

Prevalence and mechanisms of evolutionary contingency in human influenza H3N2 neuraminidase

Ruipeng Lei¹, Timothy J.C. Tan², Andrea Hernandez Garcia¹, Yiquan Wang¹, Chuyun Teo¹,
Gopika Gopan³, Zahra Tavakoli Dargani¹, Qi Wen Teo^{1,4}, Claire S. Graham¹, Satish K. Nair^{1,2,4},
Nicholas C. Wu^{1,2,4,5,§}

¹ Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

² Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

³ Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

⁴ Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

⁵ Carle Illinois College of Medicine, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[§] To whom correspondence may be addressed. Email: nicwu@illinois.edu

ABSTRACT

Neuraminidase (NA) of human influenza H3N2 virus has evolved rapidly and been accumulating mutations for more than half-century. However, biophysical constraints that govern the evolutionary trajectories of NA remain largely elusive. Here, we show that among 70 natural mutations that are present in the NA of a recent human H3N2 strain, >10% are deleterious for an ancestral strain. By mapping the permissive mutations using combinatorial mutagenesis and next-generation sequencing, an extensive epistatic network is revealed. Biophysical and structural analyses further demonstrate that certain epistatic interactions can be explained by non-additive stability effect, which in turn modulates membrane trafficking and enzymatic activity of NA. Additionally, our results suggest that other biophysical mechanisms also contribute to epistasis in NA evolution. Overall, these findings not only provide mechanistic insights into the evolution of human influenza NA and elucidate its sequence-structure-function relationship, but also have important implications for the development of next-generation influenza vaccines.

INTRODUCTION

Influenza A virus, which causes around 300,000 to 650,000 deaths yearly over the world (Iuliano et al., 2018), continues to be a major global health concern. Influenza A virus has two surface glycoproteins, hemagglutinin (HA) and neuraminidase (NA), that act as the major antigens. Because of the need to constantly escape from herd immunity, both HA and NA of human influenza A virus evolve rapidly. Most studies on the evolution of human influenza A virus focus on HA, due to its dominant role in vaccine development (Wu and Wilson, 2017). Nevertheless, recent studies have shown that NA immunity also contributes to protection against influenza infection (Couch et al., 2013; Memoli et al., 2016; Monto et al., 2015; Weiss et al., 2020). These findings suggest that NA can be an effective vaccine target (Krammer et al., 2018). However, there is a lack of knowledge about the evolutionary biology of NA.

NA is a type II transmembrane protein that forms a homotetramer. The natural function of NA is to cleave the sialic acids from cellular receptors to facilitate virus release and from newly synthesized HA to prevent virion aggregation (McAuley et al., 2019). After entering the human population in 1968, H3N2 has accumulated around 75 amino acid mutations in NA, which account for 16% of the entire protein (Wang et al., 2021). Such an extensive evolution of NA can at least be partly attributed to epistasis, where the fitness of a mutation depends on genetic background. For example, our recent study indicates that the emergence of mutations in an antigenic region of human H3N2 NA is contingent on the evolution of other parts of the sequence (Wang et al., 2021). Similarly, the emergence of oseltamivir resistance in seasonal H1N1 virus is contingent on

permissive mutations (Abed et al., 2011; Bloom et al., 2010; Duan et al., 2014). Although epistasis is central to the understanding of evolutionary trajectories (Lyons and Luring, 2018; Miton and Tokuriki, 2016), the prevalence of epistasis in the evolution of human influenza NA remains elusive.

Mechanisms of epistasis can be categorized into two classes, specific and non-specific (Starr and Thornton, 2016). Specific epistasis, which describes a nonlinear mapping from sequence to physical property, typically involves mutations that are spatially proximal on the protein structure and physically interacting with each other. In contrast, non-specific epistasis, which describes a nonlinear mapping from physical property to phenotype, depends less on the spatial distance between mutations. Both types of epistasis are involved in the natural evolution of human influenza virus. For example, specific epistasis is observed in a network of HA mutations that coordinately modulate the evolution of receptor-binding mode (Wu et al., 2018, 2020). On the other hand, non-specific epistasis is observed in the evolution of influenza nucleoprotein (NP), in which the emergence of destabilizing mutations is contingent on the presence of stabilizing mutations to maintain a melting temperature of at least 43 °C for optimum transcriptional activity (Gong et al., 2013). As a result, characterization of epistasis can provide mechanistic insights into the biophysical constraints that govern the evolutionary trajectories of influenza proteins (Wu et al., 2018).

In this study, we systematically identified natural mutations in human H3N2 NA that were unfit in an ancestral strain but subsequently emerge. We further mapped an epistatic network among

natural mutations in human H3N2 NA using combinatorial mutagenesis and next-generation sequencing. Microscopic and flow cytometry analyses showed that membrane trafficking efficiency and enzymatic activity contribute to epistasis between NA mutations. Using biophysical assays and X-ray crystallography, we also demonstrated that non-additive stability effect led to specific epistasis in NA evolution. Overall, this study represents an in-depth analysis of epistasis in the natural evolution of human H3N2 NA and the underlying mechanisms.

RESULTS

Prevalence of evolutionary contingency in human H3N2 NA

Positive epistasis can lead to evolutionary contingency, where a deleterious mutation in an ancestral strain becomes neutral and emerges after one or more permissive mutations arise (Starr and Thornton, 2016). To probe the prevalence of evolutionary contingency in human H3N2 NA, we constructed 70 single mutants of A/Hong Kong/1968 (HK68) NA, each representing an amino-acid mutation from HK68 to A/Victoria/361/2011 (Vic11) (**Figure S1A**). The fitness effects of these HK68 NA mutants were then measured by a virus rescue experiment (**Figure 1A**). Eight out of the 70 (11%) mutants had at least two-log drop in titer as compared to the wild type (WT), indicating the presence of evolutionary contingency. Except H155Y, which located at the protomer interface, the other seven mutations were all far away from the protomer interface and were solvent exposed (**Figure 1B**). For our downstream experiments, we focused on the two mutations with the most extreme phenotypes, namely N336H and N387K, both of which did not yield any detectable titer in the virus rescue experiment. These mutational fitness effects were unlikely due

to the difference in HA, as replacing the HK68 HA with Vic11 HA did not affect the WT virus titer (Figure S2).

Natural emergence of compatibility with N336H and N387K

To narrow down the time period when N336H and N387K became neutral to the virus, we introduced N336H and N387K into the NAs of other H3N2 strains from different years, including A/Bilthoven/17938/1969 (Bil69), A/Bilthoven/21438/1971 (Bil71), A/Albany/1/1976 (Alb76), A/Bangkok/1/1979 (BK79), A/Beijing/353/1989 (Bei89), A/Shandong/9/1993 (SD93), A/Moscow/10/1999 (Mos99), and A/Wyoming/3/2003 (Wy03). While N387K was neutral in Wy03, it attenuated Mos99 and reduced the virus titer of Bk79, Bei89, and SD93 to below the detection limit (Figure 1C). This observation indicates that the permissive mutations for N387K emerged between 1993 and 2003. Similarly, our result implies that the permissive mutations for N336H emerged between 1969 and 1971, since mutating residue 336 to His was neutral in Bil71 and Alb76 but reduced the virus titer of Bil69 to below the detection limit (Figure 1D).

K385N is a permissive mutation for N387K

Next, we aimed to identify the permissive mutations for N387K. Since N387K decreased the titer of Mos99 by around two-log but was neutral in Wy03, we postulated that the difference in the NA sequences between Mos99 and Wy03 should include at least one permissive mutation for N387K. The NA head domains of Mos99 and Wy03 differed by five mutations (Figure 2A and Figure S1B). One of these mutations, K385N (Lys in Mos99 and Asn in Wy03), was adjacent to N387K.

In fact, the double mutant K385N/N387K in Mos99 yielded a WT-like titer (**Figure 2B**), demonstrating that K385N was a permissive mutation for N387K. Of note, K385N alone was neutral to the Mos99 (**Figure S3**). Nevertheless, mutating residue 385 to Asn could not restore the fitness of N387K in earlier strains, including HK68, Bk79, Bei89, and SD93 (**Figure 2B**), suggesting that additional permissive mutations are required for the natural emergence of N387K.

Mapping additional permissive mutations for N387K

Although N387K decreased the fitness of Mos99, it reduced the titer of SD93 to below the detection limit (**Figure 1C**). As a result, the difference in the NA sequences between SD93 and Mos99 should include additional permissive mutations for N387K. The NA head domains of SD93 and Mos99 differed by 18 mutations (**Figure S1C**). Since multiple mutations may be required simultaneously to restore the fitness of N387K in SD93, we have decided to use a high-throughput approach that coupled combinatorial mutagenesis and next-generation sequencing (Wang et al., 2021; Wu et al., 2018, 2020) to identify permissive mutations for N387K. Briefly, we constructed a mutant library for SD93 that contained all possible combinations of 10 mutations ($n = 2^{10} = 1024$ variants), namely E248G, R249K, I265T, Y336H, R338L, N339D, K344E, S346G, E369K, and G381E (**Figure 2C**), in the genetic background of N387K. Those 10 mutations were selected based on their distances to N387K in the protein structure and the primary sequence. The viral mutant library was then rescued and passaged once in MDCK-SIAT1 cells, which could minimize the emergence of cell-adaptive mutation (Tamura et al., 2013). Mutations E248G, R249K, Y336N, K344E, and E369K were enriched among variants with a frequency of >0.1% in the post-

passaged mutant libraries of both biological replicates (**Figure 2D**), suggesting that they could restore the fitness of N387K in SD93. Of note, Y336N was not included in the construction of the mutant library, and hence a de novo mutation that emerged during viral passaging. Interestingly, Y336N represented a reversion to the ancestral strain HK68 (**Figure S1A**).

