Supporting Information

Shen et al. 10.1073/pnas.1113107108

SI Text

Previous reports have indicated that all three mutants form monomers in gel filtration (1, 2), and that R416A is defective in RNA binding (3). The WT NP was shown by electron microscopy to form trimers even when bound to a 24-mer synthetic RNA (1). Using analytical ultracentrifugation (AUC), we found that it exists as trimers in the free form, but forms hexamers predominantly (with some trimers coexisting) in the presence of the same synthetic RNA (Fig. S1 *A* and *B*). The difference is likely caused by the very dilute condition used in EM studies. We then used AUC to examine the effect of mutations on the oligomerization properties of NP. The results indicate that E339A, R416A, Δ402–428, and E339A/Δ402–428 mutants of NP exist in monomers instead of trimers in the absence of RNA (Fig. S1 *C*, *E*, *G*, and *I*). The secondary structures of the monomers, however, were not perturbed substantially, as the circular dichroism (CD) spectra of the mutants are nearly identical to that of the WT NP (Fig. S1*K*). Interestingly, when mixed with the 24-mersynthetic RNA, E339A, R416A, Δ402–428, and E339A/Δ402–428 mutants remained as monomers (Fig. S1 *E*, *F*, *H*, and *J*) whereas wild type NP formed a mixture of trimers and hexamers (Figs. S1*B*). This result suggests that the RNA binding properties of the mutants were significantly perturbed.

SI Materials and Methods. ***Materials.*** The sources of the compounds in the library were described in our recent report (4). Compound **3**, **7**, **12**, and **23** were purchased from Molport, AMRI, Life Chemicals and Enamine, respectively. Compounds were prepared at 10 mM in 100% dimethyl sulfoxide (DMSO). Tail-loop peptide, peptide 1, 2, and 3 were synthesized by the peptide synthesis lab (Institute of Biological Chemistry, Academia Sinica), GL Biochem, and Dr. R. P. Cheng's lab (National Taiwan University), and the purity of peptides was determined by HPLC and mass spectrometry to be >90%.

Viruses and cells. HEK293T and MDCK cells were grown in DMEM containing 10% FBS. Cells were maintained at 37 °C and 5% CO_2. The NP WT and mutant MDCK stable cell lines were established by the Retro-X™ Universal Packaging System (Clontech Laboratories Inc). The system includes the GP2-293 cell line, which has the viral gag and pol genes incorporated in its genome. Influenza virus A/WSN/33 (H1N1) was provided by Dr. Shin-Ru Shih (Chang Gung University, Taiwan). Virus was propagated in MDCK cells and embryonated hen eggs. The virus titer was determined by plaque assays.

Plasmids. Plasmids pcDNA-PB1, -PB2, -PA, -NP, pClneo-NP-HA and -NP-FLAG, encoding, respectively, the polymerase basic protein 1 (PB1), basic protein 2 (PB2), acidic protein (PA), and nucleoprotein (NP) proteins of the WSN virus have been described previously (5) and obtained from Dr. K. S. Jeng. NP Δ402–428, R416A, E339A and E339A/Δ402–428 were generated by QuickChange Site-Directed Mutagenesis Kit (Stratagene). The DNA sequences of desired mutations were confirmed by sequencing. Plasmid pEGFP-tail-loop was constructed with the WSN virus tail-loop sequence (amino acids 402–428) inserted into pEGFP-C3 vector (BD Biosciences Clontech). The pPOLI-Luc-RT plasmid (from Dr. K. S. Jeng) contained the firefly luciferase open reading frame in negative orientation flanked by the noncoding regions of nonstructural (NS) sequence of the WSN virus and by the human RNA polymeraseI promoter and the mouse RNA polymeraseI terminator (6). The Retro-X™ Universal Packaging System contains the pLNCX and pVSV-G plasmids. Plasmid pLNCX is designed for retrovial gene delivery and expression. Plasmid pVSV-G can express pantropic vesicular stomatitis virus envelope glycoproteins (VSV-G). The pLNCX-NP-V5 plasmid contains NP gene that was inserted into the pLNCX vector and the V5 tag is in the C-terminal.

