into virions (Bolte et al., 2019; Dadonaite et al., 2019; Fournier et al., 2012; Gavazzi et al., 2013; Miyamoto et al., 2022; Shafiuddin and Boon, 2019; Williams et al., 2018). However, outside of the role of regulating the packaging of the eight gene segments, not much is known about other potential roles of RNA structures within the vRNAs in IAV biology.

Techniques such as PAR-CLIP, which identifies the RNA sequences associated with NP in infected cells, have demonstrated that NP is distributed unevenly along the vRNA, allowing RNA structures to form within the segment (Williams et al., 2018). In addition, recently developed techniques such as SHAPE-MaP have shown that RNA structures can form in the vRNA within vRNP complexes- the

relevant context in which vRNAs are found in infected cells (Dadonaite et al., 2019). The role of RNA structures within the IAV segment can also be tested indirectly by utilizing programs, such as the CodonShuffle program (Jorge et al., 2015), that introduce silent substitutions within the vRNA to modify RNA features such as GC content, codon bias, and RNA structures without altering the protein coding sequence.

Utilizing the SHAPE-MaP technique, I generated a comprehensive RNA structural profile of the eight gene segments of the PR8 (NP:WT/F346S) strains and used the data to probe the roles of those structures in NP-dependent regulation of NA expression (discussed in Chapter 2) and gene segment expression levels. Additionally, using the CodonShuffle program (Jorge et al., 2015) to introduce synonymous mutations into the NA segment, I probed the role of RNA sequence structures and motifs in modulating baseline expression levels of the NA segment. Together, these results demonstrate a novel role for RNA structures and motifs within the vRNAs in regulating gene segment expression.

3.2 Results

SHAPE-MaP-predicted structures in the PR8 NA ORF play no role in NP-dependent regulation

To probe the role of RNA structures within the non-packaging signal region of the PR8 NA ORF, I used the CodonShuffle program to introduce synonymous mutations into the NA ORF and demonstrated that these mutations had no effect on the susceptibility of the NA segment to NP-dependent regulation (see chapter 2.2, Fig. 2.3B). Even though saturating the NA ORF sequence with silent

substitutions would presumably change the RNA structural profile of that region, to more definitively rule out the role of RNA structures in the non-packaging signal region of the NA ORF, I used SHAPE-MaP to elucidate the RNA structural profiles of the eight segments in the context of vRNPs for the PR8 and PR8 NP:F346S viruses. During the SHAPE-MaP protocol, purified virions are incubated with a reagent called 1-Methyl-7-nitroisatoic anhydride (1M7) which selectively reacts with unpaired ribonucleotides (Dadonaite et al., 2019; Smola et al., 2015). During the RT-PCR step, mutations are introduced at 1M7 labeled sites which allows the basepairing probabilities at each nucleotide to be inferred based on their mutational frequencies (Dadonaite et al., 2019; Smola et al., 2015). This information can be utilized to build a comprehensive RNA structural profile of each segment at single nucleotide resolution (Dadonaite et al., 2019; Smola et al., 2015).

One potential hypothesis that we had for the mechanism of NP-dependent regulation of NA expression (see chapter 2) was that if the NP:F346S substitution altered the ability of NP:F346S to bind to and oligomerize along the viral RNA, then this could alter the likelihood of formation of RNA secondary structures potentially important for regulation of NA gene segment expression along the NA segment. However, we previously found that NP:F346S does affect NP RNA binding or oligomerization (Fig. 2.7), and our comparison of the NA RNA structural profile in the presence of NP:WT and NP:F346S revealed that the NA RNA structural profiles were highly similar (Fig. 3.1A), suggesting that NP:F346S does not significantly alter the NA RNA structural profile. Additionally, we found

that mutating the RNA structures within the NA vRNA that were conserved in the presence of both NP:WT and NP:F346S (NA SHAPE-MaP Mut) did not affect the susceptibility of the NA segment to NP-dependent regulation (Fig. 3.1B).

Together, these observations demonstrate that the SHAPE-MaP predicted RNA structures within the NA ORF do not play a role in modulating susceptibility of the segment to NP-dependent regulation.

Evaluation of the role of intrasegment interactions within the NA segment in NP-dependent regulation

In addition to the role of local RNA structures within the NA gene segment in NP-dependent regulation, I was also interested in exploring whether long-range intrasegment interactions within the NA gene segment played a role in NP-dependent regulation. Long-range intrasegment interactions as well as intersegment interactions can be identified using a technique called sequencing of psoralen-cross-linked, ligated and selected hybrids (SPLASH) which involves sequencing interacting sequences of RNA in crosslinked vRNP complexes (Dadonaite et al., 2019). Using a published SPLASH dataset for the PR8 strain (Dadonaite et al., 2019), I identified a long-range intrasegment interaction within the PR8 NA vRNA between nucleotides 365-379 and 1229-1244 (Fig. 3.2A).

To probe the role of this intrasegment interaction within NA in NP-dependent regulation, I generated viruses containing NA segments with silent substitutions that disrupted this putative interaction via two different methods. The first involved introducing silent substitutions to disrupt the predicted NA RNA secondary structure present within nucleotides 1229-1244 of the NA vRNA (NA

Stem Mut) (Fig. 3.2B). The second involved introducing silent substitutions to directly disrupt the interaction between nucleotides 1229-1244 and 365-379 (NA Loop mut) (Fig. 3.2B). Disrupting this intrasegment interaction either by mutating the RNA secondary structure in one portion of the interacting region or directly disrupting the basepairing interactions between the two regions had minimal effect on the susceptibility of the NA segment to NP-dependent regulation relative to the WT NA segment (Fig. 3.2C). However, the expression level of the NP gene segment in the NA loop mutant virus in the presence of NP:F346S was about 40% of the level in the presence of NP:WT (Fig. 3.2C). Although there was a decrease in the abundance of NP vRNA in the presence of NP:F346S and the NA Loop Mut segment, there was no corresponding decrease in NP protein expression (Fig. 3.2D). A possible explanation is that the amount of NP vRNA expressed was sufficient to sustain maximum NP protein production levels, as it has previously been observed that in NP:F346S infected cells, the decrease in NA protein expression levels is several fold less than the decrease in NA vRNA levels (3-fold and 10-fold respectively (Brooke et al., 2014)). Potentially, the decrease in NP vRNA abundance observed in infected cells is the result of a packaging defect which would decrease effective gene dosage of NP and, therefore, its vRNA abundance in infected cells.

Evaluating the role of RNA structures and motifs within the NA gene segment in regulating its baseline expression level

Although none of the RNA structures and intrasegment interactions that I identified in the NA segment modulated susceptibility of the segment to NP-

dependent regulation, another possible, and previously underexplored role, for these RNA features could be regulating the baseline expression level of the NA segment. To explore this possibility, I examined the relative abundance of several of the mutant NA segments that I designed to explore the role of RNA features in NP-dependent regulation (ex.) Codon Shuffled NA (chapter 2), NA SHAPE-MaP Mut, NA Stem Mut, & NA Loop Mut) relative to the WT PR8 NA segment in the presence of NP:WT. The mutations designed to mutate RNA structures in the NA vRNA identified via SHAPE-MaP (NA SHAPE-MaP mut) or the intrasegment interaction within the NA vRNA identified via SPLASH (NA Stem Mut & NA Loop Mut) decreased the baseline expression level to around half of the level observed for the WT NA segment (Fig. 3.3A). Meanwhile, the Codon Shuffled NA that contained significantly more silent mutations in the NA ORF relative to the NA SHAPE-MaP, NA Stem Mut, and NA Loop Mut segments (227 mutations versus 93, 10, and 9 mutations respectively) exhibited an approximately 20-fold decrease in expression relative to the WT NA segment (Fig. 3.3B). As the silent mutations in the Codon Shuffled NA segment extend into the traditional packaging signals for the N1 NA segment (Fujii et al., 2003), part of this decreased expression level could be explained by decreased packaging of the Codon Shuffled NA segment. A Codon Shuffled NA segment that had the silent substitutions removed from the regions of the NA ORF containing the traditional packaging signals (Codon Shuffled NA [183,157]), exhibited a smaller (~4-fold) decrease in its baseline expression level relative to the WT NA (Fig. 3.3B). Overall, these results indicate that even small numbers of

silent substitutions relative to the full length of the PR8 NA gene segment (ex.) 14% of the segment [Codon Shuffled NA] and lower) can play a large role in regulating its baseline expression level.

3.3 Discussion

Overall, I was unable to identify any role for the RNA structures within the NA segment ORF in NP-dependent regulation. However, my observation that mutating some of the predicted RNA structures in the NA segment or introducing large stretches of silent substitutions to disrupt other RNA sequence features or motifs reduced its expression level relative to the WT NA segment suggested that there may be a role for RNA structures in regulating the expression of both NA and the remaining segments. Therefore, by utilizing the SHAPE-MaP data for PR8, which revealed the RNA structural profile of each gene segment, and the CodonShuffle program to account for changes in RNA features such as GC content and codon bias as a result of silent substitutions, I can design gene segments with specific RNA structures eliminated via silent substitutions and probe their role in regulating gene segment expression.

Moving beyond evaluating the roles of these RNA structures in regulating the expression level of each segment, I was also interested in probing their roles in other viral replication processes such as gene segment packaging and viral RNA synthesis. My data demonstrating that disrupting an intrasegment interaction within the NA segment decreased the expression of another segment, NP, in the presence of NP:F346S (Fig. 3.2C), in addition to other studies which previously

demonstrated the roles of RNA structures in regulating gene segment packaging (Bolte et al., 2019; Dadonaite et al., 2019; Fournier et al., 2012; Gavazzi et al., 2013; Miyamoto et al., 2022; Shafiuddin and Boon, 2019; Williams et al., 2018), suggest that it would be of interest to probe the roles of the SHAPE-MaP predicated RNA structures in gene segment packaging. Additionally, it has been suggested that RNA structures or other sequence elements may influence the synthesis of defective viral RNA products such as defective interfering RNAs (DI RNAs) and mini viral RNAs (mvRNAs) which exhibit large internal deletions formed as a result of the dissociation from and then reassociation of the viral RNA-dependent RNA polymerase with the RNA template (Alnaji et al., 2021; Alnaji and Brooke, 2020; Davis et al., 1980; te Velthuis et al., 2018). Therefore, using the RNA structural profiles for each segment identified via SHAPE-MaP and the CodonShuffle program as well as our dataset detailing the location of the major deletion junctions in the PR8 strain (Alnaji et al., 2019), I can generate mutant segments with silent mutations introduced into these junction locations to disrupt predicted RNA structures or different types of RNA sequence features (GC content, dinucleotide repeats, codon bias etc...) to see if they play any role in DI RNA or mvRNA formation.

3.4 Methods

Plasmids

The A/Puerto Rico/8/34 reverse genetics plasmids were gifts from Adolfo Garcia-Sastre (Icahn School of Medicine at Mt. Sinai). The PR8 NA SHAPE-MaP Mut, PR8 NA Stem Mut, PR8 NA Loop Mut, Codon Shuffled NA, and Codon Shuffled

NA (183,157) segments were synthesized by Bio Basic Inc. *Lgul* restriction sites were added to both ends of the segments for cloning into the pDZ vector. The segments were then digested with the *Lgul* restriction enzyme (Thermo Fisher; FD1934), ligated into the pDZ vector (New England BioLabs [NEB]; T4 DNA ligase; M0202S), and transformed into chemically competent *Escherichia coli* cells (NEB; Turbo C2984H) via the heat shock method. Insert sequences were confirmed via Sanger sequencing.

Cells

Madin-Darby canine kidney (MDCK), MDCK-SIAT1 cells, and 293T cells were obtained from Jonathan Yewdell (NIH), Jesse Bloom (Fred Hutchinson Cancer Research Center), and Joanna Shisler (NSF), respectively, and were maintained in Gibco's minimal essential medium (MEM) with GlutaMax (Life Technologies) supplemented with 8.3% fetal bovine serum (Seradigm) (MEM+FBS) and incubated at 37°C with 5% CO₂. MDCK and MDCK-SIAT1 cells were used for virus propagation and infection experiments due to their ability to support robust replication of IAV and their tolerance to the trypsin necessary to convert the HA protein into its fusion-competent state. 293T cells were used for transfections to generate viruses via the reverse genetics system or to express viral proteins due to their high transfection efficiency.

Viruses

Recombinant A/Puerto Rico/8/1934 (H1N1) (rPR8) viruses were generated using an 8-plasmid reverse genetics system. The rPR8 clones differ from the published

sequence (GenBank accession no. AF389115 to AF389122) at two positions: PB1 A549C (K175N) and HA A651C (I207L) (numbering from initiating Met). Viruses containing the mutated NA segments were generated by rescuing the viruses using plasmids containing those specific segments (generated as described above).

Viruses were rescued by transfecting 293T cells with 500 ng each of the relevant reverse genetics plasmids using JetPrime (Polyplus) according to the manufacturer's instructions. At 18 to 24 h post-transfection, the medium was replaced with viral growth medium (MEM, 1 mM HEPES, 1 µg/mL of tosylsulfonyl phenylalanyl chloromethyl ketone [TPCK] trypsin [Worthington Biochemical Corporation, Lakewood, NJ], 50 µg/mL of gentamicin) containing 2×10^5 MDCK cells. Transfection supernatants were collected 24 h after medium change.

To generate the seed stocks of the viruses, transfection supernatants were plaque, and a single plaque was used to infect a single well of MDCK-SIAT1 cells in a 6-well plate. Viral growth was performed in viral growth medium (MEM, 1 mM HEPES, 1 µg/mL of TPCK trypsin, 50 µg/mL of gentamicin). Seed stocks were harvested and clarified (14,000 rpm, 15 min, 4°C) between 24 and 72 h post-infection. Virus titers were determined by TCID₅₀ assay on MDCK-SIAT1 cells.

SHAPE-MaP

To grow up enough of the PR8 and PR8 NP:F346S viruses for virion purification, sixteen T175 flasks of MDCK cells per virus were infected with the PR8 or PR8

NP:F346S viruses at an MOI of 0.1 in viral growth medium (MEM, 1 mM HEPES, 1 µg/mL of TPCK trypsin, 50 µg/mL of gentamicin). The viral supernatant was harvested 48hpi and clarified (3500 rpm, 15 min, 4°C). Virions were purified from the viral supernatants using a three-step ultracentrifugation procedure: Per ~38 mL (25 x 89 mm) Beckman Coulter Polypropylene Centrifuge Tubes (326823), 10 mL of 20% sucrose solution was added to the bottom of the tube and then overlaid with 28mL of viral supernatant. The tubes were added to buckets in a SW28 rotor and samples were ultracentrifuged at 27,000 rpm for 2 h at 4°C. The viral supernatant and then 20% sucrose were aspirated, and then 300 µL of (+/+) PBS was added to the viral pellets and incubated at 4°C overnight. The viral pellets were resuspended in the (+/+) PBS and combined for each virus. Per ~12 mL (14 x 89 mm) Beckman Coulter Polypropylene Centrifuge Tubes (331372), 5 mL of 60% sucrose was added to the bottom of the tube and then gently overlaid with 5 mL of 15% sucrose followed by 1.2 mL of the viral solution from the first ultracentrifugation step. The tubes were added to buckets in a SW41 rotor and samples were then ultracentrifuged at 34,200 rpm for 2 h at 4°C. The 15% sucrose was aspirated and the virion band at the interface between the 15% and 60% sucrose solutions was collected. The virions preps were then washed as follows: The virion preps for each prep were combined together into a single ~38 mL (25 x 89 mm) Beckman Coulter Polypropylene Centrifuge Tube (326823) and the virions were resuspended in (+/+) PBS up to a volume of 38mL. The tubes were added to buckets in a SW28 rotor and ultracentrifuged at 27,000 rpm for 2 h at 4°C. The (+/+) PBS was aspirated and then 300 µL of (+/+) PBS was

added to each virion pellet and incubated at 4°C overnight. The virion pellet was resuspended the next day and immediately used for the subsequent 1M7 labeling step.

Virions were pre-incubated at 37°C for 5 min, 10 mM of 1-methyl-7-nitroisatoic anhydride (1M7) reagent (Tocris 6602, CAS no. 73043-80-8) (two replicates per virus) or anhydrous dimethylsulfoxide (DMSO) (Invitrogen D12345) (negative control; two replicates per virus) was added to each sample, and then the reaction was allowed to proceed at 37°C for 75 s. Samples were treated with a Proteinase K reaction mixture (0.25mg/mL Proteinase K (NEB P81075), 0.5% SDS, 1x Proteinase K Buffer (100mM Tris pH=7.5, 50mM NaCl, 10mM EDTA) Final Concentration) for 40 min at 37°C with shaking (300 rpm) to digest virion protein.