To validate the findings from our high-throughput screening, we performed a series of virus rescue experiments with individually constructed SD93 mutants (**Figure 2E**). The pentamutant E248G/R249K/Y336N/K344E/E369K was indeed able to partially restore the fitness of N387K. The fitness of E248G/R249K/Y336N/K344E/E369K/N387K could be further increased by the addition of K385N. There seemed to be a synergistic effect among these permissive mutations, because none of E248G/R249K, Y336N, K344E, E369K, and K385N alone could restore the fitness of N387K to a detectable level (**Figure 2B and E**). We also tested whether Y336H, which was a mutation in our mutant library, instead of Y336N, which was a de novo mutation that emerged during passaging of the mutant library (see above), could act as a permissive mutation for N387K. However, both Y336H alone and the pentamutant E248G/R249K/Y336H/K344E/E369K could not restore the fitness of N387K (**Figure 2E**). Nonetheless, addition of K385N could partially restore the fitness of E248G/R249K/Y336H/K344E/E369K/N387K. Together, these results demonstrate the complexity of epistatic networks that led to the natural emergence of N387K in human H3N2 NA.

N356D is a permissive mutation for N336H

Besides N387K, N336H was another highly deleterious mutation in earlier strains (**Figure 1D**), and yet emerged during the evolution of human H3N2 NA. Mutations between the NA head domains of Bil69 and Bil71 were identified as candidate permissive mutations for N336H (**Figure S1D**), since these two strains have drastically different compatibility with His336 (**Figure 1D**). In particular, we focused on eight mutations that were spatially proximal to residue 336, namely R249K, D286G, I302V, M307V, D329N, K334S, N356D, and L370S (**Figure 3A**). To identify which of these eight mutations were permissive mutations for N336H, we used the same high-throughput approach that was described above. Our Bil69 mutant library contained all possible combinations of the eight mutations of interest ($n = 2^8 = 256$ variants), in the background of N336H. Among variants that were enriched by at least 2-fold in the post-passaged mutant libraries of both biological replicates, N356D was highly conserved (**Figure 3B**). Our virus rescue experiment further validated that N356D was a permissive mutation for N336H and was neutral when being introduced alone (**Figure 3C**). While N336H represented a mutation from HK68 to Vic11, Bil71 has a Tyr at residue 336 (**Figure S1D**). Same as N336H, N336Y also reduced the titer of Bil69 to below the detection limit and N356D could restore it to the WT level (**Figure 3C**). These results show that N356D is a permissive mutation for the natural evolution of residue 336, which in turn participates in a more extensive epistatic network that involve N387K (**Figure 2D-E**).

Enzymatic activities of H3N2 NA mutants

To understand the mechanisms of epistasis that involved N336H/Y and N387K, we first measured the NA activities on the surface of cells that were transiently transfected with different NA mutants.

N387K drastically reduced the cell surface NA activity of Mos99 and SD93 (**Figure 4A-B**), whereas both N336H and N336Y abolished the cell surface NA activity of Bil69 (**Figure 4C**). The reduced cell surface NA activity of N387K in Mos99 and SD93 could be partially restored by K385N and E248G/R249K/Y336N/K344E/E369K, respectively (**Figure 4A-B**). Similarly, the reduced cell surface NA activity of N336H and N336Y in Bil69 could be fully restored by N356D (**Figure 4C**). Therefore, the cell surface NA activities of different mutants were largely consistent with their effects on virus replication fitness (**Figures 2B, 2E and 3C**).

We further measured the NA activities of different mutants using recombinant proteins (**Figure 4D-F**). While the NA activities of recombinant proteins showed an overall similar pattern as cell surface NA activities, there were two notable differences. First, although N387K almost completely abolished the cell surface NA activity of Mos99, it still retained around 40% NA activity as a recombinant protein. Similarly, although N336H and N336Y completely abolished the cell surface NA activity of Bil69, they still retained around 75% and 40% NA activities, respectively, as recombinant proteins. These results indicate that protein enzymatic activity can only partly explain the epistasis among natural mutations in human H3N2 NA.

N387K lowers the membrane trafficking efficiency of NA

Since a high cell surface NA activity would require efficient membrane trafficking of NA, we hypothesized that membrane trafficking also plays a role in the epistasis among NA mutations. Flow cytometry analysis showed that N387K decreased the cell surface expression of Mos99 NA,

while the addition of K385N restored it (**Figure 5A**). Microscopy analysis further illustrated that N387K increased the endoplasmic reticulum (ER) retention of Mos99 NA (**Figure 5B**). In contrast, none of Mos99 NA WT, N387K, and K385N/N387K displayed strong Golgi localization (**Figure S4**). These results demonstrate that membrane trafficking efficiency also contributes to the epistasis in the evolution of human H3N2 NA.

Specific epistasis due to non-additive protein stability

To dissect the biophysical basis of epistasis, we quantified the protein thermostability of different NA mutants by measuring their melting temperatures (T_m) using thermal shift assay. Our results showed that N387K reduced the T_m of Mos99 NA and SD93 NA by around 5 °C (**Figure 6A**). The destabilizing effect of N387K could be partially compensated by K385N. However, K385N alone was not stabilizing, suggesting that the epistasis between K385N and N387K was specific. These results suggest that the defect of N387K in trafficking and enzymatic activity can be attributed to the decrease in protein stability.

In contrast to N387K, N336H did not reduce the T_m of Bil69 NA. However, the first differential curve of the thermal shift assay for Bil69 N336H had an atypical shape (**Figure S5**). Similarly, the first differential curve for Bil69 N336Y showed two peaks. These results indicate the presence of multiple oligomeric states, hence aggregations (Gao et al., 2020). The shapes of their first differential curves returned to WT-like when N356D was introduced. As a result, the biophysical basis of epistasis between K385N and N387K is different from that between N336H/Y and N356D.

221

222 We further aimed to examine the structural mechanism of epistasis between K385N and N387K.

223 We determined the crystal structure of NAs from Bil69, SD93, and Mos99 to 1.54 Å, 1.65 Å, and

224 1.40 Å, respectively (**Table S1**). Of note, the structure of SD93 NA was determined in complex

225 with zanamivir. Asn387 in Mos99 NA is sandwiched by two positively charged residues, Lys385

226 and Lys389 (**Figure 6B**). Consequently, the positively charged mutation N387K would result in an

227 unfavorable electrostatic interaction with Lys385 and Lys389, which explains the destabilizing

228 effect of N387K in Mos99 NA. The local backbone conformation of the loop that contains residues

229 385 to 389 is highly conserved between NAs from Bil69, SD93, and Mos99 (**Figure 6C**) (Vavricka

230 et al., 2011). However, the backbone of this loop has a ~3 Å shift in the NAs of more recent human

231 H3N2 strains, namely A/Perth/16/2009 (Perth09) and A/Tanzania/205/2010 (Tan10) (Hadházi et

232 al., 2018; Zhu et al., 2012) (**Figure 6D-E**). This shift can be explained by the double mutant

233 N387K/K385N, which is naturally present in Perth09 NA and Tan10 NA. K385N alone is unlikely

234 to cause a change in backbone conformation, since Bil69 NA and Mos99 NA carry different amino

235 acids at residue 385 and yet have the same backbone conformation in this region. Nevertheless,

236 K385N removes a positively charged side chain, which enables Lys387 to adapt a different

237 backbone conformation and point away from Lys389 to minimize unfavorable electrostatic

238 interaction. Therefore, our structural analysis explains why K385N is a permissive mutation for

239 N387K and exemplifies how specific epistasis in the natural evolution of human H3N2 NA

240 modulates its structural conformation.

241

DISCUSSION

Epistatic interactions between mutations is a key determinant of adaptability and evolutionary trajectory (Lyons and Luring, 2018; Sanjuán et al., 2005; Starr and Thornton, 2016). Through a systematic analysis, this study described the prevalence of epistasis in the evolution of human H3N2 NA and identified an extensive epistatic network that contains residues 248, 249, 336, 344, 356, 369, 385, 387 (**Figure 6F**). In-depth characterization of the epistatic network further revealed the underlying mechanisms. The complexity of epistasis in human H3N2 NA evolution is demonstrated by the number of mutations in the epistatic network, their widespread locations on the protein, as well as the diverse biophysical mechanisms.

Several residues in the epistatic network are within or immediately adjacent to known antigenic regions. For example, residue 369 is at antigenic region III (Colman et al., 1983, 1987), whereas residues 248 and 249 are immediately adjacent to a recently identified epitope (Wan et al., 2019). Moreover, mutations at residue 344, which locates at antigenic region II, can lead to antibody escape (Air et al., 1985; Colman et al., 1983). These observations substantiate the involvement of epistasis in NA antigenic drift, which corroborates with our recent study on a seven-residue antigenic region that contains residues 344 and 369 (Wang et al., 2021). Specifically, our recent study shows that charge balancing gives rise to epistasis in the antigenic evolution of NA (Wang et al., 2021). Such result is highly parallel to the specific epistasis between K385N and N387K in the present study, which is governed by non-additive stability effect due to electrostatic interaction. Nevertheless, the epistasis between N336H/Y and N356D has a clearly different mechanism from

that between K385N and N387K. Furthermore, wild type strains with very different NA stability can have similar virus replication fitness (**Figure 1C and Figure 6A**), suggesting that protein stability is only one of the biophysical determinants in NA for virus replication fitness. As a result, the biophysical constraints of NA evolution and antigenic drift remain to be fully comprehended.

