Interactions between influenza A virus nucleoprotein and gene segment UTRs facilitate selective modulation of viral gene expression

Short Title: Regulation of influenza virus gene expression

Meghan Diefenbacher[1], Timothy JC Tan[2], David LV Bauer[3], Beth Stadtmueller[4,5], Nicholas C. Wu[2,4,5,6], Christopher B. Brooke[1,6*]

[1]Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States of America

[2]Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States of America

[3]RNA Virus Replication Laboratory, The Francis Crick Institute, London, United Kingdom

[4]Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States of America

[5]Department of Biomedical and Translational Sciences, Carle Illinois College of Medicine, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States of America

[6]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States of America

*Corresponding author: Christopher B. Brooke

E-mail: cbrooke@illinois.edu

Abstract

The influenza A virus (IAV) genome is divided into eight negative-sense, single-stranded RNA segments. Each segment exhibits a unique level and temporal pattern of expression, however the exact mechanisms underlying the patterns of individual gene segment expression are poorly understood. We previously demonstrated that a single substitution in the viral nucleoprotein (NP:F346S) selectively modulates neuraminidase (NA) gene segment expression while leaving other segments largely unaffected. Given what is currently known about NP function, there is no obvious explanation for how changes in NP can selectively modulate the replication of individual gene segments. We found that the specificity of this effect for the NA segment is virus strain specific and depends on the UTR sequences of the NA segment. While the NP:F346S substitution did not significantly alter the RNA binding or oligomerization activities of NP *in vitro*, it specifically decreased the ability of NP to promote NA segment vRNA synthesis. In addition to NP residue F346, we identified two other adjacent aromatic residues in NP (Y385 & F479) capable of similarly regulating NA gene segment expression, suggesting a larger role for this domain in gene-segment specific regulation. Our findings reveal a new role for NP in selective regulation of viral gene segment replication and demonstrate how the expression patterns of individual viral gene segments can be modulated during adaptation to new host environments.

Author summary

Influenza A virus (IAV) is a respiratory pathogen that remains a significant source of morbidity and mortality. Escape from host immunity or emergence into new host species often requires mutations that modulate the functional activities of the viral glycoproteins hemagglutinin (HA) and neuraminidase (NA) which are responsible for virus attachment to and release from host cells, respectively. Maintaining the functional balance between the activities of HA and NA is required for fitness across multiple host systems. Thus, selective modulation of viral gene expression patterns may be a key determinant of viral immune escape and cross-species transmission potential. We identified a novel mechanism by which the viral nucleoprotein (NP) gene can selectively modulate NA segment replication and gene expression through interactions with the segment UTR. Our work highlights an unexpected role for NP in selective regulation of expression from the individual IAV gene segments.

Introduction

Influenza A virus (IAV) is a major respiratory pathogen that causes seasonal epidemics and occasional pandemics that result in substantial morbidity and mortality (1). The genome of IAV is divided into eight negative sense, single-stranded RNA segments that encode one or more viral proteins (2). These negative-sense genomic 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 (2). The individual IAV gene segments vary in both overall expression levels and timing, but the specific factors that govern this variation are poorly understood (3,4).

Each IAV gene segment consists of one or more open reading frames (ORFs) flanked by untranslated regions (UTRs) (2). 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 (2). Previous studies have established roles for segment-specific sequences within the UTRs in modulating gene expression in a segment-specific manner (5–10). 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) (2).

NP is a highly conserved (11) and multi-functional protein. To perform its integral role in vRNP formation, NP has two major known activities: RNA binding and oligomerization. NP binds RNA non-specifically through a positively charged groove located between its head and body domains (12,13). Oligomerization of individual NP protomers occurs through the insertion of a C-terminal tail loop into the receptor groove of the neighboring protomer (12,14,15). 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 (<100nts) can be generated in its absence or in the presence of binding/oligomerization-deficient NP mutants (15). NP facilitates the import and export of vRNPs from the nucleus (16–18). Finally, NP is critically involved in the selective packaging of the viral genome segments – both directly through specific amino acid residues (19,20) and indirectly though determining the accessibility of RNA structures important for packaging (21,22).

We previously identified an NP substitution (NP:F346S) that was sufficient to significantly enhance the replication and transmissibility of the A/Puerto Rico/8/1834 (PR8) strain of IAV in guinea pigs while selectively decreasing the expression of the neuraminidase (NA) gene segment (23,24). This finding suggested (a) that NP plays an unappreciated role in selectively regulating the expression of individual viral genes, and (b) 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.

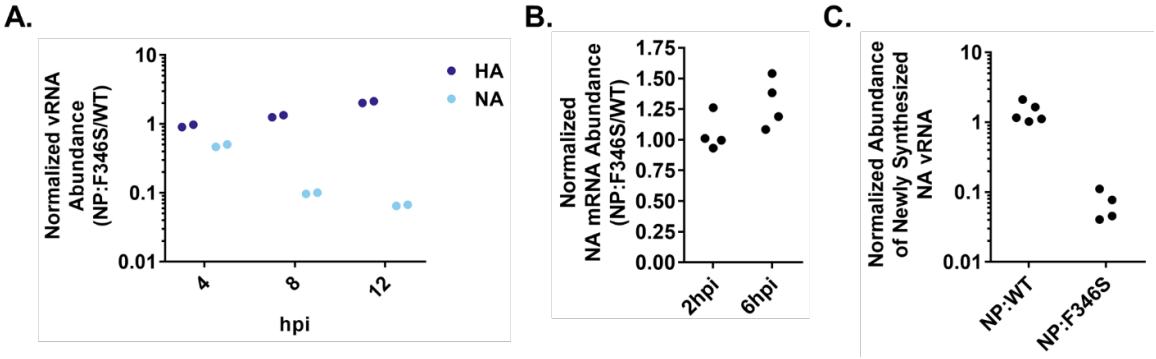
Here, we dissect the mechanism by which specific residues in NP selectively modulate NA segment expression. In addition, we pinpoint the specific determinants within the NA genomic RNA that are required for susceptibility to selective regulation by NP. Altogether, these results illuminate a new mode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

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 hours post-infection (hpi) (24), while leaving HA expression largely unaffected (**Fig 1A**). We previously showed that NP:F346S also decreased NA mRNA abundance, raising the possibility that this substitution directly

119 affected all NA segment-derived RNA species (24). To determine whether the effects of
120 NP:F346S are specific for vRNA, we compared levels of primary NA mRNA synthesis
121 between NP:WT and NP:F346S in the presence of 100µg/mL cycloheximide.
122 Cycloheximide blocks translation of the viral replicase machinery needed for vRNA
123 synthesis, thus only allowing primary transcription of viral mRNAs from incoming vRNPs
124 (25). We observed no differences in the NA mRNA levels between WT and NP:F346S
125 in the presence of cycloheximide, indicating that NP:F346S has no effects on primary
126 mRNA transcription (**Fig 1B**).
127



128

129 ***Fig 1. NP:F346S affects NA vRNA replication but not primary transcription. A.)***
130 *Abundances of NA and HA vRNA (measured by RT-qPCR on total cellular RNA) at the*
131 *indicated timepoints following infection of MDCK cells at an MOI of 0.1 NP-expressing*
132 *units (NPEU)/cell under single cycle conditions. Data represent values obtained during*
133 *infection with PR8-NP:F346S, normalized to values obtained during infection with PR8-*
134 *NP:WT. The data shown are individual cell culture well replicates representative of the*
135 *data obtained through two similar experiments.* ***B.)*** *NA mRNA abundances (measured*
136 *by RT-qPCR on total cellular RNA) in PR8-NP:F346S-infected MDCK-SIAT1 cells,*
137 *normalized to values obtained during infection with PR8-NP:WT. Infections were*
138 *initiated at MOI=5 $TCID_{50}$/cell in the presence of 100µg/mL cycloheximide. Data points*
139 *indicate individual cell culture well replicates pooled from two independent experiments.*
140 ***C.)*** *Abundances of newly synthesized NA vRNA in PR8-NP:F346S and PR8-NP:WT*
141 *infected cells, as measured by 4-thiouridine (4SU) pulse labeling. MDCK cells were*
142 *infected with PR8-NP:WT or PR8-NP:F346S at an MOI of 5 $TCID_{50}$/cell for 7hrs,*
143 *followed by 1hr pulse with 500µM 4SU. Cellular RNA was then harvested and the*
144 *abundance of 4SU-labeled viral RNAs were determined by RT-qPCR using a universal,*
145 *vRNA-sense specific primer for the RT reaction followed by segment-specific primers for*
146 *the qPCR. Data points indicate individual cell culture well replicates pooled from two*
147 *independent experiments.*
148

149 To determine whether the effect of NP:F346S on NA vRNA abundance is due to
150 reduced synthesis, as opposed to a decrease in stability, we pulsed infected cells with
151 4-thiouridine (4SU) for one hour, and measured the amount of vRNA synthesized during
152 the pulse by performing RT-qPCR on 4SU-labeled RNAs using a universal, vRNA-
153 specific primer for the RT reaction followed by segment-specific primers for the qPCR,
154 or a tagged, vRNA and segment-specific primer for the RT reaction followed by a
155 combination of a tag-specific and segment-specific primer for the qPCR step (**Figs 1C**

and S1 Fig.). In both cases, levels of the 4SU-containing newly synthesized NA vRNA were over 10-fold lower during NP:F346S infection compared with NP:WT (**Figs 1C and S1 Fig.**). 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 genome segments, how can substitutions in NP selectively reduce synthesis of the NA RNA while leaving the other segments largely unaffected? 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.

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 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 this chimeric 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 2B**).

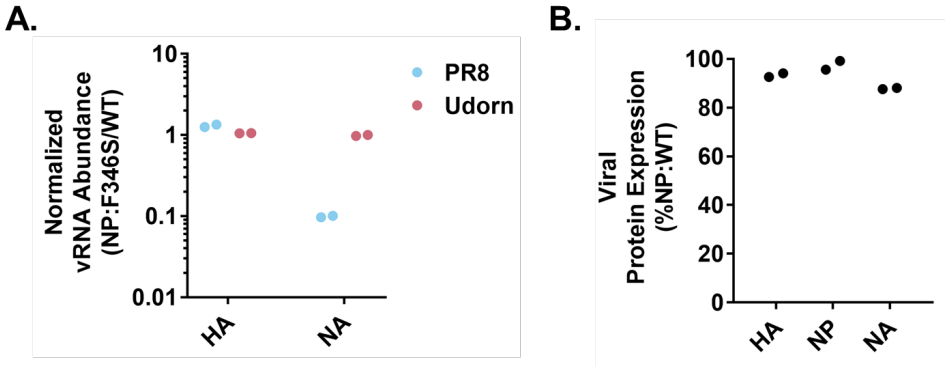


Fig 2. Susceptibility to the effects of NP:F346S is NA segment genotype specific. A.) *Normalized vRNA abundances as determined by qRT-PCR in PR8-NP:F346S or Udorn-NP:F346S infected MDCK cells (MOI=0.1 NPEU/cell, 8hpi) expressed as fraction of PR8 NP:WT or Udorn NP:WT respectively. Secondary infection was blocked via the addition of ammonium chloride at 3hpi. The data points represent individual cell culture well replicates representative of the data obtained through two similar experiments.* ***B.)*** *Viral protein expression levels as determined by geometric mean fluorescence intensity (GMFI) in rPR8 Udorn HA/NA NP:F346S infected MDCK cells (MOI=0.03 $TCID_{50}$/cell, 16hpi) expressed as a percentage of rPR8 Udorn HA/NA NP:WT. The data shown are individual cell culture well replicates representative of the data obtained through two similar experiments.*

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 confer susceptibility to selective modulation by NP:F346S, we first divided the NA segment into three broad functional regions: (a) the portion of the NA ORF that does not overlap known packaging signals, (b) the portions of the NA ORF that do overlap known packaging signals, and (c) 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 (26)) were important for the effects of NP:F346S, we used the Codon Shuffle package (27) to introduce 227 silent substitutions within the region encompassing nucleotides 38-1319 of the NA segment while minimizing effects on codon frequency or di-nucleotide content (**Fig 3A**). The codon shuffled NA segment exhibited a similar decrease in its expression level as NA WT in the presence of NP:F346S, indicating the effects of NP:F346S on NA replication do not require motifs within the non-packaging signal region of the NA ORF (**Fig 3B**).