Viral RNA was extracted from 1M7/DMSO and Proteinase K-treated virions as follows: Nuclease-free water (Ambion) was added to each sample to obtain a total volume of 250 µL, then 750 µL of TRIzol LS reagent (Invitrogen) was added and then the mixture was incubated at room temperature for 5 min. The mixtures were transferred to 2-mL Quanta Bio 5PRIME Phase Lock Heavy Tubes, an additional 200 µL of chloroform was added to each tube, and the contents in the tube were mixed by vigorous shaking for 15 s. The phases were then separated by incubating the tubes at room temperature for 2-3 min and centrifuging at 12,000 x g for 15 min at 4°C. The aqueous phase (top layer) was transferred to a new 1.5 mL microcentrifuge tube, 0.5 mL isopropanol and 2 µL of 20 mg/mL

glycogen (Thermo Scientific R0561) were added to each sample, the sample was mixed and then incubated at room temperature for 10 min. The RNA was pelleted by centrifuging at 12,000 x g for 10 min at 4°C. The supernatants were then removed and the RNA pellets were washed with 1 mL of 75% ethanol (7,500 x g, 5 min, 4°C). The ethanol was removed and the RNA pellets were allowed to air dry for 5-10 min. The pellets were resuspended in 20 µL of Nuclease-free water (Ambion) and incubated at 55°C for 10 min to aid in dissolving the pellet. The RNA concentration was quantified using the Qubit RNA HS Kit.

Virion RNA was reverse transcribed into cDNA as follows: 150 ng of RNA and 1.35 µL of 50 µM Random Nonamer primers (IDT) were combined and resuspended in Nuclease-free water (Ambion) up to a total volume of 11 µL, and the reactions were incubated at 65°C for 5 min followed by 4°C for 1 min. RT Buffer was prepared by mixing 90 mM MnCl₂ with 3x MaP Pre-buffer (150 mM Tris-HCl pH=8, 225 mM KCl, 30 mM DTT, 1.5 mM dNTP Mix [NEB; N0447S]) at a 1:5 ratio. 8 µL of RT Buffer was added to each reaction and the reactions were incubated at 25°C for 2 min. 1 µL of Superscript II Reverse Transcriptase (ThermoFisher) was added to each reaction and the reactions were incubated at 25°C for 10 min, 42°C for 3 h, 70°C for 15 min, and then held at 4°C.

Performed second-strand synthesis from the cDNA as follows: Resuspended the cDNA up to a volume of 50 µL in Nuclease-free water (Ambion) and purified the cDNA from the reverse transcription reactions using G-25 Micro Spin Columns

(GE Healthcare, Cat. No. 27-5325-01) according to the manufacturer's instructions. Diluted the cDNA up to a total volume of 68 μ L in Nuclease-free water (Ambion), added 8 μ L of NEBNext Second Strand Synthesis Reaction Buffer (10X) (NEB 6113A) and 4 μ L of NEBNext Second Strand Synthesis Enzyme Mix (NEB 6112A), and incubated the reactions at 16°C for 2.5 h. The dsDNA was purified from the reaction using PureLink PCR Micro Spin Columns according to the manufacturer's instructions and eluted in 10 μ L of Nuclease-free water (Ambion). The sample concentrations were quantified using the Qubit dsDNA HS Kit. The samples were then subjected to next-generation sequencing using the Illumina Hi-seq platform.

Quantification of single replication cycle viral gene expression levels

MDCK-SIAT1 cells were infected with the PR8, PR8 NA SHAPE-MaP Mut, PR8 NA Stem Mut, PR8 NA Loop Mut, and PR8 NA (183,157) Codon Shuffled NP:WT/F346S viruses at an MOI of 0.1 TCID₅₀/cell in a 24-well plate for 1 h at 37°C. At 1 h post-infection, the virus supernatant was replaced with 0.5 mL of MEM+FBS. At 3 h post-infection, the MEM+FBS was replaced with 0.5 mL of NH₄Cl medium (MEM, 50 mM HEPES buffer, 20 mM NH₄Cl [pH 7.2]) to prevent viral spread (Sun et al., 2020; Sun and Brooke, 2018). The cell monolayers were harvested 8 hpi, and cellular RNA was extracted using the RNeasy kit or E.Z.N.A Total RNA Kit (Qiagen). Reverse transcription was performed using the Superscript III reverse transcriptase kit (Thermo Fisher). The reactions were set up as follows: 4 μ L of RNA, 0.5 μ L of 10 μ M PR8 RT_4A primer (5'-AGCAAAAGCAGG-3'), 1 μ L of 10 mM dNTP (NEB; N0447S), and 6.5 μ L of

Nuclease-free water (Ambion) were incubated at 65°C for 5 min and then 4°C for 1 min. Four microliters of 5× first-strand RNA buffer, 1 µL of 0.1 M DTT, 2 µL of SupraseIN RNase inhibitor (Thermo Fisher), and 1 µL of Superscript III reverse transcriptase (Thermo Fisher) were added to the previous reaction and incubated at 55°C for 60 min and 70°C for 15 min and held at 4°C. cDNA was stored at -20°C. Quantitative real-time PCR on cDNA was carried out using Power SYBR Green PCR master mix (MM) (Thermo Fisher) on a QuantStudio 3 thermal cycler (Thermo Fisher). qPCR primer sets used are listed in Table 3.1. Reactions were set up as follows: 2 µL of cDNA, 10 µL of 2× Power SYBR Green PCR Master Mix (Applied Biosystems 4367659), 1 µL of 10 µM forward primer, 1 µL of 10 µM reverse primer, and 6 µL of nuclease-free water (Ambion). The cycling conditions were as follows: 95°C for 10 min, and then 95°C for 15 s, followed by 60°C for 1 min repeated 40 times. For the PR8 NA Codon Shuffled NP:WT/F346S infection and RT-qPCR protocol see section 2.4.

Flow cytometry to detect viral protein expression in singly infected cells

MDCK cells were infected at an MOI of 0.03 TCID₅₀/cell in a 6-well plate. At 1 hpi, the virus was replaced with 3 mL of MEM+FBS. At 3 hpi, the MEM+FBS was replaced with 3 mL of NH₄Cl medium (MEM, 50 mM HEPES buffer, 20 mM NH₄Cl [pH 7.2]) to prevent secondary infection (Sun et al., 2020; Sun and Brooke, 2018). At 16 hpi, the cells were harvested, fixed, and permeabilized with foxP3 fix/perm buffer (eBioscience). Cells were subsequently stained with the following antibodies: PR8 NA: NA2-1C1 PacO, PR8 NP: HB65 AF647, and PR8 HA: H36-26 AF488. The cells were run on a BD LSR II flow cytometer and analyzed using

FlowJo version 10.1 (Tree Star, Inc.). Viral protein expression levels were determined from the geometric mean fluorescence intensity (GMFI) of the fluorophore associated with each protein.

3.5 Figures and Table

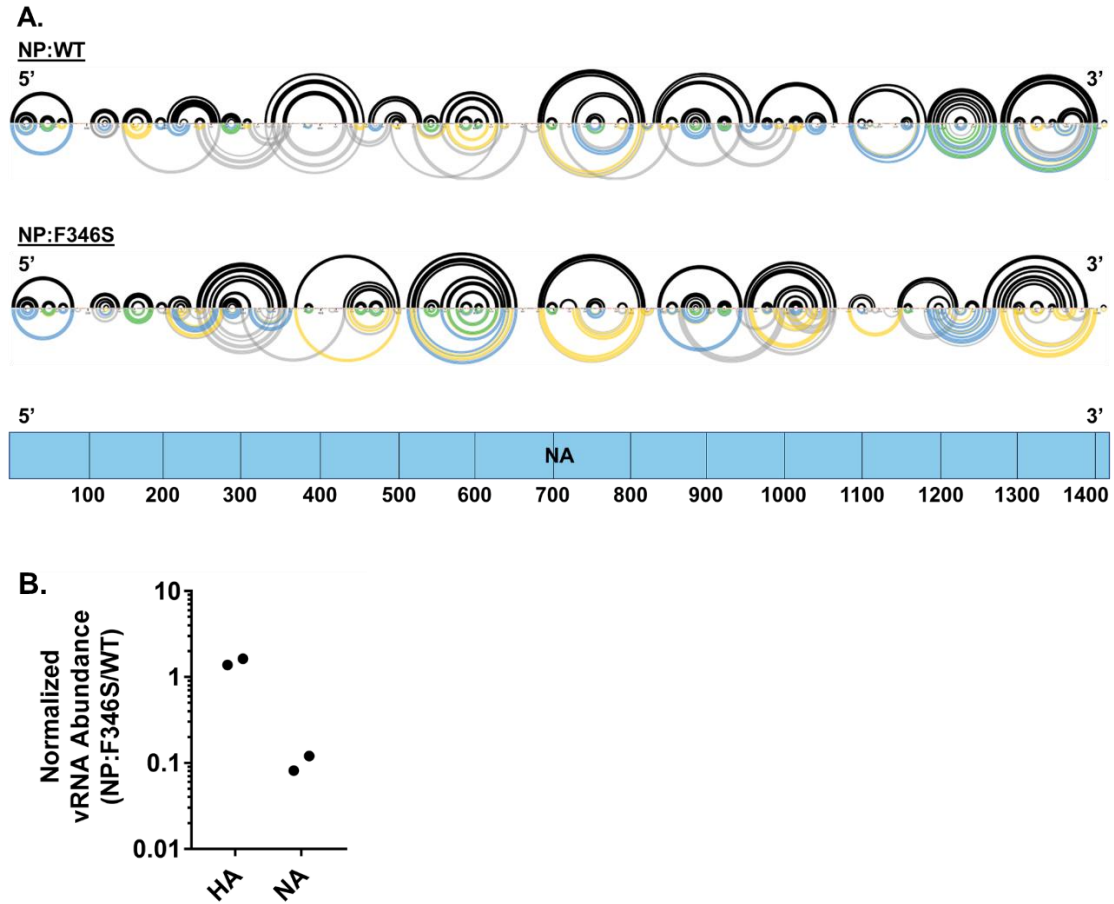


Fig. 3.1. SHAPE-MaP-predicted RNA structures within the NA ORF do not modulate susceptibility to regulation by NP:F346S. (A) Arc plots of the NA vRNA in the presence of NP:WT or NP:F346S. The black arcs above the axis represent the maximum expected accuracy RNA structures in which the ends of each arc represent the sequence origin of the two nucleotides participating in that basepairing interaction, and the green, blue, yellow, and gray arcs below the axis represent the probability of that basepairing interaction occurring (>80%, 30-80%, 10-30%, and <10% respectively). The numbering starts from the 5' end of the negative sense vRNA. (B) NP-normalized vRNA abundances in PR8 NA SHAPE-MaP Mut NP:F346S virus infected MDCK-SIAT1 cells (MOI=0.1 TCID₅₀/cell, 8hpi) relative to PR8 NA SHAPE-MaP Mut NP:WT as determined by RT-qPCR. Each data point represents a single cell culture well replicate from a single experiment.

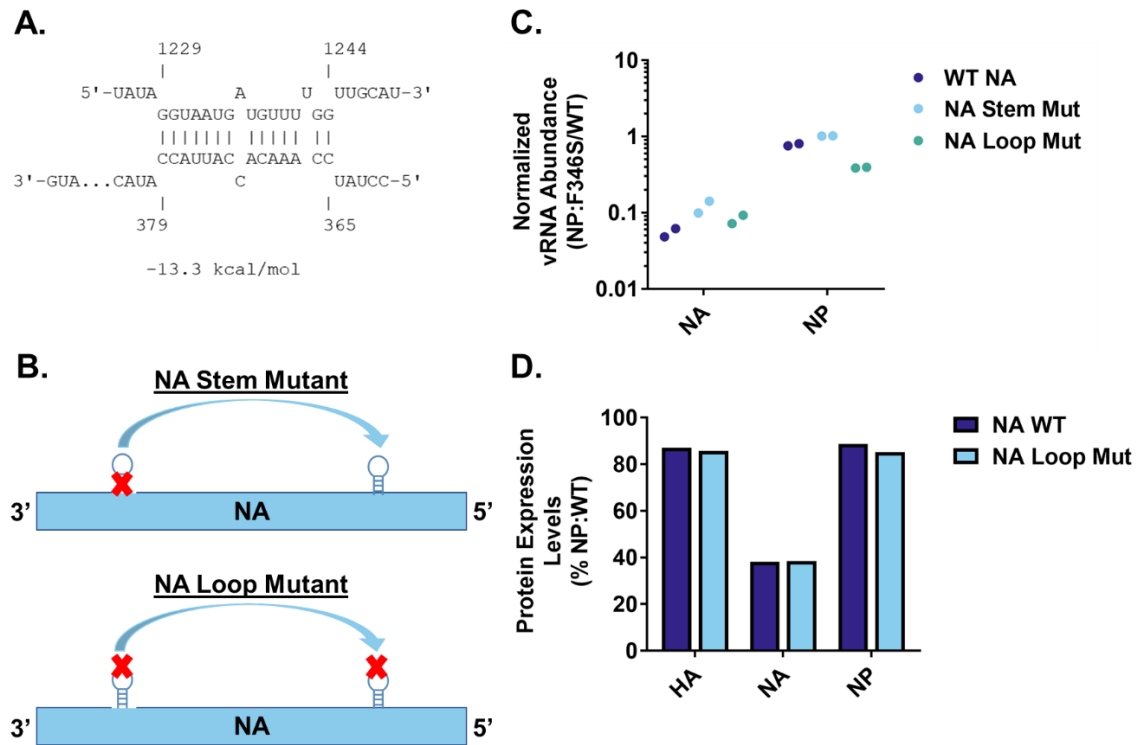


Fig. 3.2. Evaluating the role of an intrasegment RNA interaction within the NA vRNA in modulating susceptibility to NP:F346S. (A) Predicted basepairing interactions between nucleotides 365-379 and 1229-1244 of the NA vRNA. (B) Schematic depiction of the mutations introduced to disrupt either the stem loop structure presenting the basepairing nucleotides (NA Stem Mut) or to directly disrupt the basepairing interactions between the two regions of the NA vRNA (NA Loop Mut). (C) HA-normalized vRNA abundances in PR8 NP:F346S, PR8 NA Stem Mut NP:F346S, and PR8 NA Loop Mut NP:F346S infected MDCK-SIAT1 cells (MOI=0.1 TCID₅₀/cell, 8hpi) relative to the NP:WT-encoding versions of the same viruses as determined by RT-qPCR. (D) Viral protein expression levels as determined by flow cytometry (MOI=0.03 TCID₅₀/cell, 16hpi) in MDCK-SIAT1 cells. Viral spread was blocked 3hpi via the addition of 20mM NH₄Cl (Sun et al., 2020; Sun and Brooke, 2018). Each data point represents a single cell culture well replicate from a single experiment.

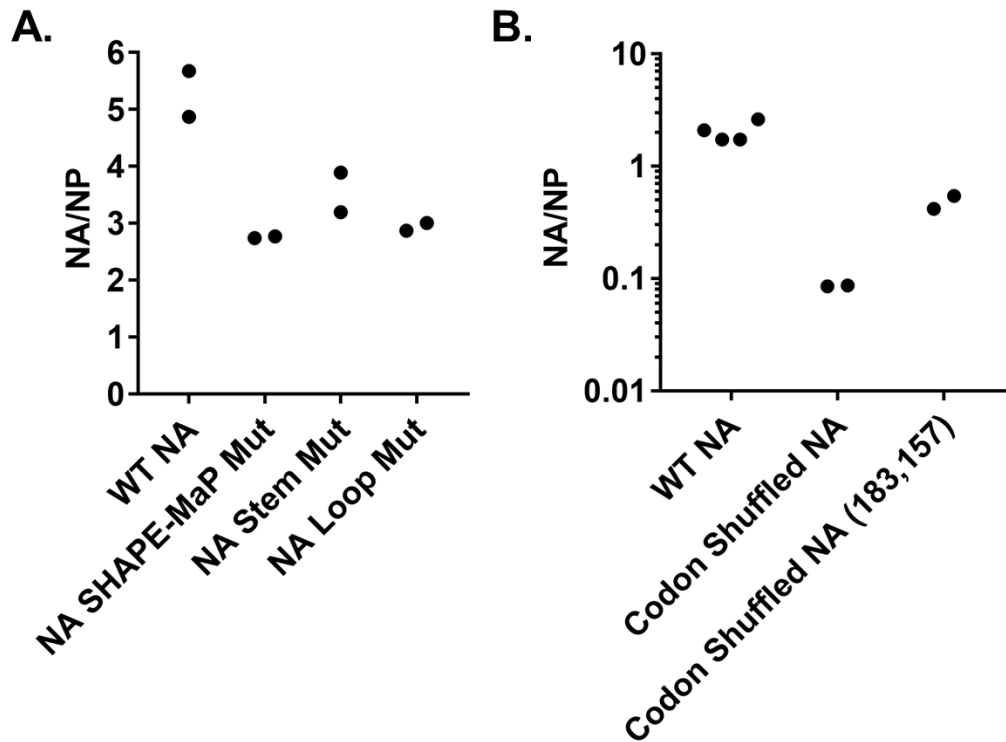


Fig. 3.3. Effect of silent substitutions within the NA ORF on baseline expression levels. (A) NP-normalized NA vRNA abundances as determined by RT-qPCR from cellular RNA extracted from MDCK-SIAT1 cells infected with the PR8, PR8 NA SHAPE-MaP Mut, PR8 NA Stem Mut, and PR8 NA Loop Mut NP:WT viruses (MOI=0.1 TCID₅₀/cell, 8hpi). (B) NP-normalized NA vRNA abundances as determined by RT-qPCR from cellular RNA extracted from MDCK/MDCK-SIAT1 cells infected with the PR8 NA Codon Shuffle and PR8 NA (183,157) Codon Shuffle NP:WT viruses respectively (MOI=0.1 TCID₅₀/cell, 8hpi). Each data point represents a single cell culture well replicate from a single experiment.