An interesting observation in this study is that while SD93 NA K385N/N387K mutant has a WT-like thermostability, which indicates proper folding of the protomer, its enzymatic activity is minimal. Since tetramerization of NA is required for enzymatic activity (McAuley et al., 2019; Saito et al., 1995), it is possible that the head domain of the SD93 NA K385N/N387K mutant failed to form a proper tetramer, despite all recombinant NA proteins in this study were fused to a tetramerization domain and purified as a tetramer by size exclusion column. Consistently, our preliminary analysis on SD93 NA using tryptophan fluorescence spectroscopy indicates that the conformations of N387K and K385N/N387K are different from WT and the two mutants with WT-like enzymatic activity (**Figure S6**). Because the surface loop that contains residues 385 and 387 is distal from the tetramerization interface, future studies are needed to explore whether mutations on the NA surface can modulate the tetramerization process through long-range interaction.

Understanding the biophysical constraints of NA evolution not only provides mechanistic insights into influenza antigenic drift, but also facilitates next-generation immunogen design. Development of NA-based immunogens has largely focused on the tetramerization domain (Deng et al., 2021; Gao et al., 2021; Strohmeier et al., 2021). In contrast, immunogen designs for other viruses, such

as HIV and SARS-CoV-2, mainly focus on introducing mutations to the antigen to increase its stability and expression yield (Hsieh et al., 2020; Krarup et al., 2015; Pallesen et al., 2017; Sanders and Moore, 2021; Sanders et al., 2013). The feasibility of incorporating stabilizing mutations into NA-based immunogen design is demonstrated by a mutation in the center of an N9 NA that improves its tetramerization (Zhu et al., 2019). As NA is an emerging target for the development of a universal influenza vaccine (Eichelberger and Monto, 2019; Krammer et al., 2018; Stadlbauer et al., 2019), knowledge of the sequence-structure-function relationship for NA is becoming more important than ever.

MATERIALS AND METHODS

Virus rescue experiments

All H3N2 viruses generated in this study were based on the influenza A/WSN/33 (H1N1) eight-plasmid reverse genetics system (Neumann et al., 1999). Chimeric 6:2 reassortments were employed with six internal segments (PB2, PB1, PA, NP, M, and NS) from A/WSN/33. For HA, the ectodomain was from HK68, whereas the non-coding region, N-terminal secretion signal, C-terminal transmembrane domain, and cytoplasmic tail were from A/WSN/33 (Wu et al., 2018). For NA, the entire coding region was from H3N2 strains of interest, whereas the non-coding region of NA was from A/WSN/33 (Wu et al., 2018). Transfection was performed in a co-culture of HEK 293T and MDCK-SIAT1 cells (ratio of 6:1) at 60% confluence using Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's instructions. At 24 h post-transfection, cells were washed with phosphate-buffered saline (PBS) and cell culture medium was replaced with

OPTI-MEM medium supplemented with 0.8 $\mu\text{g mL}^{-1}$ TPCK-trypsin. Virus was harvested at 72 h post-transfection. To measure virus titer by TCID₅₀ assay, MDCK-SIAT1 cells were washed with PBS before the addition of virus in OPTI-MEM medium supplemented with 0.8 $\mu\text{g mL}^{-1}$ TPCK-trypsin.

Construction of the SD93 NA combinatorial mutant library

The SD93 NA-encoding plasmid of the influenza eight plasmid reverse genetics system was used as the template for insert and vector PCRs. Primers are shown in **Table S2**. To generate the insert, two independent PCRs were performed. The first PCR used SD93lib-I265T-F and SD93lib-336-346-5mut-R as primers. The second PCR used SD93lib-336-346-5mut-F and SD93lib-E369K-R as primers. The products of these two PCRs were mixed and used as templates for an overlapping PCR with SD93lib-E248G-R249K-F and SD93lib-G381E-R as primers. The product of the overlapping PCR was the complete insert. To generate the vector, SD93lib-N387K-VF and SD93lib-VR were used as primers. All PCRs were performed using PrimeSTAR Max polymerase (Takara Bio) according to the manufacturer's instructions. PCR products were purified using Monarch DNA Gel Extraction Kit (New England Biolabs). Both the vector and the complete insert were digested by BsmBI (New England Biolabs) and ligated using T4 DNA ligase (New England Biolabs). The ligated product was transformed into MegaX DH10B T1R cells (Thermo Fisher Scientific). At least half million colonies were collected. Plasmid mutant libraries were purified from the bacteria colonies using PureLink HiPure Expi Plasmid Purification Kit (Thermo Fisher Scientific).

326

327 **Construction of the Bil69 NA combinatorial mutant library**

328 The Bil69 NA-encoding plasmid of the influenza eight plasmid reverse genetics system was used
 329 as the template for insert and vector PCRs. Primers are shown in **Table S3**. To generate the insert,
 330 three independent PCRs were performed. The first PCR (PCR #1) used Bil69lib-R249K-F and
 331 Bil69lib-D286G-R as primers. The second PCR (PCR #2) used Bil69lib-I302V-M307V-F as the
 332 forward primer and a mixture of Bil69lib-D329N-K334S-1-R and Bil69lib-D329N-K334S-2-R at
 333 equal molar ratio as the reverse primer. The third PCR (PCR #3) used a mixture of Bil69lib-
 334 D329N-K334S-1-F and Bil69lib-D329N-K334S-2-F at equal molar ratio as the forward primer, and
 335 Bil69lib-N356D-R as the reverse primer. The product of PCR #1 was used as a template for
 336 another PCR (PCR #4) using Bil69lib-R249K-F and Bil69lib-I302V-M307V-R as primers. The
 337 products of PCRs #2, #3 and #4 were mixed and used as templates for an overlapping PCR with
 338 Bil69lib-R249K-F and Bil69lib-L370S-R as primers. The product of the overlapping PCR was the
 339 complete insert. To generate the vector, Bil69lib-VF and Bil69lib-VR were used as primers. All
 340 PCRs were performed using PrimeSTAR Max polymerase (Takara Bio) according to the
 341 manufacturer's instructions. PCR products were purified using Monarch DNA Gel Extraction Kit
 342 (New England Biolabs). Both the vector and the complete insert were digested by BsmBI (New
 343 England Biolabs) and ligated using T4 DNA ligase (New England Biolabs). The ligated product
 344 was transformed into MegaX DH10B T1R cells (Thermo Fisher Scientific). At least half million
 345 colonies were collected. Plasmid mutant libraries were purified from the bacteria colonies using
 346 PureLink HiPure Expi Plasmid Purification Kit (Thermo Fisher Scientific).

347

348 **Characterizing the fitness of individual variants in the mutant libraries**

349 Virus mutant libraries were rescued in a co-culture of HEK 293T and MDCK-SIAT1 cells (ratio of
350 6:1) at 60% confluence in a T75 flask (75 cm²) using Lipofectamine 2000 (Thermo Fisher Scientific)
351 according to the manufacturer's instructions. At 24 h post-transfection, cells were washed twice
352 with PBS and cell culture medium was replaced with OPTI-MEM medium supplemented with 0.8
353 µg mL⁻¹ TPCK-trypsin. Virus mutant libraries were harvested at 72 h post-transfection. Virus
354 mutant libraries were titered by TCID₅₀ assay using MDCK-SIAT1 cells and stored at -80 °C until
355 used. To passage the virus mutant libraries, MDCK-SIAT1 cells in a T75 flask were washed twice
356 with PBS and infected with half of the virus stock in OPTI-MEM medium supplemented with 0.8
357 µg mL⁻¹ TPCK-trypsin. At 2 h post-infection, infected cells were washed twice with PBS and fresh
358 OPTI-MEM medium supplemented with 0.8 µg mL⁻¹ TPCK-trypsin was added to the cells. At 24
359 h post-infection, supernatant was harvested. Each replicate was transfected and passaged
360 independently. Viral RNA was then extracted from the supernatant using QIAamp Viral RNA Mini
361 Kit (Qiagen). The extracted RNA was reverse transcribed to cDNA using Superscript III reverse
362 transcriptase (Thermo Fisher Scientific). The plasmid mutant libraries and the cDNA from the
363 post-infection viral mutant libraries were amplified by PCR to add part of the adapter sequence
364 required for Illumina sequencing. For SD93 mutant library, SD93lib-recover-F and SD93lib-
365 recover-R were used as primers (**Table S2**). For Bil69 mutant library, Bil69lib-recover-F and
366 Bil69lib-recover-R were used as primers (**Table S3**). A second PCR was carried out to add the
367 rest of the adapter sequence and index to the amplicon using primers: 5'-AAT GAT ACG GCG

ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT-3' and 5'-CAA GCA GAA GAC
GGC ATA CGA GAT XXX XXX GTG ACT GGA GTT CAG ACG TGT GCT-3'. Positions annotated
by an "X" represented the nucleotides for the index sequence. The final PCR products were
submitted for the next-generation sequencing using Illumina MiSeq PE250.

Sequencing data analysis

Sequencing data were obtained in FASTQ format. Forward and reverse reads were merged by
PEAR (Zhang et al., 2014). Using the SeqIO module in BioPython (Cock et al., 2009), primer
sequences were trimmed from the merged reads. If the length of the trimmed reads did not match
with that of the reference nucleotide sequence, the reads would be discarded. The trimmed reads
were then translated to amino acid sequences. Amino acid mutations were called by comparing
the translated reads to the translated reference sequence. For Bil69 mutant library, frequency of
mutant i in sample s was computed for each replicate as follows:

$$frequency_{i,s} = \frac{read\ count_{i,s} + 1}{\sum_{k \in s} read\ count_{k,s}} \quad (1)$$

A pseudocount of 1 was added to the read counts to avoid division by zero in subsequent steps.
Subsequently, the enrichment of mutant i was computed for each replicate as follows:

$$enrichment_i = \frac{frequency_{post-passaged\ library}}{frequency_{plasmid\ mutant\ library}} \quad (2)$$

Bil69 mutants with a read count of less than 10 in the plasmid mutant library were discarded. For
SD93 mutant library, the frequency of each mutant was calculated using the same method, except
that the pseudocount was not added.

