www.sciencemag.org/cgi/content/full/328/5983/1272/DC1

Supporting Online Material for

Permissive Secondary Mutations Enable the Evolution of Influenza Oseltamivir Resistance

Jesse D. Bloom, Lizhi Ian Gong, David Baltimore*

*To whom correspondence should be addressed. E-mail: baltimo@caltech.edu

Published 4 June 2010, *Science* **328**, 1272 (2010)
DOI: 10.1126/science.1187816

This PDF file includes:

Materials and methods

Plasmids

NAs from the following H1N1 viral strains were studied: A/WSN/33 (WSN), A/PR/8/34 (PR8), A/Texas/36/1991 (TX91), A/New Caledonia/20/1999 (NC99), A/Solomon Islands/3/2006 (SI06), and A/California/4/2009 (CA09). The protein sequences of these NAs, as well as the viral RNA (vRNA) sequences for WSN and NC99 NA segments, are shown in Table S2. The bidirectional polymerase I / polymerase II plasmids (*1*) encoding all gene segments from WSN (pHW181-PB2, pHW182-PB1, pHW183-PA, pHW184-HA, pHW185-NP, pHW186-NA, pHW187-M, pHW188-NS) and PR8 (of which pHW196-NA was used here), as well as the empty cloning plasmid pHW2000, were kind gifts from Robert Webster and Richard Webby of St. Jude Children's Research Hospital. The TX91 virus, a NC99 HA/NA reassortant with PR8, a SI06 HA/NA reassortant with PR8, and the CA09 virus were obtained from the Biodefense and Emerging Infections Resource Repository (BEI Resources, catalog numbers of NR-3223, NR-9692, NR-9693, and NR-13658, respectively). Total RNA was extracted from these viruses, and used as a template for primer specific reverse-transcriptase PCR to amplify the NA gene segments from all four strains, and all gene segments from TX91. Any BsmBI sites internal to the segments were removed by strand overlap extension (SOE-ing) mutagenesis PCR (*2*), and the segments were then cloned into the BsmBI sites of pHW2000 to create the following plasmids: pHWTX91-PB2, pHWTX91-PB1, pHWTX91-PA, pHWTX91-HA, pHWTX91-NP, pHWTX91-NA, pHWTX91-M, pHWTX91-NS, pHWNC99-NA, pHWSI06-NA, and pHWCA09-NA.

The H274Y mutation and the other mutations studied here were introduced into the NAs by SOE-ing PCR. The H274Y mutation is named in accordance with the N2 numbering scheme (to maintain consistency with prior literature), while all other mutations are named according to sequential numbering of the relevant NA beginning with one at the N-terminal methionine. The locations of the mutations in the different numbering schemes is summarized in Table S1. The activity assays in all figure panels except Figure 3C used NAs expressed from the relevant pHW2000-derived plasmid.

To construct the C-terminal epitope tagged plasmids used in the assays of Figure 3C, the NC99 NA protein coding sequences were cloned into an expression plasmid (HDM) containing a CMV promoter, 5' EcoRI and 3' NotI cloning sites, followed by an internal ribosome entry site (IRES) expressing mCherry. For the control untagged plasmids, the wildtype or H274Y NA coding sequences were cloned directly into the EcoRI / NotI cloning sites. For the epitope-tagged plasmids, an HA epitope tag (YPYDVPDYA) was added at the C-terminus of the NA.

The PB1flank-eGFP viral RNA was encoded in the polymerase I cassette plasmid pHH21 (*3*), a kind gift from Yoshihiro Kawaoka of the University of Wisconsin. The design of this viral RNA is shown in Figure S3; it contains the untranslated regions and the 80 coding nucleotides from each termini of the WSN PB1 gene segment. All potential start codons in the 80 coding nucleotides from the N-terminal end of the

PB1 coding sequence were mutated. The enhanced green fluorescent protein (eGFP) gene was placed between the two PB1 coding nucleotide regions with flanking BamHI and EcoRI sites. This design was inspired by earlier work (*4, 5*) indicating that the untranslated and terminal 80 coding nucleotides of PB1 were sufficient to ensure efficient packaging. The resulting plasmid was named pHH-PB1flank-eGFP. The red fluorescent protein mCherry version of this plasmid, pHH-PB1flank-mCherry, was constructed by replacing the eGFP with mCherry using the BamHI and EcoRI sites.

We created a plasmid to make lentivirus capable of transducing cells to constitutively express the WSN PB1 protein. The PB1 protein coding sequence was cloned into pHAGE2-CMV-W, the backbone plasmid for a second-generation self-inactivating lentiviral vector (Alex Balazs et al, to be described elsewhere). This plasmid could be used to create lentivirus with PB1 under the CMV promoter when transfected into 293T cells with appropriate helper plasmids. In order to prevent expression of the potentially toxic PB1-F2 peptide, we introduced a stop codon after the eighth residue of PB1-F2 as described in (*6*).