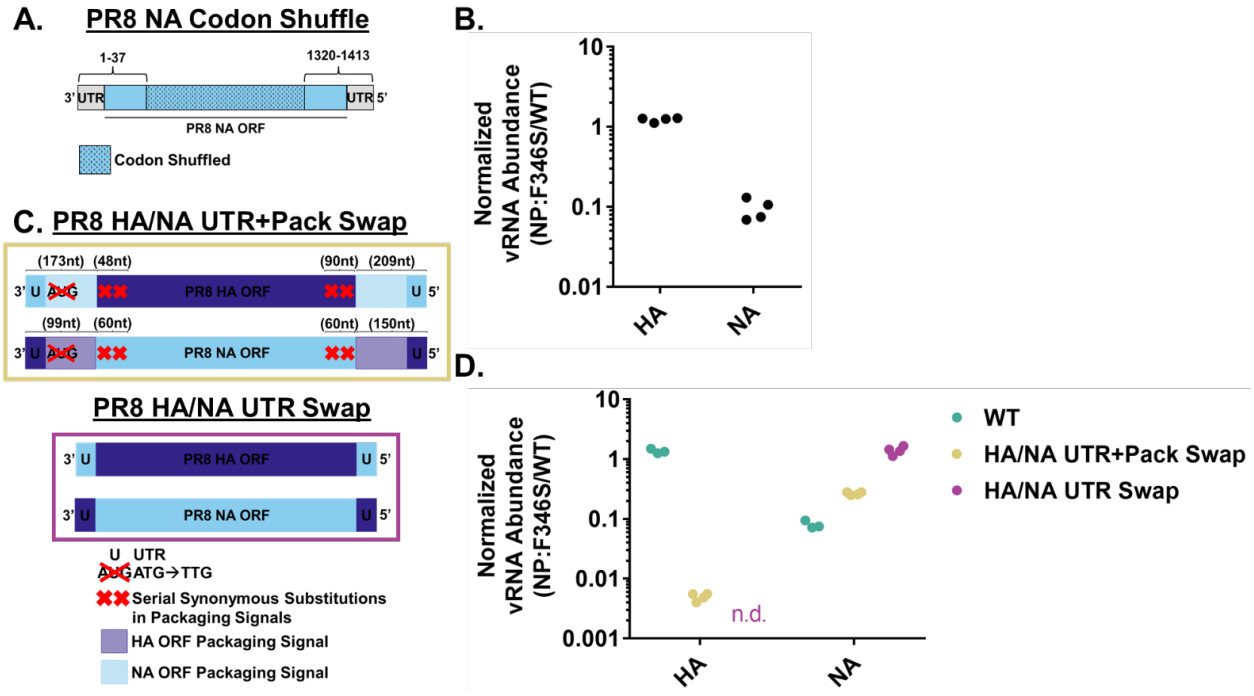


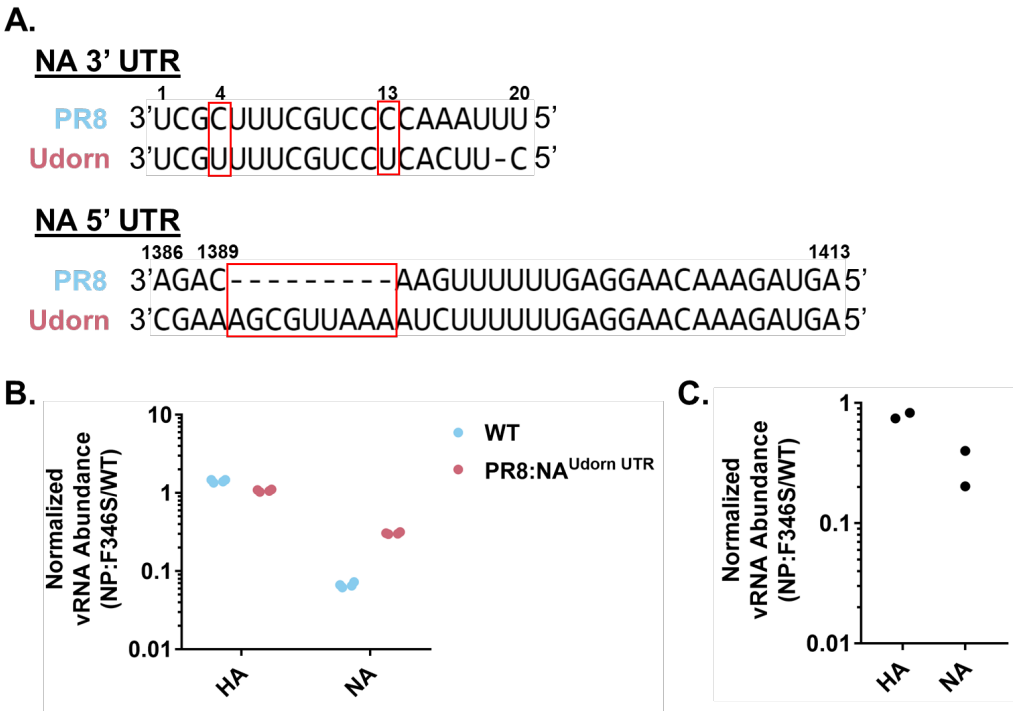
Fig 3. Susceptibility to NP-dependent regulation maps to the UTRs of the NA segment. A.) *Schematic depiction of the codon shuffled PR8 NA construct. The Codon Shuffle program was used to introduce 227 silent mutations within the region encompassing nucleotides 38-1319 of the PR8 NA segment to alter features of the RNA sequence while minimizing changes in codon frequencies or dinucleotide content.* ***B.)*** *Relative 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, 8hpi) as determined by*

RT-qPCR on cellular RNA expressed as a fraction of PR8 NA Codon Shuffle NP:WT respectively. Each data point represents a cell culture well replicate pooled from two separate experiments. ***C.)*** *Schematic depictions of the PR8 HA/NA UTR+Pack Swap and PR8 HA/NA UTR Swap gene segments. The PR8 HA/NA UTR+Pack Swap segments were generated by replacing the UTRs and packaging signal regions of one segment (HA/NA) with those of the other segment (NA/HA). The start codon of the newly appended packaging signal for each segment was mutated to prevent the expression of any protein encoded by the packaging signal sequence. The packaging signals within the native ORFs were disrupted via the addition of silent substitutions to all codons to prevent duplication of the packaging signals in the swapped segments. The PR8 HA/NA UTR Swap gene segments were generated by swapping the UTRs of the PR8 HA/NA segments.* ***D.)*** *Relative 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, 8hpi) 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. Each data point represents a cell culture well replicate pooled from two separate experiments.*

Not surprisingly, attempts to use codon shuffling to mutagenize the regions of the NA ORF that overlap the packaging signals resulted in non-viable viruses. 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 where 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 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 (28). 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 encoding the HA ORF with the NA UTRs exhibited a severe packaging deficiency (**S2A Fig.**), and the abundance of the segment in infected cells was below the limit of detection for the qPCR assay.

We infected MDCK cells with these recombinant viruses and quantified the effects of NP:F346S on intracellular HA and NA vRNA levels (**Fig 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 3D**). Similarly, while WT HA expression is 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 3D**). 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

266 changes in gene segment packaging ratios – likely a result of removing the packaging
267 signals from their native context (**S2A/B Figs.**). Further, we observed that replacing the
268 NA segment UTRs with those from HA completely eliminated the effect of NP:F346S on
269 NA vRNA abundance (**Fig 3D**). Altogether, these data indicate that the selective effects
270 of NP:F346S on vRNA synthesis depend upon the segment UTR sequences.
271
272 **Identification of specific nucleotides within the NA UTR that determine sensitivity**
273 **to NP:F346S**
274 We next sought to identify which specific elements within the NA UTR are required for
275 susceptibility to modulation by NP:F346S. To do this, we took advantage of the high
276 degree of similarity between the NA segment UTRs of PR8 (susceptible to the effects of
277 NP:F346S) and Udorn (resistant to the effects of NP:F346S) (**Figs 2 and 4A**). The 3'
278 UTRs of the PR8 and Udorn NA segments differ in (a) the identity of the nucleotides at
279 positions 4 and 13 (CC/UU for PR8/Udorn respectively) and (b) the sequence directly
280 upstream of the initiating Met codon of the NA ORF (AAAUUU/ACUUC) for PR8/Udorn
281 respectively) (**Fig 4A**). In the 5' NA segment UTR, Udorn has a 9bp insertion relative to
282 the PR8 sequence plus a few additional nucleotide substitutions (**Fig 4A**).
283



284

285 ***Fig 4. The UTRs of the Udorn NA segment confer resistance to regulation by***
286 ***NP:F346S. A.)*** *Alignment of the PR8 NA and Udorn NA 3' & 5' UTRs using the M-Coffee*
287 *alignment algorithm on the T-Coffee web server (29). Regions of interest are boxed in*
288 *red. PR8 NA nucleotide numbering is shown.* ***B.)*** *Relative abundances of the HA and*
289 *NA segments in MDCK cells infected with the PR8 NP:F346S, PR8:NA*$^{Udorn\ UTR}$
290 *NP:F346S (MOI=0.1 NPEU/cell, 8hpi) viruses as determined by RT-qPCR expressed as*
291 *a fraction of PR8 NP:WT and PR8:NA*$^{Udorn\ UTR}$ *NP:WT respectively. Each data point*
292 *represents a cell culture well replicate pooled from two independent experiments.* ***C.)***

Relative abundances of the HA and NA segments in MDCK cells infected with PR8:Udorn HA, NA$^{PR8\ UTR}$ NP:F346S (MOI=0.03 NPEU/cell, 8hpi) virus as determined by RT-qPCR expressed as a fraction of PR8:Udorn HA, NA$^{PR8\ UTR}$ NP:WT. Each data point represents a cell culture well replicate from a single experiment.

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 4A and 5**). The effects of NP:F346S on PR8:NA$^{Udorn\ UTR}$ were reduced compared with 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 4B**). We also attempted to generate viruses where 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 10^4-10^5 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% v. 5-10% of NP:WT respectively), further substantiating the role of the NA UTRs in regulation by NP:F346S (**Fig 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 (**S2C Fig.**).

We next generated a panel of recombinant viruses with chimeric PR8-Udorn NA UTR sequences (**Fig 5**).

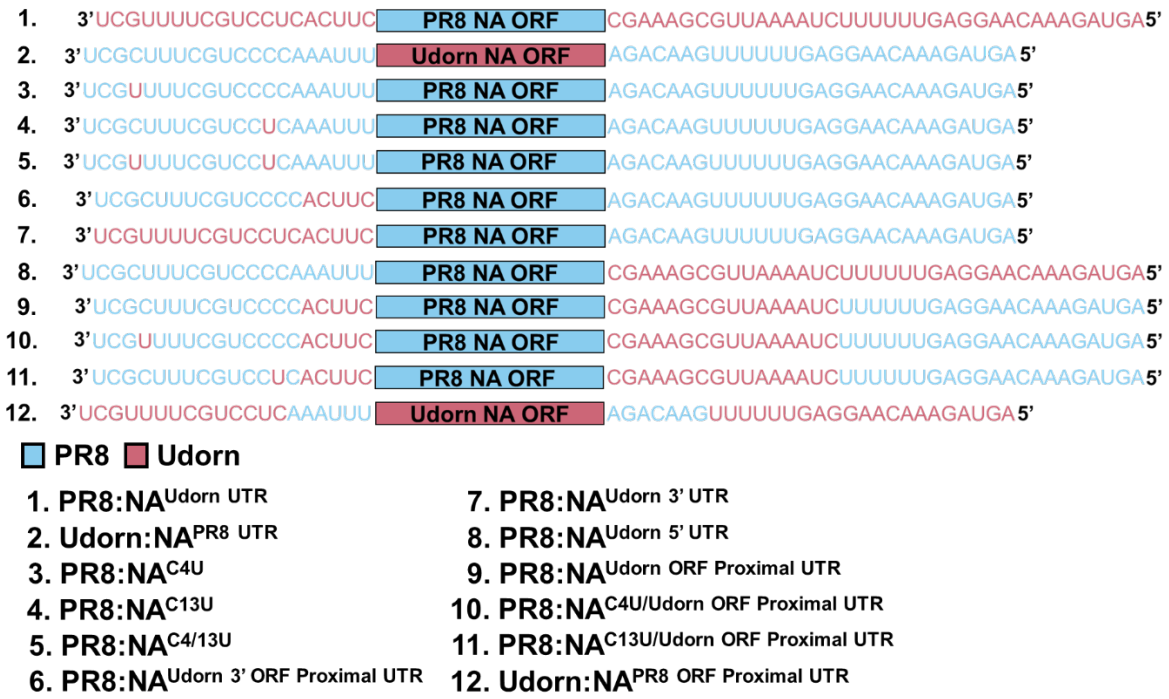


Fig 5. UTR sequences of the PR8-Udorn NA UTR chimeric constructs. *The sequences derived from PR8 and Udorn NA are colored blue and pink, respectively. Sequences shown in negative sense, 3'->5'.*

We first examined both the polymorphisms at positions 4 and 13 of the promoter and 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 6A**). Introducing both C4U and C13U together decreased the relative effect of NP:F346S on NA vRNA levels by roughly 6x (**Fig 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$^{\text{Udorn 3' ORF Proximal UTR}}$) slightly reduced the extent to which NP:F346S decreased NA vRNA levels, to an extent comparable with the individual C4U or C13U mutations (**Fig 6A**). Replacing the entire PR8 NA 3' UTR with that of Udorn (PR8:NA$^{\text{Udorn 3' UTR}}$) (Has U4/13 and Udorn ORF Proximal UTR sequence) decreased the relative effect of NP:F346S to approximately the same level as the dual C4/13U mutations (**Fig 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.

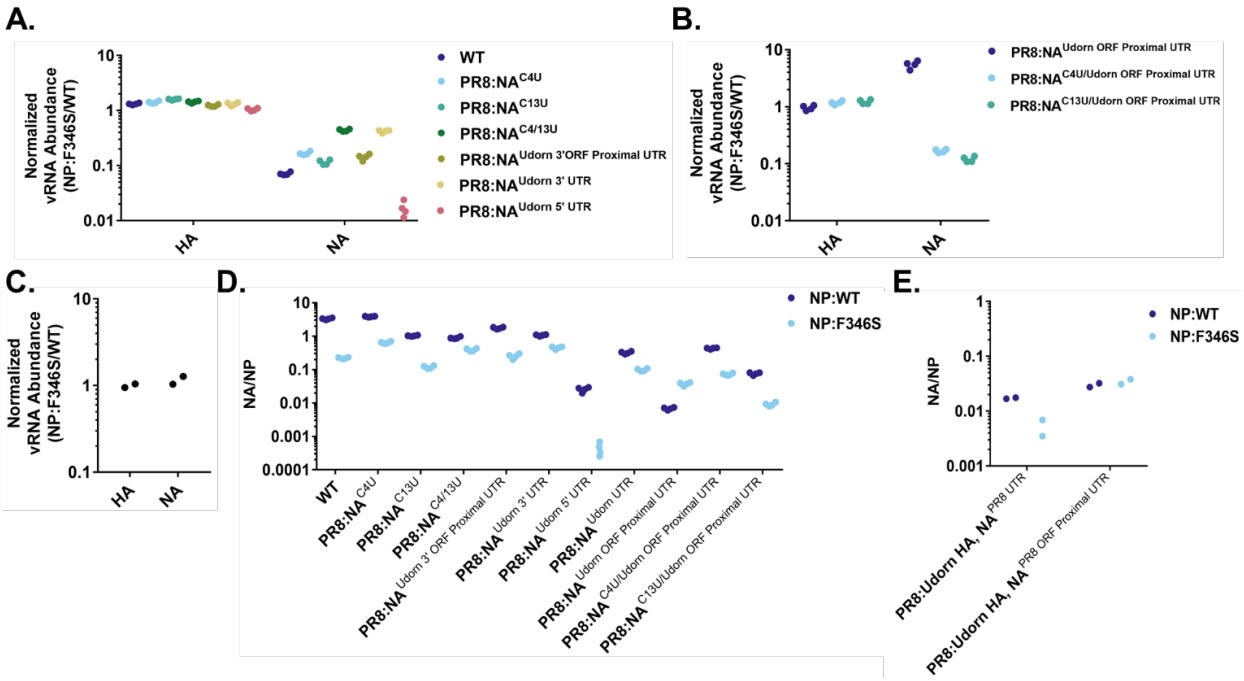
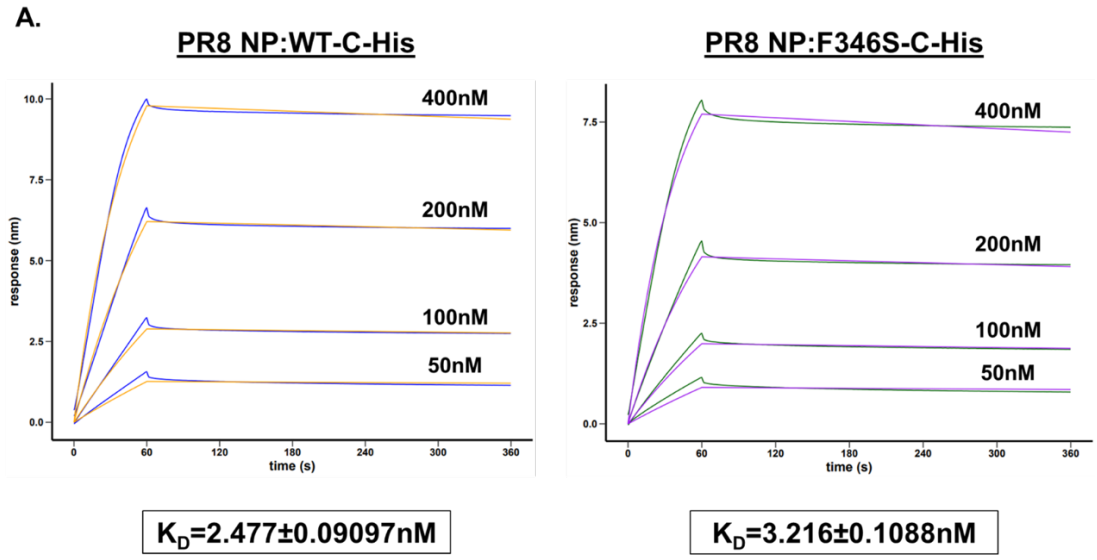