Protein expression and purification.

The NA head domains, which contain residues 82 to 469, were fused to an N-terminal gp67 signal peptide, 6×His-tag, a vasodilator-stimulated phosphoprotein (VASP) tetramerization domain, and a thrombin cleavage site. Recombinant bacmid DNA that carried the NA ectodomain from the strain of interest was generated using the Bac-to-Bac system (Thermo Fisher Scientific). Baculovirus was generated by transfecting purified bacmid DNA into adherent Sf9 cells using Cellfectin reagent (Thermo Fisher Scientific) according to the manufacturer's instructions. The baculovirus was further amplified by passaging in adherent Sf9 cells at a multiplicity of infection (MOI) of 1. Recombinant NA head domains were expressed by infecting 1 L of suspension Sf9 cells at an MOI of 1. On day 3 post-infection, Sf9 cells were pelleted by centrifugation at 4,000 ×g for 25 min, and soluble recombinant NA was purified from the supernatant by affinity chromatography using Ni Sepharose excel resin (Cytiva) and then size exclusion chromatography on a HiLoad 16/100 Superdex 200 prep grade column (Cytiva) in 20 mM Tris-HCl pH 8.0, 100 mM NaCl, and 10 mM CaCl₂. For crystallography, recombinant NA was further digested by thrombin (MilliporeSigma) for three weeks in 4 °C using 15 U thrombin per mg of recombinant NA. The thrombin-digested recombinant NA was incubated with TALON metal affinity resin (Takara) for 2 h. The thrombin-digested recombinant NA in the flow-through and 10 mM imidazole wash was purified by size exclusion chromatography on a HiLoad 16/100 Superdex 200 prep grade column (Cytiva) in 20 mM Tris pH 8.0, 100 mM NaCl, and 10 mM CaCl₂.

Cell surface NA activity assay

NA from the strain of interest was cloned into pHCMV3 vector. pHCMV3-NA was transfected into 293T cells using Lipofectamine 2000 (Thermo Fisher Scientific). At 24 h post-transfection, the cells were washed with PBS and then resuspended in 200 μ L PBS. Cells were then mixed with fluorogenic substrate methylumbelliferyl- α -D-N-acetylneuraminic acid (MUNANA) (MilliporeSigma) at a final concentration of 800 μ M in PBS that contained 33 mM MES pH 6.5 and 4.0 mM CaCl_2 . The sample was then transferred to a 96-well half area plate (Corning). NA enzymatic kinetics at 37 $^{\circ}\text{C}$ were measured by a SpectraMax M2 microplate reader (Molecular Devices) with an excitation at 365 nm and an emission at 450 nm. Fluorescence was recorded for 50 min at 30 s interval. NA activity was quantified by the rate of increase in fluorescence signal. Background signal, which was determined by untransfected cell, was subtracted. NA activity was normalized to the cell concentration, which was determined by cell counting.

Recombinant NA activity assay

In a 96-well half area plate (Corning), 2.5 ng of recombinant NA was incubated with 100 μ M MUNANA in PBS that contained 33 mM MES pH 6.5 and 4 mM CaCl_2 . NA enzymatic kinetics at 37 $^{\circ}\text{C}$ were measured by a SpectraMax M2 microplate reader (Molecular Devices) with an excitation at 365 nm and an emission at 450 nm. Fluorescence was recorded for 30 min at 30 s interval. NA activity was quantified by the rate of increase in fluorescence signal. Background signal, which was determined in the absence of recombinant NA, was subtracted.

Flow cytometry analysis of NA expression

4×10⁵ HEK 293T cells were seeded on each well of 6-well plates and grown overnight at 37 °C and 5% CO₂ in a humidified incubator. Cells were transfected with 2 µg of pHCMV3 plasmids encoding the C-terminal HA-tagged NA from the strain of interest using Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's instructions. At 24 h post-transfection, medium was discarded and cells were washed once with PBS. Cells were subsequently detached with versene (Thermo Fisher Scientific) and resuspended in fluorescence-activated single cell sorting (FACS) buffer (50 mM EDTA, 2% v/v FBS in DMEM with high glucose and HEPES, without phenol red). Cells were then pelleted via centrifugation at 300 ×g for 5 min at 4 °C and the supernatant was aspirated. In the subsequent steps, cells and all reagents were kept on ice.

For surface staining, cell pellets were resuspended in 1 mL FACS buffer. Cells were incubated with 1 µg mL⁻¹ PE-anti HA.11 (BioLegend) at 4 °C with rocking for 1 h. Subsequently, cells were pelleted via centrifugation at 300 ×g for 5 min at 4 °C and the supernatant was aspirated. Cells were washed with 1 mL FACS buffer, pelleted, and resuspended in 1 mL FACS buffer.

For intracellular staining, cell pellets were fixed in 300 µL of 1% v/v paraformaldehyde (in PBS) and incubated at 4 °C with rocking for 30 min. Cells were pelleted, washed with 1 mL FACS buffer, permeabilized in 300 µL of 0.1% v/v Triton X-100 (in PBS) and incubated at 4 °C with rocking for 30 min. Cells were pelleted, washed with 1 mL FACS buffer and incubated with 1 µg mL⁻¹ PE-anti HA.11 (BioLegend) at 4 °C with rocking for 1 h. Then, cells were pelleted, washed with 1 mL FACS buffer, and resuspended in 1 mL FACS buffer.

452

453 Cells were analyzed using an Accuri C6 flow cytometer (BD Biosciences) with a 488 nm laser and
454 a 585/40 bandpass filter. Data were collected using Accuri C6 software. 10^4 singlets were
455 collected for each sample. Data were analyzed using FCS Express 6 software (De Novo
456 Software).

457

458 **Confocal microscopy**

459 7.5×10^4 HeLa cells were seeded on 12 mm #1.5 glass cover slips in 24-well plates and grown
460 overnight at 37 °C with 5% CO₂ in a humidified incubator. Then, cells were transfected with 500
461 ng of plasmid encoding the indicated protein and 1.5 µL of Lipofectamine 2000 (Thermo Fisher
462 Scientific) according to the manufacturer's instructions. At 24 h post-transfection, medium was
463 discarded and cells were washed with PBS. Cells were fixed with 4% paraformaldehyde for 10
464 min at room temperature and then washed with PBS. Cells were permeabilized with ice-cold 0.1%
465 v/v Triton X-100 (in PBS) for 20 min at 4 °C and subsequently washed with PBS. Cells were
466 blocked in 5% v/v normal donkey serum (in PBS) for 1 h at room temperature. Then, cells were
467 incubated overnight at 4 °C with the indicated primary antibody, namely mouse anti-HA tag
468 (Thermo Fisher Scientific) 1:500, rabbit anti-calnexin (Thermo Fisher Scientific) 1:500, or rabbit
469 anti-GM130 (Thermo Fisher Scientific) 1:200, and then washed with PBS. Then, cells were
470 incubated for 1 h at room temperature with secondary antibody, namely Alexa Fluor 488 donkey
471 anti-rabbit IgG (Abcam) 1:200 or Alexa Fluor 555 donkey anti-mouse IgG (Thermo Fisher
472 Scientific) 1:500, as well as DAPI (to a final concentration of 1 µg ml⁻¹), and then washed with

PBS. All primary antibodies, secondary antibodies, and DAPI were diluted in 5% v/v normal donkey serum (Abcam). Subsequently, cover slips were mounted on slides with ProLong Diamond Antifade Mountant (Thermo Fisher Scientific) and sealed with nail polish.

Images were captured on an LSM 700 microscope and a Plan Apochromat 63X/1.4 Oil DIC objective using Zen software (Zeiss). Excitation laser lines 405 nm, 488 nm and 555 nm were used for acquiring images stained with DAPI, Alexa Fluor 488 and Alexa Fluor 555, respectively. Individual channels of each image were linearly adjusted using Zen software. Transfected cells were manually outlined as regions of interest in Fiji. Cytofluorogram data and Pearson correlation coefficients of transfected cells were calculated with fluorescence intensity thresholds automatically adjusted using the JACoP plug-in (Bolte and Cordelières, 2006).

Thermal shift assay

200 ng protein was mixed with 5× SYPRO orange (Thermo Fisher Scientific) in 20 mM Tris-HCl pH 8.0, 100 mM NaCl, and 10 mM CaCl₂ at a final volume of 25 µL. The sample mixture was then transferred into an optically-clear PCR tube (VWR). SYPRO orange fluorescence data in relative fluorescence unit (RFU) was collected from 10 °C to 95 °C using CFX Connect Real-Time PCR Detection System (Bio-Rad). The temperature corresponding to the lowest point of the first derivative $-d(\text{RFU})/dT$ was determined to be the T_m .

Crystallization and structural determination

Crystallization screening was performed using the JCSG Core Suites I-IV (Rigaku) with thrombin-digested NA at 7 mg mL⁻¹. Sitting drop for crystallization screening was set up by equal volume of precipitant and protein solution using the Crystal Gryphon (Art Robbins Instruments). Crystallization screens were incubated at 18 °C. Initial hits were further optimized using the sitting drop method at 18 °C, with 350 µL reservoir solution and 1:1 ratio of precipitant and protein solution. The crystallization conditions were as follows:

Mos99 WT: 2.0 M ammonium sulfate, 0.1 M acetate pH 4.6

SD93 WT: 8% PEG-8000, 0.1 M Tris pH 8.5

Bil69 WT: 10% PEG-6000, 5% MPD, 0.1 M HEPES pH 6.5

Crystals were soaked in precipitant solution containing the cryoprotectant prior to vitrification in liquid nitrogen. The cryoprotectants were as follows:

Mos99 WT: 15% ethylene glycol

SD93 WT: 25% PEG 200

Bil69 WT: 20% ethylene glycol

Data were collected at the Advanced Photon Source (APS) at Argonne National Laboratory via the Life Science Collaborative Access Team (LS-CAT) at beamlines 21-ID-D, 21-ID-G, and 21-ID-F. Initial diffraction data were indexed, integrated, and scaled using autoPROC (Vonrhein et al.,

2011). The structure was solved by molecular replacement using Phaser-MR included in the Phenix suite (Adams et al., 2010), using as PDB 2AEP as the replacement model (Venkatramani et al., 2006). The structure was further refined using REFMAC5 (Murshudov et al., 2011) and was manually built in COOT (Emsley et al., 2010). Ramachandran statistics were calculated using MolProbity (Chen et al., 2010).