NA activity and expression assays

The NA expression and activity assays shown in Figures 1A, 1B, 1C, 2A, 2B, 3B, and 3C all utilized 293T cells. For each assay, fresh preps of plasmid DNA were prepared, and quantified using a NanoDrop 1000 (ThermoScientific). 293T cells were seeded in 12-well dishes at 2.5×10^5 cells/well, and the next day were

transfected with 1 µg of plasmid DNA using the BioT transfection reagent (Bioland Scientific). At the indicated times, the cells were collected by a very brief treatment with trypsin-EDTA, which was quickly neutralized by the addition of serum, and then removed by centrifigution. The cells were resuspended in the assay buffer described below. In this assay, potential sources of variation include: differences in numbers of cells per well, differences in transfection efficiency, and differences in cell collection efficiency. Therefore, great care was taken to perform the cell seeding, transfections, and collections consistently. Every data point shown on the graphs represents the mean and SEM of at least three independent transfections. As can be seen from these graphs, the SEMs are small compared to the means, indicating that the external factors were well controlled. For the assays shown in Figure 3C, we also possessed an internal transfection control, since the plasmids contained an IRES-mCherry (see above). We therefore were able to quantify the mCherry expression by flow cytometry as a proxy for transfection efficiency, and use this as a normalization factor for the measured activities / staining intensities. However, in practice there was very little variation in the percentage and intensity of mCherry positive cells, indicating reproducible transfections.

NA activity was assayed using the fluorogenic 2'-(4-Methylumbelliferyl)-alpha-D-N-acetylneuraminic acid (MUNANA) substrate. The MUNANA assays used a protocol based on that posted by the Neuraminidase Inhibitor Susceptibility Network at http://www.nisn.org/documents/A.Hurt_Protocol_for_NA_fluorescence.pdf. Because we were specifically interested in monitoring the activity of cell surface NA,

we modified the protocol to use a non-lysing buffer (isotonic, without detergents, and containing serum to maintain cell viability). This assay buffer consisted of 15 mM MOPS, 145 mM sodium chloride, 2.7 mM potassium chloride, and 4.0 mM calcium chloride, adjusted to pH 7.4. Immediately before use, the assay buffer was supplemented with 2% heat-inactivated fetal bovine serum. Wells of 96-well black Costar plates containing 100 µl of cells diluted in assay buffer were mixed with 50 µl of 0.3 mM MUNANA (Sigma M8639) in assay buffer, and incubated for one hour at 37 degrees C. The reaction was then quenched by adding 100 µl of 150 mM sodium hydroxide in 83% ethanol. The fluorescence was read in a Tecan Safire 2 plate reader (excitation 360 nm, slit width 5 nm; emission 448 nM, slit width 20 nm). The activities were quantified as the fluorescence above the background from untransfected cells. In all plots, the activities are given in arbitrary units such that the wildtype NA from that strain has a value of 100. The activity units among different plots are not comparable since each is normalized separately to its own wildtype; in addition, the raw readings themselves would not be comparable since the plate reader automatically adjusted the gain for optimal sensitivity.

For the antibody staining in Figure 1B, we used the NA2-1C1 mouse monoclonal IgG1-kappa antibody against the folded PR8 NA tetramer (BEI Resources, NR4540). Cells in assay buffer were incubated with a 1:20 dilution of NA2-1C1, washed twice, incubated with a goat anti-mouse IgGI TRI-COLOR conjugate secondary antibody (Invitrogen, M32006), then washed twice more, before performing flow cytometry. The mean fluorescent intensity above untransfected control cells was quantified.

Figure 3C shows the antibody staining of cells transfected with NAs containing C-terminal epitope tags. These cells were stained using a 1:100 dilution of a FITC-conjugated mouse monoclonal anti-HA antibody (Santa Cruz Biotechnology, HA-probe F-7, sc-7392 FITC). Cells were gated both on mCherry (to detect transfected cells) and on FITC. Staining was quantified as the FITC mean fluorescent intensity of the mCherry positive cells above the background of mCherry positive cells transfected with NA lacking the C-terminal epitope tag.

Cells, viruses, and viral growth experiments

The cells utilized here were the 293T human kidney line (ATCC, CRL-11268), the MDCK canine kidney line (ATCC, CCL-34), the A549 human lung carcinoma line (ATCC, CCL-185), and the MDCK-SIAT1 line (a variant of MDCK cells that overexpress the 2-6 sialic acid linkage preferred by human influenza (*7*); HPA Cultures, 05071502). To make variants of these cells that constitutively expressed the PB1 protein, all were transduced by the pHAGE2-CMV-PB1-W lentivirus described above. A control parallel transduction with a ZsGreen lentivirus induced fluorescence in essentially all cells. The resulting cell lines were named 293T-CMV-PB1, MDCK-CMV-PB1, A549-CMV-PB1, and MDCK-SIAT1-CMV-PB1. In order to reduce variability in subsequent viral growth assays, the A549-CMV-PB1 and MDCK-SIAT1-CMV-PB1 infected lines were expanded from a single cell by limiting dilution cloning. All cell lines were maintained in D10 media (DMEM supplemented with

10% heat-inactivated fetal bovine serum, 2 mM L-glutamine, 100 U/ml of penicillin, and 100 μg/ml of streptomycin) at 37 degrees C and 5% carbon dioxide.

