Discovery of Potent Sialic Acid Inhibitors Against the Hemagglutinin of Influenza A Virus

Yu-Jen Chang
The Ph.D. Program of Biotechnology and Biomedical industry, China Medical University, Taichung
Cheng-Yun Yeh
The Ph.D. Program of Biotechnology and Biomedical industry, China Medical University, Taichung
Ju-Chien Cheng
Department of Medical Laboratory Science and Biotechnology, China Medical University, Taichung
Yu-Qi Huang
Department of Medical Laboratory Science and Biotechnology, China Medical University, Taichung
Kai-Cheng Hsu
Graduate Institute of Cancer Biology and Drug Discovery, Taipei Medical University, Taipei
Yu-Feng Lin
Department of Medical Laboratory Science and Biotechnology, Asia University, Taichung
Chih-Hao Lu (chlu@mail.cmu.edu.tw)
The Ph.D. Program of Biotechnology and Biomedical industry, China Medical University, Taichung

Research Article

Keywords: Influenza A virus, Virtual screening, Hemagglutinin, Sialic acid inhibitors

Posted Date: December 18th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-126042/v1

Version of Record: A version of this preprint was published at Scientific Reports on April 21st, 2021. See the published version at https://doi.org/10.1038/s41598-021-87845-0.

Discovery of potent sialic acid inhibitors against the hemagglutinin of influenza A virus

Yu-Jen Chang[1,*], Cheng-Yun Yeh[1,*], Ju-Chien Cheng[2], Yu-Qi Huang[2], Kai-Cheng Hsu[3], Yu-Feng Lin[4] & Chih-Hao Lu[1,2,5]

[1]The Ph.D. Program of Biotechnology and Biomedical industry, China Medical University, Taichung, Taiwan

[2]Department of Medical Laboratory Science and Biotechnology, China Medical University, Taichung, Taiwan

[3]Graduate Institute of Cancer Biology and Drug Discovery, Taipei Medical University, Taipei, Taiwan

[4]Department of Medical Laboratory Science and Biotechnology, Asia University, Taichung, Taiwan

[5]Graduate Institute of Biomedical Sciences, China Medical University, Taichung, Taiwan

***These authors contributed equally to this work.**

Correspondence

Chih-Hao Lu, The Ph.D. Program of Biotechnology and Biomedical industry, China Medical University, No.91, Hsueh-Shih Road, Taichung, 40402, Taiwan

Email: chlu@mail.cmu.edu.tw

Abstract

Influenza A virus (IAV) is hard to eradicate due to its genetic drift and reassortment ability. Neuraminidase (NA), frequently the target protein, cleaves the sialic acid (SA) and discharges virions to complete the infectious cycle. However, the increasing use of NA inhibitors aroused drug resistance in recent years. Hemagglutinin (HA), on the opposite, initiates the infectious cycle and sticks virions to cells by connecting to the host SA so that HA might be a tempting target. In this study, HA was chosen for SA-binding site model preparation to screen more than 200,000 compounds by molecular docking method *in silico*. According to the post-screening analysis by iGemdock and SiMMap, nine of the top twenty compounds based on the ranking score were purchased and evaluated. NSC85561 was initially identified from the compounds through the bioassays. Next, the twelve derivatives of NSC85561 were selected to consolidate the primary results. NSC85561 and two of its derivatives were finally identified as potent HA inhibitors by cell proliferation, plaque reduction, and hemagglutination inhibition assays *in vitro*. These results suggested that virtual screening may be a powerful tool to concise the compounds from the massive database and reduce the complicated bioassays.