Expression and purification of NP and mutants. The NP gene from the WSN virus coding a 498-residue protein was cloned into vector pET15b and expressed in BL21-CodonPlus® (DE3)-RIPL cell (Stratagene). The purification of NP and its mutants for AUC analyses followed a previously reported procedure (1). In brief, the expressed NP proteins in *Escherichia coli* were purified using Chelating Sepharose FF, Hiprep heparin FF and HiLoad Superdex 200 16 × 60 columns (GE Healthcare). The purification buffer contains 50 mM Tris, pH 8.0, 200 mM NaCl, 2 mM EDTA, and 2 mM β-mercaptoethanol. The NP obtained at this stage contained endogenous RNA derived from *E. coli*. To obtain the NP protein free of RNA, the purified proteins were incubated with ribonuclease A. Purified NP proteins were confirmed by SDS-PAGE and MS analysis.

Luciferase-based reporter assays. Approximately 4×10^5 HEK293T cells in six-well plates were transfected with pPOLI-Luc-RT, pcDNA-PB1, -PB2, -PA and -NP (1 μg each) by jetPEI transfection reagent (PolyPlus). At 30 h after transfection, the cell extracts were examined for firefly luciferase levels with luciferase assay system (Promega) and measured by luminescence reader.

Western blot analysis. NP WT and mutant MDCK stable cells with expressed NP protein or HEK 293T cells were transfected with 2 μg plasmid. After 24 h, the cells were incubated with DMEM containing 0.2% BSA and 25 mM Hepes, and infected with influenza virus A/WSN/33 at a multiplicity of infection (MOI) of 0.2 for 9 h. Cells were collected and lysed with 50 μL of lysis buffer (0.2% Triton X-100 in PBS). The supernatant was added with the SDS sample buffer. Aliquots of the supernatant were fractionated on a 10% NuPAGE Novex Bis-Tris gel (Invitrogen) and electroblotted onto BioTrace PVDF polyvinylidene fluoride transfer membrane (Pall Corporation). Membranes were blocked with skim milk for 1 h at room temperature, then incubated with specific antibodies and visualized by a Western Lightening chemiluminescence reagent plus kit (PerkinElmer Life Science).

Plaque assay. NP WT and mutant MDCK stable cells with expressed NP protein or HEK 293T cells were transfected with 2 μg plasmid. After 24 h, the cells were incubated with medium A (DMEM containing 0.2% BSA and 25 mM Hepes) and 25 mM Hepes and were infected with influenza virus A/WSN/33 at a MOI of 0.2 for 9 h. The cell culture supernatant was collected. Monolayer of 10^6 MDCK cells in six-well plates was inoculated with 1 mL of virus dilution. Serial (powers of ten) dilutions were made from cell culture supernatant. After 1 h, the inoculums were removed and the cells were washed twice with PBS. The cells were covered with 2 mL of agar medium (100 mL of 2 × medium A and 100 mL of 2% agar). After 3 days, cells were fixed with 10% formaldehyde in PBS for 30 min. The agar was cleaned and 1% crystal violet in 20% ethanol was added to facilitate plaque counting.

Analytical ultracentrifuge analysis. Sedimentation velocity (SV) experiment was performed by a Beckman–Coulter XL-I analytical ultracentrifuge (Fullerton). Samples and buffers were loaded into 12-mm standard double-sector Epon charcoal-filled centrepieces and mounted in an An-60 Ti rotor. SV experiments were performed at rotor speed of 40,000 rpm at 20 °C. The signals of samples were monitored at 280 nm. The partial specific volume of influenza A NP is 0.7256. The raw experimental data were analyzed by Sedfit (http://www.analyticalultracentrifugation.com/default.htm) and the plots were generated by MATLAB (MathWork, Inc.). The calculated $c(s,fr)$ distribution was shown in two dimensions with grid lines representing the s and fr grids in the thermograph. Below this $c(s,fr)$ surface a contour plot of the distribution was projected into the s-fr plane, where the magnitude of $c(s,fr)$ was indicated by contour lines at constant $c(s,fr)$ in equidistant intervals of c. Contour plots were transformed from the calculated $c(s,fr)$ distribution and were shown as $c(s,M)$ distributions. The dotted lines indicate lines of f_r (frictional ratio). The signal of the $c(s,M)$ distribution is indicated by the color temperature. The insert grayscale bars in the right panels indicate the residuals bitmap of each fit. All samples were visually checked for clarity after ultracentrifugation, and no precipitation was observed.