To create the WSN-PB1flank-eGFP/mCherry viruses, cocultures of 293T-CMV-PB1 and MDCK- CMV-PB1 cells (3:1 ratio) were co-transfected with pHW181-PB1, pHW183-PA, pHW184-HA, pHW185-NP, pHW186-NA (with relevant secondary mutations), pHW187-M, pHW188-NS, and either pHH-PB1flank-eGFP or pHH-PB1flank-mCherry. After 10-12 hours, the cells were washed once with PBS, and the media was changed to influenza growth media (Opti-MEM I supplemented with 0.3% bovine serum albumin, 0.01% heat-inactivated fetal bovine serum, 100 U/ml penicillin, 100 μg/ml streptomycin, and 100 μg/ml calcium chloride) containing 2 μg/ml TPCK-treated trypsin. The viruses were allowed to grow for 48 before the supernatant was harvested through a 0.45 μm filter. After 48 hours, essentially all cells were brightly fluorescent, and substantial cytopathic effect was visible. Cytopathic effect was also visible in cells transfected with all of the plasmids except pHH-PB1flank-eGFP/mCherry, although it was less pronounced. Figure S3 shows microscope images of fluorescent cells, as well as limiting dilution assays indicating that nearly all viruses capable of causing cytopathic effects also carry the fluorescent protein gene. The PB1flank-eGFP/mCherry segments were stably maintained through five passages in A549-CMV-PB1 cells. Flow cytometry titering of viruses collected 48 hours post-transfection indicated titers on the order of 10^5 infectious particles per μl of collected supernatant.

The viruses carrying the NC99 NAs (from pHWNC99-NA) had their remaining six gene segments derived from TX91 (from pHWTX91-PB1, pHWTX91-PA, pHWTX91-HA, pHWTX91-NP, pHWTX91-M, pHWTX91-NS); the CMV-PB1 in the cells and the PB1flank-eGFP/mCherry were still derived from WSN. This was done simply because we lacked access to the full genome of NC99 virus; TX91 has greater than 98% protein identity to NC99. These viruses were produced as with the WSN-PB1flank-eGFP/mCherry viruses, except the cocultures utilized MDCK-SIAT1-CMV-PB1 cells rather than MDCK-CMV-PB1 cells, and the influenza growth media was supplemented with 4 μg/ml TPCK-trypsin. These viruses grew somewhat slower than the WSN-PB1flank-eGFP viruses, typically taking at least 60 hours to reach concentrations on the order of 10^5 infectious particles per μl of collected supernatant.

The viruses were titered by flow cytometry by infecting A549-CMV-PB1 or MDCK-SIAT1-CMV-PB1 cells in influenza growth media plus TPCK-trypsin with virus, and then collecting the cells after 12 hours for flow cytometry (fixing in 1% paraformaldehyde to inactivate virus). Examples of the type of flow cytometry data produced are in Figure 1D. The 12-hour timepoint was chosen for titering because there was bright fluorescence of the primary infections without substantial fluorescence from secondary infections. The titers computed by flow cytometry closely matched those measured by limiting dilution tissue-culture infectious dose 50% assays (Figure S3).

For the WSN-PB1flank viruses, the relative fitnesses of the viruses were compared in head-to-head competition growth assays by infecting an estimated 50 infectious particle of each color into 6-well dishes containing $2.5x10^5$ A549-CMV-PB1 cells in influenza growth media + 2 μg/ml TPCK-trypsin. The cells were harvested for flow cytometry after 46 to 50 hours, a timepoint at which typically 5 to 30% of cells were fluorescent. This produced the type of data shown in Figure 2D. The relative fitnesses of the viruses were calculated as the ratio of their Malthusian growth parameters (*8*).

For the NC99 NA / TX91 PB1flank viruses, we had great difficulty performing head-to-head competition assays in MDCK-SIAT1-CMV-PB1 cells (we wanted to use this cell line since it expresses high levels of the sialic acid linkage preferred by human influenza). The difficulty was in correctly timing the cell collections so a reasonable percent (5 to 30%) of cells were fluorescent. We suspect this difficulty is because of more efficient influenza growth/maturation in MDCK versus A549 cells (*9*), leading to a narrower time window when a modest number of cells are infected. Because we were rapidly exhausted by poorly timed experiments that required virus collections at 2 AM, we instead resorted to the less elegant but more straightforward approach shown in Figure 3D. In this approach, we simply infected wells containing $2.5x10^5$ MDCK-SIAT1-CMV-PB1 cells with an estimated 50 PB1flank-eGFP infectious particles, and then collected supernatant samples at the indicated times. The virus titers of the supernatants were in turn determined using flow cytometry. The growth assays were performed both in influenza growth media

with 4 μg/ml TPCK-trypsin and the same media also supplemented with 250 nM oseltamivir carboxylate (a kind gift from J. Smith and A. Perrin of F. Hoffmann-La Roche).