Fig 6. The effect of mutations in the PR8 NA UTRs on baseline expression levels and sensitivity to NP:F346S. A.) *Relative abundances of the HA and NA segments at 8hpi in MDCK cells infected with the indicated viruses encoding NP:F346S at MOI=0.1 NPEU/cell, as determined by qRT-PCR normalized to the NP:WT-encoding versions of the same viruses. Each data point represents an individual cell culture well replicate pooled from two independent experiments.* ***B.)*** *Relative abundances of the HA and NA segments in MDCK cells infected with the indicated viruses encoding NP:F346S (MOI=0.1 NPEU/cell, 8hpi), as determined by qRT-PCR normalized to the NP:WT-encoding versions of the same viruses. Each data point represents an individual cell culture well replicate pooled from two independent experiments.* ***C.)*** *Relative abundances of the HA and NA segments in MDCK cells infected with the PR8:Udorn HA, NA$^{PR8\ ORF\ Proximal\ UTR}$ NP:F346S virus (MOI=0.03 NPEU/cell, 8hpi) as determined by qRT-PCR normalized to PR8:Udorn HA, NA$^{PR8\ ORF\ Proximal\ UTR}$ NP:WT. Each data point represents an individual cell culture well replicate from a single experiment.* ***D,E.)*** *Data from experiments shown in* ***(4B and 6A,B)*** *and* ***(4C/6C)*** *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 8hpi) as determined by qRT-PCR on total cellular RNA. The data represents two cell culture well replicates pooled from either two independent experiments* ***(D)*** *or a single experiment* ***(E)****.*

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 of susceptibility to NP:F346S (**Fig 6A**). Replacing the 5' UTR of the PR8 NA with that of Udorn (PR8:NA$^{Udorn\ 5'\ UTR}$) enhanced susceptibility to NP:F346S by ~4x (**Fig 6A**) likely due to the fact that there was an additional packaging defect for the segment in the presence

371 of NP:F346S (**S2D Fig.**). Interestingly, replacing both the ORF proximal regions of the
372 PR8 NA with those from Udorn (PR8:NA$^{\text{Udorn ORF Proximal UTR}}$) (Has 4/13C and Udorn NA
373 ORF proximal sequences) made the segment resistant to NP:F346S, actually
374 increasing expression ~5-6x relative to NP:WT (**Fig 6B**).
375
376 The only differences between the PR8:NA$^{\text{Udorn UTR}}$ segment, which was partially
377 resistant to the effects of NP:F346S (**Fig 4**), and PR8:NA$^{\text{Udorn ORF Proximal UTR}}$, which was
378 completely resistant, were the identity of nucleotides 4 and 13 of the 3' UTR (**Fig 4**), so
379 we next asked whether these nucleotides were responsible for the resistance
380 phenotype observed for PR8:NA$^{\text{Udorn ORF Proximal UTR}}$). Mutating the C at position 4 or 13 of
381 the PR8:NA$^{\text{Udorn ORF Proximal UTR}}$) segment to U (PR8:NA$^{\text{C4U/Udorn ORF Proximal}}$
382 ^{UTR}/PR8:NA$^{\text{C13U/Udorn ORF Proximal UTR}}$), restored the susceptibility of the segment to NP-
383 dependent regulation, again emphasizing the importance of positions 4 and 13 of the
384 PR8 NA 3' UTR to determining the effects of NP:F346S (**Fig 6B**). Interestingly, we also
385 found that replacing the ORF proximal regions of the Udorn NA UTR with those of PR8
386 (PR8: Udorn HA,NA$^{\text{PR8 ORF Proximal UTR}}$) (Has 4/13U and PR8 NA ORF Proximal
387 Sequences) made the segment resistant to NP:F346S (**Fig 6C**). In conclusion,
388 susceptibility of a gene segment to the effects of NP:F346S depends upon a specific
389 combination of nucleotide identities at positions 4 and 13 of the 3' UTR and ORF
390 proximal sequences (4/13C and Udorn NA ORF proximal UTRs, or 4/13U and PR8 NA
391 ORF proximal UTRs).
392
393 As several of the PR8-Udorn NA UTR mutant viruses were highly attenuated, we
394 wanted to determine whether there were any compensatory mutations that may have
395 emerged that could potentially confound our results. We performed next-generation
396 sequencing on these viruses and found that the only virus with any mutations over
397 ~30% in the population was PR8:NA$^{\text{Udorn UTR}}$ NP:F346S, which had a fixed
398 nonsynonymous substitution in PB2 (E191G). We cannot rule out the possibility that this
399 mutation affects NA segment expression.
400
401 We hypothesized that the effects of different UTR mutations on sensitivity of NA
402 expression levels to NP:F346S could arise from two distinct mechanisms: (1) abrogating
403 the selective effect of NP:F346S on NA vRNA synthesis, restoring NA vRNA levels to
404 those observed in the context of NP:WT, or (2) reducing NA levels in the context of
405 NP:WT, bringing them closer to what is observed with NP:F346S. To distinguish
406 between these possibilities, we compared the expression of the different NA UTR
407 mutant constructs to NP vRNA levels (which are unaffected by the NP:F346S
408 substitution) (**Figs 6D and 6E**). We observed that the expression of all the PR8/Udorn
409 NA UTR chimeric constructs except for PR8:NA$^{\text{C4U}}$ was reduced compared to WT NA in
410 the context of NP:WT, and in some cases, lower than the level observed for the WT NA
411 segment in the presence of NP:F346S (**Figs 6D and 6E**). Thus, none of the UTR
412 mutants tested mitigated the effects of NP:F346S by simply restoring NA levels to those
413 observed in the context of NP:WT. The second possibility, that these mutations
414 appeared to reduce sensitivity to NP:F346S because they reduced NA levels in the
415 context of NP:WT to levels associated with NP:F346S, was also not supported by these
416 data. For instance, the PR8:NA$^{\text{Udorn ORF Proximal UTR}}$ segment exhibited increased

417 expression in the presence of NP:F346S, while the PR8:NA$^{Udorn\ 5'\ UTR}$ segment exhibited
418 an even more substantial decrease in the presence of NP:F346S than the WT NA (**Fig**
419 **6D**). Altogether, our data suggest that it is impossible to cleanly separate the effects of
420 the UTR sequences on susceptibility to the effects of NP:F346S from their broader
421 effects on baseline expression levels in the context of NP:WT.
422
423 **NP:F346S has no measurable effects on NP RNA binding or oligomerization**
424 While the UTR sequences of the NA segment are clearly involved in determining the
425 segment-specificity of the effects of NP:F346S on gene expression, the specific
426 mechanisms involved are not obvious. NP:F346S is not located in any previously
427 described functional domains, thus is was not immediately apparent how the NP:F346S
428 substitution might alter NP protein function (12,14,30–34).
429
430 NP has two well-described biochemical activities that are required for the synthesis of
431 full-length viral RNA transcripts: RNA binding and oligomerization (12,15,35). To
432 determine whether F346S affects the RNA-binding activity of NP, we compared the *in*
433 *vitro* RNA binding affinities of purified C-terminal his-tagged versions of the NP:WT and
434 NP:F346S proteins using bio-layer interferometry (BLI) (**Fig 7A**). NP:F346S was
435 associated with a slightly higher K_D compared with NP:WT (3.216±0.1088nM vs.
436 2.477±0.09097nM), however, it was not clear that this difference was biologically
437 significant (**Fig 7A**). Although these data suggest that the F346S substitution has
438 minimal effects on the RNA-binding affinity of NP, our *in vitro* assay may have failed to
439 fully recapitulate conditions as they occur during infection.
440



441
442

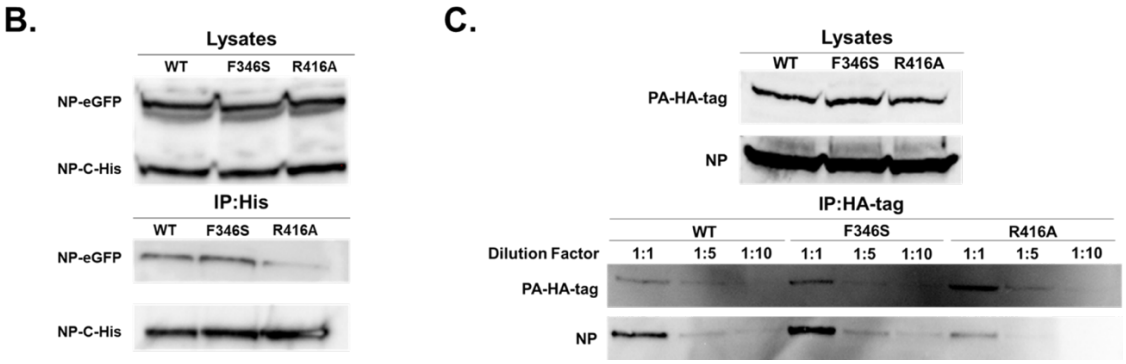
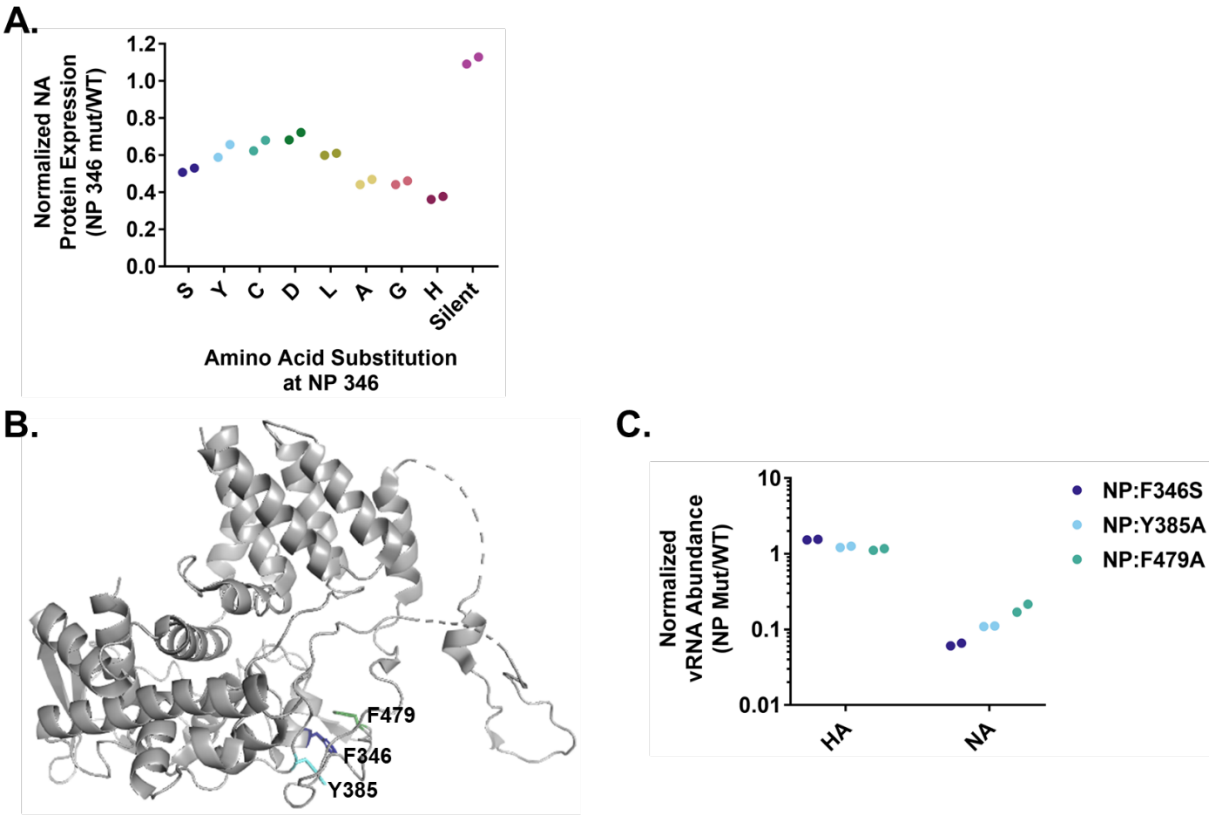


Fig 7. NP:F346S does not affect NP RNA binding or oligomerization. A.) *RNA binding kinetics of the PR8 NP:WT-C-His and PR8 NP:F346S-C-His proteins as determined by BLI. The raw data is colored blue/green and the fitted data is colored orange/purple for the NP:WT/F346S-C-His proteins, respectively.* ***B.)*** *Co-immunoprecipitation (IP) of eGFP and His-tagged versions of the indicated NP proteins. 293T cells were transfected with expression vectors encoding the eGFP- and His-tagged versions of either WT, F346S, or R416A NP proteins. Lysates were harvested after 24hrs. His-tagged NP was immunoprecipitated, and then IP samples were probed via western blot with anti-eGFP and anti-6x His antibodies. Western blots of total cell lysates stained with an anti-NP antibody are also shown.* ***C.)*** *Co-IP of vRNP-associated NP and PA. Cells were transfected with plasmids encoding the vRNP complex (PB2, PB1, PA-HA-tag, and NP (WT, F346S, or R416A) and a vRNA template (NA vRNA). Lysates were harvested 24hrs post transfection, and vRNP complexes were IP-ed using an anti-HA-tag antibody. Undiluted, 1:5 diluted, or 1:10 diluted IP-ed protein was probed with anti-NP and anti-HA-tag antibodies via western blot. Western blots of whole cell lysates shown for comparison.*

We next evaluated whether NP:F346S affects the oligomerization of NP monomers. We overexpressed both His-tagged and eGFP-tagged versions of either NP:WT or NP:F346S in 293T cells, and quantified the amount of eGFP-NP that co-immunoprecipitated with His-NP by western blot. As a positive control, we assessed the effect of the oligomerization-deficient R416A mutant (12,14,36,37) in our assay and measured a substantial reduction in pull-down efficiency (**Fig 7B**). In contrast, we did not observe any effect of F346S on the co-immunoprecipitation 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* over-expression conditions (**Fig 7B**).

Additionally, we evaluated whether NP:F346S decreases the NP content of vRNPs. We overexpressed the PB2, PB1, HA-tagged PA, NP (WT/F346S/R416A) proteins and a NA vRNA template in 293T cells to generate vRNPs and visualized the amount of NP that co-immunoprecipitated with HA-tagged PA via western blot. Mirroring our data looking at NP monomer association, there was a substantial reduction in pull-down efficiency for the oligomerization-deficient NP:R416A mutant, but no difference in the pull-down efficiencies between NP:WT and F346S (**Fig 7C**). These data suggest that the NP content of vRNPs is not affected by NP:F346S.