<u>Segment</u>	<u>Forward Primer (5'-->3')</u>	<u>Reverse Primer (5'-->3')</u>
PR8 HA	AAGGCAAACCTACTGGT CCTGTT OR AGGCAAACCTACTGGTC CTG	AATTGTTCGCATGGTAG CCTATAC OR ACTGTGTCAACAGTGTC GGT
PR8 NP	AGGCACCAAACGGTCTT ACG	TTCCGACGGATGCTCTG ATT
PR8 NA/ NA Stem Mut/NA Loop Mut	GGCTAGACTGTATGAGG CCG	CCGTCTGGCCAAGACC AATC
PR8 NA/ SHAPE- MaP NA Mut/ Codon Shuffled NA (183,157)	AAATCAGAAAATAACAAC CATTGGA	ATTCCCTATTTGCAATAT TAGGCT

Table 3.1. qPCR primer sets used for the indicated segments.

CHAPTER 4: IDENTIFICATION OF THE GENETIC DETERMINANTS OF SEMI-INFECTIOUS PARTICLE PRODUCTION IN INFLUENZA A VIRUS

4.1 Introduction

A majority of virions within IAV populations are classified as semi-infectious particles (SIPs) that lack expression of one or more gene segments (Brooke et al., 2013). SIPs are hypothesized to arise as a result of the failure of IAV to successfully complete one or more stages of the IAV replication cycle (see chapter 1 for review), however, the exact stages of the IAV replication cycle that contribute the greatest to SIP formation remain unknown (Diefenbacher et al., 2018). IAV strains are known to vary significantly in terms of SIP production, suggesting that SIP formation is governed by a genetic determinant (Brooke et al., 2013). To identify which gene segments contribute the most to SIP formation, I generated reassortants between a high and low SIP producing strain (A/Fort Monmouth/1/1947 and A/California/04/2009 respectively) and measured the ratio of total particles to fully-infectious particles for each reassortant. By determining which gene segments are responsible for the SIP production phenotype, specific stages of the IAV replication cycle can be identified and tested for their role in SIP formation.

4.2 Results

IAV strains vary substantially in their SIP production phenotype

To identify the genetic determinants behind SIP production in IAV, the approach was to identify high and low SIP producing IAV strains, make reassortants of the two strains, and then determine the SIP production phenotypes of the

reassortants. Taking advantage of the IAV strains available in our lab, I determined the particle to infectivity ratios of the IAV strains by dividing the total particle number (determined by calculating the endpoint dilution for the hemagglutination of red blood cells by viral particles) by the number of infectious particles (determined via plaque assay) (Fig. 4.1). A higher particle to infectivity ratio is indicative of a greater proportion of SIPs in the population. The SIP production phenotype varied over a log-fold between the different IAV strains with strains such as A/California/04/2009 (H1N1) (Cal04), A/California/07/2009 (H1N1) (Cal07), and A/Perth/16/2009 (H3N2) (Perth) producing low amounts of SIPs and having a particle to infectivity ratio around 1, while strains such as A/New Caledonia/20/1999 (H1N1) (NC99) and A/Fort Monmouth/1/1947 (H1N1) (FM) produced high amounts of SIPs and had particle to infectivity ratios greater than 15:1 (Fig. 4.1). The two viruses I selected to generate reassortants to probe the role of each of the individual viral gene segments in the SIP production phenotype were Cal04 and FM which had the lowest and highest level of SIP production respectively.

As the particle to infectivity ratio provides a general measure to assess the ratio of total to fully-infectious particles, it is only suggestive of a gene segment packaging or expression defect as it does not measure gene expression directly. Therefore, to substantiate that the FM strain produces a greater number of particles lacking one or more segments, and, as a result, generates more SIPs than the Cal04 strain, I infected cells with equal PFUs of the FM and Cal04 viruses and measured the percentage of cells expressing the NP gene segment.

As the cells are infected with equal infectious particles (PFUs) of each virus, if the virus has more SIPs, there should be a greater number of NP+ cells present because there are a higher proportion of non-infectious particles relative to infectious particles. In agreement with our expectations, there were more NP+ cells present after FM infection relative to Cal04 infection (Fig. 4.2A).

Additionally, there was an even greater disparity between FM and Cal04 when including the low expressing NP cells in the NP+ cell total instead of solely the high expressing NP+ cells (Fig. 4.2A). The low NP expressing cells are indicative of viral particles that lack one of the segments encoding the vRNP complex because in that situation, the remaining segments can only undergo primary transcription from their associated vRNP. Therefore, the observation that FM had a greater proportion of low NP expressing NP cells (i.e.) greater increase in total NP+ cells when combining both high and low NP expressing cells) than Cal04 further substantiates that FM produces more SIPs as the low NP expressing cells fail to express one or more of the vRNP components.

The SIP production phenotype maps to the vRNP complex genotype

After confirming the SIP production phenotypes of the FM and Cal04 strains, I next generated reassortants of the two strains to identify which gene segments contribute the greatest to the SIP production phenotype. The first set of reassortants that I generated swapped groups of segments that were functionally linked, such as the vRNP complex proteins and the glycoproteins HA and NA, between FM and Cal04 with the aim of increasing the likelihood that the viruses would be viable as well as minimizing protein incompatibilities that could interfere

with the SIP production phenotypic readout. Of the four possible reassortants, only two were viable- FM Cal04 vRNP and Cal04 FM HA/NA. Interestingly, the SIP production phenotype of these reassortants was closer to that of the Cal04 parental virus with a particle to infectivity ratio of 2:1 (Fig. 4.2B). The gene segments that are shared between the FM Cal04 vRNP and Cal04 FM HA/NA reassortant viruses are the segments encoding the Cal04 vRNP complex (PB2, PB1, PA, & NP), suggesting that the vRNP complex genotype may determine the SIP production phenotype.

To further pinpoint which gene segment encoding one of the four vRNP components is responsible for this SIP production phenotype, I generated reassortants with a FM backbone and a single Cal04 gene segment (PB2, PB1, PA, or NP). The FM virus with the Cal04 NP was unable to be rescued after several attempts. The reassortants with the FM backbone and the Cal04 PB2 or PB1 segment produced ~4-5 fold more SIPs than the FM strain, while the reassortant with the FM backbone and the Cal04 PA segment produced an equivalent number of SIPs relative to FM (Fig. 4.2B). The observation that the FM Cal04 PB2 and FM Cal04 PB1 produced more SIPs than FM is likely due to the incompatibility of the Cal04 PB2 and PB1 proteins with the remaining components of the FM RdRp or vRNP complex which would interfere with gene segment replication and transcription. Since the introduction of the Cal04 PA segment into FM did not change its SIP production phenotype, the Cal04 PA segment is likely compatible with the FM PB2/PB1/NP segments and is not responsible for the FM SIP production phenotype. In conclusion, these data

suggest that the genotype of the vRNP complex as a whole, and not its individual components, determines the SIP production phenotype.

4.3 Discussion

My data demonstrating that the SIP production phenotype varies significantly between different IAV strains suggests that SIP formation can be modulated by viral genetic determinants. Through generating reassortants of the high SIP producing FM strain and low SIP producing Cal04 strain, I identified the genetic background of the vRNP complex as the major determinant of SIP production in these two strains.

How could the genetic background of the vRNP complex affect the SIP production phenotype? One possibility is that since the vRNP complexes are required for viral gene segment replication, a vRNP complex that is less efficient at these processes would generate fewer viral genomic RNAs and transcripts, which would increase the chance of stochastic failure of subsequent steps in the viral replication cycle due to decreased gene segment abundance (Heldt et al., 2015). To test this, I plan to perform vRNP reconstitution assays in which I compare the ability of the FM and Cal04 vRNP complexes to direct the replication of a defined template. Another possibility is that the FM or Cal04 vRNP complex may induce the production of different amounts of defective RNA species such as mini viral RNAs (mvRNAs) which can activate the immune response and inhibit infection (te Velthuis et al., 2018). I plan to test this by either comparing IFN λ abundance in FM or Cal04 infected cells or by performing RNA-

seq to compare the abundance of mvRNAs in FM and Cal04 infected cells.

Finally, as the gene segments in the polymerase segments have been shown to be important for coordinating the packaging of multiple gene segments (Majarian et al., 2019; Muramoto et al., 2006; Noda et al., 2012), FM may potentially produce more SIPs than Cal04 due to a defect in the packaging of a particular segment or a group of segments. I plan to test this by quantifying the abundance of each gene segment in purified virions.

What are the potential consequences of a suboptimal vRNP complex or defect in gene segment packaging for IAV replication and evolution? IAV strains that are poorly adapted to their hosts often require co-infection to complement missing segments and increase the overall level of viral gene segment expression to compensate for suboptimal replication (Phipps et al., 2020). If a vRNP complex of a specific IAV strain has low activity or produces inflammatory defective RNA products, it could decrease the likelihood of the virus completing a full replication cycle, especially in a novel host, potentially resulting in a viral population high in SIPs which could restrict the spread of the strain to novel host populations.

Similarly, a defect in gene segment packaging due to suboptimal interactions between the gene segments encoding the vRNP components and the remaining segments would result in the production of more SIPs which would in turn require higher amounts of co-infection to complement the missing gene segments and increase expression levels. Overall, we identified the vRNP complex genotype as an important determinant of SIP production which has important implications for

the structure of IAV populations and their replication and transmission between different hosts.

4.4 Materials and Methods

Plasmids

The A/Puerto Rico/8/34 and A/Udorn/72 reverse genetics plasmids were gifts from Adolfo Garcia-Sastre (Icahn School of Medicine at Mt. Sinai) and Kanta Subbarao (Doherty Institute), respectively. The A/California/04/2009 (Cal04) reverse genetics plasmids were obtained from Jonathan Yewdell (NIH). The only differences between the Cal04 and A/California/07/2009 plasmids were two mutations: HA:A660G and NP:A335G. The A/WSN/1933 plasmids were obtained from Nicholas Wu (University of Illinois at Urbana-Champaign). The A/Perth/16/2009 (Perth) plasmids were a gift from Seema Lakdawala (University of Pittsburgh). The A/New Caledonia/20/1999 (NC99) plasmids were generated by Fadi Alnaji. The A/Fort Monmouth/1/1954 (FM) plasmids were generated by Kelsie Green, Anshika Gupta, and Kevin Gill. The FM plasmids contained the following mutations: FM PB2: T4I, N16T, R380S, L512I, and E578 silent; FM PB1: F22 silent, M171I, A453 silent, G739E, A741T, E745K, E752 silent, G754R, R757K, W758 stop (results in loss of an amino acid), and 3' UTR (mRNA sense): A→G 3nts ahead of polyA sequence and GG→AA 13-14nts ahead of polyA sequence; FM PA: 5' UTR (mRNA sense) A→G 3nts before start codon and M607I; FM NP: D290N; FM NA: 15nt deletion between and within the codons encoding amino acids N68 to N73; FM NS: W203Stop in NS1 (results in loss of 34 amino acids at the C-terminus)/G46E in NS2.

Cells

Madin-Darby canine kidney (MDCK) and 293T cells were obtained from Jonathan Yewdell (NIH) and Joanna Shisler (NSF), respectively, and were maintained in Gibco's minimal essential medium (MEM) with GlutaMax (Life Technologies) supplemented with 8.3% fetal bovine serum (Seradigm) (MEM+FBS) and incubated at 37°C with 5% CO₂. MDCK cells were used for virus propagation and infection experiments due to their ability to support robust replication of IAV and their tolerance to the trypsin necessary to convert the HA protein into its fusion-competent state. 293T cells were used for transfections to generate viruses via the reverse genetics system or to express viral proteins due to their high transfection efficiency.

Viruses

Recombinant A/Puerto Rico/08/1934 (PR8), A/New Caledonia/20/1999 (NC99), A/California/04/2009 (Cal04), A/California/07/2009 (Cal07), A/Perth/16/2009 (Perth), A/WSN/1933 (WSN), A/Fort Monmouth/1/1954 (FM) viruses, and Cal04-FM reassortants were generated using 8-plasmid reverse genetics systems. The A/Udorn/72 (Udorn) virus was generated using a 12-plasmid reverse genetics system. Viruses were rescued by transfecting 293T cells with 500 ng each of the relevant reverse genetics plasmids using JetPrime (Polyplus) according to the manufacturer's instructions. At 18 to 24 h post-transfection, the medium was replaced with viral growth medium (MEM, 1 mM HEPES, 1 µg/mL of tosylsulfonyl phenylalanyl chloromethyl ketone [TPCK] trypsin [Worthington Biochemical

Corporation, Lakewood, NJ], 50 µg/mL of gentamicin) containing 2×10^5 MDCK cells. Transfection supernatants were collected 24 h after medium change.

To generate the seed stocks of the viruses, transfection supernatants were plaque, and a single plaque was used to infect a single well of MDCK cells in a 6-well plate. Viral growth was performed in viral growth medium (MEM, 1 mM HEPES, 1 µg/mL of TPCK trypsin, 50 µg/mL of gentamicin). Seed stocks were harvested and clarified (14,000 rpm, 15 min, 4°C) between 24 and 72 h post-infection. MDCK cells in a T75 or T175 flask were then infected with the seed stocks at a multiplicity of infection (MOI) of 0.001 or 0.01 50% tissue culture infective dose (TCID₅₀)/cell, respectively, in viral growth medium, and the working stocks were harvested 24 to 72 h post-infection and clarified (3,500 rpm, 15 min, 4°C). Infectious virus titers were determined by plaque assay on MDCK cells.

The A/Malaysia/1954 (Malaysia) virus was obtained from Jonathan Yewdell (NIH).

Flow cytometry to measure NP protein expression under single-cycle infection conditions

MDCK cells were infected at an MOI of 0.03 PFU/cell in a 6-well plate. At 1 hpi, the virus was replaced with 3 mL of MEM+FBS. At 3 hpi, the MEM+FBS was replaced with 3 mL of NH₄Cl medium (MEM, 50 mM HEPES buffer, 20 mM NH₄Cl [pH 7.2]) to prevent secondary infection (Sun et al., 2020; Sun and Brooke, 2018). At 16 hpi, the cells were harvested, fixed, and permeabilized with foxP3 fix/perm buffer (eBioscience). Cells were subsequently stained with the HB65

AF647 antibody. The cells were run on a BD LSR II flow cytometer and analyzed using FlowJo version 10.1 (Tree Star, Inc.).

Hemagglutination assay

For each virus, a two-fold serial dilution series of 50 μ L was performed in (+/+) PBS from wells 2 to 11 of a single row of a round-bottom 96 well plate, leaving the first well as an undiluted virus control and the last well as a negative control ([+/+] PBS only). The plate was kept on ice for the entirety of the hemagglutination assay. Turkey red blood cells (RBCs) (Lampire Biological Laboratories Cat. #7249409) were resuspended to a concentration of 1% in (+/+) PBS, and 50 μ L of RBCs were added to each well. The plate was incubated at 4°C for 1hr and the particle count was determined by multiplying the inverse of the endpoint dilution factor by the RBC concentration.