Differential distribution of sedimentation coefficients and fictional ratios $c(s,fr)$ were calculated with sedfit using $c(s,*)$ model with Eq. **S1** (7)

$$a(r.t) = \iint c(s,fr) \times (s,D(s,fr),r,t)dsdfr. \qquad \textbf{[S1]}$$

The $c(s,fr)$ distribution could be transformed to a molar mass distribution for each s-value, called $c(s,M)$ distribution, by Eq. **S2** (7).

$$a(r,t) = \iint c(s,M) \times (s,D(s,M)r,t)dsdM. \qquad \textbf{[S2]}$$

Circular dichroism analysis. Monitoring the CD spectrum of the protein was monitored at 25 °C in a Jasco J-815 spectropolarimeter under constant N_2 flush and using a 0.1 nm path length to cell analyze the secondary structure of the enzyme. Ten repetitive scans between 250 and 180 nm were averaged. For direct comparison, all enzyme solutions were adjusted to the same protein concentration (1.0 mg mL^{-1}). Mean residue ellipticity (Φ) was obtained by the following equation:

$$\Phi = [\Phi]_{222} M_{\mathrm{MRW}}/10dc$$

in which M_{MRW} of NP is 112.9, the mean amino acid residue weight. d is the cell path in cm, and c is the concentration of NP in mg mL^{-1}.

Virtual screening. All computational work was done by using Accelrys Discovery Studio™, version 2.5, and Pipeline Pilot™, version 7.5 (Accelrys). The molecular structure of nucleoprotein NP from the influenza A virus was obtained from Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB ID code 2IQH); hydrogen atoms were added before refinement. The structures of 1.7 million compounds (the collection of the Genomics Research Center) were generated and energy optimized. The process of screening was conducted in five steps: first, pharmacophore features were developed by Phase™ based on the tail-loop and the pocket of NP structure (1), and partitioned into five subpockets (Fig. S5*A*). Second, using interaction generation protocol of Discovery Studio™, we obtained interaction features, which were clustered by type. A structure-based pharmacophore model, which contained thre hydrogen-bond donor, one hydrogen-bond acceptor, and one hydrophobic, was created by using only the center of each clustered feature (Fig. S5*B*). Third, a modified-Lipinski filter (HBA count≦2; HBD count≦3; molecular weight≦500; AlogP≦5) based on the pharmacophore model was applied to prescreening the 1.7 million for the sake of time-cost. Fourth, the resulted 200,000 compounds were grouped by structure similarity (Tanimoto > 0.8) and the centers of each group (about two thousands) were subjected to energy minimization (CHARMm force field, GBMV implicit solvent model, and a maximum of 2,000 iterations) and conformational generation (a maximum number of 255 of conformation generation) for binding evaluation. Finally, pharmacophore mapping and clustering were used to find compounds that mapped to the model and that showed structural similarity to the experimental hits from the high-throughput screening assays we performed previously (4). There were 1,050 possible ligands that mapped to the model, which fall into five clusters structurally. Representative 24 compounds from the clusters containing experimental hits were selected for further antiviral assay in this work (Fig. S3).