480 Altogether, these data suggest that F346S has minimal effects on the RNA-binding and
481 oligomerization activities of NP, at least in *in vitro* binding assays. If true in the context
482 of infection, it would suggest that the effects of this substitution on NA segment
483 replication occur through some other, uncharacterized feature of NP protein biology.
484
485 **A cluster of aromatic residues within NP governs NA segment replication**
486 Finally, we examined the effects of alternative substitutions at the NP:F346 locus on NA
487 expression. Introducing a silent substitution (T1082C) into the NP:F346 codon had no
488 effects on NA protein expression levels, indicating that the effects of NP:F346S require
489 the amino acid substitution (**Fig 8A**). Interestingly, we observed that any amino acid
490 substitution at position 346 resulted in a selective decrease in NA expression indicating
491 that a phenylalanine is required at NP position 346 for maximal NA expression (**Fig 8A**).
492 To better understand the need for a phenylalanine residue at this position, we examined
493 the surrounding protein structure. We noticed two additional aromatic residues (Y385 &
494 F479) directly adjacent to F346 that could potentially interact via π-π stacking
495 interactions (**Fig 8B**). Mutation of either Y385 or F479 to alanine resulted in a selective
496 decrease in NA abundance, though not quite as pronounced as that observed for
497 NP:F346S (**Fig 8C**). These data suggest that maximal expression of the NA gene
498 segment depends upon a cluster of aromatic residues F346, Y385, and F479 in NP.
499



500

501 ***Fig 8. A cluster of aromatic residues is involved in the regulation of NA gene***
502 ***segment expression. A.)*** *Normalized NA protein expression levels in cells infected with*
503 *the indicated PR8 NP 346 mutant viruses (MOI=0.03 $TCID_{50}$/cell, 16hpi) as determined*

by geometric mean fluorescent intensity (GMFI) expressed as a fraction of PR8 NP:WT. The data shown are 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 2IQH).* ***C.)*** *Normalized viral RNA abundance in PR8 NP:F346S, PR8 NP:Y385A and PR8 NP:F479A infected MDCK cells (MOI=0.1 $TCID_{50}$/cell, 8hpi) as determined by RT-qPCR and expressed as a fraction of PR8 NP:WT. The data shown are individual cell culture well replicates representative of the data obtained through two similar experiments.*

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 (38), 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 (15,25). Previous studies have demonstrated that these functions require both the RNA binding and oligomerization activities of NP (15,25,35). Surprisingly, we found that NP:F346S does not appreciably affect the RNA binding or oligomerization activities of NP *in vitro* (**Fig 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 positions NP F346-Y385-F479 govern interactions with other viral and/or cellular proteins involved in IAV gene segment replication (33,34,39–46). 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 svRNAs, which can modulate replication in a segment-specific manner (47,48). 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 8**), thus these residues may play a role in maintaining the structural integrity and 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

550 and Udorn: C's at positions 4 and 13 (as in WT PR8) and the Udorn-derived 3' & 5' ORF
551 proximal sequences (**Figs 5 and 6B**). Interestingly, this specific NA segment also
552 exhibited a >10-fold reduction in baseline NA expression in the absence of NP:F346S
553 (**Fig 6D**). Disrupting this pairing by mutating one of the C's at position 4 or 13, or
554 replacing one of the Udorn-derived ORF proximal regions with that from PR8 restored
555 the effects of NP:F346S on NA synthesis (**Figs 5 and 6A,B**). All NA constructs that we
556 tested that were less sensitive to the effects of NP:F346S, with the exception of
557 PR8:NA^{C4U}, also exhibited significantly lower baseline levels of NA expression under
558 WT NP conditions, indicating that sensitivity to the effects of NP:F346S cannot be
559 uncoupled from baseline NA expression levels.
560
561 For all eight genome segments, gene expression is influenced by the structure of the
562 viral promoter that is formed by base-paring interactions between the 3' and 5' UTRs
563 (49). Base-pairing between positions 4 of the 3' and 5' UTR and between positions 13 of
564 the 3' UTR and 14 of the 5' UTR influence the promoter structure and stability and have
565 been shown to be important for regulating gene segment replication and transcription
566 (6,49–52). For the NA segment, a C at position 4 in the 3' UTR (which is unable to
567 base-pair with the A at position 4 of the 5' UTR) promotes genome replication, while a U
568 at this position favors mRNA transcription (50). For PR8, NA is the only segment with
569 C's at both positions 4 and 13 of the 3' UTR, and thus has the fewest number of base-
570 pairing interactions based on the traditional panhandle structure of the IAV promoter.
571 This unique feature may make the NA segment of PR8 uniquely dependent on WT NP
572 to facilitate the stable interaction between the promoter and the viral replicase.
573
574 The ORF-proximal regions of the UTRs also influence gene segment expression,
575 however, the exact mechanism(s) remain unclear (5,7–10). Consistent with this, we
576 observed that the ORF proximal regions of the NA UTRs play important roles in both
577 regulating the baseline expression level of the NA segment and in determining
578 sensitivity to NP:F346S. Relative to Udorn and the remaining seven gene segments of
579 PR8, the PR8 NA segment UTRs harbor a unique extended stretch of base-pairing from
580 3'-nt14/5'-nt15 to 3'-nt17/5'-nt18 located within the poly U stretch of the 5' UTR. Given
581 the hypothesized role for NP in promoter escape and elongation (15), the NA segment
582 may be particularly dependent on WT NP (and thus sensitive to NP:F346S) to allow the
583 polymerase to bypass this extended base-pairing region during elongation of the
584 nascent vRNA. Altogether, our data suggest that the unique sequence and presumed
585 structure of the PR8 NA segment UTRs confer elevated sensitivity to perturbations of
586 NP function and thus, likely explain the segment specificity of the effects of NP:F346S.
587
588 Through mutations in the UTR sequences of individual segments, IAV can more finely
589 coordinate the expression of the eight individual gene segments without altering the
590 protein coding capacity of the segments or polymerase activity. Similar to our findings, a
591 recent study demonstrated that the 3' UTR of the HA segment played a role in
592 regulating HA expression in a segment-specific manner, and that this regulation was
593 only observed when the HA segment had to compete with the remaining seven
594 segments for replication/transcription (7). An additional study using a reporter system
595 also found that the UTR sequences of the segments affected the ability of the segments

596 to compete with one another for access to the viral polymerase (52). Altogether, these
597 studies highlight the importance of the individual segment UTR sequences in
598 maintaining the optimal balance in expression of the eight gene segments during
599 infection. Our results further demonstrate how slight perturbations in polymerase or NP
600 function can affect the expression of specific segments to a greater degree than others.
601
602 What are the implications of NA (and potentially other viral gene segments) being
603 sensitive to individual substitutions in NP? HA and NA facilitate viral particle attachment
604 and release respectively, and balancing these activities is necessary for maintaining
605 viral fitness (53–62). HA and NA evolve at faster rates than the rest of the IAV genome
606 due to immune selection (63). Immune escape substitutions within HA and/or NA often
607 alter glycoprotein function and require compensatory mutations to restore their
608 functional balance and viral fitness (64–72). If substitutions in NP tune NA expression, it
609 expands the number of available genetic pathways maintaining HA/NA functional
610 balance. The functional link between the NP and NA segments that we establish here
611 also has important consequences for reassortment, as the need to maintain compatible
612 NP and NA genotypes may constrain the repertoire of viable reassortant progeny when
613 heterologous viral strains mix. Finally, variation in NP-requirements between segments
614 could influence patterns of expression kinetics, as the concentration of NP within the
615 cell is dynamic over time. Altogether, our results highlight the potential of genome
616 segmentation to facilitate dynamic changes in gene expression patterns through
617 mechanisms that are not readily available to non-segmented viruses. This regulatory
618 agility may help promote viral adaptation in response to changing host environments.
619
620 In summary, we identified a novel mechanism through which interactions between NP
621 and other gene segment UTRs facilitate selectively regulation of viral gene expression.
622 Our data reveal a new functional domain in the NP protein and suggest a broader role
623 for NP in selective regulation of individual viral gene segments. These findings
624 demonstrate how the expression of individual gene segments can be modulated to
625 maximize viral fitness under different host conditions.
626
627 **Materials and Methods**
628 **Plasmids**
629 The A/Puerto Rico/8/34 and A/Udorn/72 reverse genetics plasmids were gifts from Drs.
630 Adolfo Garcia-Sastre and Kanta Subbarao, respectively. The pCI vector was provided
631 by Dr. Joanna Shisler. The lentivirus generation plasmids- pHAGE2-EF1aInt WSN HA
632 W, HDM Hgpm2, HDM tatlb, pRC CMV Rev1b, HDM VSV-G were provided by Dr.
633 Jesse Bloom. The peGFP-C1 plasmid for generating C-terminal eGFP-tagged proteins
634 was provided by Dr. Andrew Mehle.
635
636 Point mutations were introduced into the PR8 NP segment via site-directed
637 mutagenesis and *Lgu*I restriction sites were added to both ends. The inserts were then
638 digested with *Lgu*I, ligated into the pDZ vector, and transformed into chemically
639 competent *E. coli* cells via the heat-shock method. Insert sequences were confirmed via
640 sanger sequencing.
641

642 PR8 NA ORF HA UTR+Pack Swap and PR8 HA ORF NA UTR+Pack inserts were
643 generated via overlap extension PCR with primers designed to introduce the PR8 HA or
644 PR8 NA UTR+Packaging Signal regions at the ends of the PR8 NA ORF or PR8 HA
645 ORF using the pDZ PR8 NA or pDZ PR8 HA plasmid as a template respectively.
646 Primers were used to add *Lgu*I restriction sites to each end of the inserts. PR8 NA ORF
647 HA UTR and PR8 HA ORF NA UTR inserts were generated via PCR with primers
648 designed to add the PR8 HA or PR8 NA UTRs to the PR8 NA or HA ORFs using the
649 pDZ PR8 NA or pDZ PR8 HA plasmid as a template respectively. *Lgu*I restriction sites
650 were added to each end. The inserts were then digested with *Lgu*I, ligated into the pDZ
651 vector, and transformed into chemically competent *E. coli* cells via the heat-shock
652 method. Insert sequences were confirmed via sanger sequencing.
653
654 For UTR chimera constructs, inserts were generated via PCR with primers designed to
655 introduce the 4U and/or 13U mutations into the PR8 NA UTR, or to replace specific
656 regions of the PR8 NA UTR at the 3' and/or 5' ends or the Udorn NA UTR at the 3'
657 and/or 5' ends with the equivalent region(s) present in the Udorn NA UTR or PR8 NA
658 UTR respectively. For the plasmids containing a chimeric Udorn-PR8 NA segment with
659 the Udorn NA ORF, the A763C silent mutation was introduced to the Udorn NA
660 sequence to disrupt an internal *Lgu*I restriction site. *Lgu*I restriction sites were added to
661 the ends of the chimeric segments via PCR, the inserts were then digested with *Lgu*I,
662 ligated into the pDZ vector, and transformed into chemically competent *E. coli* cells via
663 the heat-shock method. Insert sequences were confirmed via sanger sequencing.
664
665 To clone IAV ORFs into the pCI mammalian expression vector, inserts were generated
666 by PCR with primers that bound to the terminal regions of the PR8 NA/HA/NP ORFs
667 and added *Eco*RI and *Sal*I restriction sites to the 5'/3' ends respectively using the pDZ
668 PR8 NA/HA/NP plasmids as templates. The PR8 PB1 ORF insert was generated with
669 primers that bound to the terminal regions of the PR8 PB1 ORF and added *Mlu*I and
670 *Kpn*I restriction sites to the 5'/3' ends respectively using the pDZ PR8 PB1 plasmid as a
671 template. Internal *Eco*RI/*Sal*I restriction sites in HA were removed via site-directed
672 mutagenesis. The PR8 NA/HA/NP and PR8 PB1 inserts were then digested with the
673 *Eco*RI/*Sal*I or *Mlu*I/*Kpn*I restriction enzymes respectively. The inserts were then ligated
674 into the pCI vector, and transformed into chemically competent *E. coli* cells via the heat-
675 shock method. Insert sequences were confirmed via sanger sequencing.
676
677 For epitope and eGFP-tagged NP expression vectors, mutations in PR8 NP were
678 introduced via site directed mutagenesis. C-terminal 6x His tags were introduced by
679 performing PCR with primers designed to add a C-terminal 6x His tag before the stop
680 codon of the PR8 NP ORF using the pDZ PR8 NP plasmid as a template. A C-terminal
681 HA-tag was added to PR8 PA by performing PCR with primers designed to add a C-
682 terminal HA-tag before the stop codon of the PR8 PA ORF using the pDZ PR8 PA
683 plasmid as a template. For cloning into the pCI plasmid, *Eco*RI/*Sal*I restriction sites
684 were introduced to the 5'/3' ends respectively of the PR8 NP-C-His (WT/F346S/R416A)
685 and PR8 PA-HA tag ORFs. For cloning into the peGFP-C1 plasmid, *Bsp*EI/*Kpn*I
686 restriction sites were introduced to the 5'/3' ends respectively of the PR8 NP
687 (WT/F346S/R416A) ORFs. The inserts were then digested with the *Eco*RI/*Sal*I (pCI

688 PR8 NP:WT/F346S/R416A C-His and pCI PR8 PA-HA tag) or *Bsp*EI/*Kpn*I (peGFP-PR8
689 NP:WT/F346S/R416A) restriction enzymes, ligated into the pCI or peGFP vectors
690 respectively, and transformed into chemically competent *E. coli* cells via the heat-shock
691 method. Insert sequences were confirmed via sanger sequencing.
692
693 For lentiviral expression vectors, inserts were generated by PCR with primers designed
694 to bind to the 5' and 3' terminal regions of the PR8 HA ORF and introduce *Bam*HI/*Not*I
695 restriction sites to the 5'/3' ends respectively using the pDZ PR8 HA plasmid as a
696 template. The insert was then digested with the *Bam*HI/*Not*I restriction enzymes, ligated
697 into the pHAGE-EF1aInt vector (generated by restriction digest of the pHAGE2-EF1aInt
698 WSN HA W plasmid with the same restriction enzymes), and transformed into
699 chemically competent *E. coli* cells via the heat-shock method. Insert sequences were
700 confirmed via sanger sequencing.
701
702 **Cells**
703 Madin-Darby canine kidney (MDCK), MDCK-SIAT1 cells, and 293T cells were obtained
704 from Drs. Jonathan Yewdell, Jesse Bloom, and Joanna Shisler respectively and were
705 maintained in Gibco's minimal essential medium (MEM) with GlutaMax (Life
706 Technologies) supplemented with 8.3% fetal bovine serum (Seradigm) (MEM+FBS) and
707 incubated at 37°C with 5% CO_2.
708
709 MDCK cells expressing the PR8 HA protein were generated via lentiviral transduction.
710 293T cells were transfected with 250ng each of the plasmids required for lentivirus
711 generation (HDM Hgpm2, HDM tatlb, pRC CMV Rev1b, HDM VSV-G), and 1µg of the
712 transfer vector pHAGE PR8 HA (generated as described above). One day post
713 transfection, the transfection media was replaced with 2mL MEM+FBS. The next day,
714 the lentiviral supernatant was collected and 1mL was used to infect three wells of
715 MDCK cells plated in a 6 well plate at 10% confluency. Two days post transduction, the
716 MDCK cells were harvested and combined, surface stained with an anti-HA antibody
717 (H36-26 AF488), and positive cells were sorted out via fluorescence activated cell
718 sorting (FACS).
719
720 **Viruses**
721 Recombinant A/Puerto Rico/8/1934 (H1N1) (PR8) and A/Udorn/72 (H3N2) (Udorn)
722 viruses were generated using 8 and 12 plasmid reverse genetics systems respectively.
723 The rPR8 clones differ from the published sequence (GenBank accession no.
724 AF389115 to AF389122) at two positions: PB1 A549C (K175N) and HA A651C (I207L)
725 (numbering from initiating Met). Viruses containing single point mutations in the NP or
726 NA segments were generated by rescuing the viruses using plasmids containing the
727 specific mutations introduced via site-directed mutagenesis. The Udorn HA segment-
728 encoding plasmids used were found to have the following mutations: A81G (N18D),
729 C129T (H34Y), G1103T (silent), T1486A (F486Y), & A1614G (N529D) relative to the
730 Udorn HA reference sequence (GenBank accession no. AX350190).
731
732 Viruses were rescued by transfecting 293T cells with 500ng each of the relevant
733 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 500ng 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. 18-24hrs post transfection, the media was replaced with viral growth media (MEM, 1 mM HEPES, 1 µg/mL TPCK trypsin (Worthington Biochemical Corporation; Lakewood, NJ, USA), 50 µg/mL gentamicin) containing 2×10^5 MDCK cells). For the viruses with the chimeric PR8 HA/NA segments, the viral growth media was modified by adding $2x10^5$ PR8 HA+ MDCK cells instead of MDCK cells. Transfection supernatants were collected 24hrs post media 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 media (MEM, 1 mM HEPES, 1 µg/mL TPCK trypsin, 50 µg/mL gentamicin). Seed stocks were harvested and clarified (14000rpm, 15min, 4°C) between 24-72hrs 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 an MOI of 0.001 or 0.01 $TCID_{50}$/cell respectively, and the working stocks were harvested 24-72hrs post infection and clarified (3500rpm, 15min, 4°C). Viral growth was performed in viral growth media.