4.5 Figures

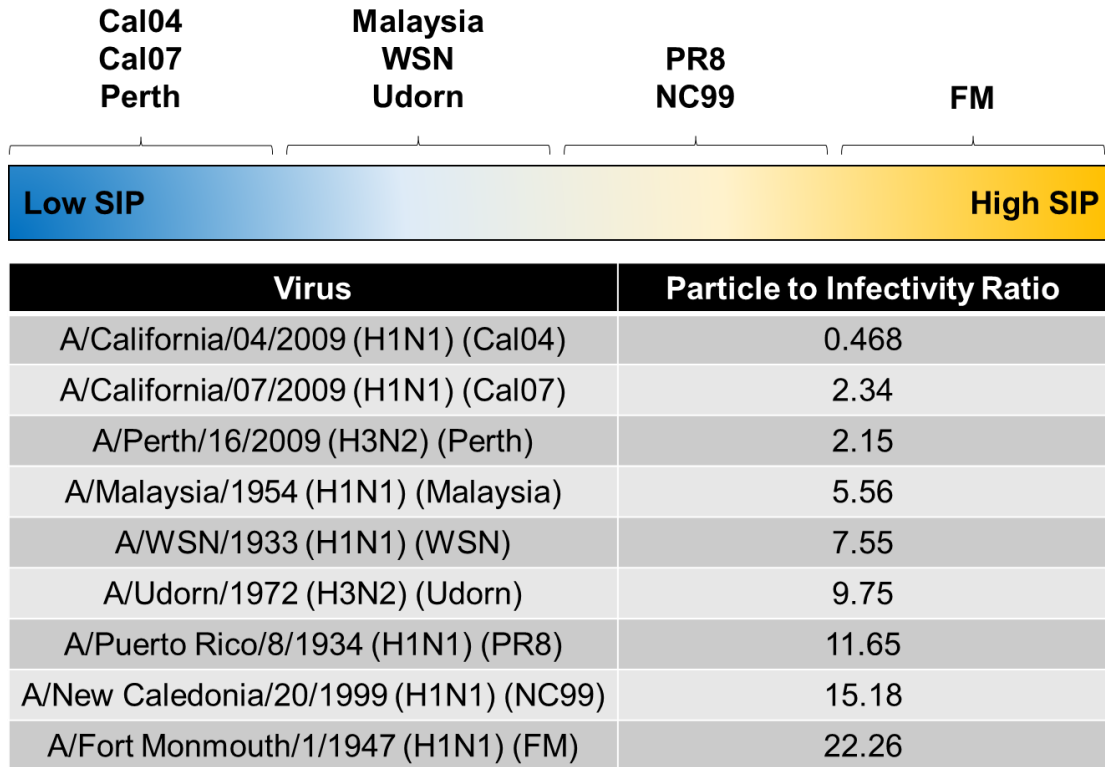


Fig. 4.1. Variation in the ratios of total to fully-infectious particles observed for different H1N1 and H3N2 IAV strains. The table lists the particle to infectivity ratios of different IAV strains determined by dividing the total particle number (determined by hemagglutination assay) by the infectious titer (determined by plaque assay [PFU]). The position of each IAV strain within the SIP production spectrum is noted above.

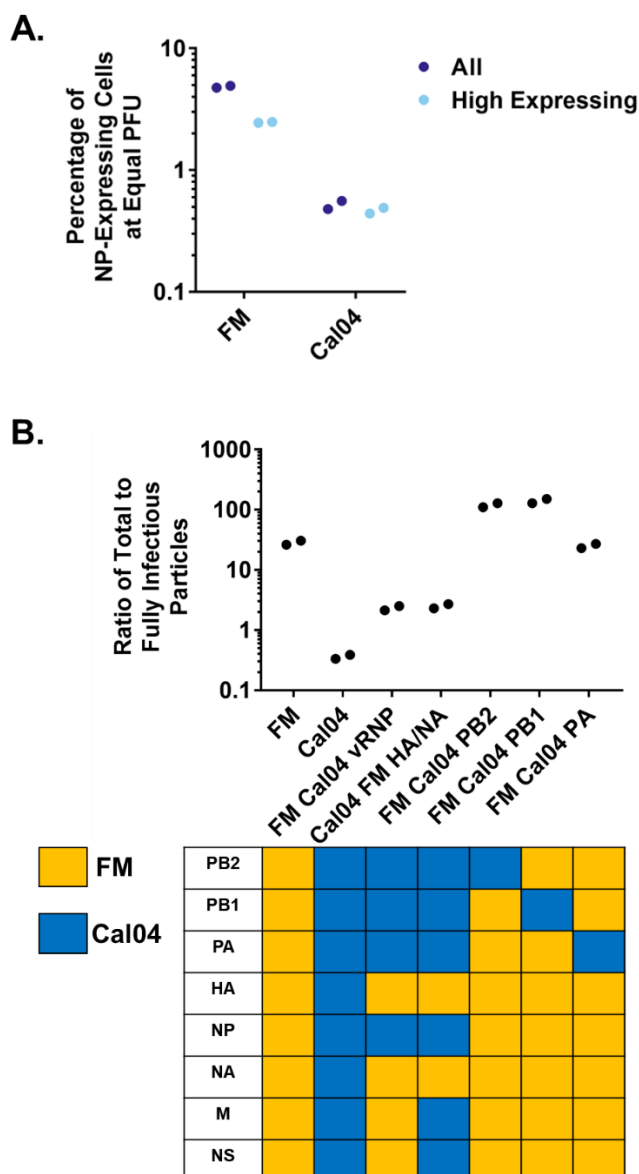


Fig. 4.2. The genetic determinant of SIP production maps to the vRNP complex. (A) Percentage of FM and Cal04 infected MDCK cells expressing high or low levels of NP as determined by flow cytometry (MOI=0.03 PFU/mL, 16hpi). Viral spread was blocked 3hpi via the addition of 20mM NH_4Cl (Sun et al., 2020; Sun and Brooke, 2018). All indicates both low and high NP expressing cells while high expressing only indicates high NP expressing cells. Data represents a single cell culture well replicate from a single experiment. (B) Particle to infectivity ratios of the indicated viruses determined by dividing the total particle number (determined by hemagglutination assay) by the infectious titer (determined by plaque assay [PFU]). The gene segment composition of each virus is indicated below. Data represents a hemagglutination assay replicate divided by the average PFU titer.

CHAPTER 5: CONCLUSIONS AND DISCUSSION

Genetic interactions within and between segments are known to play an important role in the replication and evolution of IAV populations, however, in many cases, the specific molecular mechanisms remain unknown or incomplete. In this thesis, I probed deeper into the mechanisms of epistasis at several different levels of IAV biology, ranging from the expression of individual gene segments to the genomic structure of the IAV populations.

In chapter 2 of this thesis, I explored the mechanism behind a novel form of intersegment gene-segment specific regulation in which a specific NP genotype (F346S) resulted in the selective decrease in the expression of the NA segment. In addition to residue F346S, I identified two neighboring aromatic residues (Y385 and F479) that similarly regulated NA expression, suggesting a larger role for this specific region of the NP protein in regulating gene segment expression. Utilizing sets of viruses in which I swapped regions of the sensitive NA segment with those of a resistant segment, I mapped the susceptibility determinant of NA to NP-dependent regulation to the UTRs of the segment. Both specific nucleotides or groups of nucleotides within the UTRs and pairings between those motifs in the 3' and 5' UTRs were important in determining susceptibility of the NA segment to NP-dependent regulation. Even though the F346S substitution did not affect any of the main activities of NP such as RNA binding and oligomerization, using a pair of chemical treatment and metabolic labeling protocols, I determined that NP:F346S was specifically affecting the synthesis rate of the NA vRNA via an unknown mechanism. Overall, this study provided

insights into mechanisms of how IAV can selectively regulate expression of individual segments with implications for improving our understanding of the underlying mechanisms coordinating the complex patterns of IAV gene segment expression during infection. A limitation of this study was that we were unable to determine how the NA UTR sequences and NP:F346S protein were interacting at the molecular level to result in this NP-dependent regulation. As NP:F346S specifically decreases the NA vRNA synthesis rate, further *in vitro* chemical studies examining the ability of the viral RdRp to associate with the promoter, and initiate and proceed to the elongation stage of replication in the presence of the WT or mutant NA UTRs and NP:WT or NP:F346S would be informative. To further refine the mechanism, it would be of interest to determine whether the specific motifs observed to be important for NP-dependent regulation are acting in a nucleotide sequence specific manner or through the modulation of the formation and interaction between different RNA secondary structures within the NA segment. Additionally, as a proof of concept, it would be intriguing to test whether the specific motifs we defined can affect the expression of other gene segments which would have implications for the rational design of segments with desired expression patterns. Finally, as modulating the expression of individual gene segments has the potential to alter the fine-tuned balance between the expression levels and activities of other viral gene segments and proteins respectively during infection, thereby modulating IAV replication and transmission potential, it would be of interest to see if NP residues F346, Y385, and F479 as

well as the specific motifs we identified in the UTRs are involved in adaptation of IAV to different host species or are epistatically linked.

The identification and functional validation of specific RNA structures and motifs within the ORFs of the negative-sense genomic RNAs and the evaluation of their roles in the IAV replication cycle is currently under studied and their relevance is uncertain. Building upon the work described in chapter 2, in chapter 3 of this thesis I explored the potential role of RNA sequence motifs or specific RNA structures in the NA segment ORF in determining susceptibility to NP-dependent regulation and regulating the overall expression levels of each segment.

Although I was unable to identify any RNA structures or other RNA sequence motifs in the NA ORF that regulated susceptibility of the NA segment to NP-dependent regulation, I found that introducing blocks of silent substitutions to either eliminate a specific RNA structure or RNA-RNA interaction, or generally disrupt RNA sequence elements such as dinucleotide frequencies and codon bias, reduced the expression level of the NA segment, and in one case reduced the abundance of a different segment (NP). These results are preliminary and need further experimental validation to determine if the mutations introduced successfully disrupted the predicted RNA structures and altered other RNA features such as dinucleotide content or codon bias to an extent that could influence expression levels. However, they are suggestive of a potential role for RNA structures or sequence elements within the gene segment ORF in regulating overall gene segment expression levels or RNA-RNA interactions involved in gene segment packaging. Utilizing our SHAPE-MaP data, in future

experiments, I plan to probe the role of RNA structures in other segments in gene segment packaging. Additionally, an alternative approach of introducing large blocks of silent mutations into the segments, identifying regions that exhibit a gene segment expression or packaging phenotype, and then analyzing those regions for RNA structures or motifs that are responsible would aid in the identification of and the expanding of the roles known for RNA structures or motifs in the genomic RNA ORFs in the IAV replication cycle.

Throughout the previous chapters, I demonstrated how genetic interactions within and between segments can influence gene expression levels, but how does this translate to the IAV population structure and the evolution of IAV populations as a whole? In chapter 4, I examined intersegment epistasis at a broader level, looking at how genetic interactions between segments influenced the SIP content of IAV populations. As SIPs are the majority of particles in the IAV population, they are important to consider when studying the spatial dynamics and kinetics of infection as well as the evolution of IAV populations. The observation that the SIP content of different IAV strains varied nearly a log-fold suggested that SIP production is governed by a genetic determinant, and I identified the vRNP complex genotype as a potential candidate. To substantiate these results, further mechanistic studies aimed at understanding how the vRNP complex is regulating SIP production are needed, such as measuring the relative activities of or generation of inflammatory defective RNA products by high and low SIP producing vRNP complexes, or looking at gene segment packaging efficiencies in the vRNP complex reassortants. Additionally, it is of interest to

determine if these results are transferrable to other IAV strains; i.e.) the vRNP complexes regulate SIP production in other IAV strains; and, more broadly, to determine if SIP production has consequences for process driving IAV evolution such as reassortment.

In this thesis, I demonstrated the importance of genetic interactions within and between segments (intra/intersegment epistasis respectively) in regulating aspects of IAV biology such as gene expression and the genomic structure of IAV populations. Overall, by synthesizing the mechanistic knowledge regarding how genetic interactions within and between segments regulate IAV gene segment expression and genomic heterogeneity at the population level, I aim to generate a more complete picture of how these interactions combine across multiple scales to modulate the overall pattern of IAV replication and the evolution of IAV populations.

REFERENCES

1. Abed, Y., Pizzorno, A., Bouhy, X., Boivin, G., 2011. Role of Permissive Neuraminidase Mutations in Influenza A/Brisbane/59/2007-like (H1N1) Viruses. *PLoS Pathog.* 7. <https://doi.org/10.1371/journal.ppat.1002431>.
2. Albo, C., Valencia, A., Portela, A., 1995. Identification of an RNA binding region within the N-terminal third of the influenza A virus nucleoprotein. *J. Virol.* 69, 3799–3806.
3. Ali, A., Avalos, R.T., Ponimaskin, E., Nayak, D.P., 2000. Influenza Virus Assembly: Effect of Influenza Virus Glycoproteins on the Membrane Association of M1 Protein. *J. Virol.* 74, 8709–8719.
4. Alnaji, F.G., Brooke, C.B., 2020. Influenza virus DI particles: Defective interfering or delightfully interesting? *PLOS Pathog.* 16, e1008436. <https://doi.org/10.1371/journal.ppat.1008436>.
5. Alnaji, F.G., Holmes, J.R., Rendon, G., Vera, J.C., Fields, C.J., Martin, B.E., Brooke, C.B., 2019. Sequencing Framework for the Sensitive Detection and Precise Mapping of Defective Interfering Particle-Associated Deletions across Influenza A and B Viruses. *J. Virol.* 93. <https://doi.org/10.1128/JVI.00354-19>.
6. Alnaji, F.G., Reiser, W.K., Rivera-Cardona, J., te Velhuis, A.J.W., Brooke, C.B. 2021. Influenza A Virus Defective Viral Genomes Are Inefficiently Packaged into Virions Relative to Wild-Type Genomic RNAs. *mBio* 12. <https://doi.org/10.1128/mBio.02959-21>.
7. Amorim, M.J., Bruce, E.A., Read, E.K.C., Foeglein, Á., Mahen, R., Stuart, A.D., Digard, P., 2011. A Rab11- and Microtubule-Dependent Mechanism for Cytoplasmic Transport of Influenza A Virus Viral RNA ν . *J. Virol.* 85, 4143–4156. <https://doi.org/10.1128/JVI.02606-10>.
8. Arranz, R., Coloma, R., Chichón, F.J., Conesa, J.J., Carrascosa, J.L., Valpuesta, J.M., Ortín, J., Martín-Benito, J., 2012. The Structure of Native Influenza Virion Ribonucleoproteins. *Science* 338, 1634–1637. <https://doi.org/10.1126/science.1228172>.
9. Babar, M.M., Zaidi, N.-S.S., 2015. Protein sequence conservation and stable molecular evolution reveals influenza virus nucleoprotein as a universal druggable target. *Infect. Genet. Evol.* 34, 200–210. <https://doi.org/10.1016/j.meegid.2015.06.030>.

10. Babcock, H.P., Chen, C., Zhuang, X., 2004. Using Single-Particle Tracking to Study Nuclear Trafficking of Viral Genes. *Biophys. J.* 87, 2749–2758. <https://doi.org/10.1529/biophysj.104.042234>.
11. Bae, S.-H., Cheong, H.-K., Lee, J.-H., Cheong, C., Kainosho, M., Choi, B.-S., 2001. Structural features of an influenza virus promoter and their implications for viral RNA synthesis. *Proc. Natl. Acad. Sci.* 98, 10602–10607. <https://doi.org/10.1073/pnas.191268798>.
12. Barman, S., Adhikary, L., Chakrabarti, A.K., Bernas, C., Kawaoka, Y., Nayak, D.P., 2004. Role of Transmembrane Domain and Cytoplasmic Tail Amino Acid Sequences of Influenza A Virus Neuraminidase in Raft Association and Virus Budding. *J. Virol.* 78, 5258–5269. <https://doi.org/10.1128/JVI.78.10.5258-5269.2004>.
13. Baudin, F., Petit, I., Weissenhorn, W., Ruigrok, R.W.H., 2001. In Vitro Dissection of the Membrane and RNP Binding Activities of Influenza Virus M1 Protein. *Virology* 281, 102–108. <https://doi.org/10.1006/viro.2000.0804>.
14. Benton, D.J., Martin, S.R., Wharton, S.A., McCauley, J.W., 2015. Biophysical Measurement of the Balance of Influenza A Hemagglutinin and Neuraminidase Activities. *J. Biol. Chem.* 290, 6516–6521. <https://doi.org/10.1074/jbc.M114.622308>.
15. Bhatt, S., Holmes, E.C., Pybus, O.G., 2011. The Genomic Rate of Molecular Adaptation of the Human Influenza A Virus. *Mol. Biol. Evol.* 28, 2443–2451. <https://doi.org/10.1093/molbev/msr044>.
16. Bloom, J.D., 2014. An Experimentally Determined Evolutionary Model Dramatically Improves Phylogenetic Fit. *Mol. Biol. Evol.* 31, 1956–1978. <https://doi.org/10.1093/molbev/msu173>.
17. Bloom, J.D., Gong, L.I., Baltimore, D., 2010. Permissive Secondary Mutations Enable the Evolution of Influenza Oseltamivir Resistance. *Science* 328, 1272–1275. <https://doi.org/10.1126/science.1187816>.
18. Bolte, H., Rosu, M.E., Hagelauer, E., García-Sastre, A., Schwemmle, M., 2019. Packaging of the Influenza Virus Genome Is Governed by a Plastic Network of RNA- and Nucleoprotein-Mediated Interactions. *J. Virol.* 93, e01861-18. <https://doi.org/10.1128/JVI.01861-18>.
19. Böttcher-Friebertshäuser, E., Garten, W., Matrosovich, M., Klenk, H.D., 2014. The Hemagglutinin: A Determinant of Pathogenicity 3–34. https://doi.org/10.1007/82_2014_384.