Keywords: *Influenza A virus, Virtual screening, Hemagglutinin, Sialic acid inhibitors*

Introduction

Influenza viruses have posed a severe threat to human health, causing up to 650,000 deaths and economic loss annually. Since the 2009 H1N1 flu pandemics, the 2018 flu season has recently been the worst catastrophe. Among the most deaths are the people age 65 or older in industrialized countries[1]. Another study shows that children younger than five years attribute to influenza deaths, with 99% occurring in developing countries[2]. Over the years, seasonal vaccines are predicted to protect humans from infection but frequently failed because of influenza's genetic drift and reassortment ability. Yearly vaccination is the primary means of preventing and controlling influenza[3], but it is inadequate to apply every scenario. There are three FDA-approved influenza antiviral drugs recommended by CDC for use against recently circulating influenza viruses. Rapivab (peramivir), Relenza (zanamivir), and Tamiflu (oseltamivir), which target neuraminidase (NA) as inhibitors, interfere with the release of progeny influenza virus from infected host cells. This process prevents the infection of new host cells and halts infection spread in the respiratory tract[4]. However, those NA inhibitors, which seem to be powerful agents to treat the influenza virus, recently arose some sporadic drug-resistance events. In Japan, oseltamivir-resistant H1N1 viruses were isolated from 7 of 43 patients (16%)[5]. About 27% of patients infected with H1N1 viruses showed resistance to oseltamivir in the UK[6]. Owing to the urgent situation that new targets must be explored, so we switch our direction to the hemagglutinin (HA).

HA plays a pivotal role during the Influenza A virus (IAV) life cycle and possesses a unique structure. HA takes the role in the attachment and penetration to the host cells. Nevertheless, there are no HA inhibitor drugs offered to clinical patients. The main reason is that the variety of HA maybe even is strengthened by antigenic drift and shift[7,8]. A report showed that HA globular head glycosylation might be an essential step for viruses to gain virulence and antigenic properties[9]. HA's specific structure is a homotrimer whose globular domain contains a receptor-binding site (RBS) specific for sialic acid (SA) expressed by the host cell glycoproteins and glycolipids, which terminates host glycans[10]. When the RBS of influenza binds to the host SA, the viruses can quickly enter the host cells. To prevent IAV from obtaining entry to the host cells, the RBS might be a potential target.

The homotrimer is constructed by the un-cleaved HA0 subunits, which are unable to fuse the host membranes. HA0 inhibitors interrupt the correct folding of HAs in progeny influenza viruses; therefore, they elicit nonfunctional HA conformation to block viral entry[11]. One example is AF4H1K1, which could block the immature HA0 cleavage and inhibit the infectivity of IAV[12]. The HA1 folds into a jelly-roll motif of eight stranded antiparallel β-sheets and a shallow pocket at the distal tip, acting as an RBS surrounded by antigenic sites[13]. The mechanism of HA1 inhibitors mainly lies in blocking receptor binding so that it prevents virus infective initiation. The scientists found that a specific plant from Borneo was traditionally to treat symptoms of influenza infection. They extracted the plant and discovered that the useful substances were not limited to SA-like compounds but non-SA-like

components that might act against other viral proteins apart from HA and NA[14]. Besides, the synthetic SA-mimic peptides showed a high affinity to the RBS to block the initial infection in another study[15]. The HA2 mediates the host endosomal membrane's fusion with the viral membrane, allowing viral ribonucleoprotein entry into the host cell[16]. In the process of HA-mediated fusion, metastable HA2 subunits undergo irreversible rearrangement at acidic pH, which causes membrane fusion and the completion of viral entry[11]. One HA2 inhibitor is Arbidol (ARB, Umifenovir), which interacts with HA2 directly, and can prevent acidification-induced conformation change of HA2 and impeding fusion in endosomes[17].

From the overview of the IAV structure, our study focused on the RBS of the HA1 subunit. The compounds were from the NCI database to identify potential inhibitors. Over 200,000 compounds were screened by the molecular docking method. Next, the docked compounds were used to recognize interaction preference by analyzing the RBS with interacting residues and specific physical-chemical properties. With these models, we could identify inhibitors with novel scaffolds. Some HA inhibitors were identified and validated by a series of bioassays, including cell proliferation, plaque reduction, and hemagglutination inhibition. To refine our analysis and find more compounds targeting HA1, we selected the initial compound's derivatives to enter the same experimental tests. Some attractive and tempting outcomes were finally unveiled. This study provided the comprehensive framework of efficient screening of HA1 inhibitor leading compounds for further drug design and development.