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 1mL of transfection supernatant for 1hr at 37°C with rocking, and then the transfection supernatant was removed and replaced with 3mL of viral growth media with 0.5 µg/mL TPCK trypsin and left to incubate for up to 48hrs. The infection supernatants were harvested and clarified (14,000rpm, 15min, 4°C), and 1mL 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 24-48hrs post infection. The PR8 HA/NA UTR swap NP:WT virus had a mutation in the PR8 HA ORF NA UTR segment (C33A→L5I).

To generate 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, MDCK cells in a single well of a 6 well plate were infected with 1mL of transfection supernatant for 1hr at 37°C with rocking, and then the transfection supernatant was removed and replaced with 3mL of viral growth media and left to incubate for up to 48hrs. The infection supernatants were harvested and clarified (14,000rpm, 15min, 4°C) once CPE was observed.

780
781 Virus titers were determined by $TCID_{50}$ assay on MDCK cells, or by determining the
782 fraction of viral particles expressing NP (NPEU)(24) via flow cytometry on infected
783 MDCK cells using the anti-NP AF647 (HB65) antibody.
784
785 **Next generation sequencing of viruses**
786 Viral RNA was extracted from 140µL of the viral infection supernatant using the QIAamp
787 Kit (Qiagen) and eluted in 60µL of Nuclease-free water (Ambion). Contaminating DNA
788 was removed using the Qiagen RNase-free DNase Set, and then the RNA was cleaned
789 using the RNeasy Kit (Qiagen) and eluted in 30µL of Nuclease-free water (Ambion).
790 cDNA was synthesized using the Superscript III Reverse Transcriptase Kit
791 (ThermoFisher) as follows: 1µL of 2µM MBTUni-12 primer (5'-
792 ACGCGTGATCAGCRAAAGCAGG-3') + 1µL 10mM dNTPs Mix (NEB #N0447S) + 8µL
793 Nuclease-free water (Ambion) were added to 3µL of RNA and then incubated at 65°C
794 for 5 min and then 4°C for 2 min. 4µL of 5x First Strand cDNA Synthesis Buffer + 1µL
795 0.1M DTT + 1µL SUPERase-In RNase Inhibitor (Invitrogen #AM2696) + 1µL
796 Superscript III RT (Invitrogen #18080-044) were added to the reaction and then the
797 reaction was incubated at 45°C for 50 min. The cDNA was then stored at -20°C. The
798 PCR reaction to simultaneously amplify all eight gene segments was performed using
799 Phusion polymerase (NEB #M0530L) as follows: 2.5µL of 10µM MBTUni-12 primer +
800 2.5µL of 10µM MBTUni-13 primer (5'-ACGCGTGATCAGTAGAAACAAGG-3') + 10µL 5x
801 HF Phusion Buffer + 1µL 10mM dNTPs mix (NEB #N0447S) + 0.5µL Phusion
802 Polymerase + 28.5µL of Nuclease-free water (Ambion) was added to 5µL of cDNA. The
803 cycling conditions for the PCR were as follows: 98°C for 30sec, (98°C for 10sec/ 57°C
804 for 30sec/ 72°C for 1min 30sec) x 25, 72°C for 5min, 4°C Hold). The PCR products
805 were cleaned using the Invitrogen PureLink PCR Purification Kit (Invitrogen #K310002)
806 using the Buffer for the <300bp cutoff and eluted in 30µL of Nuclease-free water
807 (Ambion). The PCR products were then subjected to next generation sequencing using
808 the Illumina NovaSeq or MiSeq platforms.
809
810 **Time-course infections**
811 MDCK cells were infected with the PR8 NP:WT/F346S or Udorn NP:WT/F346S viruses
812 at an MOI of 0.1 NPEU/cell in a 24 well plate for 1hr at 37°C. 1hr post infection, the
813 virus supernatant was replaced with 0.5mL MEM+FBS. 3hr post infection, the
814 MEM+FBS was replaced with 0.5mL NH_4Cl media (MEM, 50mM HEPES Buffer, 20mM
815 NH_4Cl, pH=7.2) to prevent viral spread. The cell monolayers were harvested 4, 8, and
816 12hrs post infection, and cellular RNA was extracted using the RNeasy kit (Qiagen).
817 Reverse transcription was performed using the Verso cDNA Synthesis Kit
818 (ThermoFisher). The reactions were set up as follows: 4µL RNA + 4µL 5x cDNA
819 Synthesis Buffer + 2µL dNTP Mix + 1µL 10µM PR8 RT_4A primer (5'-
820 AGCAAAAGCAGG-3') + 1µL RT Enhancer + 1µL Verso Enzyme Mix + 7µL Nuclease-
821 free water, and incubated at 45°C for 50min, 95°C for 2min, and held at 4°C. cDNA was
822 stored at -20°C. Quantitative real-time PCR on cDNA was carried out using Power
823 SYBR green PCR Master Mix (Thermo Fisher) on a QuantStudio 3 thermal cycler
824 (Thermo Fisher). The strand-specific forward and reverse primers for quantitative real-
825 time PCR for PR8 HA, NP, and NA were 5'-AAGGCAAACCTACTGGTCCTGTT-3' & 5'-

826 AATTGTTCGCATGGTAGCCTATAC-3', 5'-AGGCACCAAACGGTCTTACG-3' & 5'-
827 TTCCGACGGATGCTCTGATT-3', and 5'-AAATCAGAAAATAACAACCATTGGA-3' & 5'-
828 ATTCCCTATTTGCAATATTAGGCT-3' respectively. The strand-specific forward and
829 reverse primers for Udorn HA, NP, and NA were 5'-GACTATCATTGCTTTGAGC-3' &
830 5'- CACTAGTGTTCCGTTTGGC-3' and 5'-CGGTCTTATGAACAGATGG-3' & 5'-
831 TCGTCCAATTCCATCAATC-3' and 5'-AACAATTGGCTCTGTCTCTC-3' & 5'-
832 GTCGCACTCATATTGCTTG-3' respectively. Reactions were set up as follows: 2µL
833 cDNA + 10µL 2x Power SYBR Green MM + 1µL 10µM Forward Primer + 1µL 10µM
834 Reverse Primer + 6µL Nuclease-free water. The cycling conditions were as follows:
835 50°C for 2min, 95°C for 10min, and then 95°C for 15 sec followed by 60°C for 1min
836 repeated 40x.
837
838 **Analysis of primary viral transcription in infected cells**
839 MDCK-SIAT1 cells were infected with the PR8 NP:WT/F346S viruses at an MOI of 5
840 $TCID_{50}$/cell in a 6 well plate in the presence of 100µg/mL of cycloheximide (Sigma-
841 Aldrich). Infected cells were harvested at 2 and 6hrs post-infection, and cellular RNA
842 was extracted using the RNeasy Kit (Qiagen). Reverse transcription was performed
843 using the Superscript III Reverse Transcriptase Kit (ThermoFisher). The reactions were
844 set up as follows: 4µL RNA + 0.5µL 100µM Oligo dT_{20} primer (IDT) + 1µL 10mM dNTP
845 + 6.5µL Nuclease-free water incubated at 65°C for 5min and then 4°C for 1min. 4µL 5x
846 First Strand RNA Buffer + 1µL 0.1M DTT + 2µL SuperaseIN RNase Inhibitor
847 (ThermoFisher) + 1µL of Superscript III Reverse Transcriptase was added to the
848 previous reaction and incubated at 50°C for 60min, 70°C for 15min, and held at 4°C.
849 cDNA was stored at -20°C. Quantitative real-time PCR on cDNA was carried out using
850 Power SYBR green PCR Master Mix (Thermo Fisher) on a QuantStudio 3 thermal
851 cycler (Thermo Fisher). The strand-specific forward and reverse primers for quantitative
852 real-time PCR for PR8 HA and NA were as follows: PR8 HA 5'-
853 AAGGCAAACCTACTGGTCCTGTT-3' & 5'-AATTGTTCGCATGGTAGCCTATAC-3' and
854 PR8 NA 5'-AAATCAGAAAATAACAACCATTGGA-3' & 5'-
855 ATTCCCTATTTGCAATATTAGGCT-3'. Reactions were set up as follows: 1.5µL cDNA
856 + 10µL 2x Power SYBR Green MM + 1µL 10µM Forward Primer + 1µL 10µM Reverse
857 Primer + 6.5µL Nuclease-free water. The cycling conditions were as follows: 50°C for
858 2min, 95°C for 10min, and then 95°C for 15 sec followed by 60°C for 1min repeated
859 40x.
860
861 **4SU RNA Pulse**
862 60-70% confluent MDCK cells were infected with the PR8 NP:WT/F346S viruses at an
863 MOI of 5 $TCID_{50}$/cell and incubated at 37°C for 1hr. The virus was aspirated and then
864 3mL of MEM+FBS was added to each well. 7hpi the MEM+FBS was replaced with 1mL
865 of fresh MEM+FBS containing 500µM 4-thiouridine (4SU) (Tri-Link Biotechnologies N-
866 1025). The cells were kept in the dark during the labeling process to prevent cross-
867 linking of 4SU to cellular proteins. 1hr post labeling the cells were harvested and cellular
868 RNA was extracted using the RNeasy Kit (Qiagen).
869
870 The cellular RNA was then biotinylated by performing a reaction with EZ-Link HPDP-
871 Biotin (ThermoScientific). Reactions conditions were as follows: 10µg RNA and Biotin-

HPDP (0.2µg/µL final concentration) were added to Biotinylation Buffer (10mM Tris-HCl pH=7.5, 1mM EDTA) resulting in a total reaction volume of 250µL, and then incubated at room temperature with end-over-end rotation for 2hrs protected from light. RNA was then extracted using the chloroform: isoamyl alcohol procedure performed as follows: 400µL of chloroform: isoamyl alcohol (49:1 ratio) was added to each reaction, mixed, and then added to a 2mL Quanta Bio 5PRIME Phase Lock Heavy Tube. The phases were separated by centrifugation (Full speed, 5min, 4°C), and the aqueous layer (top) was transferred to a new 1.5mL microcentrifuge tube. The previous steps were then repeated one additional time. RNA was precipitated as follows: One volume of isopropanol, one-tenth volume of 5M NaCl, and 1µL of 15µg/mL GlycoBlue Coprecipitant (Invitrogen) were added to each sample and mixed. The samples were then frozen at -70°C overnight. The next day the samples were thawed and RNA was pelleted via centrifugation (Full speed, 20min, 4°C). The pellet was then washed 2x with 400µL of 80% ethanol (Full speed, 5min, 4°C). The pellet was then air-dried at room temperature for 5min and resuspended in 20µL of Nuclease-free water.

Biotinylated RNAs were then selectively purified using the µMACS Streptavidin Kit (Miltenyi Biotec) as follows: 15µL of RNA was added to 85µL of Nuclease-free water and denatured by incubating at 65°C for 10min followed by cooling on ice for 5min. 100µL of Miltenyi streptavidin beads were added to each reaction and incubated at room temperature for 15min with end-over-end rotation. Meanwhile, the µMACS columns were equilibrated by adding 100µL of the Equilibration Buffer for Nucleic Acid Applications and then washed with 1mL of wash buffer (100mM Tris-HCl pH=7.5, 10mM EDTA, 1M NaCl). The biotinylated RNA-streptavidin bead solution was then added to the columns, and then the columns were washed with 0.9mL of 65°C wash buffer 3x followed by 0.9mL of room temperature wash buffer 3x. The biotinylated RNA was then eluted with 150µL of 0.1M DTT. The RNA was stored at -70°C.

Reverse transcription was performed using the Superscript III Reverse Transcriptase Kit (ThermoFisher). A universal vRNA-specific primer or a tagged, segment-specific, vRNA-specific primer was used. The reactions were set up as follows: 2µL RNA + 0.5µL 10µM primer (Universal: PR8 RT_4A primer (5'-AGCAAAAGCAGG-3') or segment specific: NA vRNA-24 tag (5'-GGCCGTCATGGTGGCGAATAATCCAAATCAGAAAATAACAACC-3') or 10µM HA vRNA-36 tag primer (5'-GGCCGTCATGGTGGCGAATAAGGCAAACCTACTGGTCCTGTT-3')) + 1µL 10mM dNTP + 8.5µL Nuclease-free water incubated at 65°C for 5min and then 4°C for 1min. 4µL 5x First Strand RNA Buffer + 1µL 0.1M DTT + 2µL SuperaseIN RNase Inhibitor (ThermoFisher) + 1µL of Superscript III Reverse Transcriptase were added to the previous reaction, and the reactions were incubated at 50°C for 60min, 70°C for 15min, 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). The strand-specific forward and reverse primers for quantitative real-time PCR for PR8 HA and NA vRNA were: For RT reaction using universal primer (PR8 HA 5'-AAGGCAAACCTACTGGTCCTGTT-3' & 5'-AATTGTTCGCATGGTAGCCTATAC-3' and PR8 NA 5'-

AAATCAGAAAATAACAACCATTGGA-3' & 5'-ATTCCCTATTTGCAATATTAGGCT-3'), and for RT reaction using segment-specific primers (vtag (5'-GGCCGTCATGGTGGCGAAT-3') & PR8 HA qPCR 3' (5'-AATTGTTCGCATGGTAGCCTATAC-3'), and vtag (5'-GGCCGTCATGGTGGCGAAT-3') & PR8 NA qPCR 3' (5'- ATTCCCTATTTGCAATATTAGGCT-3'). Reactions were set up as follows: 1.5µL cDNA + 10µL 2x Power SYBR Green MM + 1µL 10µM Forward Primer + 1µL 10µM Reverse Primer + 6.5µL Nuclease-free water. The cycling conditions were as follows: 95°C for 10min, and then 95°C for 15 sec followed by 54/57°C for PR8 NA/HA vRNA respectively for 1min repeated 40x.