20. Brooke, C.B., 2017. Population Diversity and Collective Interactions during Influenza Virus Infection. *J. Virol.* 91, e01164-17. <https://doi.org/10.1128/JVI.01164-17>.
21. Brooke, C.B., 2014. Biological activities of “noninfectious” influenza A virus particles. *Future Virol.* 9, 41–51. <https://doi.org/10.2217/fvl.13.118>.
22. Brooke, C.B., Ince, W.L., Wei, J., Bennink, J.R., Yewdell, J.W., 2014. Influenza A virus nucleoprotein selectively decreases neuraminidase gene-segment packaging while enhancing viral fitness and transmissibility. *Proc. Natl. Acad. Sci. U. S. A.* 111, 16854–16859. <https://doi.org/10.1073/pnas.1415396111>.
23. Brooke, C.B., Ince, W.L., Wrammert, J., Ahmed, R., Wilson, P.C., Bennink, J.R., Yewdell, J.W., 2013. Most Influenza A Virions Fail To Express at Least One Essential Viral Protein. *J. Virol.* 87, 3155–3162. <https://doi.org/10.1128/JVI.02284-12>.
24. Brunotte, L., Flies, J., Bolte, H., Reuther, P., Vreede, F., Schwemmler, M., 2014. The Nuclear Export Protein of H5N1 Influenza A Viruses Recruits Matrix 1 (M1) Protein to the Viral Ribonucleoprotein to Mediate Nuclear Export. *J. Biol. Chem.* 289, 20067–20077. <https://doi.org/10.1074/jbc.M114.569178>.
25. Byrd-Leotis, L., Cummings, R.D., Steinhauer, D.A., 2017. The Interplay between the Host Receptor and Influenza Virus Hemagglutinin and Neuraminidase. *Int. J. Mol. Sci.* 18, 1541. <https://doi.org/10.3390/ijms18071541>.
26. Campbell, P.J., Danzy, S., Kyriakis, C.S., Deymier, M.J., Lowen, A.C., Steel, J., 2014. The M Segment of the 2009 Pandemic Influenza Virus Confers Increased Neuraminidase Activity, Filamentous Morphology, and Efficient Contact Transmissibility to A/Puerto Rico/8/1934-Based Reassortant Viruses. *J. Virol.* 88, 3802–3814. <https://doi.org/10.1128/JVI.03607-13>.
27. Casalegno, J.-S., Ferraris, O., Escuret, V., Bouscambert, M., Bergeron, C., Linès, L., Excoffier, T., Valette, M., Frobert, E., Pillet, S., Pozzetto, B., Lina, B., Ottmann, M., 2014. Functional Balance between the Hemagglutinin and Neuraminidase of Influenza A(H1N1)pdm09 HA D222 Variants. *PLOS ONE* 9, e104009. <https://doi.org/10.1371/journal.pone.0104009>.
28. CDC, 2020. Benefits of Flu Vaccination During 2018-2019 Flu Season [WWW Document]. *Cent. Dis. Control Prev.* URL <https://www.cdc.gov/flu/about/burden-averted/2019-2020.htm> (accessed 3.28.22).

29. Chan, W.-H., Ng, A.K.-L., Robb, N.C., Lam, M.K.-H., Chan, P.K.-S., Au, S.W.-N., Wang, J.-H., Fodor, E., Shaw, P.-C., 2010. Functional Analysis of the Influenza Virus H5N1 Nucleoprotein Tail Loop Reveals Amino Acids That Are Crucial for Oligomerization and Ribonucleoprotein Activities. *J. Virol.* 84, 7337–7345. <https://doi.org/10.1128/JVI.02474-09>.
30. Chen, B.J., Leser, G.P., Morita, E., Lamb, R.A., 2007. Influenza Virus Hemagglutinin and Neuraminidase, but Not the Matrix Protein, Are Required for Assembly and Budding of Plasmid-Derived Virus-Like Particles. *J. Virol.* 81, 7111–7123. <https://doi.org/10.1128/JVI.00361-07>.
31. Chen, K.-Y., Santos Afonso, E.D., Enouf, V., Isel, C., Naffakh, N., 2019. Influenza virus polymerase subunits co-evolve to ensure proper levels of dimerization of the heterotrimer. *PLoS Pathog.* 15. <https://doi.org/10.1371/journal.ppat.1008034>.
32. Chen, L.-M., Davis, C.T., Zhou, H., Cox, N.J., Donis, R.O., 2008. Genetic Compatibility and Virulence of Reassortants Derived from Contemporary Avian H5N1 and Human H3N2 Influenza A Viruses. *PLoS Pathog.* 4. <https://doi.org/10.1371/journal.ppat.1000072>.
33. Chenavas, S., Estrozi, L.F., Slama-Schwok, A., Delmas, B., Di Primo, C., Baudin, F., Xinping Li, Crépin, T., Ruigrok, R.W.H., 2013. Monomeric Nucleoprotein of Influenza A Virus. *PLoS Pathog.* 9, 1–10. <https://doi.org/10.1371/journal.ppat.1003275>.
34. Chlanda, P., Schraidt, O., Kummer, S., Riches, J., Oberwinkler, H., Prinz, S., Kräusslich, H.-G., Briggs, J.A.G., 2015. Structural Analysis of the Roles of Influenza A Virus Membrane-Associated Proteins in Assembly and Morphology. *J. Virol.* 89, 8957–8966. <https://doi.org/10.1128/JVI.00592-15>.
35. Choppin, P.W., Murphy, J.S., Tamm, I., 1960. STUDIES OF TWO KINDS OF VIRUS PARTICLES WHICH COMPRISE INFLUENZA A2 VIRUS STRAINS. *J. Exp. Med.* 112, 945–952.
36. Chou, Y., Heaton, N.S., Gao, Q., Palese, P., Singer, R., Lionnet, T., 2013. Colocalization of Different Influenza Viral RNA Segments in the Cytoplasm before Viral Budding as Shown by Single-molecule Sensitivity FISH Analysis. *PLoS Pathog.* 9. <https://doi.org/10.1371/journal.ppat.1003358>.
37. Chu, C.M., Dawson, I.M., Elford, W.J., 1949. FILAMENTOUS FORMS ASSOCIATED WITH NEWLY ISOLATED INFLUENZA VIRUS. *The Lancet* 253, 602. [https://doi.org/10.1016/S0140-6736\(49\)91699-2](https://doi.org/10.1016/S0140-6736(49)91699-2).

38. Cohen-Daniel, L., Zakay-Rones, Z., Resnick, I.B., Shapira, M.Y., Dorozhko, M., Mador, N., Greenbaum, E., Wolf, D.G., 2009. Emergence of oseltamivir-resistant influenza A/H3N2 virus with altered hemagglutination pattern in a hematopoietic stem cell transplant recipient. *J. Clin. Virol.* 44, 138–140. <https://doi.org/10.1016/j.jcv.2008.11.014>.
39. Dadonaite, B., Gilbertson, B., Knight, M.L., Trifkovic, S., Rockman, S., Laederach, A., Brown, L.E., Fodor, E., Bauer, D.L.V., 2019. The structure of the influenza A virus genome. *Nat. Microbiol.* 1. <https://doi.org/10.1038/s41564-019-0513-7>.
40. Dahmani, I., Ludwig, K., Chiantia, S., 2019. Influenza A matrix protein M1 induces lipid membrane deformation via protein multimerization. *Biosci. Rep.* 39. <https://doi.org/10.1042/BSR20191024>.
41. Dai, M., Du, W., Martínez-Romero, C., Leenders, T., Wennekes, T., Rimmelzwaan, G.F., van Kuppeveld, F.J.M., Fouchier, R.A.M., Garcia-Sastre, A., de Vries, E., de Haan, C.A.M. 2021. Analysis of the Evolution of Pandemic Influenza A(H1N1) Virus Neuraminidase Reveals Entanglement of Different Phenotypic Characteristics. *mBio* 12. <https://doi.org/10.1128/mBio.00287-21>.
42. Das, S.R., Hensley, S.E., David, A., Schmidt, L., Gibbs, J.S., Puigbò, P., Ince, W.L., Bennink, J.R., Yewdell, J.W., 2011. Fitness costs limit influenza A virus hemagglutinin glycosylation as an immune evasion strategy. *Proc. Natl. Acad. Sci.* 108, E1417–E1422. <https://doi.org/10.1073/pnas.1108754108>.
43. Das, S.R., Hensley, S.E., Ince, W.L., Brooke, C.B., Subba, A., Delboy, M.G., Russ, G., Gibbs, J.S., Bennink, J.R., Yewdell, J.W., 2013. Defining Influenza A Virus Hemagglutinin Antigenic Drift by Sequential Monoclonal Antibody Selection. *Cell Host Microbe* 13, 314–323. <https://doi.org/10.1016/j.chom.2013.02.008>.
44. Davis, A.R., Hiti, A.L., Nayak, D.P., 1980. Influenza defective interfering viral RNA is formed by internal deletion of genomic RNA. *Proc. Natl. Acad. Sci.* 77, 215–219.
45. de Castro Martin, I.F., Fournier, G., Sachse, M., Pizarro-Cerda, J., Risco, C., Naffakh, N., 2017. Influenza virus genome reaches the plasma membrane via a modified endoplasmic reticulum and Rab11-dependent vesicles. *Nat. Commun.* 8. <https://doi.org/10.1038/s41467-017-01557-6>.
46. Deng, T., Vreede, F.T., Brownlee, G.G., 2006. Different De Novo Initiation Strategies Are Used by Influenza Virus RNA Polymerase on Its cRNA and Viral RNA Promoters during Viral RNA Replication. *J. Virol.* 80, 2337–2348. <https://doi.org/10.1128/JVI.80.5.2337-2348.2006>.

47. Desselberger, U., Racaniello, V.R., Zazra, J.J., Palese, P., 1980. 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. [https://doi.org/10.1016/0378-1119\(80\)90007-4](https://doi.org/10.1016/0378-1119(80)90007-4).
48. Diefenbacher, M., Sun, J., Brooke, C.B., 2018. The parts are greater than the whole: the role of semi-infectious particles in influenza A virus biology. *Curr. Opin. Virol.*, Virus vector interactions • Special Section: Multicomponent viral systems 33, 42–46. <https://doi.org/10.1016/j.coviro.2018.07.002>.
49. Dou, D., Hernández-Neuta, I., Wang, H., Östbye, H., Qian, X., Thiele, S., Resa-Infante, P., Kouassi, N.M., Sender, V., Hentrich, K., Mellroth, P., Henriques-Normark, B., Gabriel, G., Nilsson, M., Daniels, R., 2017. Analysis of IAV Replication and Co-infection Dynamics by a Versatile RNA Viral Genome Labeling Method. *Cell Rep.* 20, 251–263. <https://doi.org/10.1016/j.celrep.2017.06.021>.
50. Dou, D., Revol, R., Östbye, H., Wang, H., Daniels, R., 2018. Influenza A Virus Cell Entry, Replication, Virion Assembly and Movement. *Front. Immunol.* 9. <https://doi.org/10.3389/fimmu.2018.01581>.
51. Doud, M.B., Hensley, S.E., Bloom, J.D., 2017. Complete mapping of viral escape from neutralizing antibodies. *PLoS Pathog.* 13. <https://doi.org/10.1371/journal.ppat.1006271>.
52. Duan, S., Govorkova, E.A., Bahl, J., Zaraket, H., Baranovich, T., Seiler, P., Prevost, K., Webster, R.G., Webby, R.J., 2014. Epistatic interactions between neuraminidase mutations facilitated the emergence of the oseltamivir-resistant H1N1 influenza viruses. *Nat. Commun.* 5, 5029. <https://doi.org/10.1038/ncomms6029>.
53. DuBois, R.M., Zaraket, H., Reddivari, M., Heath, R.J., White, S.W., Russell, C.J., 2011. Acid Stability of the Hemagglutinin Protein Regulates H5N1 Influenza Virus Pathogenicity. *PLoS Pathog.* 7. <https://doi.org/10.1371/journal.ppat.1002398>.
54. Eisfeld, A.J., Kawakami, E., Watanabe, T., Neumann, G., Kawaoka, Y., 2011. RAB11A Is Essential for Transport of the Influenza Virus Genome to the Plasma Membrane. *J. Virol.* 85, 6117. <https://doi.org/10.1128/JVI.00378-11>.
55. Eisfeld, A.J., Neumann, G., Kawaoka, Y., 2015. At the centre: influenza A virus ribonucleoproteins. *Nat. Rev. Microbiol.* 13, 28–41. <https://doi.org/10.1038/nrmicro3367>.

56. Elde, N.C., Child, S.J., Eickbush, M.T., Kitzman, J.O., Rogers, K.S., Shendure, J., Geballe, A.P., Malik, H.S., 2012. Poxviruses deploy genomic accordions to adapt rapidly against host antiviral defenses. *Cell* 150, 831–841. <https://doi.org/10.1016/j.cell.2012.05.049>.
57. Elton, D., Medcalf, E., Bishop, K., Digard, P., 1999a. Oligomerization of the Influenza Virus Nucleoprotein: Identification of Positive and Negative Sequence Elements. *Virology* 260, 190–200. <https://doi.org/10.1006/viro.1999.9818>.
58. Elton, D., Medcalf, L., Bishop, K., Harrison, D., Digard, P., 1999b. Identification of Amino Acid Residues of Influenza Virus Nucleoprotein Essential for RNA Binding. *J. Virol.* 73, 7357–7367.
59. Elton, D., Simpson-Holley, M., Archer, K., Medcalf, L., Hallam, R., McCauley, J., Digard, P., 2001. Interaction of the Influenza Virus Nucleoprotein with the Cellular CRM1-Mediated Nuclear Export Pathway. *J. Virol.* 75, 408–419. <https://doi.org/10.1128/JVI.75.1.408-419.2001>.
60. Fan, H., Walker, A.P., Carrique, L., Keown, J.R., Martin, I.S., Karia, D., Sharps, J., Hengrung, N., Pardon, E., Steyaert, J., Grimes, J.M., Fodor, E., 2019. Structures of influenza A virus RNA polymerase offer insight into viral genome replication. *Nature* 573, 287–290. <https://doi.org/10.1038/s41586-019-1530-7>.
61. Ferhadian, D., Contrant, M., Printz-Schweigert, A., Smyth, R.P., Paillart, J.-C., Marquet, R., 2018. Structural and Functional Motifs in Influenza Virus RNAs. *Front. Microbiol.* 9. <https://doi.org/10.3389/fmicb.2018.00559>.
62. Fodor, E., Velthuis, A.J.W. te, 2019. Structure and Function of the Influenza Virus Transcription and Replication Machinery. *Cold Spring Harb. Perspect. Med.* a038398. <https://doi.org/10.1101/cshperspect.a038398>.
63. Fonville, J.M., Marshall, N., Tao, H., Steel, J., Lowen, A.C., 2015. Influenza Virus Reassortment Is Enhanced by Semi-infectious Particles but Can Be Suppressed by Defective Interfering Particles. *PLoS Pathog.* 11. <https://doi.org/10.1371/journal.ppat.1005204>.
64. Fournier, E., Moules, V., Essere, B., Paillart, J.-C., Sirbat, J.-D., Isel, C., Cavalier, A., Rolland, J.-P., Thomas, D., Lina, B., Marquet, R., 2012. A supramolecular assembly formed by influenza A virus genomic RNA segments. *Nucleic Acids Res.* 40, 2197–2209. <https://doi.org/10.1093/nar/gkr985>.