Primer extension assay. MDCK stable cells were infected with the WSN virus (MOI = 2) and then harvested 6 h post infection. Total RNA was extracted by the TRIzol reagent (Invitrogen). RNA samples were mixed with each ^{32}P-labeled primer and denatured at 95 °C for 5 min. The mixture was cooled on ice and then incubated at 45 °C for 10 min, and added with the reverse transcription buffer and enzyme (Toyobo Life Science) to start the reverse transcription reaction. Two neuraminidase (NA) gene-specific primers (8) and one canis 16S ribosomal RNA primer were used: NA_negative: 5′-TGGACTAGTGGGAGCATCAT-3′ (to detect vRNA, 122 nt), NA_positive: 5′-TCCAGTATGGTTTTGATTTCCG-3′ (to detect mRNA, >161 nt, and cRNA, 161 nt) and canis 16S_118-99: 5′-TACTATCTCTATCGCTCCAA-3′ (to detect canis rRNA, 118 nt). The reaction was stopped by addition of 8 μL 90% formamide and heating at 99 °C for 5 min, and analyzed on 6% polyacrylamide gels containing 7 M urea in TBE (Tris-borate-EDTA buffer). Transcription products were detected by autoradiography.

In vitro transcription. Approximately 10^7 plaque forming units (pfu) of the WSN virus were incubated in a buffer containing 100 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$,100 mM KCl, 1 mM DTT, 0.25% Triton N-101 and 0–4 mM peptides or compounds at 25 °C. After 1 h, the samples were incubated with 200 μM ApG, RNase-inhibitor 1 U/μL, 100 μM ATP, 50 μM CTP, 50 μM UTP, 1 μM GTP, 5 μCi [3H] UTP for 30 min at 30 °C. The RNA synthesized was precipitated by 10% trichloroacetic acid (TCA) on ice for 1 h. Glass microfiber filters (GF/C) were rinsed with 10% TCA and placed onto the vacuum filter device. The samples were spotted on the labeled GF/C filters, washed with 10% TCA and dried with 95% ethanol for 1 min, and then were counted in a scintillation counter.

Preparation of isogenic recombinant influenza viruses. The isogenic recombinant influenza viruses were prepared as described in our recent report (4). The recombinant influenza pairs isogenic at the 52nd, 289th or 52nd/289th amino acid of NP was produced.

Antiviral assay. In a 96-well plate, 1×10^4 MDCK cells were seeded per well and incubated for 24 h at 37 °C. The MDCK cells were then incubated with the compound, then inoculated with medium alone or the WSN virus (MOI of 0.001) for 48 h at 35 °C. The number of metabolically viable cells was determined by the MTS assay (Promega) or Cell-Titer Glo® (Promega).

Cytotoxicity assay. In a 96-well plate, 1×10^4 MDCK cells were seeded per well and incubated for 24 h at 37 °C. The MDCK cells were then incubated with compounds for 48 h at 35 °C. The

number of metabolically viable cells was determined by the MTS assay (Promega).

Indirect immunoflourescence staining and confocal microscopy. MDCK cells were grown on coverslips and then were infected with the WSN virus (MOI = 1) and incubated with compounds at 6 h post infection. After 8 h post infection, cells were fixed, and immunostained with anti-NP (mouse), and then with anti-mouse conjugated fluorescein isothiocyanate (FITC) and with 4′,6-diamidino-2-phenylindole (DAPI). Immunofluorescence images were obtained by using a Leica TCS-SP2 laser scanning confocal microscope (Leica Microsystems GmbH).

1. Ye Q, Krug RM, Tao YJ (2006) The mechanism by which influenza A virus nucleoprotein forms oligomers and binds RNA. *Nature* 444:1078–1082.
2. Chan WH, et al. (2010) Functional analysis of the influenza virus H5N1 nucleoprotein tail loop reveals amino acids that are crucial for oligomerization and ribonucleoprotein activities. *J Virol* 84:7337–7345.
3. Elton D, Medcalf L, Bishop K, Harrison D, Digard P (1999) Identification of amino acid residues of influenza virus nucleoprotein essential for RNA binding. *J Virol* 73: 7357–7367.
4. Su CY, et al. (2010) High-throughput identification of compounds targeting influenza RNA-dependent RNA polymerase activity. *Proc Natl Acad Sci USA* 107: 19151–19156.
5. Fodor E, et al. (2002) A single amino acid mutation in the PA subunit of the influenza virus RNA polymerase inhibits endonucleolytic cleavage of capped RNAs. *J Virol* 76: 8989–9001.
6. Neumann G, Zobel A, Hobom G (1994) RNA polymerase I-mediated expression of influenza viral RNA molecules. *Virology* 202:477–479.
7. Brown PH, Schuck P (2006) Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. *Biophys J* 90:4651–4661.
8. Vreede FT, Brownlee GG (2007) Influenza virion-derived viral ribonucleoproteins synthesize both mRNA and cRNA in vitro. *J Virol* 81:2196–2204.