Phylogenetic tree

The phylogenetic tree shown in Figure 3A was built using the coding nucleotide sequences for the NAs from TX91, NC99, and all full-length human seasonal H1N1 influenza A viruses (excluding those from the swine-origin 2009 pandemic H1N1 strain) isolated in 2006 or later, excluding lab strains and removing sequences encoding for identical proteins, as taken from the Influenza Virus Resource (*10*) on August-31-2009. These sequences were aligned using MUSCLE (*11*). A maximum-likelihood tree was then constructed using RAxML version 7.0.4 (*12*) with 100 rapid-bootstrap replicates (*13*), using the general time reversible (GTR) model of nucleotide substitution with the gamma model of rate heterogeneity. This tree was rooted to TX91, which forms a clear outgroup since it predates all of the other sequences used by at least eight years in the human seasonal H1N1 lineage. After using it to root the tree, the TX91 sequence was removed from the tree image displayed in Figure 3A. Bootstrap values were provided by RAxML from the 100 rapid-bootstrap replicates, and bootstrap percentages for the two key nodes are shown in Figure 3A.

The maximum-likelihood reconstruction of sequences at interior nodes on the phylogenetic tree was performed using the PAML program, version 4.4 (*14*). The

reconstruction was done at the level of the protein sequences, using the Poisson model assuming equal rates for any amino acid substitution and using equal amino acid distances. All of the reconstructed residue identities at interior nodes had probabilities of >99% (the short branch lengths and high density of sequences makes for relatively straightforward reconstructions at interior nodes). In the reconstructed tree, it was not possible to distinguish the relative order of the D382N/V234M and the R222Q/E214G. The actual image shown in Fig. 3A was drawn using FigTree (http://beast.bio.ed.ac.uk/FigTree).

Prediction of R194G WSN NA mutation

The R194G mutation was predicted using the bioinformatics approach described in detail in (*15*). The prediction was made using a modified version of the PIPS computer program (available upon request) that differed from the one described in (*15*) primarily in using more sophisticated optimization routines. The predictions utilized regularizing priors. After excluding mutations to sites not present in the N1 NA crystal stucture or found in antigenic sites (which may be under positive antigenic selection, confounding the PIPS approach), R194G was the top predicted mutation. The PIPS program was originally developed for predicting stabilizing mutations; R194G appears to improve some aspect of folding rather than stability to irreversible thermal denaturation. We nonetheless chose to test a PIPS predicted mutation because we felt that the approach of PIPS made it likely to be able to predict mutations that improved folding efficiency as well as stability. It is certainly

clear from the experimental results shown here that PIPS made a very good prediction when it came to R194G. However, once we identified R194G we did not test any further predictions, since our goal was simply to find a proof-of-principle secondary mutations to WSN could rescue both surface-expressed NA activity and viral fitness. Our results therefore do not, at this time, represent an unbiased systematic test of the performance of PIPS for predicting NA mutations that compensate for H274Y.

Table S1. The locations of mutations in different NA numbering schemes. The first column gives the name used to refer to the mutation in the text. The second column gives the number of that residue in sequential numbering of the NA in which it occurs, beginning with one for the N-terminal methionine. The third column gives the number of that residue in the N2 numbering scheme, which is the scheme used in the 2HU0 PDB structure of the H5N1 NA that is used in Figure S2.

Name of mutation in text	**Number in sequential numbering of NA**	**Number in N2 numbering scheme**
H274Y to WSN NA	259 in WSN	274
H274Y to PR8 NA	260 in WSN	274
H274Y to NC99 NA	275 in NC99	274
H274Y to TX91 NA	275 in TX91	274
H274Y to SI06 NA	275 in SI06	274
H274Y to CA09 NA	275 in CA09	274
R194G to WSN NA	194 in WSN	209
R222Q to NC99 NA	222 in NC99	221
V234M to NC99 NA	234 in NC99	233
V48I to NC99 NA	48 in NC99	52
E214G to NC99 NA	214 in NC99	213
D382N to NC99 NA	382 in NC99	385

Table S2. NA protein and viral RNA sequences used in this study.

>WSN NA wildtype protein sequence used in this study
mnpnqkiitigsicmvvgiislilqigniisiwishsiqtgnqnhtgicnqgsitykvvagqdstsviltgnsslcpirgwaihskdngirigskgdvfvir
epfiscshlecrtffltqgallndkhsrgtfkdrspyralmscpvgeapspynsrfesvawsasachdgmgwltigisgpddgavavlkynriitetikswr
knilrtqesectcvngscftimtdgpsdglasykifkiekgkvtksielnapnshyeecscypdtgkvmcvcrdnwhgsnrpwvsfdqnldykigyicsgvf
gdnprpkdgtgscgpvsadgangvkgfsykygngvwigrtksdssrhgfemiwdpngwtetdsrfsmrqdvvaitnrsgysgsfvqhpeltgldcmrpcfwv
elirglpeedaiwtsgsiisfcgvnsdtvdwswpdgaelpftidk

>PR8 NA wildtype protein sequence used in this study
mnpnqkiitigsiclvvglislilqigniisiwishsiqtgsqnhtgicnqniityknstwvkdttsviltgnsslcpirgwaiyskdnsirigskgdvfvi
repfiscshlecrtffltqgallndkhssgtvkdrspyralmscpvgeapspynsrfesvawsasachdgmgwltigisgpdngavavlkyngiitetiksw
rkkilrtqesecacvngscftimtdgpsdglasykifkiekgkvtksielnapnshyeecscypdtgkvmcvcrdnwhgsnrpwvsfdqnldygigyicsgv
fgdnprpedgtgscgpvyvdgangvkgfsyrygngvwigrtkghssrhgfemiwdpngwtetdskfsvrqdvvamtdwsgysgsfvqhpeltgldcmrpcfw
velirgrpkektiwtsassisfcgvnsdtvdwswpdgaelpfsidk