Fluorescent emission spectroscopy

The steady-state tryptophan emission spectra were measured using a fluorometer (FP8300, JASCO). Samples for the emission spectra measurements were prepared by diluting the protein stock in SEC buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 10 mM CaCl₂) to an OD₂₈₀ = 0.32 as measured by UV-Vis spectroscopy (UV-1800, Shimadzu). The Trp residues in the protein samples were excited at 295 nm and the emission spectra were collected from 290 to 450 nm with 2.5 nm bandwidths. The emission spectra data were then normalized to a range of 0 to 1 and plot for analysis.

Data availability

Raw sequencing data have been submitted to the NIH Short Read Archive under accession number: BioProject PRJNA790468. The X-ray coordinates and structure factors will be deposited to the RCSB Protein Data Bank prior to publication.

Code availability

Custom python scripts for analyzing the deep mutational scanning data have been deposited to https://github.com/nicwulab/N2_evol_contingency.

ACKNOWLEDGEMENTS

This work was supported by National Institutes of Health (NIH) R00 AI139445 (N.C.W.), DP2 AT011966 (N.C.W.), and R01 AI167910 (N.C.W.). We thank Martin Gruebele for helpful discussion, the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign for assistance with next-generation sequencing, the Microscopy Core Facility of the Institute of Genomic Biology at the University of Illinois at Urbana-Champaign for microscopy use, and Spencer Anderson and colleagues at LS-CAT (Argonne National Labs) for facilitating X-ray data collection.

AUTHOR CONTRIBUTIONS

R.L. and N.C.W. conceived and designed the study. R.L. and Z.T.D. performed the virus rescue experiments. R.L. constructed the mutant libraries and prepared the sequencing libraries. Y.W. and N.C.W. analyzed the next-generation sequencing data. T.J.C.T. and Q.W.T. performed the flow cytometry and microscopy experiments. R.L., C.S.G., and C.T. expressed and purified the recombinant NA. R.L. and C.S.G. performed the MUNANA assay and the thermal shift assay. R.L., A.H.G., and S.K.N. performed protein crystallization. A.H.G. and S.K.N. collected the X-ray diffraction data as well as determined and refined the structures. R.L. and G.G. performed the tryptophan fluorescence spectroscopy. R.L. and N.C.W. wrote the paper and all authors reviewed

557 and edited the paper.

558

559 **COMPETING INTERESTS**

560 The authors declare no competing interests.

FIGURES

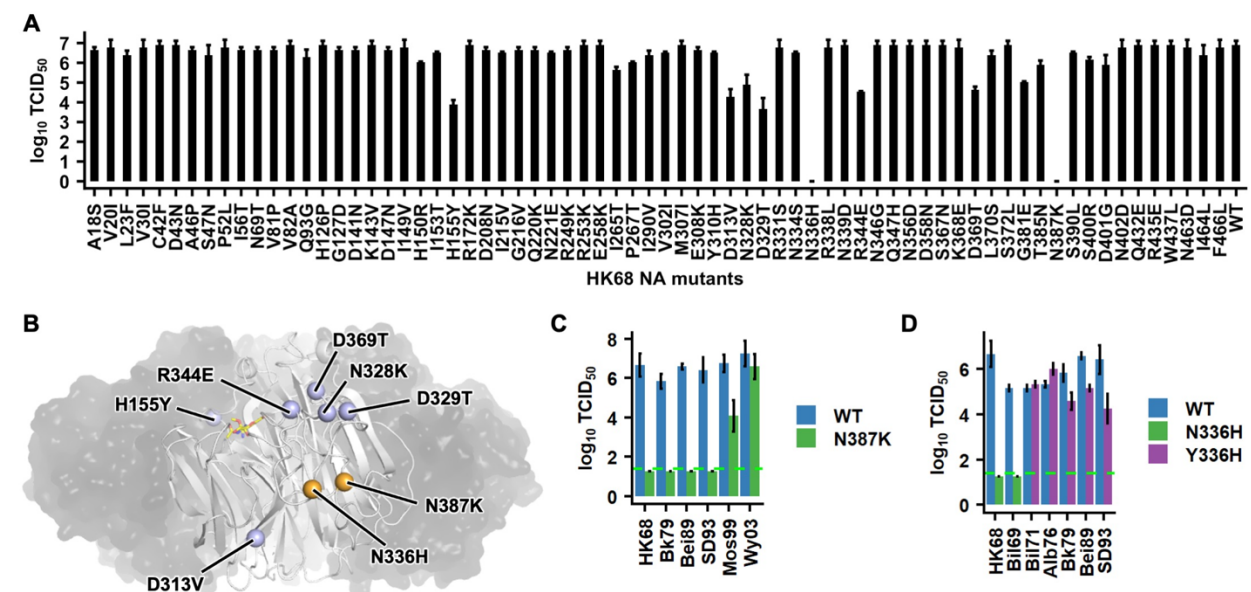


Figure 1. Virus rescue experiments of natural mutations. (A) The amino acid sequences of HK68 NA and Vic11 NA differ by 70 amino acid substitutions. These 70 mutations were individually introduced into HK68 NA. Their replication fitness was examined by a virus rescue experiment. Virus titer was measured by 50% cell culture infectious dose (TCID₅₀). Error bars indicate the standard deviation of three independent experiments. (B) The locations of the eight mutations that showed at least two-log decrease in virus titer compared to the wild type (WT) are shown as spheres on one protomer of NA that is in white cartoon representation, while the other three protomers are shown as semitransparent black surface (PDB 3TIA) (Vavricka et al., 2011). Sialic acid in the active site is shown as yellow sticks. Location of the two mutations that did not yield any detectable titer, namely N336H and N387K, are in orange. The other six mutations are in blue. (C-D) Fitness effects of (C) N387K and (D) N/Y336H on different strains were examined by a virus rescue experiment. While WTs of HK68 and Bil69 have an Asn at residue 336, WTs of Bil71, Alb76,

575 Bk79, Bei89, and SD93 have a Tyr at residue 336. Virus titer was measured by TCID₅₀. Error bars
576 indicate the standard deviation of three independent experiments. The green dashed line
577 represents the lower detection limit.

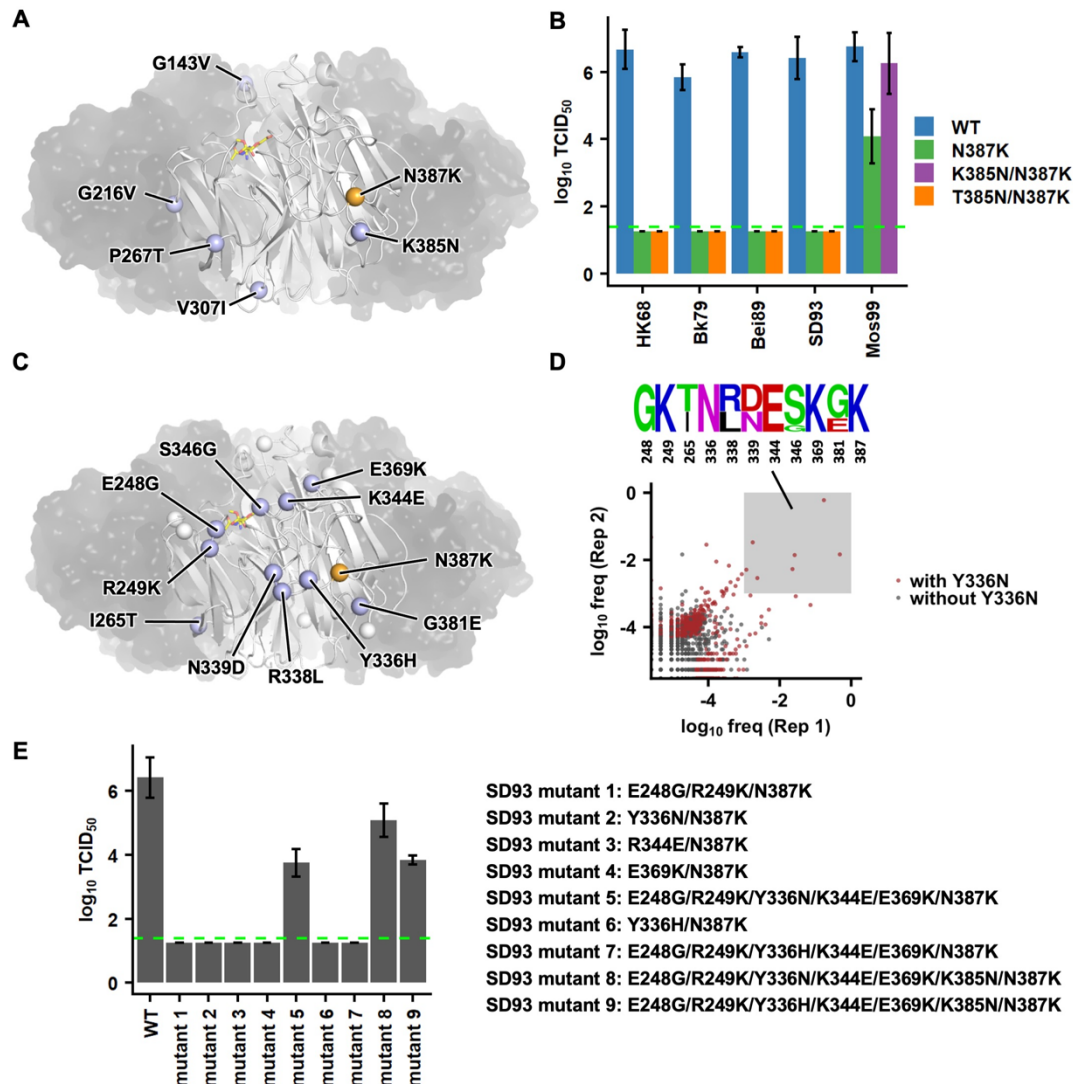


Figure 2. Permissive mutations for NA N387K in Mos99 and SD93. (A) Locations of natural mutations from Mos99 to Wy03 are shown as blue spheres on one protomer of NA that is in white cartoon representation, while the other three protomers are shown as semitransparent black surface (PDB 3TIA) (Vavricka et al., 2011). Location of N387K is shown as an orange sphere. Sialic acid in the active site is shown as yellow sticks. (B) The compensatory effects of mutating residue 385 to Asn in different chimeric H3N2 strains were examined by a virus rescue experiment. Virus titer was measured by TCID₅₀. Error bars indicate the standard deviation of three