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 or PR8:Udorn HA,NA$^{PR8\ ORF\ Proximal\ UTR}$ NP:WT/F346S viruses due to their low titer) or an MOI of 0.1 $TCID_{50}$/cell (PR8 NP:WT v. F346S v. Y385A v. F479A) in a 24 well plate. 1hpi infection, the viral supernatant was removed and replaced with $TCID_{50}$ media (MEM, 1 mM HEPES, 0.5 or 1 µg/mL TPCK trypsin, 50 µg/mL gentamicin). Cells were harvested at 8hpi 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 DNaseI (Qiagen) and cleaned using the RNeasy Kit (Qiagen). Reverse transcription was performed using the Verso cDNA Synthesis Kit (ThermoFisher). The reactions were set up as follows: 4µL RNA + 4µL 5x cDNA Synthesis Buffer + 2µL dNTP Mix + 1µL 10µM PR8 RT_4A primer (5'-AGCAAAAGCAGG-3') + 1µL RT Enhancer + 1µL Verso Enzyme Mix + 7µL Nuclease-free water, and incubated at 45°C for 50min, 95°C for 2min, 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). The strand-specific forward and reverse primers for quantitative real-time PCR for PR8 HA, NP, and NA/PR8 NA ORF HA UTRs were 5'-AAGGCAAACCTACTGGTCCTGTT-3' & 5'-AATTGTTCGCATGGTAGCCTATAC-3', 5'-AGGCACCAAACGGTCTTACG-3' & 5'-TTCCGACGGATGCTCTGATT-3', and 5'-AAATCAGAAAATAACAACCATTGGA-3' & 5'-ATTCCCTATTTGCAATATTAGGCT-3' respectively. The strand-specific forward and reverse primers for Udorn HA, NP, and NA were 5'-GACTATCATTGCTTTGAGC-3' & 5'- CACTAGTGTTCCGTTTGGC-3' and 5'-CGGTCTTATGAACAGATGG-3' & 5'-TCGTCCAATTCCATCAATC-3' and 5'-AACAATTGGCTCTGTCTCTC-3' & 5'-GTCGCACTCATATTGCTTG-3' respectively. The primers for the PR8 HA ORF NA UTR+Pack and PR8 NA ORF HA UTR+Pack segments were 5'-AAATCAGAAAATAACAACCATTGGA-3' & 5'-CAACAATACCAACAGATTAGC-3' and 5'-AAGGCAAACCTACTGGTCCTGTT-3' & 5'-ATCAGCCCTACCACGAGGC-3' respectively. The primers for the PR8 NA Codon Shuffle segment were: 5'- CAAATGGGACCGTCAAAGACCGC-3' and 5'-GATGGGGCTTCACCGACTGG-3'. Reactions were set up as follows: 2µL cDNA + 10µL 2x Power SYBR Green MM + 1µL 10µM Forward Primer + 1µL 10µM Reverse Primer + 6µL Nuclease-free water. The cycling conditions were as follows: 50°C for 2min, 95°C for 10min, and then 95°C for 15 sec followed by 60°C for 1min repeated 40x.

964
965 **Flow cytometry to detect viral protein expression in singly-infected cells**
966 MDCK cells were infected at an MOI of 0.03 $TCID_{50}$/cell in a 6 well plate, or $1.5x10^6$
967 MDCK cells per well were infected with 10^{-1} to 10^{-4} dilutions of the viral stock in a 6 well
968 plate (for NP-expressing unit (NPEU) determination). 1hpi the virus was replaced with
969 3mL of MEM+FBS. 3hpi the MEM+FBS was then replaced with 3mL of NH_4Cl media
970 (MEM, 50mM HEPES Buffer, 20mM NH_4Cl, pH=7.2) to prevent secondary infection.
971 16hpi the cells were harvested, fixed and permeabilized with foxP3 fix/perm buffer
972 (eBioscience). Cells were subsequently stained with one or multiple of the following
973 antibodies: PR8 NA: Rabbit anti-NA (08-0096-03 EXSANG 3/9/09) followed by Donkey
974 anti-Rabbit PE (711-116-152 Lot 121465), PR8 NP: HB65 AF647 or PacB, Udorn NA:
975 Goat anti-N2 NA primary followed by Donkey anti-Goat (705-116-147) PE secondary,
976 and Udorn HA: H14A2 AF647. The cells were run on a BD LSR II flow cytometer and
977 analyzed using FlowJo version 10.1 (Tree Star, Inc.). Viral protein expression levels
978 were determined from the geometric mean fluorescence intensity (GMFI) of the
979 fluorophore associated with each protein. NPEU titers were calculated by dividing the
980 number of infected cells (% NP+)(Total Number of Cells) by the dilution factor and the
981 volume of the inoculum.
982
983 **PR8 NP:WT/F346S-C-His protein purification**
984 100µg of the pCI PR8 NP:WT-C-His or pCI PR8 NP:F346S-C-His plasmids were
985 transiently transfected into 100mL cultures of HEK Expi-293-F cells with ExpiFectamine
986 according to the company protocol (Thermo Fisher). The transfected cells were pelleted
987 (1000rpm, 5min, 4°C) and resuspended in 3mL of Equilibration Buffer (PBS, 10mM
988 Imidazole, 1X cOmplete EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich)). Cells
989 were kept frozen at -70°C until the next step of the purification procedure was
990 performed. The cells were thawed and then lysed via the freeze-thaw method:
991 Incubation in a dry-ice-ethanol bath followed by 42°C water bath repeated 2x. The
992 chromosomal DNA was then sheared by passing the lysate through an 18G needle 4x.
993 The cellular debris was pelleted by centrifugation (3000xg, 15min, 4°C), and the clarified
994 lysates were transferred to new tubes. The clarified lysates were then treated with
995 RNaseA (50µg/mL final concentration) for 2hrs at room temperature.
996
997 His-tagged proteins were then selectively purified using the HisPur Ni-NTA Spin
998 Purification Kit, 0.2mL (ThermoFisher). To improve the purity of the eluted protein
999 fractions and increase the protein yield, the eluate fractions for the PR8 NP:WT/F346S
1000 C-His proteins respectively were combined, the buffer for the eluate fractions was
1001 exchanged to Equilibration Buffer using Pierce Protein Concentrators (PES, 10kDa
1002 MWCO, 0.5mL) (ThermoFisher Scientific), and the HisPur Ni-NTA column purification
1003 was repeated once more. The eluate fractions for each protein were then combined,
1004 added to a Slide-A-Lyzer Dialysis Cassette (Extra Strength, 10kDa MWCO, 0.5-3mL
1005 capacity) (Thermo Scientific), and dialyzed in PBS at 4°C overnight. The dialyzed
1006 protein samples were concentrated using Pierce Concentrators (PES, 10kDa MWCO,
1007 0.5mL) (Thermo Scientific), and the protein concentrations were determined using the
1008 Pierce Coomassie Plus Bradford Assay Kit (Thermo Scientific).
1009

1010 **Biolayer interferometry (BLI)**
1011 Binding kinetics of PR8 NP:WT/F346S C-His proteins to biotinylated single-stranded
1012 RNA (ssRNA) was determined by biolayer interferometry using an Octet Red96e
1013 instrument (FortéBio) at room temperature. Anti-streptavidin biosensors (FortéBio) were
1014 washed in 180µl 1X kinetics buffer (0.002% v/v Tween 20 in 1X PBS, pH 7.4) for 30min
1015 at room temperature. The experiment consisted of five steps: (1) baseline: 60s with 1X
1016 kinetics buffer; (2) loading: 300s with 5' biotinylated 24nt ssRNA (5'-
1017 UUUGUUACACACACACACGCUGUG-3') (IDT) at 1µM; (3) baseline: 60s with 1X
1018 kinetics buffer; (4) association: 60s with 50, 100, 200, and 400nM of PR8 NP:WT/F346S
1019 C-His protein; (5) dissociation: 300s with 1X kinetics buffer. Octet Data Acquisition
1020 (version 11.1, FortéBio) software was used to obtain biolayer interferometry data. To
1021 calculate the dissociation constant (KD) via curve fitting, a 1:1 binding model was used.
1022 Octet Data Analysis (version 11.1, FortéBio) software was used to analyze binding
1023 kinetics. Negative controls were set up with 1X kinetics buffer replacing 1µM of
1024 biotinylated ssRNA in the loading step.
1025
1026 **Transfection protocols for eGFP/His-tagged NP protein co-immunoprecipitation**
1027 **and HA-tagged vRNP complex pulldown experiments**
1028 For the eGFP/His-tagged NP protein co-immunoprecipitation experiment: 293T cells in
1029 a 10cm dish were transfected with 5µg each of one of the following pairs of plasmids
1030 using JetPrime (Polyplus) according to the manufacturer's instructions: pCI PR8
1031 NP:WT-C-His & peGFP-PR8 NP:WT, pCI PR8 NP:F346S-C-His & peGFP-PR8
1032 NP:F346S, or pCI PR8 NP:R416A-C-His & peGFP-PR8 NP:R416A. For the HA-tagged
1033 vRNP complex pulldown experiment: 293T cells in a 10cm dish were transfected with
1034 2µg each of the following plasmids (pCI PR8 PB2, pCI PR8 PB1, pCI PR8 PA-HA tag,
1035 pCI PR8 NP:WT/F346S/R416A, & pHH21 PR8 NA) using JetPrime (Polyplus) according
1036 to the manufacturer's instructions.
1037
1038 **Co-immunoprecipitation**
1039 24hrs post transfection, the cells were lysed in MOPS Co-IP Lysis Buffer (20mM MOPS
1040 pH=7.5, 150mM NaCl, 0.5% Igepal CA-630,1x cOmplete EDTA-free Protease Inhibitor
1041 Cocktail (Sigma-Aldrich)) and clarified via centrifugation (20,000xg, 15min, 4°C).
1042 Mouse-anti-His (HIS.H8) antibody (Invitrogen) (for eGFP/His-tagged NP protein co-
1043 immunoprecipitation) or Mouse-anti-HA tag (2-2.2.14) antibody (Invitrogen) (for HA-
1044 tagged vRNP complex pulldown experiment) was added to the clarified lysates to a final
1045 dilution of 1:100, and the lysates were incubated with the antibody with end-over-end
1046 rotation overnight at 4°C. Antigen-antibody complexes were selectively purified using
1047 Pierce Protein A Agarose (ThermoFisher Scientific) according to the manufacturer's
1048 instructions.
1049
1050 **Western blot**
1051 For eGFP/His-tagged NP protein co-immunoprecipitation experiment: Western blots
1052 were performed on the immunoprecipitated protein samples using the mouse-anti-eGFP
1053 (F56-6A.1.2.3) primary antibody (Invitrogen) (1:1,000) followed by the rat-anti-mouse-
1054 HRP conjugated (187.1) secondary antibody (BD Biosciences) (1:500) to detect co-
1055 immunoprecipitated PR8 NP:WT/F346S/R416A eGFP tagged proteins, or the 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 NP proteins in transfected cells.

For HA-tagged vRNP-complex pulldown experiment: Western blots were performed on the immunoprecipitated protein samples using the 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 the rabbit anti-NP polyclonal antibody (GeneTex GTX125989) (1:1,000) followed by the goat-anti-rabbit-HRP conjugated (G-21234) secondary antibody (Invitrogen) (1:10,000) to detect co-immunoprecipitated NP.

Proteins were visualized using the SuperSignal Pico West Plus Chemiluminescent Substrate (ThermoFisher Scientific) and imaged using the iBright CL1000 Imaging System (Invitrogen).

Quantifying gene segment ratios in viral stocks

140µL of the viral stock was treated with 0.25µg of RNaseA for 30min at 37°C. The viral RNA was then extracted using the QIAamp Viral RNA Extraction Kit (Qiagen) and eluted in 60µL of nuclease-free water (Ambion). RNA was treated with RNase-free DNaseI (Qiagen) and cleaned using the RNeasy Kit (Qiagen). Reverse transcription was performed using the Verso cDNA Synthesis Kit (ThermoFisher). The reactions were set up as follows: 4µL RNA + 4µL 5x cDNA Synthesis Buffer + 2µL dNTP Mix + 1µL 10µM PR8 RT_4A primer (5'-AGCAAAAGCAGG-3') + 1µL RT Enhancer + 1µL Verso Enzyme Mix + 7µL Nuclease-free water, and incubated at 45°C for 50min, 95°C for 2min, 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). The strand-specific forward and reverse primers for quantitative real-time PCR for PR8 HA, NP, and NA/PR8 NA ORF HA UTRs were 5'-AAGGCAAACCTACTGGTCCTGTT-3' & 5'-AATTGTTCGCATGGTAGCCTATAC-3', 5'-AGGCACCAAACGGTCTTACG-3' & 5'-TTCCGACGGATGCTCTGATT-3', and 5'-AAATCAGAAAATAACAACCATTGGA-3' & 5'-ATTCCCTATTTGCAATATTAGGCT-3' respectively. The strand-specific forward and reverse primers for Udorn HA, NP, and NA were 5'-GACTATCATTGCTTTGAGC-3' & 5'- CACTAGTGTTCCGTTTGGC-3' and 5'-CGGTCTTATGAACAGATGG-3' & 5'-TCGTCCAATTCCATCAATC-3' and 5'-AACAATTGGCTCTGTCTCTC-3' & 5'-GTCGCACTCATATTGCTTG-3' respectively. The primers for the PR8 HA ORF NA UTR+Pack and PR8 NA ORF HA UTR+Pack segments were 5'-AAATCAGAAAATAACAACCATTGGA-3' & 5'-CAACAATACCAACAGATTAGC-3' and 5'-AAGGCAAACCTACTGGTCCTGTT-3' & 5'-ATCAGCCCTACCACGAGGC-3' respectively. The primers for the PR8 HA ORF NA UTR segment were 5'-CAGGAGTGCCAAATTGAGGATGG-3' and 5'- CCGGCAATGGCTCCAAATAGACC-3'.

Reactions were set up as follows: 2µL cDNA + 10µL 2x Power SYBR Green MM + 1µL 10µM Forward Primer + 1µL 10µM Reverse Primer + 6µL Nuclease-free water. The cycling conditions were as follows: 50°C for 2min, 95°C for 10min, and then 95°C for 15 sec followed by 60°C for 1min repeated 40x.

Acknowledgments

This research was funded in whole, or in part, by the Wellcome Trust [FC011104]. For the purpose of Open Access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. We would like to thank Tongyu Liu and Jiayi Sun for providing plasmids and members of the Brooke lab for their feedback on this manuscript. Additionally, we would like to thank Sonya Kumar Bharathkar for assistance with mammalian protein expression in HEK Expi-293-F cells and Sarah Leonard for helpful discussions regarding NP-RNA binding experiments. We are also grateful to Drs. Ervin Fodor and Andy Mehle for helpful discussions. 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), and the Wellcome Trust (FC011104), and startup funds from the University of Illinois.

References

1. Krammer F, Smith GJD, Fouchier RAM, Peiris M, Kedzierska K, Doherty PC, et al. Influenza. Nat Rev Dis Primer. 2018 Jun 28;4(1):3.

2. Fodor E, Velthuis AJW te. Structure and Function of the Influenza Virus Transcription and Replication Machinery. Cold Spring Harb Perspect Med. 2019 Dec 23;a038398.

3. Shapiro GI, Gurney T, Krug RM. Influenza virus gene expression: control mechanisms at early and late times of infection and nuclear-cytoplasmic transport of virus-specific RNAs. J Virol. 1987 Mar 1;61(3):764–73.