>TX91 NA wildtype protein sequence used in this study
mnpnqkiitigsismaigiislilqigniisiwashsiqtgsqnhtgicnqriityenstwvnqtyvninntnvvagkdktsvtlagnsslcpirgwaiytk
dnsirigskgdvfvirepfiscshlecrtffltqgallndkhsngtvkdrspyralmscplgeapspynsrfesvawsasachdgmgwltigisgpdngava
vlkyngiitetikswkkrilrtqesecvcvngscftimtdgpsngaasyrifkiekgkvtksieldapnyhyeecscypdtgtvmcvcrdnwhgsnrpwvsf
nqnldyqigyicsgvfgdnprpkdgegscnpvtvdgadgvkgfsyrygngvwigrtksnrlrkgfemiwdpngwtdtdsdfsvkqdvvamtdwsgysgsfvq
hpeltgldcmrpcfwvelirgrprenttiwtsgssisfcgvnsdtanwswpdgaelpftidk

>NC99 NA wildtype protein sequence used in this study
mnpnqkiitigsisiaigiislmlqigniisiwashsiqtgsqnhtgvcnqriityenstwvnhtyvninntnvvagkdktsvtlagnsslcsisgwaiytk
dnsirigskgdvfvirepfiscshlecrtffltqgallndkhsngtvkdrspyralmscplgeapspynskfesvawsasachdgmgwltigisgpdngava
vlkyngiitetikswkkrilrtqesecvcvngscftimtdgpsngaasykifkiekgkvtksielnapnfhyeecscypdtgtvmcvcrdnwhgsnrpwvsf
nqnldyqigyicsgvfgdnprpkdgegscnpvtvdgadgvkgfsykygngvwigrtksnrlrkgfemiwdpngwtdtdsdfsvkqdvvaitdwsgysgsfvq
hpeltgldcirpcfwvelvrglprenttiwtsgssisfcgvnsdtanwswpdgaelpftidk

>SI06 NA wildtype protein sequence used in this study
mnpnqkiitigsisiaigiislilqigniisiwashsiqtgsqnhtgicnqriityenstwvnntyvninntnvvaekdktsvtlagnsslcsisgwaiytk
dnsirigskgdvfvirepfiscshlecrtffltqgallndkhsngtvkdrspyralmscplgeapspynsrfesvawsasachdgmgwltigisgpdngava
vlkyngiitetikswkkrilrtqesecvcmngscftimtdgpsngaasykifkiekgkvtktielnapnfhyeecscypdtgtvmcvcrdnwhgsnrpwvsf
nqnldyqigyicsgvfgdnprpkdgegscnpvtvdgadgvkgfsykygngvwigrtksnrlrkgfemiwdpngwtntdsdfsvkqdvvaitdwsgysgsfvq
hpeltgldcirpcfwvelvrglprenttiwtsgssisfcgvnsgtanwswpdgaelpftidk

>CA09 NA wildtype protein sequence used in this study
mnpnqkiitigsvcmtigmanlilqigniisiwishsiqlgnqnqietcnqsvityenntwvnqtyvnisntnfaagqsvvsvklagnsslcpvsgwaiysk
dnsvrigskgdvfvirepfiscsplecrtffltqgallndkhsngtikdrspyrtlmscpigevpspynsrfesvawsasachdginwltigisgpdngava
vlkyngiitdtikswrnnilrtqesecacvngscftvmtdgpsngqasykifriekgkivksvemnapnyhyeecscypdsseitcvcrdnwhgsnrpwvsf
nqnleyqigyicsgifgdnprpndktgscgpvssngangvkgfsfkygngvwigrtksissrngfemiwdpngwtgtdnnfsikqdivginewsgysgsfvq
hpeltgldcirpcfwvelirgrpkentiwtsgssisfcgvnsdtvgwswpdgaelpftidk