independent experiments. The green dashed line represents the lower detection limit. **(C)** Locations of natural mutations from SD93 to Mos99 are shown as blue or white spheres on one protomer of NA that is in white cartoon representation, while the other three protomers are shown as semitransparent black surface (PDB 3TIA) (Vavricka et al., 2011). Blue spheres represent the mutations that were included in the SD93 combinatorial mutant library. **(D)** High-throughput identification of permissive mutations for N387K in SD93. The occurrence frequency of each variant in the post-passaged mutant library is shown. A de novo mutation Y336N was emerged during passaging. Data points that represent mutants with Y336N are in red. The amino acid sequences of variants with >0.1% frequencies in the post-passaged mutant libraries of both biological replicates are represented by a sequence logo. **(E)** The replication fitness of different SD93 mutants was examined by a virus rescue experiments. Error bars indicate the standard deviation of three independent experiments. Virus titer was measured by TCID₅₀. The green dashed line represents the lower detection limit.

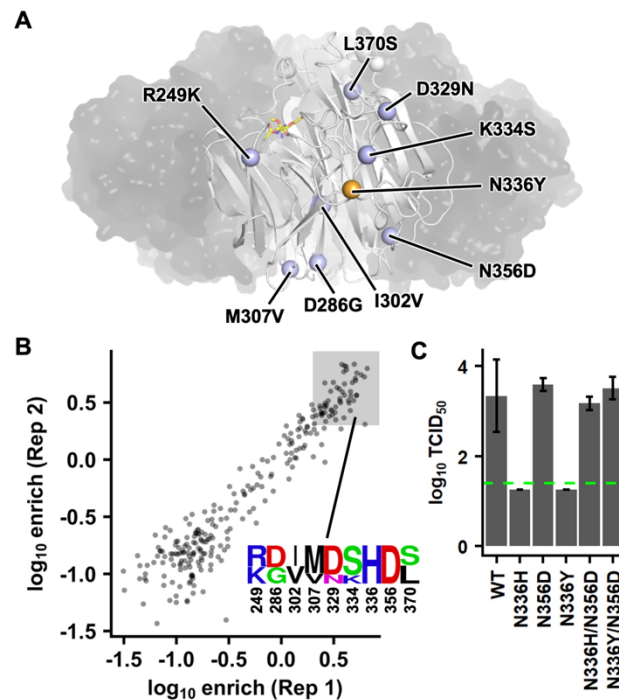


Figure 3. Permissive mutations for NA N336Y in Bil69. (A) Locations of natural substitutions from Bil69 to Bil71 are shown as blue spheres on one protomer of NA that is in white cartoon representation, while the other three protomers are shown as semitransparent black surface (PDB 3TIA) (Vavricka et al., 2011). The location of N336Y is shown as an orange sphere. Sialic acid in the active site is shown as yellow sticks. **(B)** High-throughput identification of permissive mutations for N336Y in Bil69. The frequency enrichment of each variant in the post-passaged mutant library is shown. The amino acid sequences of variants with at least 2-fold enrichment (\log_{10} enrichment ≥ 0.3) in the post-passaged mutant libraries of both biological replicates are represented by a sequence logo. **(C)** The replication fitness of different Bil69 mutants was examined by a virus rescue experiment. Error bars indicate the standard deviation of three independent experiments. Virus titer was measured by TCID₅₀. The green dashed line represents the lower detection limit.

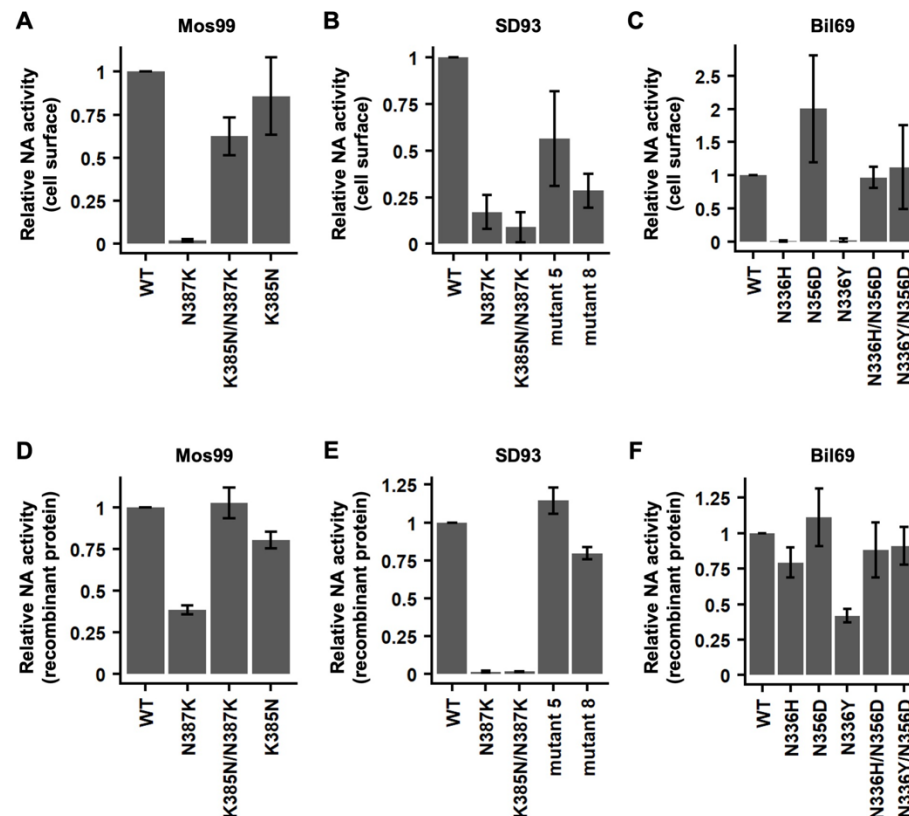


Figure 4. Functional characterization of different NA mutants. (A-C) Relative cell surface NA activities of **(A)** Mos99 mutants, **(B)** SD93 mutants, and **(C)** Bil69 mutants. **(D-F)** Relative NA activities of recombinant proteins of **(D)** Mos99 mutants, **(E)** SD93 mutants, and **(F)** Bil69 mutants. **(A-F)** Error bars indicate the standard deviation of three independent experiments. Relative NA activities of the mutants were normalized to that of their corresponding WT.