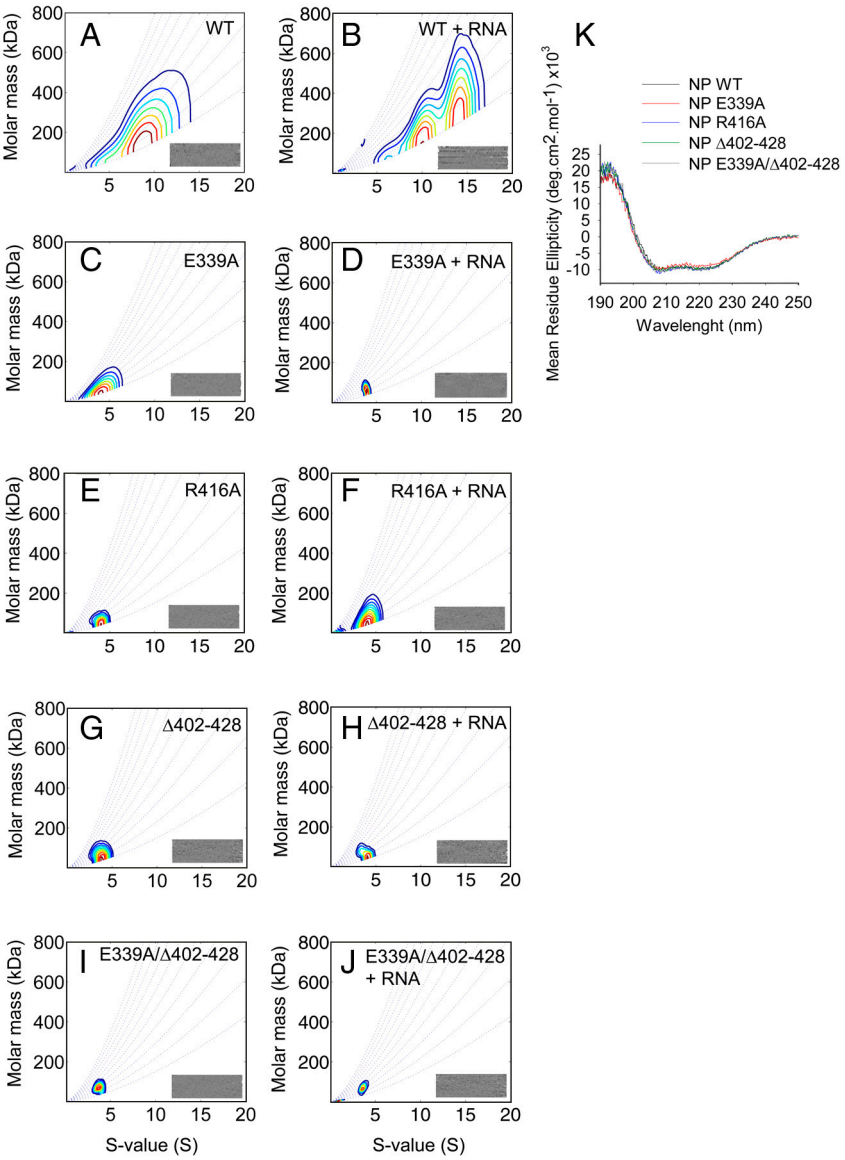

Fig. S1. Analyses of the structural properties of NP proteins by AUC (*A–J*) and CD (*K*). (*A–J*) The samples in the *Left* panels are the free form and in the *Right* panels are in the presence of one equivalent of synthetic 24-mer RNA (5′-UUUGUUACACACACACACGCUGUG-3′). Contour plots of $c(s,M)$ were transformed from the calculated $c(s,fr)$ distribution (see *SI Materials and Methods*). The dotted lines indicate lines of f_r (fictional ratio). The signal of the $c(s,M)$ distribution is indicated by the color temperature. The grayscale bar indicates the residuals bitmap of each fit. The concentrations of NP protein and RNA were both 3 μM. (*K*) The CD spectra of NP WT and mutants, recorded at 25 °C by a Jasco J-815 spectropolarimeter under constant N_2 flush and using a cell of 0.01 cm path length. All protein solutions were 2 mg/mL. The results of CD ensure that the secondary structures of the NP mutants have not been perturbed relative to WT NP.