65. Fujii, K., Fujii, Y., Noda, T., Muramoto, Y., Watanabe, T., Takada, A., Goto, H., Horimoto, T., Kawaoka, Y., 2005. Importance of both the Coding and the Segment-Specific Noncoding Regions of the Influenza A Virus NS Segment for Its Efficient Incorporation into Virions. *J. Virol.* 79, 3766–3774. <https://doi.org/10.1128/JVI.79.6.3766-3774.2005>.
66. Fujii, Y., Goto, H., Watanabe, T., Yoshida, T., Kawaoka, Y., 2003. Selective incorporation of influenza virus RNA segments into virions. *Proc. Natl. Acad. Sci.* 100, 2002–2007. <https://doi.org/10.1073/pnas.0437772100>.
67. Galloway, S.E., Reed, M.L., Russell, C.J., Steinhauer, D.A., 2013. Influenza HA Subtypes Demonstrate Divergent Phenotypes for Cleavage Activation and pH of Fusion: Implications for Host Range and Adaptation. *PLoS Pathog.* 9. <https://doi.org/10.1371/journal.ppat.1003151>.
68. Gavazzi, C., Yver, M., Isel, C., Smyth, R.P., Rosa-Calatrava, M., Lina, B., Moulès, V., Marquet, R., 2013. A functional sequence-specific interaction between influenza A virus genomic RNA segments. *Proc. Natl. Acad. Sci.* 110, 16604–16609. <https://doi.org/10.1073/pnas.1314419110>.
69. Gaymard, A., Le Briand, N., Frobert, E., Lina, B., Escuret, V., 2016. Functional balance between neuraminidase and haemagglutinin in influenza viruses. *Clin. Microbiol. Infect.* 22, 975–983. <https://doi.org/10.1016/j.cmi.2016.07.007>.
70. Ginting, T.E., Shinya, K., Kyan, Y., Makino, A., Matsumoto, N., Kaneda, S., Kawaoka, Y., 2012. Amino Acid Changes in Hemagglutinin Contribute to the Replication of Oseltamivir-Resistant H1N1 Influenza Viruses. *J. Virol.* 86, 121–127. <https://doi.org/10.1128/JVI.06085-11>.
71. Gong, L.I., Suchard, M.A., Bloom, J.D., 2013. Stability-mediated epistasis constrains the evolution of an influenza protein. *eLife* 2. <https://doi.org/10.7554/eLife.00631>.
72. Goto, H., Muramoto, Y., Noda, T., Kawaoka, Y., 2013. The Genome-Packaging Signal of the Influenza A Virus Genome Comprises a Genome Incorporation Signal and a Genome-Bundling Signal. *J. Virol.* 87, 11316–11322. <https://doi.org/10.1128/JVI.01301-13>.
73. Götz, V., Magar, L., Dornfeld, D., Giese, S., Pohlmann, A., Höper, D., Kong, B.-W., Jans, D.A., Beer, M., Haller, O., Schwemmle, M., 2016. Influenza A viruses escape from MxA restriction at the expense of efficient nuclear vRNP import. *Sci. Rep.* 6. <https://doi.org/10.1038/srep23138>.

74. Guo, H., Rabouw, H., Slomp, A., Dai, M., Vegt, F. van der, Lent, J.W.M. van, McBride, R., Paulson, J.C., Groot, R.J. de, Kuppeveld, F.J.M. van, Vries, E. de, Haan, C.A.M. de, 2018. Kinetic analysis of the influenza A virus HA/NA balance reveals contribution of NA to virus-receptor binding and NA-dependent rolling on receptor-containing surfaces. *PLOS Pathog.* 14, e1007233. <https://doi.org/10.1371/journal.ppat.1007233>.
75. Harris, A., Cardone, G., Winkler, D.C., Heymann, J.B., Brecher, M., White, J.M., Steven, A.C., 2006. Influenza virus pleiomorphy characterized by cryoelectron tomography. *Proc. Natl. Acad. Sci. U. S. A.* 103, 19123–19127. <https://doi.org/10.1073/pnas.0607614103>.
76. Hatada, E., Hasegawa, M., Mukaigawa, J., Shimizu, K., Fukuda, R., 1989. Control of Influenza Virus Gene Expression: Quantitative Analysis of Each Viral RNA Species in Infected Cells. *J. Biochem. (Tokyo)* 105, 537–546. <https://doi.org/10.1093/oxfordjournals.jbchem.a122702>.
77. Heldt, F.S., Kupke, S.Y., Dorl, S., Reichl, U., Frensing, T., 2015. Single-cell analysis and stochastic modelling unveil large cell-to-cell variability in influenza A virus infection. *Nat. Commun.* 6. <https://doi.org/10.1038/ncomms9938>.
78. Hensley, S.E., Das, S.R., Bailey, A.L., Schmidt, L.M., Hickman, H.D., Jayaraman, A., Viswanathan, K., Raman, R., Sasisekharan, R., Bennink, J.R., Yewdell, J.W., 2009. Hemagglutinin Receptor Binding Avidity Drives Influenza A Virus Antigenic Drift. *Science* 326, 734–736. <https://doi.org/10.1126/science.1178258>.
79. Hensley, S.E., Das, S.R., Gibbs, J.S., Bailey, A.L., Schmidt, L.M., Bennink, J.R., Yewdell, J.W., 2011. Influenza A Virus Hemagglutinin Antibody Escape Promotes Neuraminidase Antigenic Variation and Drug Resistance. *PLOS ONE* 6, e15190. <https://doi.org/10.1371/journal.pone.0015190>.
80. Herfst, S., Schrauwen, E.J.A., Linster, M., Chutinimitkul, S., de Wit, E., Munster, V.J., Sorrell, E.M., Bestebroer, T.M., Burke, D.F., Smith, D.J., Rimmelzwaan, G.F., Osterhaus, A.D.M.E., Fouchier, R.A.M., 2012. Airborne Transmission of Influenza A/H5N1 Virus Between Ferrets. *Science* 336, 1534–1541. <https://doi.org/10.1126/science.1213362>.
81. Huang, S., Chen, Jingjing, Chen, Q., Wang, H., Yao, Y., Chen, Jianjun, Chen, Z., 2013. A Second CRM1-Dependent Nuclear Export Signal in the Influenza A Virus NS2 Protein Contributes to the Nuclear Export of Viral Ribonucleoproteins. *J. Virol.* 87, 767–778. <https://doi.org/10.1128/JVI.06519-11>.

82. Hutchinson, E.C., Curran, M.D., Read, E.K., Gog, J.R., Digard, P., 2008. Mutational Analysis of cis-Acting RNA Signals in Segment 7 of Influenza A Virus. *J. Virol.* 82, 11869–11879. <https://doi.org/10.1128/JVI.01634-08>.
83. Hutchinson, E.C., Wise, H.M., Kudryavtseva, K., Curran, M.D., Digard, P., 2009. Characterisation of influenza A viruses with mutations in segment 5 packaging signals. *Vaccine* 27, 6270–6275. <https://doi.org/10.1016/j.vaccine.2009.05.053>.
84. Ince, W.L., Gueye-Mbaye, A., Bennink, J.R., Yewdell, J.W., 2013. Reassortment Complements Spontaneous Mutation in Influenza A Virus NP and M1 Genes To Accelerate Adaptation to a New Host. *J. Virol.* 87, 4330–4338. <https://doi.org/10.1128/JVI.02749-12>.
85. Iuliano, A.D., Roguski, K.M., Chang, H.H., Muscatello, D.J., Palekar, R., Tempia, S., Cohen, C., Gran, J.M., Schanzer, D., Cowling, B.J., Wu, P., Kyncl, J., Ang, L.W., Park, M., Redlberger-Fritz, M., Yu, H., Espenhain, L., Krishnan, A., Emukule, G., van Asten, L., da Silva, S.P., Aungkulanon, S., Buchholz, U., Widdowson, M.-A., Bresee, J.S., 2018. Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet Lond. Engl.* 391, 1285–1300. [https://doi.org/10.1016/S0140-6736\(17\)33293-2](https://doi.org/10.1016/S0140-6736(17)33293-2).
86. Jiang, L., Liu, P., Bank, C., Renzette, N., Prachanronarong, K., Yilmaz, L.S., Caffrey, D.R., Zeldovich, K.B., Schiffer, C.A., Kowalik, T.F., Jensen, J.D., Finberg, R.W., Wang, J.P., Bolon, D.N.A., 2016. A Balance between Inhibitor Binding and Substrate Processing Confers Influenza Drug Resistance. *J. Mol. Biol.* 428, 538–553. <https://doi.org/10.1016/j.jmb.2015.11.027>.
87. Jiang, T., Nogales, A., Baker, S.F., Martinez-Sobrido, L., Turner, D.H., 2016. Mutations Designed by Ensemble Defect to Misfold Conserved RNA Structures of Influenza A Segments 7 and 8 Affect Splicing and Attenuate Viral Replication in Cell Culture. *PLoS ONE* 11. <https://doi.org/10.1371/journal.pone.0156906>.
88. Johnson, N.P.A.S., Mueller, J., 2002. Updating the accounts: global mortality of the 1918-1920 “Spanish” influenza pandemic. *Bull. Hist. Med.* 76, 105–115.
89. Jorge, D.M. de M., Mills, R.E., Lauring, A.S., 2015. CodonShuffle: a tool for generating and analyzing synonymously mutated sequences. *Virus Evol.* 1. <https://doi.org/10.1093/ve/vev012>.

90. Karakus, U., Pohl, M.O., Stertz, S., 2020. Breaking the Convention: Sialoglycan Variants, Coreceptors, and Alternative Receptors for Influenza A Virus Entry. *J. Virol.* 94, e01357-19. <https://doi.org/10.1128/JVI.01357-19>.
91. Kobayashi, M., Toyoda, T., Adyshev, D.M., Azuma, Y., Ishihama, A., 1994. Molecular dissection of influenza virus nucleoprotein: deletion mapping of the RNA binding domain. *J. Virol.* 68, 8433–8436.
92. Kobayashi, Y., Dadonaite, B., Doremalen, N. van, Suzuki, Y., Barclay, W.S., Pybus, O.G., 2016. Computational and molecular analysis of conserved influenza A virus RNA secondary structures involved in infectious virion production. *RNA Biol.* 13, 883–894. <https://doi.org/10.1080/15476286.2016.1208331>.
93. Koelle, K., Cobey, S., Grenfell, B., Pascual, M., 2006. Epochal Evolution Shapes the Phylodynamics of Interpandemic Influenza A (H3N2) in Humans. *Science*. <https://doi.org/10.1126/science.1132745>.
94. Kosik, I., Angeletti, D., Gibbs, J.S., Angel, M., Takeda, K., Kosikova, M., Nair, V., Hickman, H.D., Xie, H., Brooke, C.B., Yewdell, J.W., 2019. Neuraminidase inhibition contributes to influenza A virus neutralization by anti-hemagglutinin stem antibodies. *J. Exp. Med.* 216, 304–316. <https://doi.org/10.1084/jem.20181624>.
95. Kosik, I., Ince, W.L., Gentles, L.E., Oler, A.J., Kosikova, M., Angel, M., Magadán, J.G., Xie, H., Brooke, C.B., Yewdell, J.W., 2018. Influenza A virus hemagglutinin glycosylation compensates for antibody escape fitness costs. *PLoS Pathog.* 14. <https://doi.org/10.1371/journal.ppat.1006796>.
96. Krammer, F., Smith, G.J.D., Fouchier, R.A.M., Peiris, M., Kedzierska, K., Doherty, P.C., Palese, P., Shaw, M.L., Treanor, J., Webster, R.G., García-Sastre, A., 2018. Influenza. *Nat. Rev. Dis. Primer* 4, 3. <https://doi.org/10.1038/s41572-018-0002-y>.
97. Lai, J.C.C., Karunaratna, H.M.T.K., Wong, H.H., Peiris, J.S.M., Nicholls, J.M., 2019. Neuraminidase activity and specificity of influenza A virus are influenced by haemagglutinin-receptor binding. *Emerg. Microbes Infect.* 8, 327–338. <https://doi.org/10.1080/22221751.2019.1581034>.
98. Lakdawala, S.S., Wu, Y., Wawrzusin, P., Kabat, J., Broadbent, A.J., Lamirande, E.W., Fodor, E., Altan-Bonnet, N., Shroff, H., Subbarao, K., 2014. Influenza A Virus Assembly Intermediates Fuse in the Cytoplasm. *PLoS Pathog.* 10. <https://doi.org/10.1371/journal.ppat.1003971>.

99. Lampejo, T., 2020. Influenza and antiviral resistance: an overview. *Eur. J. Clin. Microbiol. Infect. Dis.* 1–8. <https://doi.org/10.1007/s10096-020-03840-9>.
100. Lee, J.M., Huddleston, J., Doud, M.B., Hooper, K.A., Wu, N.C., Bedford, T., Bloom, J.D., 2018. Deep mutational scanning of hemagglutinin helps predict evolutionary fates of human H3N2 influenza variants. *Proc. Natl. Acad. Sci. U. S. A.* 115, E8276–E8285. <https://doi.org/10.1073/pnas.1806133115>.
101. Lee, K.H., Seong, B.L. 1998. The position 4 nucleotide at the 3' end of the influenza virus neuraminidase vRNA is involved in temporal regulation of transcription and replication of neuraminidase RNAs and affects the repertoire of influenza virus surface antigens. *J. Gen. Virol.* 79(8), 1923–1934. <https://doi.org/10.1099/0022-1317-79-8-1923>.
102. Lee, M.-K., Bae, S.-H., Park, C.-J., Cheong, H.-K., Cheong, C., Choi, B.-S., 2003. A single-nucleotide natural variation (U4 to C4) in an influenza A virus promoter exhibits a large structural change: implications for differential viral RNA synthesis by RNA-dependent RNA polymerase. *Nucleic Acids Res.* 31, 1216–1223.
103. Leser, G.P., Lamb, R.A., 2005. Influenza virus assembly and budding in raft-derived microdomains: A quantitative analysis of the surface distribution of HA, NA and M2 proteins. *Virology* 342, 215–227. <https://doi.org/10.1016/j.virol.2005.09.049>.
104. Liang, Y., Hong, Y., Parslow, T.G., 2005. cis-Acting Packaging Signals in the Influenza Virus PB1, PB2, and PA Genomic RNA Segments. *J. Virol.* 79, 10348–10355. <https://doi.org/10.1128/JVI.79.16.10348-10355.2005>.
105. Lin, Y.-C., Jeng, K.-S., Lai, M.M.C., 2017. CNOT4-Mediated Ubiquitination of Influenza A Virus Nucleoprotein Promotes Viral RNA Replication. *mBio* 8. <https://doi.org/10.1128/mBio.00597-17>.
106. Long, J.S., Mistry, B., Haslam, S.M., Barclay, W.S., 2019. Host and viral determinants of influenza A virus species specificity. *Nat. Rev. Microbiol.* 17, 67. <https://doi.org/10.1038/s41579-018-0115-z>.
107. Lowe, A.R., Siegel, J.J., Kalab, P., Siu, M., Weis, K., Liphardt, J.T., 2010. Selectivity Mechanism of the Nuclear Pore Complex Characterized by Single Cargo Tracking. *Nature* 467, 600–603. <https://doi.org/10.1038/nature09285>.
108. Lubeck, M.D., Palese, P., Schulman, J.L., 1979. Nonrandom association of parental genes in influenza A virus recombinants. *Virology* 95, 269–274. [https://doi.org/10.1016/0042-6822\(79\)90430-6](https://doi.org/10.1016/0042-6822(79)90430-6).

109. Luo, W., Zhang, J., Liang, L., Wang, G., Li, Q., Zhu, P., Zhou, Y., Li, J., Zhao, Y., Sun, N., Huang, S., Zhou, C., Chang, Y., Cui, P., Chen, P., Jiang, Y., Deng, G., Bu, Z., Li, C., Jiang, L., Chen, H., 2018. Phospholipid scramblase 1 interacts with influenza A virus NP, impairing its nuclear import and thereby suppressing virus replication. *PLoS Pathog.* 14. <https://doi.org/10.1371/journal.ppat.1006851>.
110. Lyons, D.M., Llaure, A.S., 2018. Mutation and Epistasis in Influenza Virus Evolution. *Viruses* 10. <https://doi.org/10.3390/v10080407>.
111. Ma, J., Liu, K., Xue, C., Zhou, J., Xu, S., Ren, Y., Zheng, J., Cao, Y., 2013. Impact of the segment-specific region of the 3'-untranslated region of the influenza A virus PB1 segment on protein expression. *Virus Genes* 47, 429–438. <https://doi.org/10.1007/s11262-013-0969-0>.
112. Maeda, Y., Goto, H., Horimoto, T., Takada, A., Kawaoka, Y., 2004. Biological significance of the U residue at the –3 position of the mRNA sequences of influenza A viral segments PB1 and NA. *Virus Res.* 100, 153–157. <https://doi.org/10.1016/j.virusres.2003.11.013>.
113. Majarian, T.D., Murphy, R.F., Lakdawala, S.S., 2019. Learning the sequence of influenza A genome assembly during viral replication using point process models and fluorescence in situ hybridization. *PLoS Comput. Biol.* 15. <https://doi.org/10.1371/journal.pcbi.1006199>.
114. Marklund, J.K., Ye, Q., Dong, J., Tao, Y.J., Krug, R.M., 2012. Sequence in the Influenza A Virus Nucleoprotein Required for Viral Polymerase Binding and RNA Synthesis. *J. Virol.* 86, 7292–7297. <https://doi.org/10.1128/JVI.00014-12>.
115. Marsh, G.A., Hatami, R., Palese, P., 2007. Specific Residues of the Influenza A Virus Hemagglutinin Viral RNA Are Important for Efficient Packaging into Budding Virions. *J. Virol.* 81, 9727–9736. <https://doi.org/10.1128/JVI.01144-07>.
116. Martin, B.E., Harris, J.D., Sun, J., Koelle, K., Brooke, C.B., 2020. Cellular co-infection can modulate the efficiency of influenza A virus production and shape the interferon response. *PLoS Pathog.* 16. <https://doi.org/10.1371/journal.ppat.1008974>.
117. Martínez-Alonso, M., Hengrung, N., Fodor, E., 2016. RNA-Free and Ribonucleoprotein-Associated Influenza Virus Polymerases Directly Bind the Serine-5-Phosphorylated Carboxyl-Terminal Domain of Host RNA Polymerase II. *J. Virol.* 90, 6014–6021. <https://doi.org/10.1128/JVI.00494-16>.