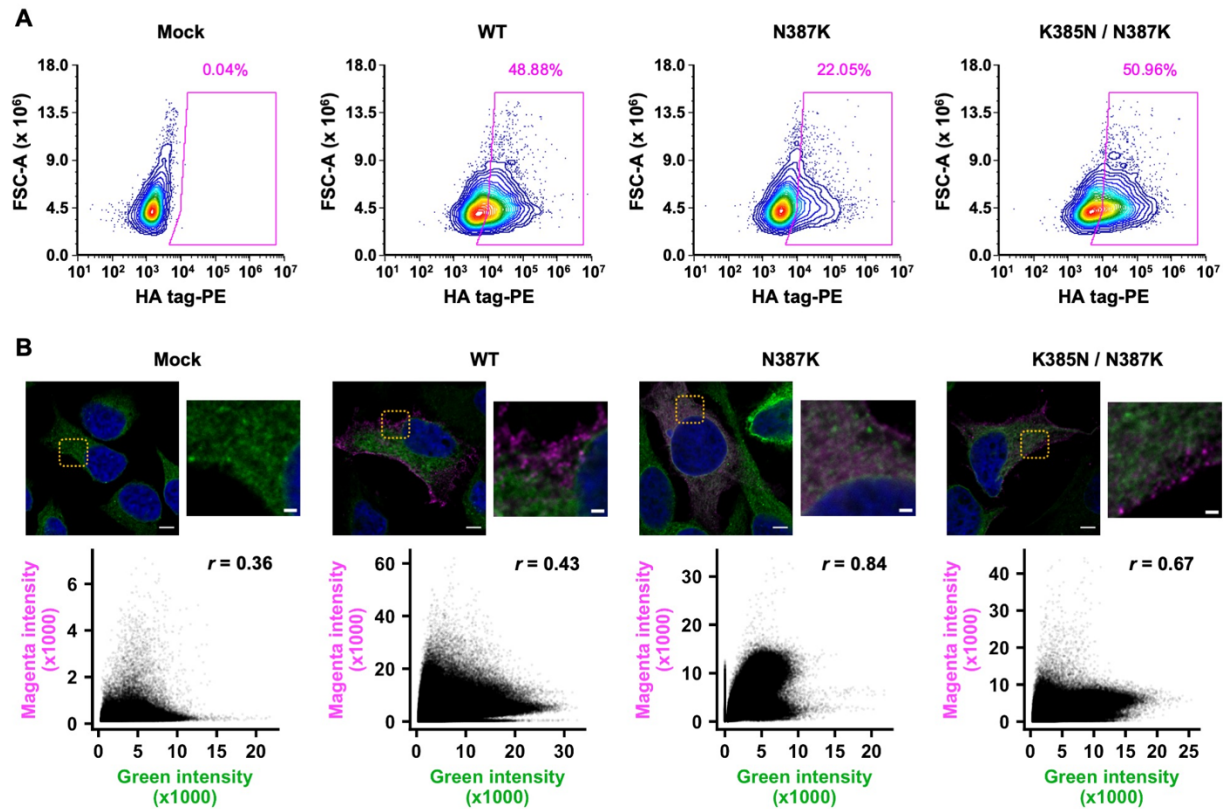


Figure 5. Cell surface protein expression and intracellular localization of NA Mos99 mutants. (A) Cell surface expressions of HA-tagged Mos99 mutants were analyzed by flow cytometry. (B) Confocal microscopy result of HA-tagged Mos99 mutants. Blue (DAPI), Green (Calnexin, ER), Magenta (NA). The orange box highlights the zoomed in region. Scale bar for the large image is 5 μ m and for zoomed in image is 2 μ m. Below each micrograph is a cytofluorogram along with the Pearson correlation coefficient.

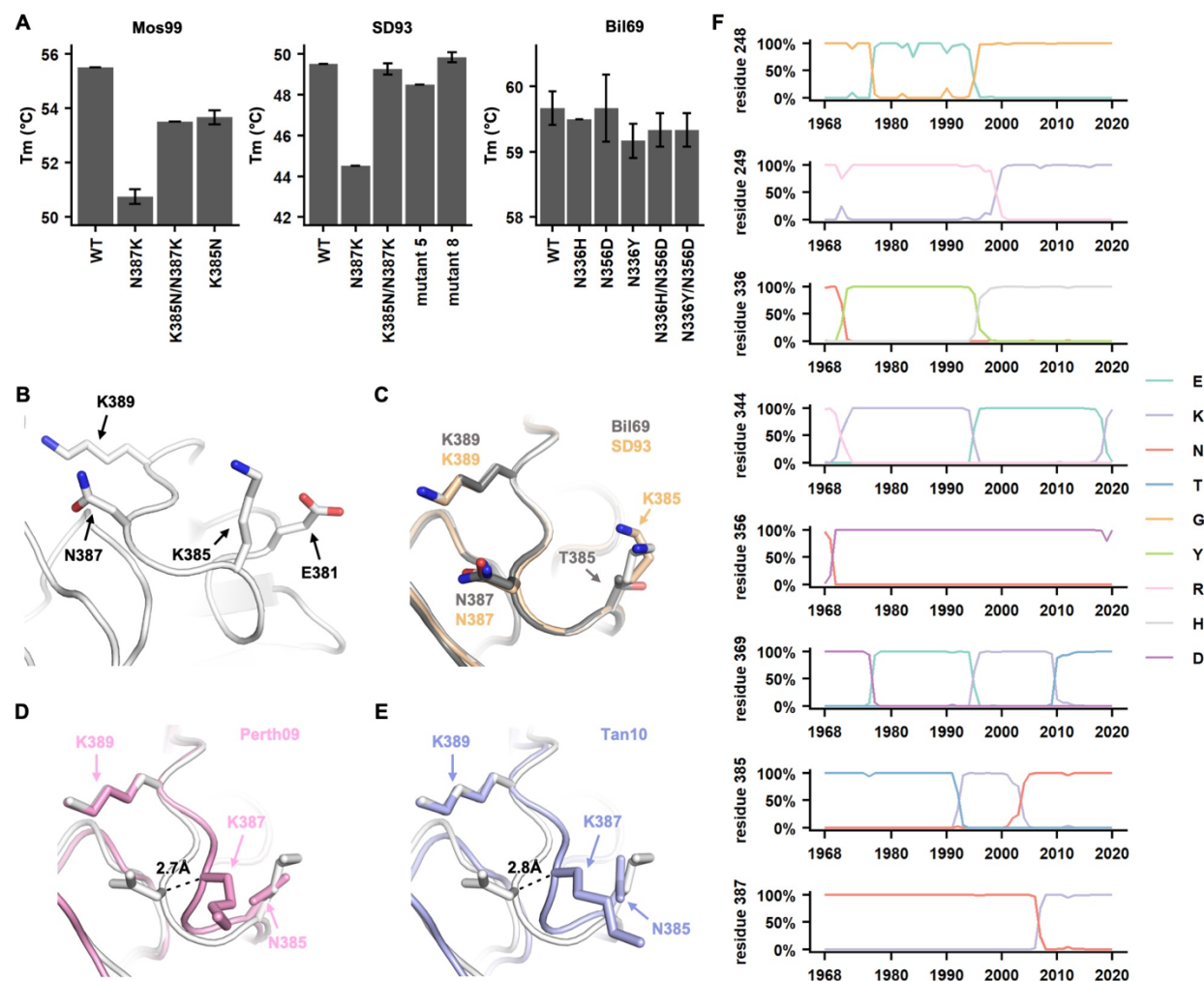


Figure 6. Structure and stability of H3N2 neuraminidase. (A) Melting temperature (T_m) of Mos99, SD93, and Bil69 NA mutants was measured by thermal shift assay. Error bars indicate the standard deviation of six independent experiments. **(B)** The locations of the residues of interest on Mos99 NA are shown. **(C-E)** The local structure around residue 387 of Mos99 NA (white) is compared with that of **(C)** Bil69 NA and SD93 NA, **(D)** H3N2 A/Perth/16/2009 (Perth09) NA (pink, PDB 6BR5) (Hadházi et al., 2018), and **(E)** H3N2 A/Tanzania/205/2010 (Tan10) NA (blue, PDB 4GZO) (Zhu et al., 2012). **(F)** Natural occurrence frequencies of the amino acid variants that have a natural occurrence at >50% in any given year at the residues of interest are

634 shown.

635

REFERENCES

- Abed, Y., Pizzorno, A., Bouhy, X., and Boivin, G. (2011). Role of permissive neuraminidase mutations in influenza A/Brisbane/59/2007-like (H1N1) viruses. *PLOS Pathog.* 7, e1002431.
- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66, 213–221.
- Air, G.M., Els, M.C., Brown, L.E., Laver, W.G., and Webster, R.G. (1985). Location of antigenic sites on the three-dimensional structure of the influenza N2 virus neuraminidase. *Virology* 145, 237–248.
- Bloom, J.D., Gong, L.I., and Baltimore, D. (2010). Permissive secondary mutations enable the evolution of influenza oseltamivir resistance. *Science* 328, 1272–1275.
- Bolte, S., and Cordelières, F.P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* 224, 213–232.
- Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 66, 12–21.
- Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25, 1422–1423.
- Colman, P.M., Varghese, J.N., and Laver, W.G. (1983). Structure of the catalytic and antigenic sites in influenza virus neuraminidase. *Nature* 303, 41–44.
- Colman, P.M., Laver, W.G., Varghese, J.N., Baker, A.T., Tulloch, P.A., Air, G.M., and Webster, R.G. (1987). Three-dimensional structure of a complex of antibody with influenza virus neuraminidase. *Nature* 326, 358–363.
- Couch, R.B., Atmar, R.L., Franco, L.M., Quarles, J.M., Wells, J., Arden, N., Niño, D., and Belmont, J.W. (2013). Antibody correlates and predictors of immunity to naturally occurring influenza in humans and the importance of antibody to the neuraminidase. *J. Infect. Dis.* 207, 974–981.
- Deng, X., Wang, Q., Liu, M., Zheng, Q., Wu, F., and Huang, J. (2021). Tetrameric neuraminidase of influenza A virus is required to induce protective antibody responses in mice. *Front. Microbiol.* 12, 729914.
- Duan, S., Govorkova, E.A., Bahl, J., Zaraket, H., Baranovich, T., Seiler, P., Prevost, K., Webster,

669 R.G., and Webby, R.J. (2014). Epistatic interactions between neuraminidase mutations facilitated
670 the emergence of the oseltamivir-resistant H1N1 influenza viruses. *Nat. Commun.* 5, 5029.

671 Eichelberger, M.C., and Monto, A.S. (2019). Neuraminidase, the forgotten surface antigen,
672 emerges as an influenza vaccine target for broadened protection. *J. Infect. Dis.* 219, S75–S80.

673 Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot.
674 *Acta Crystallogr. D Biol. Crystallogr.* 66, 486–501.

675 Gao, J., Klenow, L., Parsons, L., Malik, T., Phue, J.-N., Gao, Z., Withers, S.G., Cipollo, J., Daniels,
676 R., and Wan, H. (2021). Design of the recombinant influenza neuraminidase antigen is crucial for
677 its biochemical properties and protective efficacy. *J. Virol.* 95, e0116021.

678 Gao, K., Oerlemans, R., and Groves, M.R. (2020). Theory and applications of differential scanning
679 fluorimetry in early-stage drug discovery. *Biophys. Rev.* 12, 85–104.

680 Gong, L.I., Suchard, M.A., and Bloom, J.D. (2013). Stability-mediated epistasis constrains the
681 evolution of an influenza protein. *eLife* 2, e00631.

682 Hadházi, Á., Li, L., Bailly, B., Maggioni, A., Martin, G., Dirr, L., Dyason, J.C., Thomson, R.J., Gao,
683 G.F., Borbás, A., et al. (2018). A sulfonozanamivir analogue has potent anti-influenza virus activity.
684 *ChemMedChem* 13, 785–789.