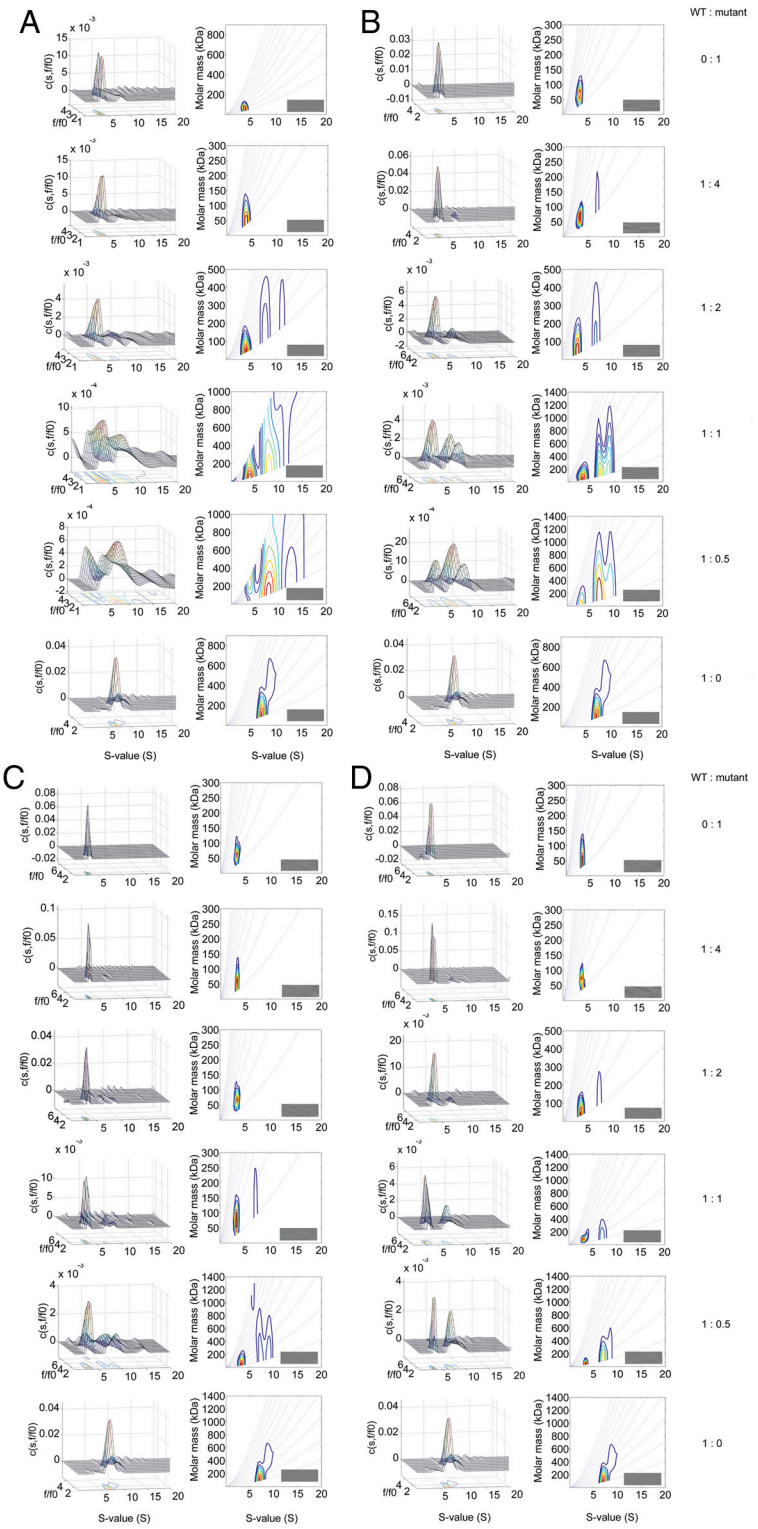

Fig. S2. AUC analyses of the WT-mutant interactions by titrating the WT NP (trimer) with mutant NP (monomer). The *Left* panels are $c(s,fr)$ distribution plots and the *Right* panels are $c(s,M)$ distributions. Detailed information of AUC data calculation and plot generation are described in *SI Materials and Methods*. The concentration of NP WT protein was 3 μM. The NP WT protein was mixed with (*A*) the E339A mutant; (*B*) the R416A mutant; (*C*) the Δ402–428 deletion mutant; and (*D*) the E339A/Δ402–428 double mutant, at molar ratios of 1:0, 1:0.5, 1:1, 1:2, 1:4, and 0:1.

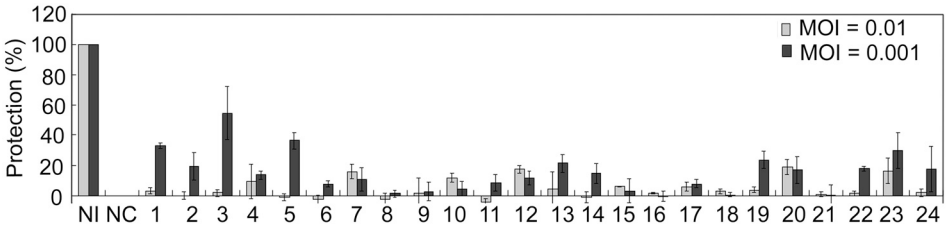

Fig. S3. Antiviral assays of the 24 compounds selected from virtual screening. NI; no infection. NC; no compound. In a 96-well plate, 1×10^4 MDCK cells were seeded per well and incubated for 24 h at 37 °C. The MDCK cells were then incubated with the compounds (10 μM), then inoculated with medium alone or the WSN virus (MOI = 0.01 or 0.001) for 48 h at 35 °C. The number of metabolically viable cells was determined by the MTS [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium] assay.