>viral RNA sequence for WSN NA segment used in this study
aguagaaacaaggaguuuuuugaacaaacuacuugucaauggugaacggcaacucagcaccgucuggccaagaccaaucuacaguaucacuauucacaccac
aaaaagaaaugaugcucccacuaguccagauugcgucccuccucagguagccccugauuaaauucaacccagaagcaaggccucauacagucuagcccuguua
gcucaggauguugaacgaaacuuccgcuguacccugaccgauuaguuauugccacaacaucuugucucauagagaaccuacuaucagucucuguccauccau
uaggaucccaaaucaucucaaacccaugucuggaacugucacuuuuaguccuuccuauccaaacaccauugccauacuuauaugaaaaucccuuuacucccgu
uugcuccaucagcagacacugggccacagcugccuguuccaucuuugggacgcggguugucaccgaaaacccacugcagauguauccuauuuuauaaacua
gguuuuggucgaaggacacccauggucgguucgaaccgugccaauugucucugcacacacacaucacuuugccgguaucaggguaacaggaacauuccucgu
agugagaauuaggugcauucaacucuaucgauuuaguaaccuucccccuucucgaucuugaaaauuuuguacgaggccagcccaucacuuggggccaucgguca
uuaugguaaaacaugaaccauuuacacagguacauucagacucuuguguucucaauauauucuuccuccaacuuuuuaugguuucaguuauuauugcgguugu
auuuuaauacagccacugcuccaucaucuggaccagaaauuccgauuguuagccagcccauuccaucaugacaugcacuugcugaccaagcaaccgauucaa
accuugaauuguacggggacggagcuucaccgacagggcagcucauuaaggcccuauaaagggcuucuguccuuaaaggucccccuugaaugcuugucauuca
guaaggcgccuugagucagaaaaaaggcuccugcauucaagugagaacaugaaaauaaaaggcucucuuaugacaaaaacgucuccuuuggaaccaauucuua
ugccauugucuuugcuguguauagcccacccacggaugggacaaagagaugaauugccgguuaauaucacugaaguugaguccugcccagcaacaacuuuau
agguaaugcugccuugguugcauauuccaguaugguuuugauuuccgguuugaauugaauggcuaauccauauugagauuauauuuccuauuugcaauauua
ggcuaauuauuccgacuaccauacagauugacccaaugguuauuauuuucugguuuggauucauuuaaacuccugcuuucgcu

>viral RNA sequence for NC99 NA segment used in this study
aguagaaacaaggaguuuuuucaacgaacuacuugucaauggugaacggcaacucagcaccgucuggccaagaccaguuugcaguaucacuauuuacgccac
aaaaagaaaugcugcucccacuaguccagauuguuguauuuucucuaggcaguccucugacuaacuacaacccagaagcaaggucuuauacaguccaauccug
uuaacucaggauguugaacgaaacuuccgcuguacccugaccaaucaguuauugccacaacaucccuguuucacugagaaaucacugucgguaucuguccauc
cauuaggaucccaaaucaucaaacccuuucuaaagucuguuacuuuuagucccuucccuaucaaaaccauauuuguaugaaaacccccuuuacuc
cgucugcuccaucaacagucacuggauuacagcugcccucuccaucuuugggacgcggauugucaccgaacaccccacugcagauguauccuauuugauaaau
ccagguuuugauuaaaagacacccaaggucgauuugaaccaugccaguugucccugcauacacacaucacugugccagugucuggguaacaggaacauuccu
cauaaugaaaauugggugcauucaacucuauugauuuaguaacccuucccuuuucgaucuugaagauuuuguacgaggcggcccccauuacucgggccaucgg
ucauuauggugaaacaugacccguucacacagacacauucagacucuuguguucuuaauauucgcuuuuucaacuuuuuaugguuucaguuauuaugccgu
uguauuuuaguacagccacagcuccauugucuggaccagaaauuccgauuguuaaccagcccaugccaucaugcaugcgcuugcugaccaugcaacugauu
caaacuuugaauuguauggggacggagcuucaccuagaggacagcucauuaaggcccuauaaggacuucugcuccuuaacggucccauuugaauguuugucau
uuaauagagcaccuugggucagaaaaaaggguucugcauucaaguagagaacaugauaugaaaggguucucuuaugacaaaaacaucucccuuuggagccaauuc
uuaugcuguugucuuuuguguauauagcccauccacugauagaacaaagagaugaauugccggccaaugucacugaaguuuugcccuuuccagcaacaacau
uaguguuguuaaauauuaacauaugugugauucacccaggugcuguuuucauaugugaugauucuuuggguugcauacuccaguggguuuugacuuccaguu
ggauugagugacuagcccauauugaaauaauauuuccuauuugcaacauuagacuaauuauuccgauugcuauacugauugauccaaugguuauuauuuuuu
gauuuggauucauuuuaaacuccugcuuucgcu