4. Phan T, Fay EJ, Lee Z, Aron S, Hu W-S, Langlois RA. Segment-Specific Kinetics of mRNA, cRNA, and vRNA Accumulation during Influenza Virus Infection. J Virol [Internet]. 2021 Apr 26;95(10). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8139675/

5. Zheng H, Palese P, García-sastre A. Nonconserved Nucleotides at the 3’ and 5’ Ends of an Influenza A Virus RNA Play an Important Role in Viral RNA Replication. Virology. 1996 Mar 1;217(1):242–51.

6. Wang J, Li J, Zhao L, Cao M, Deng T. Dual Roles of the Hemagglutinin Segment-Specific Noncoding Nucleotides in the Extended Duplex Region of the Influenza A Virus RNA Promoter. J Virol [Internet]. 2016 Dec 16;91(1). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5165193/

7. Y X, W Z, M P, Dlv B, Y B, M C, et al. A synergistic effect between 3’ terminal noncoding and adjacent coding regions of influenza A virus HA segment on template preference [Internet].

1142 Journal of virology. 2021 [cited 2021 Aug 24]. Available from:
1143 https://pubmed.ncbi.nlm.nih.gov/34190596/

1144 8. J W, Y P, L Z, M C, T H, T D. Influenza A virus utilizes a suboptimal Kozak sequence to fine-tune virus
1145 replication and host response [Internet]. The Journal of general virology. 2015 [cited 2021 Jul 30].
1146 Available from: https://pubmed.ncbi.nlm.nih.gov/25519170/

1147 9. J M, K L, C X, J Z, S X, Y R, et al. Impact of the segment-specific region of the 3'-untranslated region
1148 of the influenza A virus PB1 segment on protein expression [Internet]. Virus genes. 2013 [cited
1149 2021 Jul 30]. Available from: https://pubmed.ncbi.nlm.nih.gov/23949786/

1150 10. Maeda Y, Goto H, Horimoto T, Takada A, Kawaoka Y. Biological significance of the U residue at the
1151 −3 position of the mRNA sequences of influenza A viral segments PB1 and NA. Virus Res. 2004 Mar
1152 15;100(2):153–7.

1153 11. Babar MM, Zaidi N-SS. Protein sequence conservation and stable molecular evolution reveals
1154 influenza virus nucleoprotein as a universal druggable target. Infect Genet Evol. 2015 Aug
1155 1;34(Supplement C):200–10.

1156 12. Ye Q, Krug RM, Tao YJ. The mechanism by which influenza A virus nucleoprotein forms oligomers
1157 and binds RNA. Nature. 2006 Dec 21;444(7122):1078–82.

1158 13. Ng AK-L, Zhang H, Tan K, Li Z, Liu J, Chan PK-S, et al. Structure of the influenza virus A H5N1
1159 nucleoprotein: implications for RNA binding, oligomerization, and vaccine design. FASEB J. 2008
1160 Oct;22(10):3638–47.

1161 14. Chan W-H, Ng AK-L, Robb NC, Lam MK-H, Chan PK-S, Au SW-N, et al. Functional Analysis of the
1162 Influenza Virus H5N1 Nucleoprotein Tail Loop Reveals Amino Acids That Are Crucial for
1163 Oligomerization and Ribonucleoprotein Activities. J Virol. 2010 Jul 15;84(14):7337–45.

1164 15. Turrell L, Lyall JW, Tiley LS, Fodor E, Vreede FT. The role and assembly mechanism of nucleoprotein
1165 in influenza A virus ribonucleoprotein complexes. Nat Commun. 2013;4:1591.

1166 16. Wang P, Palese P, O'Neill RE. The NPI-1/NPI-3 (karyopherin alpha) binding site on the influenza a
1167 virus nucleoprotein NP is a nonconventional nuclear localization signal. J Virol. 1997
1168 Mar;71(3):1850–6.

1169 17. Tome-Amat J, Ramos I, Amanor F, Fernández-Sesma A, Ashour J. Influenza A Virus Utilizes Low-
1170 Affinity, High-Avidity Interactions with the Nuclear Import Machinery To Ensure Infection and
1171 Immune Evasion. J Virol. 2019 Jan 1;93(1):e01046-18.

1172 18. Eisfeld AJ, Neumann G, Kawaoka Y. At the centre: influenza A virus ribonucleoproteins. Nat Rev
1173 Microbiol. 2015 Jan;13(1):28–41.

1174 19. Moreira ÉA, Weber A, Bolte H, Kolesnikova L, Giese S, Lakdawala S, et al. A conserved influenza A
1175 virus nucleoprotein code controls specific viral genome packaging. Nat Commun [Internet]. 2016
1176 [cited 2017 Mar 23];7. Available from: https://www-ncbi-nlm-nih-
1177 gov.proxy2.library.illinois.edu/pmc/articles/PMC5035998/

1178 20. Bolte H, Rosu ME, Hagelauer E, García-Sastre A, Schwemmle M. Packaging of the Influenza Virus
1179 Genome Is Governed by a Plastic Network of RNA- and Nucleoprotein-Mediated Interactions. J
1180 Virol. 2019 Feb 15;93(4):e01861-18.

1181 21. Dadonaite B, Gilbertson B, Knight ML, Trifkovic S, Rockman S, Laederach A, et al. The structure of
1182 the influenza A virus genome. Nat Microbiol. 2019 Jul 22;1.

1183 22. Williams GD, Townsend D, Wylie KM, Kim PJ, Amarasinghe GK, Kutluay SB, et al. Nucleotide
1184 resolution mapping of influenza A virus nucleoprotein-RNA interactions reveals RNA features
1185 required for replication. Nat Commun [Internet]. 2018 Jan 31;9. Available from:
1186 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5792457/

1187 23. Ince WL, Gueye-Mbaye A, Bennink JR, Yewdell JW. Reassortment Complements Spontaneous
1188 Mutation in Influenza A Virus NP and M1 Genes To Accelerate Adaptation to a New Host. J Virol.
1189 2013 Apr 15;87(8):4330–8.

1190 24. Brooke CB, Ince WL, Wei J, Bennink JR, Yewdell JW. Influenza A virus nucleoprotein selectively
1191 decreases neuraminidase gene-segment packaging while enhancing viral fitness and
1192 transmissibility. Proc Natl Acad Sci U S A. 2014 Nov 25;111(47):16854–9.

1193 25. Vreede FT, Jung TE, Brownlee GG. Model Suggesting that Replication of Influenza Virus Is
1194 Regulated by Stabilization of Replicative Intermediates. J Virol. 2004 Sep;78(17):9568–72.

1195 26. Alnaji FG, Holmes JR, Rendon G, Vera JC, Fields CJ, Martin BE, et al. Sequencing Framework for the
1196 Sensitive Detection and Precise Mapping of Defective Interfering Particle-Associated Deletions
1197 across Influenza A and B Viruses. J Virol [Internet]. 2019 May 15;93(11). Available from:
1198 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6532088/

1199 27. Jorge DM de M, Mills RE, Lauring AS. CodonShuffle: a tool for generating and analyzing
1200 synonymously mutated sequences. Virus Evol [Internet]. 2015 Oct 2;1(1). Available from:
1201 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5014483/

1202 28. Zheng A, Sun W, Xiong X, Freyn AW, Peukes J, Strohmeier S, et al. Enhancing Neuraminidase
1203 Immunogenicity of Influenza A Viruses by Rewiring RNA Packaging Signals. J Virol [Internet]. 2020
1204 Jul 30;94(16). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7394900/

1205 29. Di Tommaso P, Moretti S, Xenarios I, Orobitg M, Montanyola A, Chang J-M, et al. T-Coffee: a web
1206 server for the multiple sequence alignment of protein and RNA sequences using structural
1207 information and homology extension. Nucleic Acids Res. 2011 Jul;39(Web Server issue):W13-17.

1208 30. Elton D, Medcalf L, Bishop K, Harrison D, Digard P. Identification of Amino Acid Residues of
1209 Influenza Virus Nucleoprotein Essential for RNA Binding. J Virol. 1999 Sep;73(9):7357–67.

1210 31. Albo C, Valencia A, Portela A. Identification of an RNA binding region within the N-terminal third of
1211 the influenza A virus nucleoprotein. J Virol. 1995 Jun;69(6):3799–806.

1212 32. Kobayashi M, Toyoda T, Adyshev DM, Azuma Y, Ishihama A. Molecular dissection of influenza virus
1213 nucleoprotein: deletion mapping of the RNA binding domain. J Virol. 1994 Dec;68(12):8433–6.

1214 33. Turrell L, Hutchinson EC, Vreede FT, Fodor E. Regulation of Influenza A Virus Nucleoprotein
1215 Oligomerization by Phosphorylation. J Virol. 2015 Jan 15;89(2):1452–5.

1216 34. Mondal A, Potts GK, Dawson AR, Coon JJ, Mehle A. Phosphorylation at the Homotypic Interface
1217 Regulates Nucleoprotein Oligomerization and Assembly of the Influenza Virus Replication
1218 Machinery. PLOS Pathog. 2015 Apr 13;11(4):e1004826.

1219 35. Vreede FT, Ng AK-L, Shaw P-C, Fodor E. Stabilization of Influenza Virus Replication Intermediates Is
1220 Dependent on the RNA-Binding but Not the Homo-Oligomerization Activity of the Viral
1221 Nucleoprotein ▿. J Virol. 2011 Nov;85(22):12073–8.

1222 36. Elton D, Medcalf E, Bishop K, Digard P. Oligomerization of the Influenza Virus Nucleoprotein:
1223 Identification of Positive and Negative Sequence Elements. Virology. 1999 Jul 20;260(1):190–200.

1224 37. Tarus B, Bakowiez O, Chenavas S, Duchemin L, Estrozi LF, Bourdieu C, et al. Oligomerization paths
1225 of the nucleoprotein of influenza A virus. Biochimie. 2012 Mar 1;94(3):776–85.

1226 38. Shu LL, Bean WJ, Webster RG. Analysis of the evolution and variation of the human influenza A
1227 virus nucleoprotein gene from 1933 to 1990. J Virol. 1993 May 1;67(5):2723–9.

1228 39. Robb NC, Chase G, Bier K, Vreede FT, Shaw P-C, Naffakh N, et al. The Influenza A Virus NS1 Protein
1229 Interacts with the Nucleoprotein of Viral Ribonucleoprotein Complexes. J Virol. 2011 May
1230 15;85(10):5228–31.

1231 40. Newcomb LL, Kuo R-L, Ye Q, Jiang Y, Tao YJ, Krug RM. Interaction of the Influenza A Virus
1232 Nucleocapsid Protein with the Viral RNA Polymerase Potentiates Unprimed Viral RNA Replication. J
1233 Virol. 2009 Jan 1;83(1):29–36.

1234 41. Marklund JK, Ye Q, Dong J, Tao YJ, Krug RM. Sequence in the Influenza A Virus Nucleoprotein
1235 Required for Viral Polymerase Binding and RNA Synthesis. J Virol. 2012 Jul;86(13):7292–7.

1236 42. Morris AK, Wang Z, Ivey AL, Xie Y, Hill PS, Schey KL, et al. Cellular mRNA export factor UAP56
1237 recognizes nucleic acid binding site of influenza virus NP protein. Biochem Biophys Res Commun
1238 [Internet]. 2020 Feb 19; Available from:
1239 http://www.sciencedirect.com/science/article/pii/S0006291X20303272

1240 43. Momose F, Basler CF, O'Neill RE, Iwamatsu A, Palese P, Nagata K. Cellular Splicing Factor RAF-
1241 2p48/NPI-5/BAT1/UAP56 Interacts with the Influenza Virus Nucleoprotein and Enhances Viral RNA
1242 Synthesis. J Virol. 2001 Feb;75(4):1899–908.

1243 44. Mayer D, Molawi K, Martínez-Sobrido L, Ghanem A, Thomas S, Baginsky S, et al. Identification of
1244 Cellular Interaction Partners of the Influenza Virus Ribonucleoprotein Complex and Polymerase
1245 Complex Using Proteomic-Based Approaches. J Proteome Res. 2007 Feb 1;6(2):672–82.

1246 45. Lin Y-C, Jeng K-S, Lai MMC. CNOT4-Mediated Ubiquitination of Influenza A Virus Nucleoprotein
1247 Promotes Viral RNA Replication. mBio [Internet]. 2017 May 23;8(3). Available from:
1248 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5442456/

1249 46. Minakuchi M, Sugiyama K, Kato Y, Naito T, Okuwaki M, Kawaguchi A, et al. Pre-mRNA Processing
1250 Factor Prp18 Is a Stimulatory Factor of Influenza Virus RNA Synthesis and Possesses Nucleoprotein
1251 Chaperone Activity. J Virol. 2017 Feb 1;91(3):e01398-16.

1252 47. Perez JT, Zlatev I, Aggarwal S, Subramanian S, Sachidanandam R, Kim B, et al. A Small-RNA
1253 Enhancer of Viral Polymerase Activity. J Virol. 2012 Dec;86(24):13475–85.

1254 48. Perez JT, Varble A, Sachidanandam R, Zlatev I, Manoharan M, García-Sastre A, et al. Influenza A
1255 virus-generated small RNAs regulate the switch from transcription to replication. Proc Natl Acad
1256 Sci U S A. 2010 Jun 22;107(25):11525–30.

1257 49. Lee M-K, Bae S-H, Park C-J, Cheong H-K, Cheong C, Choi B-S. A single-nucleotide natural variation
1258 (U4 to C4) in an influenza A virus promoter exhibits a large structural change: implications for
1259 differential viral RNA synthesis by RNA-dependent RNA polymerase. Nucleic Acids Res. 2003 Feb
1260 15;31(4):1216–23.

1261 50. Kh L, Bl S. The position 4 nucleotide at the 3’ end of the influenza virus neuraminidase vRNA is
1262 involved in temporal regulation of transcription and replication of neuraminidase RNAs and affects
1263 the repertoire of influenza virus surface antigens [Internet]. The Journal of general virology. 1998
1264 [cited 2021 Nov 13]. Available from: https://pubmed.ncbi.nlm.nih.gov/9714240/

1265 51. Wang L, Lee C-W. Sequencing and mutational analysis of the non-coding regions of influenza A
1266 virus. Vet Microbiol. 2009 Mar 30;135(3):239–47.

1267 52. Widjaja I, de Vries E, Rottier PJM, de Haan CAM. Competition between Influenza A Virus Genome
1268 Segments. PLoS ONE [Internet]. 2012 Oct 11;7(10). Available from:
1269 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3469491/

1270 53. McAuley JL, Gilbertson BP, Trifkovic S, Brown LE, McKimm-Breschkin JL. Influenza Virus
1271 Neuraminidase Structure and Functions. Front Microbiol [Internet]. 2019 Jan 29;10. Available
1272 from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6362415/

1273 54. Byrd-Leotis L, Cummings RD, Steinhauer DA. The Interplay between the Host Receptor and
1274 Influenza Virus Hemagglutinin and Neuraminidase. Int J Mol Sci. 2017 Jul 17;18(7):1541.

1275 55. Benton DJ, Martin SR, Wharton SA, McCauley JW. Biophysical Measurement of the Balance of
1276 Influenza A Hemagglutinin and Neuraminidase Activities. J Biol Chem. 2015 Mar 6;290(10):6516–
1277 21.

1278 56. Wagner R, Matrosovich M, Klenk H-D. Functional balance between haemagglutinin and
1279 neuraminidase in influenza virus infections. Rev Med Virol. 2002 May 1;12(3):159–66.