118. Matrosovich, M., Tuzikov, A., Bovin, N., Gambaryan, A., Klimov, A., Castrucci, M.R., Donatelli, I., Kawaoka, Y., 2000. Early Alterations of the Receptor-Binding Properties of H1, H2, and H3 Avian Influenza Virus Hemagglutinins after Their Introduction into Mammals. *J. Virol.* 74, 8502–8512.
119. Mayer, D., Molawi, K., Martínez-Sobrido, L., Ghanem, A., Thomas, S., Baginsky, S., Grossmann, J., García-Sastre, A., Schwemmle, M., 2007. Identification of Cellular Interaction Partners of the Influenza Virus Ribonucleoprotein Complex and Polymerase Complex Using Proteomic-Based Approaches. *J. Proteome Res.* 6, 672–682. <https://doi.org/10.1021/pr060432u>.
120. McAuley, J.L., Gilbertson, B.P., Trifkovic, S., Brown, L.E., McKimm-Breschkin, J.L., 2019. Influenza Virus Neuraminidase Structure and Functions. *Front. Microbiol.* 10. <https://doi.org/10.3389/fmicb.2019.00039>.
121. McDonald, S.M., Nelson, M.I., Turner, P.E., Patton, J.T., 2016. Reassortment in segmented RNA viruses: mechanisms and outcomes. *Nat. Rev. Microbiol.* 14, 448. <https://doi.org/10.1038/nrmicro.2016.46>.
122. Minakuchi, M., Sugiyama, K., Kato, Y., Naito, T., Okuwaki, M., Kawaguchi, A., Nagata, K., 2017. Pre-mRNA Processing Factor Prp18 Is a Stimulatory Factor of Influenza Virus RNA Synthesis and Possesses Nucleoprotein Chaperone Activity. *J. Virol.* 91, e01398-16. <https://doi.org/10.1128/JVI.01398-16>.
123. Miyamoto, S., Muramoto, Y., Shindo, K., Fujita-Fujiharu, Y., Morikawa, T., Tamura, R., Gilmore, J.L., Nakano, M., Noda, T., 2022. Contribution of RNA-RNA Interactions Mediated by the Genome Packaging Signals for the Selective Genome Packaging of Influenza A Virus. *J. Virol.* <https://doi.org/10.1128/jvi.01641-21>.
124. Momose, F., Basler, C.F., O'Neill, R.E., Iwamatsu, A., Palese, P., Nagata, K., 2001. Cellular Splicing Factor RAF-2p48/NPI-5/BAT1/UAP56 Interacts with the Influenza Virus Nucleoprotein and Enhances Viral RNA Synthesis. *J. Virol.* 75, 1899–1908. <https://doi.org/10.1128/JVI.75.4.1899-1908.2001>.
125. Momose, F., Sekimoto, T., Ohkura, T., Jo, S., Kawaguchi, A., Nagata, K., Morikawa, Y., 2011. Apical Transport of Influenza A Virus Ribonucleoprotein Requires Rab11-positive Recycling Endosome. *PLoS ONE* 6. <https://doi.org/10.1371/journal.pone.0021123>.

126. Mondal, A., Potts, G.K., Dawson, A.R., Coon, J.J., Mehle, A., 2015. Phosphorylation at the Homotypic Interface Regulates Nucleoprotein Oligomerization and Assembly of the Influenza Virus Replication Machinery. *PLOS Pathog.* 11, e1004826. <https://doi.org/10.1371/journal.ppat.1004826>.
127. Moreira, É.A., Weber, A., Bolte, H., Kolesnikova, L., Giese, S., Lakdawala, S., Beer, M., Zimmer, G., García-Sastre, A., Schwemmle, M., Juozapaitis, M., 2016. A conserved influenza A virus nucleoprotein code controls specific viral genome packaging. *Nat. Commun.* 7. <https://doi.org/10.1038/ncomms12861>.
128. Morris, A.K., Wang, Z., Ivey, A.L., Xie, Y., Hill, P.S., Schey, K.L., Ren, Y., 2020. Cellular mRNA export factor UAP56 recognizes nucleic acid binding site of influenza virus NP protein. *Biochem. Biophys. Res. Commun.* <https://doi.org/10.1016/j.bbrc.2020.02.059>.
129. Moss, W.N., Priore, S.F., Turner, D.H., 2011. Identification of potential conserved RNA secondary structure throughout influenza A coding regions. *RNA* 17, 991–1011. <https://doi.org/10.1261/rna.2619511>.
130. Moulès, V., Terrier, O., Yver, M., Riteau, B., Moriscot, C., Ferraris, O., Julien, T., Giudice, E., Rolland, J.-P., Erny, A., Bouscambert-Duchamp, M., Frobert, E., Rosa-Calatrava, M., Pu Lin, Y., Hay, A., Thomas, D., Schoehn, G., Lina, B., 2011. Importance of viral genomic composition in modulating glycoprotein content on the surface of influenza virus particles. *Virology* 414, 51–62. <https://doi.org/10.1016/j.virol.2011.03.011>.
131. Muramoto, Y., Takada, A., Fujii, K., Noda, T., Iwatsuki-Horimoto, K., Watanabe, S., Horimoto, T., Kida, H., Kawaoka, Y., 2006. Hierarchy among Viral RNA (vRNA) Segments in Their Role in vRNA Incorporation into Influenza A Virions. *J. Virol.* 80, 2318–2325. <https://doi.org/10.1128/JVI.80.5.2318-2325.2006>.
132. Myers, J.L., Wetzell, K.S., Linderman, S.L., Li, Y., Sullivan, C.B., Hensley, S.E., 2013. Compensatory Hemagglutinin Mutations Alter Antigenic Properties of Influenza Viruses. *J. Virol.* 87, 11168–11172. <https://doi.org/10.1128/JVI.01414-13>.
133. Nakatsu, S., Sagara, H., Sakai-Tagawa, Y., Sugaya, N., Noda, T., Kawaoka, Y., 2016. Complete and Incomplete Genome Packaging of Influenza A and B Viruses. *mBio* 7, e01248-16. <https://doi.org/10.1128/mBio.01248-16>.

134. Neumann, G., Hughes, M.T., Kawaoka, Y., 2000. Influenza A virus NS2 protein mediates vRNP nuclear export through NES-independent interaction with hCRM1. *EMBO J.* 19, 6751–6758. <https://doi.org/10.1093/emboj/19.24.6751>.
135. Neverov, A.D., Lezhnina, K.V., Kondrashov, A.S., Bazykin, G.A., 2014. Intr subtype Reassortments Cause Adaptive Amino Acid Replacements in H3N2 Influenza Genes. *PLoS Genet.* 10. <https://doi.org/10.1371/journal.pgen.1004037>.
136. Newcomb, L.L., Kuo, R.-L., Ye, Q., Jiang, Y., Tao, Y.J., Krug, R.M., 2009. Interaction of the Influenza A Virus Nucleocapsid Protein with the Viral RNA Polymerase Potentiates Unprimed Viral RNA Replication. *J. Virol.* 83, 29–36. <https://doi.org/10.1128/JVI.02293-07>.
137. Ng, A.K.-L., Zhang, H., Tan, K., Li, Z., Liu, J., Chan, P.K.-S., Li, S.-M., Chan, W.-Y., Au, S.W.-N., Joachimiak, A., Walz, T., Wang, J.-H., Shaw, P.-C., 2008. Structure of the influenza virus A H5N1 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. *FASEB J.* 22, 3638–3647. <https://doi.org/10.1096/fj.08-112110>.
138. Nobusawa, E., Sato, K., 2006. Comparison of the Mutation Rates of Human Influenza A and B Viruses. *J. Virol.* 80, 3675–3678. <https://doi.org/10.1128/JVI.80.7.3675-3678.2006>.
139. Noda, T., Murakami, S., Nakatsu, S., Imai, H., Muramoto, Y., Shindo, K., Sagara, H., Kawaoka, Y., 2018. Importance of the 1+7 configuration of ribonucleoprotein complexes for influenza A virus genome packaging. *Nat. Commun.* 9, 54. <https://doi.org/10.1038/s41467-017-02517-w>.
140. Noda, T., Sagara, H., Yen, A., Takada, A., Kida, H., Cheng, R.H., Kawaoka, Y., 2006. Architecture of ribonucleoprotein complexes in influenza A virus particles. *Nature* 439, 490–492. <https://doi.org/10.1038/nature04378>.
141. Noda, T., Sugita, Y., Aoyama, K., Hirase, A., Kawakami, E., Miyazawa, A., Sagara, H., Kawaoka, Y., 2012. Three-dimensional analysis of ribonucleoprotein complexes in influenza A virus. *Nat. Commun.* 3, 639. <https://doi.org/10.1038/ncomms1647>.
142. Nshogozabahizi, J.C., Dench, J., Aris-Brosou, S., 2017. Widespread Historical Contingency in Influenza Viruses. *Genetics* 205, 409–420. <https://doi.org/10.1534/genetics.116.193979>.
143. Pappas, C., Viswanathan, K., Chandrasekaran, A., Raman, R., Katz, J.M., Sasisekharan, R., Tumpey, T.M., 2010. Receptor Specificity and Transmission of H2N2 Subtype Viruses Isolated from the Pandemic of 1957. *PLoS ONE* 5. <https://doi.org/10.1371/journal.pone.0011158>.

144. Parvin, J.D., Moscona, A., Pan, W.T., Leider, J.M., Palese, P., 1986. Measurement of the mutation rates of animal viruses: influenza A virus and poliovirus type 1. *J. Virol.* 59, 377–383.
145. Past Seasons Vaccine Effectiveness Estimates | CDC [WWW Document], 2022. URL <https://www.cdc.gov/flu/vaccines-work/past-seasons-estimates.html> (accessed 3.28.22).
146. Paterson, D., Fodor, E., 2012. Emerging Roles for the Influenza A Virus Nuclear Export Protein (NEP). *PLoS Pathog.* 8. <https://doi.org/10.1371/journal.ppat.1003019>.
147. Pauly, M.D., Procaro, M.C., Luring, A.S., 2017. A novel twelve class fluctuation test reveals higher than expected mutation rates for influenza A viruses. *eLife* 6, e26437. <https://doi.org/10.7554/eLife.26437>.
148. Pereira, C.F., Read, E.K.C., Wise, H.M., Amorim, M.J., Digard, P., 2017. Influenza A Virus NS1 Protein Promotes Efficient Nuclear Export of Unspliced Viral M1 mRNA. *J. Virol.* 91. <https://doi.org/10.1128/JVI.00528-17>.
149. Perez, J.T., Varble, A., Sachidanandam, R., Zlatev, I., Manoharan, M., García-Sastre, A., tenOever, B.R., 2010. Influenza A virus-generated small RNAs regulate the switch from transcription to replication. *Proc. Natl. Acad. Sci. U. S. A.* 107, 11525–11530. <https://doi.org/10.1073/pnas.1001984107>.
150. Perez, J.T., Zlatev, I., Aggarwal, S., Subramanian, S., Sachidanandam, R., Kim, B., Manoharan, M., tenOever, B.R., 2012. A Small-RNA Enhancer of Viral Polymerase Activity. *J. Virol.* 86, 13475–13485. <https://doi.org/10.1128/JVI.02295-12>.
151. Phan, T., Fay, E.J., Lee, Z., Aron, S., Hu, W.-S., Langlois, R.A., 2021. Segment-Specific Kinetics of mRNA, cRNA, and vRNA Accumulation during Influenza Virus Infection. *J. Virol.* 95. <https://doi.org/10.1128/JVI.02102-20>.
152. Phipps, K.L., Ganti, K., Jacobs, N.T., Lee, C.-Y., Carnaccini, S., White, M.C., Manandhar, M., Pickett, B.E., Tan, G.S., Ferreri, L.M., Perez, D.R., Lowen, A.C., 2020. Collective interactions augment influenza A virus replication in a host-dependent manner. *Nat. Microbiol.* 5, 1158–1169. <https://doi.org/10.1038/s41564-020-0749-2>.
153. Poon, L.L.M., Pritlove, D.C., Fodor, E., Brownlee, G.G., 1999. Direct Evidence that the Poly(A) Tail of Influenza A Virus mRNA Is Synthesized by Reiterative Copying of a U Track in the Virion RNA Template. *J. Virol.* 73, 3473–3476.

154. Prachanronarong, K.L., Canale, A.S., Liu, P., Somasundaran, M., Hou, S., Poh, Y.-P., Han, T., Zhu, Q., Renzette, N., Zeldovich, K.B., Kowalik, T.F., Kurt-Yilmaz, N., Jensen, J.D., Bolon, D.N.A., Marasco, W.A., Finberg, R.W., Schiffer, C.A., Wang, J.P., 2019. Mutations in Influenza A Virus Neuraminidase and Hemagglutinin Confer Resistance against a Broadly Neutralizing Hemagglutinin Stem Antibody. *J. Virol.* 93, e01639-18. <https://doi.org/10.1128/JVI.01639-18>.
155. Rabadan, R., Levine, A.J., Krasnitz, M., 2008. Non-random reassortment in human influenza A viruses. *Influenza Other Respir. Viruses* 2, 9–22. <https://doi.org/10.1111/j.1750-2659.2007.00030.x>.
156. Read, E.K.C., Digard, P., 2010. Individual influenza A virus mRNAs show differential dependence on cellular NXF1/TAP for their nuclear export. *J. Gen. Virol.* 91, 1290–1301. <https://doi.org/10.1099/vir.0.018564-0>.
157. Robb, N.C., Chase, G., Bier, K., Vreede, F.T., Shaw, P.-C., Naffakh, N., Schwemmle, M., Fodor, E., 2011. The Influenza A Virus NS1 Protein Interacts with the Nucleoprotein of Viral Ribonucleoprotein Complexes. *J. Virol.* 85, 5228–5231. <https://doi.org/10.1128/JVI.02562-10>.
158. Robb, N.C., Smith, M., Vreede, F. T., Fodor, E. 2009. NS2/NEP protein regulates transcription and replication of the influenza virus genome. *J. Gen. Virol.* 90(6), 1398-1407. <https://doi.org/10.1099/vir.0.009639-0>.
159. Rossman, J.S., Jing, X., Leser, G.P., Lamb, R.A., 2010. Influenza Virus M2 Protein Mediates ESCRT-Independent Membrane Scission. *Cell* 142, 902–913. <https://doi.org/10.1016/j.cell.2010.08.029>.
160. Rutkowski, A.J., Dölken, L., 2017. High-Resolution Gene Expression Profiling of RNA Synthesis, Processing, and Decay by Metabolic Labeling of Newly Transcribed RNA Using 4-Thiouridine, in: *Eukaryotic Transcriptional and Post-Transcriptional Gene Expression Regulation, Methods in Molecular Biology*. Humana Press, New York, NY, pp. 129–140. https://doi.org/10.1007/978-1-4939-6518-2_10.
161. Schelker, M., Mair, C.M., Jolmes, F., Welke, R.-W., Klipp, E., Herrmann, A., Flöttmann, M., Sieben, C., 2016. Viral RNA Degradation and Diffusion Act as a Bottleneck for the Influenza A Virus Infection Efficiency. *PLoS Comput. Biol.* 12. <https://doi.org/10.1371/journal.pcbi.1005075>.
162. Schroeder, C., Heider, H., Möncke-Buchner, E., Lin, T.-I., 2005. The influenza virus ion channel and maturation cofactor M2 is a cholesterol-binding protein. *Eur. Biophys. J.* 34, 52–66. <https://doi.org/10.1007/s00249-004-0424-1>.