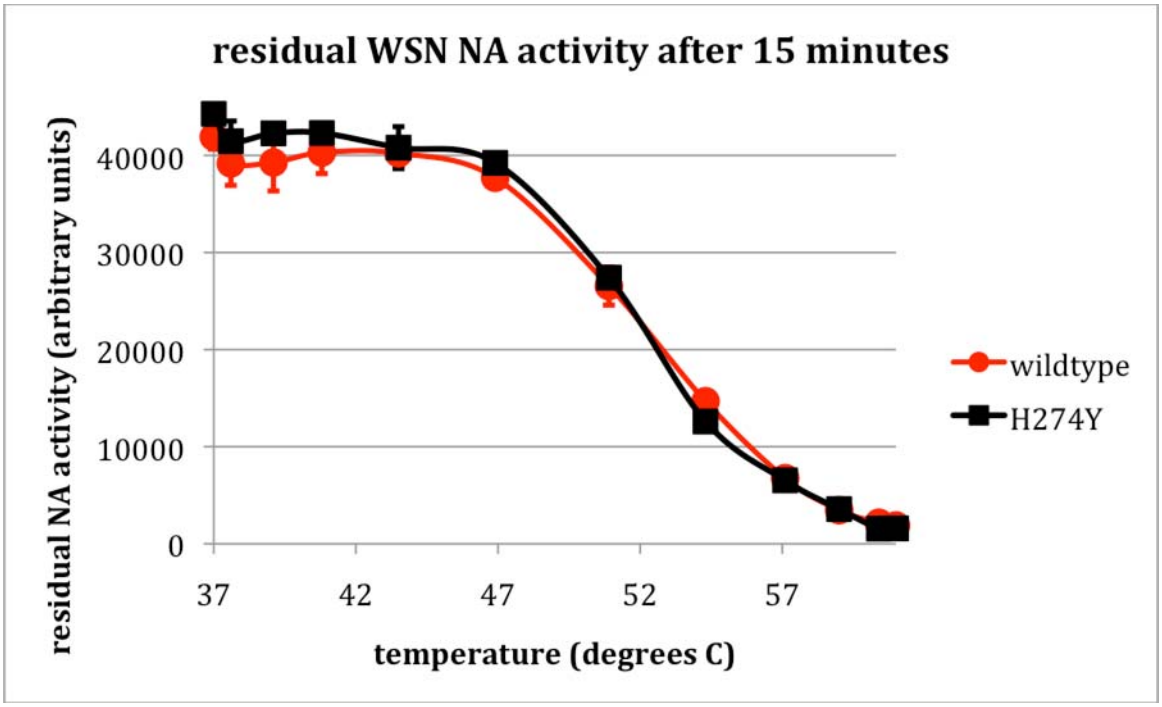

Fig. S1. Wildtype and H274Y WSN NAs are equally stable to irreversible thermal denaturation. Wildtype and H274Y WSN-PB1flank-eGFP virus was produced in 293T-CMV-PB1 / MDCK-CMV-PB1 cocultures. The viruses were adjusted to concentrations such that they possessed about equal NA activities in the MUNANA assay. Viruses in assay buffer were then heated to the indicated temperatures for 15 minutes using the gradient setting on a PCR cycler, and cooled to room temperature. The residual NA activity above background was assessed using the MUNANA assay.

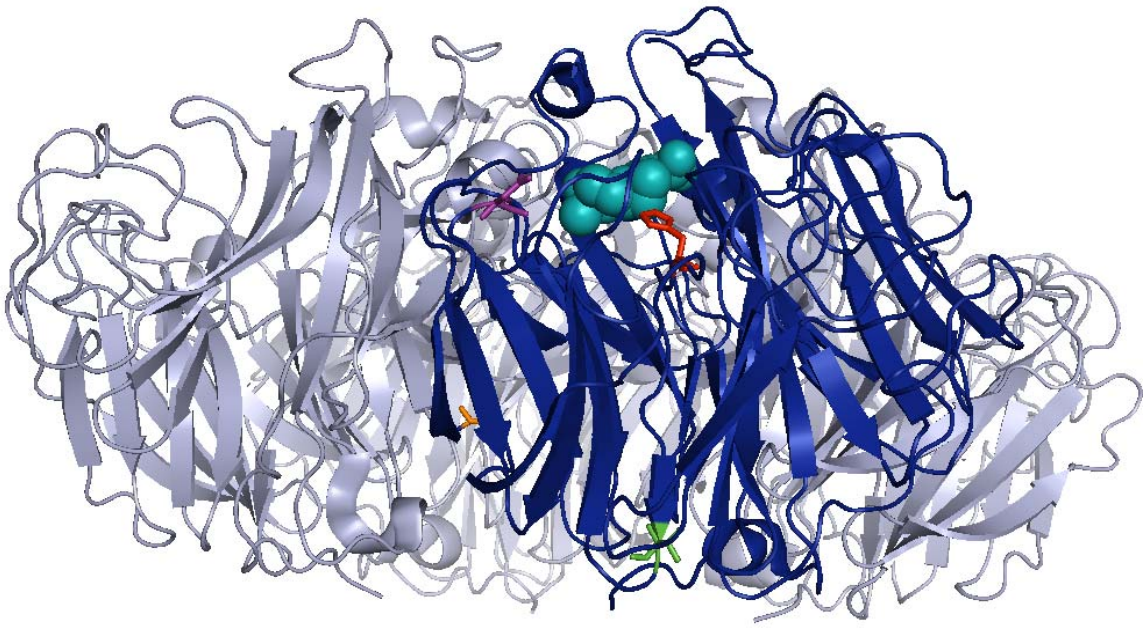

Fig. S2. Location of mutations on the N1 NA structure. Shown is the crystal structure of oseltamivir bound to NA from the avian H5N1 influenza strain A/Vietnman/1203/04 from PDB code 2HU0 (*16*). One monomer of the NA tetramer is shown in dark blue ribbons, while the remaining three monomers are shown in slate ribbons. The oseltamivir molecule is shown in teal spheres. Residue H274 is shown in red sticks. The site homologous to R222Q in NC99 NA is shown in purple sticks, the site homologous to V234M in NC99 NA is shown in green sticks, and the site homologous to R194G in WSN NA is shown in orange sticks. The residue mapping between the numbers used in the text and those in this PDB structure are given in Table S1. The image was drawn using PyMol.