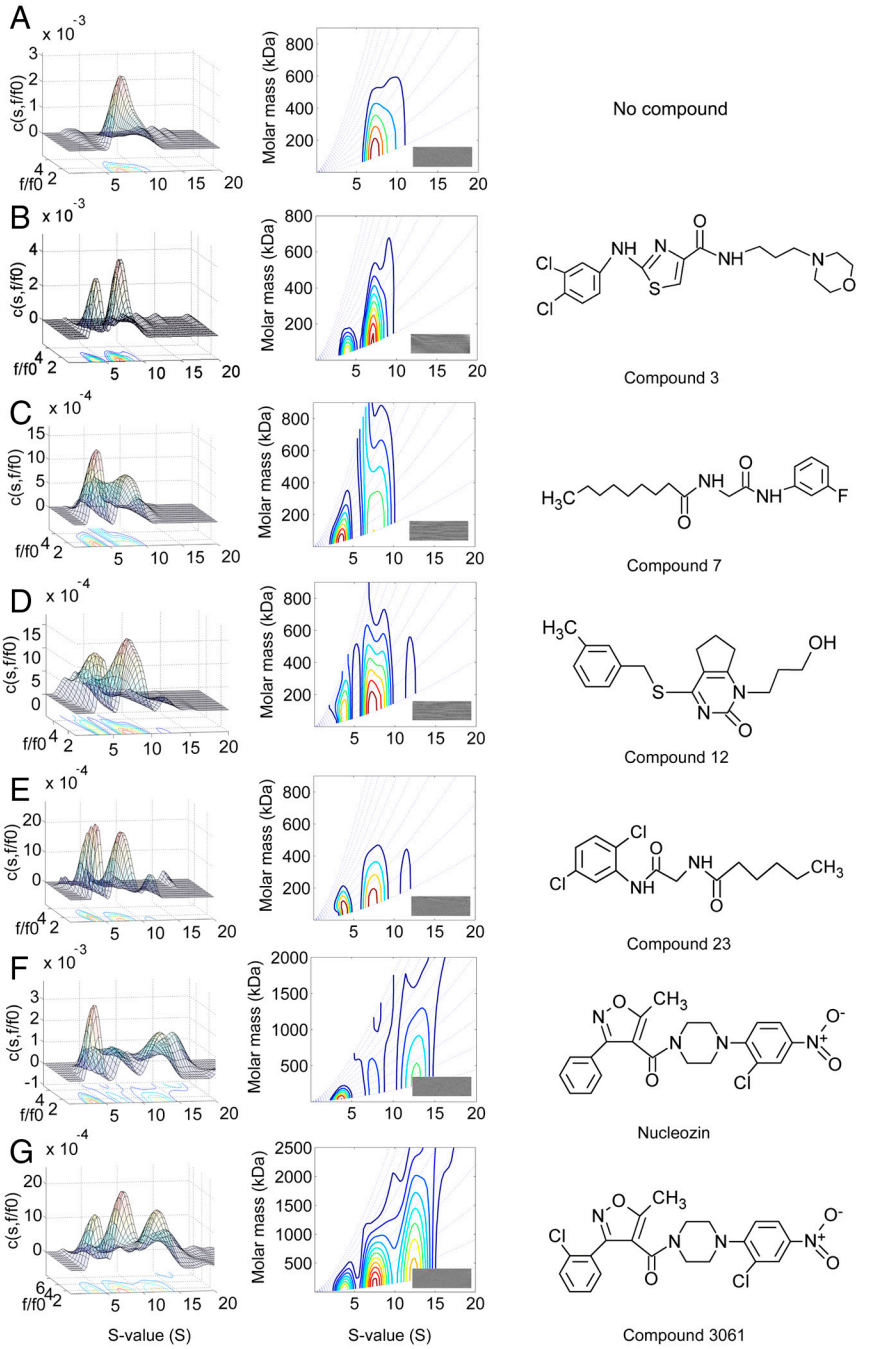

Fig. S4. The effects of compounds for NP WT trimer by AUC. Each sample contained a mixture of 3 μM WT NP with (*A*) none; (*B*) 7.5 μM compound 3; (*C*) 7.5 μM compound 7; (*D*) 7.5 μM compound 12; (*E*) 7.5 μM compound 23; (*F*) 1.5 μM nucleozin; and (*G*) 1.5 μM compound 3061. The *Left* panels are $c(s,fr)$ distribution plots and the *Right* panels are $c(s,M)$ distributions.