163. Shafiuddin, M., Boon, A.C.M., 2019. RNA Sequence Features Are at the Core of Influenza A Virus Genome Packaging. *J. Mol. Biol.* <https://doi.org/10.1016/j.jmb.2019.03.018>.
164. Shapiro, G.I., Gurney, T., Krug, R.M., 1987. Influenza virus gene expression: control mechanisms at early and late times of infection and nuclear-cytoplasmic transport of virus-specific RNAs. *J. Virol.* 61, 764–773.
165. Shi, Y., Wu, Y., Zhang, W., Qi, J., Gao, G.F., 2014. Enabling the “host jump”: structural determinants of receptor-binding specificity in influenza A viruses. *Nat. Rev. Microbiol.* 12, 822–831. <https://doi.org/10.1038/nrmicro3362>.
166. Shimizu, T., Takizawa, N., Watanabe, K., Nagata, K., Kobayashi, N., 2011. Crucial role of the influenza virus NS2 (NEP) C-terminal domain in M1 binding and nuclear export of vRNP. *FEBS Lett.* 585, 41–46. <https://doi.org/10.1016/j.febslet.2010.11.017>.
167. Shu, L.L., Bean, W.J., Webster, R.G., 1993. Analysis of the evolution and variation of the human influenza A virus nucleoprotein gene from 1933 to 1990. *J. Virol.* 67, 2723–2729.
168. Sicard, A., Yvon, M., Timchenko, T., Gronenborn, B., Michalakakis, Y., Gutierrez, S., Blanc, S., 2013. Gene copy number is differentially regulated in a multipartite virus. *Nat. Commun.* 4, 2248. <https://doi.org/10.1038/ncomms3248>.
169. Simon, L.M., Morandi, E., Luganini, A., Gribaudo, G., Martinez-Sobrido, L., Turner, D.H., Oliviero, S., Incarnato, D., 2019. In vivo analysis of influenza A mRNA secondary structures identifies critical regulatory motifs. *Nucleic Acids Res.* 47, 7003–7017. <https://doi.org/10.1093/nar/gkz318>.
170. Smola, M.J., Rice, G.M., Busan, S., Siegfried, N.A., Weeks, K.M., 2015. Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile, and accurate RNA structure analysis. *Nat. Protoc.* 10, 1643–1669. <https://doi.org/10.1038/nprot.2015.103>.
171. Soszynska-Jozwiak, M., Michalak, P., Moss, W.N., Kierzek, R., Kierzek, E., 2015. A Conserved Secondary Structural Element in the Coding Region of the Influenza A Virus Nucleoprotein (NP) mRNA Is Important for the Regulation of Viral Proliferation. *PLOS ONE* 10, e0141132. <https://doi.org/10.1371/journal.pone.0141132>.

172. Suárez, P., Valcárcel, J., Ortín, J., 1992. Heterogeneity of the mutation rates of influenza A viruses: isolation of mutator mutants. *J. Virol.* 66, 2491–2494.
173. Suárez-López, P., Ortín, J. An estimation of the nucleotide substitution rate at defined positions in the influenza virus haemagglutinin gene. *J. Gen. Virol.* 75(2), 389-393. <https://doi.org/10.1099/0022-1317-75-2-389>.
174. Sun, J., Brooke, C.B., 2018. Influenza A Virus Superinfection Potential Is Regulated by Viral Genomic Heterogeneity. *mBio* 9. <https://doi.org/10.1128/mBio.01761-18>.
175. Sun, J., Vera, J.C., Drnevich, J., Lin, Y.T., Ke, R., Brooke, C.B., 2020. Single cell heterogeneity in influenza A virus gene expression shapes the innate antiviral response to infection. *PLoS Pathog.* 16. <https://doi.org/10.1371/journal.ppat.1008671>.
176. Świerczyńska, M., Mirowska-Guzel, D.M., Pindelska, E., 2022. Antiviral Drugs in Influenza. *Int. J. Environ. Res. Public. Health* 19. <https://doi.org/10.3390/ijerph19053018>.
177. Takeda, M., Leser, G.P., Russell, C.J., Lamb, R.A., 2003. Influenza virus hemagglutinin concentrates in lipid raft microdomains for efficient viral fusion. *Proc. Natl. Acad. Sci. U. S. A.* 100, 14610–14617. <https://doi.org/10.1073/pnas.2235620100>.
178. Tarus, B., Bakowicz, O., Chenavas, S., Duchemin, L., Estrozi, L.F., Bourdieu, C., Lejal, N., Bernard, J., Moudjou, M., Chevalier, C., Delmas, B., Ruigrok, R.W.H., Di Primo, C., Slama-Schwok, A., 2012. Oligomerization paths of the nucleoprotein of influenza A virus. *Biochimie* 94, 776–785. <https://doi.org/10.1016/j.biochi.2011.11.009>.
179. Tarus, Bogdan, Chevalier, C., Richard, C.-A., Delmas, B., Di Primo, C., Slama-Schwok, A., 2012. Molecular Dynamics Studies of the Nucleoprotein of Influenza A Virus: Role of the Protein Flexibility in RNA Binding. *PLoS ONE* 7. <https://doi.org/10.1371/journal.pone.0030038>.
180. Taubenberger, J.K., Kash, J.C., 2010. Influenza Virus Evolution, Host Adaptation and Pandemic Formation. *Cell Host Microbe* 7, 440–451. <https://doi.org/10.1016/j.chom.2010.05.009>.
181. te Velthuis, A.J.W., Long, J.C., Bauer, D.L.V., Fan, R.L.Y., Yen, H.-L., Sharps, J., Siegers, J.Y., Killip, M.J., French, H., Oliva-Martín, M.J., Randall, R.E., de Wit, E., van Riel, D., Poon, L.L.M., Fodor, E., 2018. Mini viral RNAs act as innate immune agonists during influenza virus infection. *Nat. Microbiol.* 3, 1234–1242. <https://doi.org/10.1038/s41564-018-0240-5>.

182. Tome-Amat, J., Ramos, I., Amanor, F., Fernández-Sesma, A., Ashour, J., 2019. Influenza A Virus Utilizes Low-Affinity, High-Avidity Interactions with the Nuclear Import Machinery To Ensure Infection and Immune Evasion. *J. Virol.* 93, e01046-18. <https://doi.org/10.1128/JVI.01046-18>.
183. Tumpey, T.M., Maines, T.R., Hoeven, N.V., Glaser, L., Solórzano, A., Pappas, C., Cox, N.J., Swayne, D.E., Palese, P., Katz, J.M., García-Sastre, A., 2007. A Two-Amino Acid Change in the Hemagglutinin of the 1918 Influenza Virus Abolishes Transmission. *Science*. <https://doi.org/10.1126/science.1136212>.
184. Turrell, L., Hutchinson, E.C., Vreede, F.T., Fodor, E., 2015. Regulation of Influenza A Virus Nucleoprotein Oligomerization by Phosphorylation. *J. Virol.* 89, 1452–1455. <https://doi.org/10.1128/JVI.02332-14>.
185. Turrell, L., Lyall, J.W., Tiley, L.S., Fodor, E., Vreede, F.T., 2013. The role and assembly mechanism of nucleoprotein in influenza A virus ribonucleoprotein complexes. *Nat. Commun.* 4, 1591. <https://doi.org/10.1038/ncomms2589>.
186. Vahey, M.D., Fletcher, D.A., 2019. Influenza A virus surface proteins are organized to help penetrate host mucus. *eLife* 8, e43764. <https://doi.org/10.7554/eLife.43764>.
187. Veit, M., Thaa, B., 2011. Association of Influenza Virus Proteins with Membrane Rafts. *Adv. Virol.* 2011. <https://doi.org/10.1155/2011/370606>.
188. Vreede, F.T., Brownlee, G.G., 2007. Influenza Virion-Derived Viral Ribonucleoproteins Synthesize both mRNA and cRNA In Vitro. *J. Virol.* 81, 2196–2204. <https://doi.org/10.1128/JVI.02187-06>.
189. Vreede, F.T., Jung, T.E., Brownlee, G.G., 2004. Model Suggesting that Replication of Influenza Virus Is Regulated by Stabilization of Replicative Intermediates. *J. Virol.* 78, 9568–9572. <https://doi.org/10.1128/JVI.78.17.9568-9572.2004>.
190. Vreede, F.T., Ng, A.K.-L., Shaw, P.-C., Fodor, E., 2011. Stabilization of Influenza Virus Replication Intermediates Is Dependent on the RNA-Binding but Not the Homo-Oligomerization Activity of the Viral Nucleoprotein ν . *J. Virol.* 85, 12073–12078. <https://doi.org/10.1128/JVI.00695-11>.
191. Wagner, R., Herwig, A., Azzouz, N., Klenk, H.D., 2005. Acylation-Mediated Membrane Anchoring of Avian Influenza Virus Hemagglutinin Is Essential for Fusion Pore Formation and Virus Infectivity. *J. Virol.* 79, 6449–6458. <https://doi.org/10.1128/JVI.79.10.6449-6458.2005>

192. Wagner, R., Matrosovich, M., Klenk, H.-D., 2002. Functional balance between haemagglutinin and neuraminidase in influenza virus infections. *Rev. Med. Virol.* 12, 159–166. <https://doi.org/10.1002/rmv.352>.
193. Wang, J., Li, J., Zhao, L., Cao, M., Deng, T., 2016. Dual Roles of the Hemagglutinin Segment-Specific Noncoding Nucleotides in the Extended Duplex Region of the Influenza A Virus RNA Promoter. *J. Virol.* 91. <https://doi.org/10.1128/JVI.01931-16>.
194. Wang, J., Peng, Y., Zhao, L., Cao, M., Hung, T., Deng, T. 2015. Influenza A virus utilizes a suboptimal Kozak sequence to fine-tune virus replication and host response. *J. Gen. Virol.* 96(4), 756-766. <https://doi.org/10.1099/vir.0.000030>.
195. Wang, L., Lee, C.-W., 2009. Sequencing and mutational analysis of the non-coding regions of influenza A virus. *Vet. Microbiol.* 135, 239–247. <https://doi.org/10.1016/j.vetmic.2008.09.067>.
196. Wang, L., Wu, A., Wang, Y.E., Quanquin, N., Li, C., Wang, J., Chen, H.-W., Liu, S., Liu, P., Zhang, H., Qin, F.X.-F., Jiang, T., Cheng, G., 2016. Functional Genomics Reveals Linkers Critical for Influenza Virus Polymerase. *J. Virol.* 90, 2938–2947. <https://doi.org/10.1128/JVI.02400-15>.
197. Wang, P., Palese, P., O'Neill, R.E., 1997. The NPI-1/NPI-3 (karyopherin alpha) binding site on the influenza a virus nucleoprotein NP is a nonconventional nuclear localization signal. *J. Virol.* 71, 1850–1856.
198. Ward, M.J., Lycett, S.J., Avila, D., Bollback, J.P., Leigh Brown, A.J., 2013. Evolutionary interactions between haemagglutinin and neuraminidase in avian influenza. *BMC Evol. Biol.* 13, 222. <https://doi.org/10.1186/1471-2148-13-222>.
199. White, M.C., Lowen, A.C., 2018. Implications of segment mismatch for influenza A virus evolution. *J. Gen. Virol.* 99, 3–16. <https://doi.org/10.1099/jgv.0.000989>.
200. Widjaja, I., de Vries, E., Rottier, P.J.M., de Haan, C.A.M., 2012. Competition between Influenza A Virus Genome Segments. *PLoS ONE* 7. <https://doi.org/10.1371/journal.pone.0047529>.
201. Williams, G.D., Townsend, D., Wylie, K.M., Kim, P.J., Amarasinghe, G.K., Kutluay, S.B., Boon, A.C.M., 2018. Nucleotide resolution mapping of influenza A virus nucleoprotein-RNA interactions reveals RNA features required for replication. *Nat. Commun.* 9. <https://doi.org/10.1038/s41467-018-02886-w>.

202. Wu, N.C., Thompson, A.J., Xie, J., Lin, C.-W., Nycholat, C.M., Zhu, X., Lerner, R.A., Paulson, J.C., Wilson, I.A., 2018. A complex epistatic network limits the mutational reversibility in the influenza hemagglutinin receptor-binding site. *Nat. Commun.* 9. <https://doi.org/10.1038/s41467-018-03663-5>.
203. Wu, N.C., Wilson, I.A., 2019. Influenza Hemagglutinin Structures and Antibody Recognition. *Cold Spring Harb. Perspect. Med.* a038778. <https://doi.org/10.1101/cshperspect.a038778>.
204. Wu, N.C., Xie, J., Zheng, T., Nycholat, C.M., Grande, G., Paulson, J.C., Lerner, R.A., Wilson, I.A., 2017. Diversity of functionally permissive sequences in the receptor-binding site of influenza hemagglutinin. *Cell Host Microbe* 21, 742–753.e8. <https://doi.org/10.1016/j.chom.2017.05.011>.
205. Xiao, Y., Zhang, W., Pan, M., Bauer, D.L.V., Bi, Y., Cao, M., Fodor, E., Deng, T. 2021. Synergistic Effect between 3'-Terminal Noncoding and Adjacent Coding Regions of the Influenza A Virus Hemagglutinin Segment on Template Preference. *J. Virol.* 95. <https://doi.org/10.1128/JVI.00878-21>.
206. Xu, R., Zhu, X., McBride, R., Nycholat, C.M., Yu, W., Paulson, J.C., Wilson, I.A., 2012. Functional Balance of the Hemagglutinin and Neuraminidase Activities Accompanies the Emergence of the 2009 H1N1 Influenza Pandemic. *J. Virol.* 86, 9221–9232. <https://doi.org/10.1128/JVI.00697-12>.
207. Ye, Q., Krug, R.M., Tao, Y.J., 2006. The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature* 444, 1078–1082. <https://doi.org/10.1038/nature05379>.
208. Yen, H.-L., Liang, C.-H., Wu, C.-Y., Forrest, H.L., Ferguson, A., Choy, K.-T., Jones, J., Wong, D.D.-Y., Cheung, P.P.-H., Hsu, C.-H., Li, O.T., Yuen, K.M., Chan, R.W.Y., Poon, L.L.M., Chan, M.C.W., Nicholls, J.M., Krauss, S., Wong, C.-H., Guan, Y., Webster, R.G., Webby, R.J., Peiris, M., 2011. Hemagglutinin–neuraminidase balance confers respiratory-droplet transmissibility of the pandemic H1N1 influenza virus in ferrets. *Proc. Natl. Acad. Sci.* 108, 14264–14269. <https://doi.org/10.1073/pnas.1111000108>.
209. York, A., Hengrung, N., Vreede, F.T., Huiskonen, J.T., Fodor, E., 2013. Isolation and characterization of the positive-sense replicative intermediate of a negative-strand RNA virus. *Proc. Natl. Acad. Sci. U. S. A.* 110, E4238–E4245. <https://doi.org/10.1073/pnas.1315068110>.
210. Yu, M., Liu, X., Cao, S., Zhao, Z., Zhang, K., Xie, Q., Chen, C., Gao, S., Bi, Y., Sun, L., Ye, X., Gao, G.F., Liu, W., 2012. Identification and Characterization of Three Novel Nuclear Export Signals in the Influenza A Virus Nucleoprotein. *J. Virol.* 86, 4970–4980. <https://doi.org/10.1128/JVI.06159-11>.

211. Zhao, L., Peng, Y., Zhou, K., Cao, M., Wang, J., Wang, X., Jiang, T., Deng, T., 2014. New Insights into the Nonconserved Noncoding Region of the Subtype-Determinant Hemagglutinin and Neuraminidase Segments of Influenza A Viruses. *J. Virol.* 88, 11493–11503.
<https://doi.org/10.1128/JVI.01337-14>.
212. Zheng, A., Sun, W., Xiong, X., Freyn, A.W., Peukes, J., Strohmeier, S., Nachbagauer, R., Briggs, J.A.G., Krammer, F., Palese, P., 2020. Enhancing Neuraminidase Immunogenicity of Influenza A Viruses by Rewiring RNA Packaging Signals. *J. Virol.* 94.
<https://doi.org/10.1128/JVI.00742-20>.
213. Zheng, H., Palese, P., García-sastre, A., 1996. Nonconserved Nucleotides at the 3' and 5' Ends of an Influenza A Virus RNA Play an Important Role in Viral RNA Replication. *Virology* 217, 242–251.
<https://doi.org/10.1006/viro.1996.0111>.