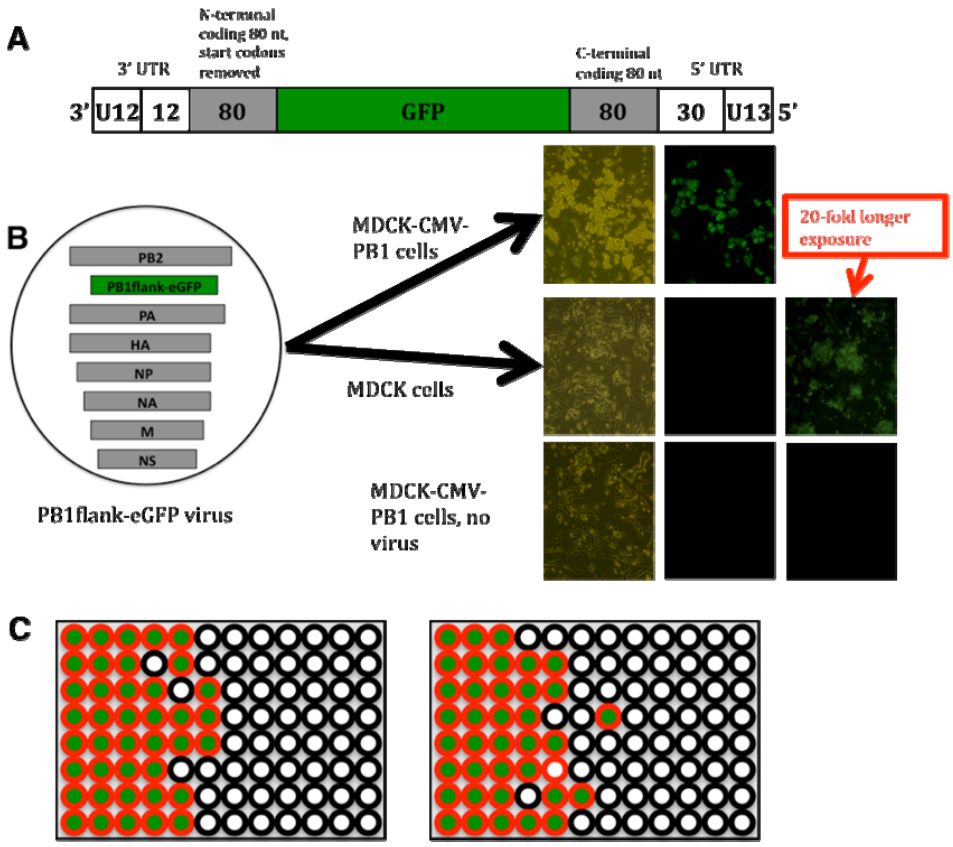

Fig. S3. WSN-PB1flank-eGFP viruses. **(A)** PB1flank-eGFP viral RNA contains the untranslated regions from the PB1 segment, plus the 80 terminal coding nucleotides. Potential start codons in the N-terminal coding region were mutated. PB1flank-mCherry RNA was identical except mCherry replaced eGFP. **(B)** When infected with PB1flank-eGFP viruses, cells constitutively expressing PB1 protein turned green. When the same viruses infected normal cells, only dim fluorescence was produced (due to PB1 proteins packaged into virions). **(C)** 96-well plates of MDCK-CMV-PB1 cells were infected with an estimated 10 PB1flank-eGFP infectious particles (as determined by flow cytometry) per well (first column) and serial 2-fold dilutions. Wells were scored for fluorescence (green fill) and cytopathic effect (red outlines). Nearly all viruses causing cytopathic effect also expressed eGFP.

References for supporting material

1. E. Hoffmann, G. Neumann, Y. Kawaoka, G. Hobom, R. G. Webster, *Proc Natl Acad Sci U S A* **97**, 6108 (2000).
2. R. Higuchi, B. Krummel, R. K. Saiki, *Nucleic Acids Res* **16**, 7351 (1988).
3. G. Neumann *et al.*, *Proc Natl Acad Sci U S A* **96**, 9345 (1999).
4. Y. Liang, Y. Hong, T. G. Parslow, *J Virol* **79**, 10348 (2005).
5. Y. Muramoto *et al.*, *J Virol* **80**, 2318 (2006).
6. W. Chen *et al.*, *Nat Med* **7**, 1306 (2001).
7. M. Matrosovich, T. Matrosovich, J. Carr, N. A. Roberts, H. D. Klenk, *J Virol* **77**, 8418 (2003).
8. A. F. Maree, W. Keulen, C. A. Boucher, R. J. De Boer, *J Virol* **74**, 11067 (2000).
9. M. Ueda *et al.*, *Virus Res* **136**, 91 (2008).
10. Y. Bao *et al.*, *J Virol* **82**, 596 (2008).
11. R. C. Edgar, *Nucleic Acids Res* **32**, 1792 (2004).
12. A. Stamatakis, *Bioinformatics* **22**, 2688 (2006).
13. A. Stamatakis, P. Hoover, J. Rougemont, *Syst Biol* **57**, 758 (2008).
14. Z. Yang, *Mol Biol Evol* **24**, 1586 (2007).
15. J. D. Bloom, M. J. Glassman, *PLoS Comput Biol* **5**, e1000349 (2009).
16. R. J. Russell *et al.*, *Nature* **443**, 45 (2006).