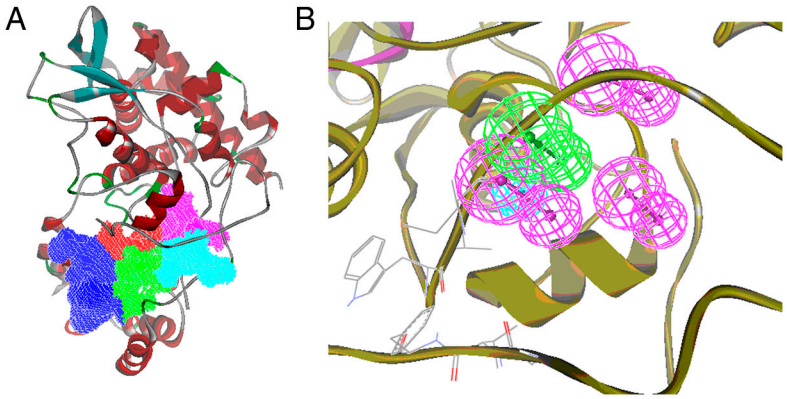

Fig. S5. Illustration of the virtual screening. (*A*) The tail-loop binding pocket of NP was partitioned into five subpockets, which were indicated in different colors. (*B*) The structure-based pharmacophore model based on the interaction of the tail-loop and the subpocket. Donors were colored in pink, acceptor in green, and hydrophobic in cyan.