1280 57. Gaymard A, Le Briand N, Frobert E, Lina B, Escuret V. Functional balance between neuraminidase
1281 and haemagglutinin in influenza viruses. Clin Microbiol Infect. 2016 Dec;22(12):975–83.

1282 58. Casalegno J-S, Ferraris O, Escuret V, Bouscambert M, Bergeron C, Linès L, et al. Functional Balance
1283 between the Hemagglutinin and Neuraminidase of Influenza A(H1N1)pdm09 HA D222 Variants.
1284 PLOS ONE. 2014 Aug 13;9(8):e104009.

1285 59. Xu R, Zhu X, McBride R, Nycholat CM, Yu W, Paulson JC, et al. Functional Balance of the
1286 Hemagglutinin and Neuraminidase Activities Accompanies the Emergence of the 2009 H1N1
1287 Influenza Pandemic. J Virol. 2012 Sep 1;86(17):9221–32.

1288 60. Yen H-L, Liang C-H, Wu C-Y, Forrest HL, Ferguson A, Choy K-T, et al. Hemagglutinin–neuraminidase
1289 balance confers respiratory-droplet transmissibility of the pandemic H1N1 influenza virus in
1290 ferrets. Proc Natl Acad Sci. 2011 Aug 23;108(34):14264–9.

1291 61. Guo H, Rabouw H, Slomp A, Dai M, Vegt F van der, Lent JWM van, et al. Kinetic analysis of the
1292 influenza A virus HA/NA balance reveals contribution of NA to virus-receptor binding and NA-
1293 dependent rolling on receptor-containing surfaces. PLOS Pathog. 2018 Aug 13;14(8):e1007233.

1294 62. Lai JCC, Karunarathna HMTK, Wong HH, Peiris JSM, Nicholls JM. Neuraminidase activity and
1295 specificity of influenza A virus are influenced by haemagglutinin-receptor binding. Emerg Microbes
1296 Infect. 2019 Feb 27;8(1):327–38.

1297 63. Bhatt S, Holmes EC, Pybus OG. The Genomic Rate of Molecular Adaptation of the Human Influenza
1298 A Virus. Mol Biol Evol. 2011 Sep;28(9):2443–51.

1299 64. Doud MB, Hensley SE, Bloom JD. Complete mapping of viral escape from neutralizing antibodies.
1300 PLoS Pathog [Internet]. 2017 Mar 13;13(3). Available from:
1301 http://www.ncbi.nlm.nih.gov/pmc/articles/PMC5363992/

1302 65. Dai M, Du W, Martínez-Romero C, Leenders T, Wennekes T, Rimmelzwaan GF, et al. Analysis of the
1303 Evolution of Pandemic Influenza A(H1N1) Virus Neuraminidase Reveals Entanglement of Different
1304 Phenotypic Characteristics. mBio. 12(3):e00287-21.

1305 66. Das SR, Hensley SE, David A, Schmidt L, Gibbs JS, Puigbò P, et al. Fitness costs limit influenza A virus
1306 hemagglutinin glycosylation as an immune evasion strategy. Proc Natl Acad Sci. 2011 Dec
1307 20;108(51):E1417–22.

1308 67. Das SR, Hensley SE, Ince WL, Brooke CB, Subba A, Delboy MG, et al. Defining Influenza A Virus
1309 Hemagglutinin Antigenic Drift by Sequential Monoclonal Antibody Selection. Cell Host Microbe.
1310 2013 Mar 13;13(3):314–23.

1311 68. Hensley SE, Das SR, Bailey AL, Schmidt LM, Hickman HD, Jayaraman A, et al. Hemagglutinin
1312 Receptor Binding Avidity Drives Influenza A Virus Antigenic Drift. Science. 2009 Oct
1313 30;326(5953):734–6.

1314 69. Hensley SE, Das SR, Gibbs JS, Bailey AL, Schmidt LM, Bennink JR, et al. Influenza A Virus
1315 Hemagglutinin Antibody Escape Promotes Neuraminidase Antigenic Variation and Drug Resistance.
1316 PLOS ONE. 2011 Feb 22;6(2):e15190.

1317 70. Prachanronarong KL, Canale AS, Liu P, Somasundaran M, Hou S, Poh Y-P, et al. Mutations in
1318 Influenza A Virus Neuraminidase and Hemagglutinin Confer Resistance against a Broadly
1319 Neutralizing Hemagglutinin Stem Antibody. J Virol. 2019 Jan 15;93(2):e01639-18.

71. Kosik I, Angeletti D, Gibbs JS, Angel M, Takeda K, Kosikova M, et al. Neuraminidase inhibition contributes to influenza A virus neutralization by anti-hemagglutinin stem antibodies. J Exp Med. 2019 Feb 4;216(2):304–16.

72. Kosik I, Ince WL, Gentles LE, Oler AJ, Kosikova M, Angel M, et al. Influenza A virus hemagglutinin glycosylation compensates for antibody escape fitness costs. PLoS Pathog [Internet]. 2018 Jan 18;14(1). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5773227/

Figure captions

Fig 1. NP:F346S affects NA vRNA replication but not primary transcription. A.) *Abundances of NA and HA vRNA (measured by RT-qPCR on total cellular RNA) at the indicated timepoints following infection of MDCK cells at an MOI of 0.1 NP-expressing units (NPEU)/cell under single cycle conditions. Data represent values obtained during infection with PR8-NP:F346S, normalized to values obtained during infection with PR8-NP:WT. The data shown are individual cell culture well replicates representative of the data obtained through two similar experiments.* ***B.)*** *NA mRNA abundances (measured by RT-qPCR on total cellular RNA) in PR8-NP:F346S-infected MDCK-SIAT1 cells, normalized to values obtained during infection with PR8-NP:WT. Infections were initiated at MOI=5 $TCID_{50}$/cell in the presence of 100µg/mL cycloheximide. Data points indicate individual cell culture well replicates pooled from two independent experiments.* ***C.)*** *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 $TCID_{50}$/cell for 7hrs, followed by 1hr pulse with 500µM 4SU. Cellular RNA was then harvested and the abundance of 4SU-labeled viral RNAs were determined by RT-qPCR using a universal, vRNA-sense specific primer for the RT reaction followed by segment-specific primers for the qPCR. Data points indicate individual cell culture well replicates pooled from two independent experiments.*

Fig 2. Susceptibility to the effects of NP:F346S is NA segment genotype specific. A.) *Normalized vRNA abundances as determined by qRT-PCR in PR8-NP:F346S or Udorn-NP:F346S infected MDCK cells (MOI=0.1 NPEU/cell, 8hpi) expressed as fraction of PR8 NP:WT or Udorn NP:WT respectively. Secondary infection was blocked via the addition of ammonium chloride at 3hpi. The data points represent individual cell culture well replicates representative of the data obtained through two similar experiments.* ***B.)*** *Viral protein expression levels as determined by geometric mean fluorescence intensity (GMFI) in rPR8 Udorn HA/NA NP:F346S infected MDCK cells (MOI=0.03 $TCID_{50}$/cell, 16hpi) expressed as a percentage of rPR8 Udorn HA/NA NP:WT. The data shown are individual cell culture well replicates representative of the data obtained through two similar experiments.*

Fig 3. Susceptibility to NP-dependent regulation maps to the UTRs of the NA segment. A.) *Schematic depiction of the codon shuffled PR8 NA construct. The Codon Shuffle program was used to introduce 227 silent mutations within the region encompassing nucleotides 38-1319 of the PR8 NA segment to alter features of the RNA sequence while minimizing changes in codon frequencies or dinucleotide content.* ***B.)*** *Relative 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, 8hpi) as determined by RT-qPCR on cellular RNA expressed as a fraction of PR8 NA Codon Shuffle NP:WT respectively. Each data point represents a cell culture well replicate pooled from two separate experiments. ***C.)*** *Schematic depictions of the PR8 HA/NA UTR+Pack Swap and PR8 HA/NA UTR Swap gene segments. The PR8 HA/NA UTR+Pack Swap segments were generated by replacing the UTRs and packaging signal regions of one segment (HA/NA) with those of the other segment (NA/HA). The start codon of the newly appended packaging signal for each segment was mutated to prevent the expression of any protein encoded by the packaging signal sequence. The packaging signals within the native ORFs were disrupted via the addition of silent substitutions to all codons to prevent duplication of the packaging signals in the swapped segments. The PR8 HA/NA UTR Swap gene segments were generated by swapping the UTRs of the PR8 HA/NA segments.* ***D.)*** *Relative 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, 8hpi) 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. Each data point represents a cell culture well replicate pooled from two separate experiments.*

Fig 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' & 5' UTRs using the M-Coffee alignment algorithm on the T-Coffee web server (29). Regions of interest are boxed in red. PR8 NA nucleotide numbering is shown.* ***B.)*** *Relative abundances of the HA and NA segments in MDCK cells infected with the PR8 NP:F346S, PR8:NA$^{Udorn\ UTR}$ NP:F346S (MOI=0.1 NPEU/cell, 8hpi) viruses as determined by RT-qPCR expressed as a fraction of PR8 NP:WT and PR8:NA$^{Udorn\ UTR}$ NP:WT respectively. Each data point represents a cell culture well replicate pooled from two independent experiments.* ***C.)*** *Relative abundances of the HA and NA segments in MDCK cells infected with PR8:Udorn HA, NA$^{PR8\ UTR}$ NP:F346S (MOI=0.03 NPEU/cell, 8hpi) virus as determined by RT-qPCR expressed as a fraction of PR8:Udorn HA, NA$^{PR8\ UTR}$ NP:WT. Each data point represents a cell culture well replicate from a single experiment.*

Fig 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. Sequences shown in negative sense, 3'->5'.*

Fig 6. The effect of mutations in the PR8 NA UTRs on baseline expression levels and sensitivity to NP:F346S. A.) *Relative abundances of the HA and NA segments at 8hpi in MDCK cells infected with the indicated viruses encoding NP:F346S at MOI=0.1 NPEU/cell, as determined by qRT-PCR normalized to the NP:WT-encoding versions of the same viruses. Each data point represents an individual cell culture well replicate pooled from two independent experiments.* ***B.)*** *Relative abundances of the HA and NA segments in MDCK cells infected with the indicated viruses encoding NP:F346S (MOI=0.1 NPEU/cell, 8hpi), as determined by qRT-PCR normalized to the NP:WT-encoding versions of the same viruses. Each data point represents an individual cell*

culture well replicate pooled from two independent experiments. ***C.)*** *Relative abundances of the HA and NA segments in MDCK cells infected with the PR8:Udorn HA, NA$^{PR8\ ORF\ Proximal\ UTR}$ NP:F346S virus (MOI=0.03 NPEU/cell, 8hpi) as determined by qRT-PCR normalized to PR8:Udorn HA, NA$^{PR8\ ORF\ Proximal\ UTR}$ NP:WT. Each data point represents an individual cell culture well replicate from a single experiment.* ***D,E.)*** *Data from experiments shown in* ***(4B and 6A,B)*** *and* ***(4C/6C)*** *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 8hpi) as determined by qRT-PCR on total cellular RNA. The data represents two cell culture well replicates pooled from either two independent experiments* ***(D)*** *or a single experiment* ***(E)****.*

Fig 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 is colored blue/green and the fitted data is colored orange/purple for the NP:WT/F346S-C-His proteins, respectively.* ***B.)*** *Co-immunoprecipitation (IP) of eGFP and His-tagged versions of the indicated NP proteins. 293T cells were transfected with expression vectors encoding the eGFP- and His-tagged versions of either WT, F346S, or R416A NP proteins. Lysates were harvested after 24hrs. His-tagged NP was immunoprecipitated, and then IP samples were probed via western blot with anti-eGFP and anti-6x His antibodies. Western blots of total cell lysates stained with an anti-NP antibody are also shown.* ***C.)*** *Co-IP of vRNP-associated NP and PA. Cells were transfected with plasmids encoding the vRNP complex (PB2, PB1, PA-HA-tag, and NP (WT, F346S, or R416A) and a vRNA template (NA vRNA). Lysates were harvested 24hrs post transfection, and vRNP complexes were IP-ed using an anti-HA-tag antibody. Undiluted, 1:5 diluted, or 1:10 diluted IP-ed protein was probed with anti-NP and anti-HA-tag antibodies via western blot. Western blots of whole cell lysates shown for comparison.*

Fig 8. A cluster of aromatic residues is involved in the regulation of NA gene segment expression. A.) *Normalized NA protein expression levels in cells infected with the indicated PR8 NP 346 mutant viruses (MOI=0.03 $TCID_{50}$/cell, 16hpi) as determined by geometric mean fluorescent intensity (GMFI) expressed as a fraction of PR8 NP:WT. The data shown are 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 2IQH).* ***C.)*** *Normalized viral RNA abundance in PR8 NP:F346S, PR8 NP:Y385A and PR8 NP:F479A infected MDCK cells (MOI=0.1 $TCID_{50}$/cell, 8hpi) as determined by RT-qPCR and expressed as a fraction of PR8 NP:WT. The data shown are individual cell culture well replicates representative of the data obtained through two similar experiments.*

SI figure legends

S1 Fig. Quantifying the abundance of newly synthesized, 4SU-labeled vRNAs using vRNA and segment-specific primers during the cDNA synthesis and qPCR steps. *Normalized abundance of 4SU-labeled NA vRNA in MDCK cells infected with PR8 NP:WT/F346S at an MOI of 5 $TCID_{50}$/cell for 7hrs and pulsed with 500μM of 4SU*

1454 *for 1hr as determined by RT-qPCR using a tagged, vRNA and segment-specific primer*
1455 *during the cDNA synthesis step, and a primer pair consisting of a tag-specific primer*
1456 *and segment-specific primer for the qPCR step. Each data point represents a single cell*
1457 *culture well replicate from a single experiment.*

1458 ***S2 Fig. Quantification of gene segment ratios in viral RNA stocks.*** *140µL of viral*
1459 *RNA supernatant was treated with 0.25µg RNaseA, viral RNA was extracted and*
1460 *DNase-treated, and then viral gene segment abundance was quantified using RT-*
1461 *qPCR.* ***A/B.)*** *Normalized HA ORF* ***(A)*** *or NA ORF* ***(B)*** *containing viral RNA abundance*
1462 *in the viral stocks of the PR8 HA/NA UTR+Pack Swap NP:WT/F346S and PR8 HA/NA*
1463 *UTR Swap NP:WT/F346S viruses.* ***C.)*** *Normalized viral RNA abundance in the viral*
1464 *stocks of the Udorn NP:F346S, PR8:Udorn HA, NA NP:F346S, PR8:Udorn HA, NA*^{PR8}
1465 ^{UTR} *NP:F346S, and PR8:Udorn HA, NA*$^{PR8\ ORF\ Proximal\ UTR}$ *NP:F346S viruses relative to*
1466 *Udorn NP:WT, PR8:Udorn HA, NA NP:WT, PR8:Udorn HA, NA:*$^{PR8\ UTR}$ *NP:WT, and*
1467 *PR8:Udorn HA, NA*$^{PR8\ ORF\ Proximal\ UTR}$ *NP:WT viruses respectively.* ***D.)*** *Normalized viral*
1468 *RNA abundance in the viral stocks of the indicated viruses with NP:F346S as*
1469 *determined by RT-qPCR normalized to the NP:WT versions of each virus. Each data*
1470 *point represents a qPCR technical replicate.*