685 Hsieh, C.-L., Goldsmith, J.A., Schaub, J.M., DiVenere, A.M., Kuo, H.-C., Javanmardi, K., Le, K.C.,
686 Wrapp, D., Lee, A.G., Liu, Y., et al. (2020). Structure-based design of prefusion-stabilized SARS-
687 CoV-2 spikes. *Science* 369, 1501–1505.

688 Iuliano, A.D., Roguski, K.M., Chang, H.H., Muscatello, D.J., Palekar, R., Tempia, S., Cohen, C.,
689 Gran, J.M., Schanzer, D., Cowling, B.J., et al. (2018). Estimates of global seasonal influenza-
690 associated respiratory mortality: a modelling study. *Lancet Lond. Engl.* 391, 1285–1300.

691 Krammer, F., Fouchier, R.A.M., Eichelberger, M.C., Webby, R.J., Shaw-Saliba, K., Wan, H.,
692 Wilson, P.C., Compans, R.W., Skountzou, I., and Monto, A.S. (2018). NAction! How can
693 neuraminidase-based immunity contribute to better influenza virus vaccines? *mBio* 9, e02332-17.

694 Krarup, A., Truan, D., Furmanova-Hollenstein, P., Bogaert, L., Bouchier, P., Bisschop, I.J.M.,
695 Widjoatmodjo, M.N., Zahn, R., Schuitemaker, H., McLellan, J.S., et al. (2015). A highly stable
696 prefusion RSV F vaccine derived from structural analysis of the fusion mechanism. *Nat. Commun.*
697 6, 8143.

698 Lyons, D.M., and Lauring, A.S. (2018). Mutation and epistasis in influenza virus evolution. *Viruses*
699 10, E407.

700 McAuley, J.L., Gilbertson, B.P., Trifkovic, S., Brown, L.E., and McKimm-Breschkin, J.L. (2019).

701 Influenza virus neuraminidase structure and functions. *Front. Microbiol.* 10, 39.

702 Memoli, M.J., Shaw, P.A., Han, A., Czajkowski, L., Reed, S., Athota, R., Bristol, T., Fargis, S.,
703 Risos, K., Powers, J.H., et al. (2016). Evaluation of antihemagglutinin and antineuraminidase
704 antibodies as correlates of protection in an influenza A/H1N1 virus healthy human challenge
705 model. *mBio* 7, e00417-00416.

706 Miton, C.M., and Tokuriki, N. (2016). How mutational epistasis impairs predictability in protein
707 evolution and design. *Protein Sci. Publ. Protein Soc.* 25, 1260–1272.

708 Monto, A.S., Petrie, J.G., Cross, R.T., Johnson, E., Liu, M., Zhong, W., Levine, M., Katz, J.M., and
709 Ohmit, S.E. (2015). Antibody to influenza virus neuraminidase: an independent correlate of
710 protection. *J. Infect. Dis.* 212, 1191–1199.

711 Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn,
712 M.D., Long, F., and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal
713 structures. *Acta Crystallogr. D Biol. Crystallogr.* 67, 355–367.

714 Neumann, G., Watanabe, T., Ito, H., Watanabe, S., Goto, H., Gao, P., Hughes, M., Perez, D.R.,
715 Donis, R., Hoffmann, E., et al. (1999). Generation of influenza A viruses entirely from cloned
716 cDNAs. *Proc. Natl. Acad. Sci. U. S. A.* 96, 9345–9350.

717 Pallesen, J., Wang, N., Corbett, K.S., Wrapp, D., Kirchdoerfer, R.N., Turner, H.L., Cottrell, C.A.,
718 Becker, M.M., Wang, L., Shi, W., et al. (2017). Immunogenicity and structures of a rationally
719 designed prefusion MERS-CoV spike antigen. *Proc. Natl. Acad. Sci. U. S. A.* 114, E7348–E7357.

720 Saito, T., Taylor, G., and Webster, R.G. (1995). Steps in maturation of influenza A virus
721 neuraminidase. *J. Virol.* 69, 5011–5017.

722 Sanders, R.W., and Moore, J.P. (2021). Virus vaccines: proteins prefer prolines. *Cell Host Microbe*
723 29, 327–333.

724 Sanders, R.W., Derking, R., Cupo, A., Julien, J.-P., Yasmeen, A., de Val, N., Kim, H.J., Blattner,
725 C., de la Peña, A.T., Korzun, J., et al. (2013). A next-generation cleaved, soluble HIV-1 Env trimer,
726 BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-
727 neutralizing antibodies. *PLoS Pathog.* 9, e1003618.

728 Sanjuán, R., Cuevas, J.M., Moya, A., and Elena, S.F. (2005). Epistasis and the adaptability of an
729 RNA virus. *Genetics* 170, 1001–1008.

730 Stadlbauer, D., Zhu, X., McMahon, M., Turner, J.S., Wohlbold, T.J., Schmitz, A.J., Strohmeier, S.,
731 Yu, W., Nachbagauer, R., Mudd, P.A., et al. (2019). Broadly protective human antibodies that
732 target the active site of influenza virus neuraminidase. *Science* 366, 499–504.

733 Starr, T.N., and Thornton, J.W. (2016). Epistasis in protein evolution. *Protein Sci. Publ. Protein*
734 *Soc.* 25, 1204–1218.

735 Strohmeier, S., Amanat, F., Zhu, X., McMahon, M., Deming, M.E., Pasetti, M.F., Neuzil, K.M.,
736 Wilson, I.A., and Krammer, F. (2021). A novel recombinant influenza virus neuraminidase vaccine
737 candidate stabilized by a measles virus phosphoprotein tetramerization domain provides robust
738 protection from virus challenge in the mouse model. *mBio* 12, e0224121.

739 Tamura, D., Nguyen, H.T., Sleeman, K., Levine, M., Mishin, V.P., Yang, H., Guo, Z., Okomo-
740 Adhiambo, M., Xu, X., Stevens, J., et al. (2013). Cell culture-selected substitutions in influenza
741 A(H3N2) neuraminidase affect drug susceptibility assessment. *Antimicrob. Agents Chemother.* 57,
742 6141–6146.

743 Vavricka, C.J., Li, Q., Wu, Y., Qi, J., Wang, M., Liu, Y., Gao, F., Liu, J., Feng, E., He, J., et al.
744 (2011). Structural and functional analysis of laninamivir and its octanoate prodrug reveals group
745 specific mechanisms for influenza NA inhibition. *PLoS Pathog.* 7, e1002249.

746 Venkatramani, L., Bochkareva, E., Lee, J.T., Gulati, U., Graeme Laver, W., Bochkarev, A., and Air,
747 G.M. (2006). An epidemiologically significant epitope of a 1998 human influenza virus
748 neuraminidase forms a highly hydrated interface in the NA-antibody complex. *J. Mol. Biol.* 356,
749 651–663.

750 Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., and
751 Bricogne, G. (2011). Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr.*
752 *D Biol. Crystallogr.* 67, 293–302.

753 Wan, H., Gao, J., Yang, H., Yang, S., Harvey, R., Chen, Y.-Q., Zheng, N.-Y., Chang, J., Carney,
754 P.J., Li, X., et al. (2019). The neuraminidase of A(H3N2) influenza viruses circulating since 2016
755 is antigenically distinct from the A/Hong Kong/4801/2014 vaccine strain. *Nat. Microbiol.* 4, 2216–
756 2225.

757 Wang, Y., Lei, R., Nourmohammad, A., and Wu, N.C. (2021). Antigenic evolution of human
758 influenza H3N2 neuraminidase is constrained by charge balancing. *eLife* 10, e72516.

759 Weiss, C.D., Wang, W., Lu, Y., Billings, M., Eick-Cost, A., Couzens, L., Sanchez, J.L., Hawksworth,
760 A.W., Seguin, P., Myers, C.A., et al. (2020). Neutralizing and neuraminidase antibodies correlate
761 with protection against influenza during a late season A/H3N2 outbreak among unvaccinated
762 military recruits. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* 71, 3096–3102.

763 Wu, N.C., and Wilson, I.A. (2017). A perspective on the structural and functional constraints for
764 immune evasion: insights from influenza virus. *J. Mol. Biol.* 429, 2694–2709.

765 Wu, N.C., Thompson, A.J., Xie, J., Lin, C.-W., Nycholat, C.M., Zhu, X., Lerner, R.A., Paulson, J.C.,

766 and Wilson, I.A. (2018). A complex epistatic network limits the mutational reversibility in the
767 influenza hemagglutinin receptor-binding site. *Nat. Commun.* 9, 1264.

768 Wu, N.C., Otwinowski, J., Thompson, A.J., Nycholat, C.M., Nourmohammad, A., and Wilson, I.A.
769 (2020). Major antigenic site B of human influenza H3N2 viruses has an evolving local fitness
770 landscape. *Nat. Commun.* 11, 1233.

771 Zhang, J., Kobert, K., Flouri, T., and Stamatakis, A. (2014). PEAR: a fast and accurate Illumina
772 Paired-End reAd mergeR. *Bioinformatics* 30, 614–620.

773 Zhu, X., McBride, R., Nycholat, C.M., Yu, W., Paulson, J.C., and Wilson, I.A. (2012). Influenza
774 virus neuraminidases with reduced enzymatic activity that avidly bind sialic acid receptors. *J. Virol.*
775 86, 13371–13383.

776 Zhu, X., Turner, H.L., Lang, S., McBride, R., Bangaru, S., Gilchuk, I.M., Yu, W., Paulson, J.C.,
777 Crowe, J.E., Ward, A.B., et al. (2019). Structural basis of protection against H7N9 influenza virus
778 by human anti-N9 neuraminidase antibodies. *Cell Host Microbe* 26, 729-738.e4.

779