Results

A Framework of anchor Construction and Post-Screening Analysis

Figure 1 gives an overview of anchors formation and post-screening analysis for the identification of novel HA inhibitors. First, the structure of the receptor-binding site (RBS) of the HA1 subunit (PDB ID: 1RUY)[18] was selected as the target protein, and 208,023 compounds from NCI were collected as the screening compound database (Fig. 1a). The molecular docking and post-screening analysis were performed by iGEMDOCK[19] (Fig. 1b). Next, SiMMap[20] anchors were constructed based on the docking score and the interaction types between top-ranked 1,000 compounds and the RBS (Fig. 1c). An anchor is composed of a binding pocket with corresponding interacting residues, moiety preference, and type of the interaction (E: electrostatic, H: hydrogen-bonding, or V: van der Waal forces). These 1,000 compounds were re-ranked based on the SiMMap score, and the top twenty compounds were selected as potential candidates (shown in Supplementary Table S1 online). Finally, nine compounds were purchased and evaluated by cell proliferation, plaque reduction, and hemagglutination inhibition assays *in vitro* (Fig. 1d).

Anchors of the Receptor-Binding Site

Three pivotal anchors, including E1, H1, and V1, (Fig. 2b) were generated by SiMMap in the RBS. The E1 anchor has the preferred interaction with sulfonic acid and carboxylic acid moieties, interacting with one basic residue, R224 lies in the 220-loop. Consequently, the compounds which have a nucleophilic tendency have the affinity to bind in the E1 anchor. The H1 anchor interacts with one acidic residue D90, a small residue G92, and one nucleophilic residue T93 located in the 90-loop. The H1 prefers to bind to amid, ketone, amine, and hydroxyl groups (the preferences decrease in this order). Last, the V1 anchor has a high tendency to bind to an aromatic and heterocyclic moiety, interacts with one amide residue N91, and one basic R224, the same as the E1 anchor. Both E1 and V1 have R224 because the weak non-bond force between guanidinium in R224 and aromatic moiety usually form planar stacking in nature, which will not interfere with the electrostatic force between R224 and acid moieties[21]. The corresponding 3D model is shown in figure 2a.

Derivative Compounds from NSC85561

Based on the SiMMap score, the nine compounds were purchased and performed the bioassays (Table S1). The MTS method was used to examine the safety of the compounds. The 100 μM compounds were taken to measure the individual OD values and insight into the MDCK cells' toxicity (Fig. 3a). NSC97307 was excluded due to its relative cell viability was lower than 0.75. Then, the 25 μM compounds were taken to the plaque reduction assay. The values of PFU (% control) treated with NSC85561 and NSC84472 were dramatically decreasing (5±5 and 40±20), so NSC85561 and NSC84472 were unveiled (Fig. 3b and Table 1). Next, the HA assay was used to validate the efficacy of NSC85561 and NSC84472. The compounds started with 100 μM by two-fold serial dilution in the HA experiments. NSC85561 formed precipitation at 12.5 μM, but NSC84472 was at 100 μM (Fig. 3c). The efficacy of NSC85561 was much higher compared with NSC84472, so we took NSC85561 as the template for generating the twelve derivative compounds according to the atom-pair fingerprint among our compound dataset. The SMILES code of these twelve derivative compounds and NSC85561 are shown in table S2.

The post-screening analysis meets the bioassay outcome

The post-screening analysis from the iGEMDOCK showed a similar outcome versus the HA assay. Figure 4A presents the hierarchical clustering results of thirteen compounds according to their interaction profile. From NSC85561 down to NSC73413, similar patterns were found, and earlier clustered (Fig. 4a). In the HA experiments, compounds started with 25 μM by two-fold serial dilution. NSC4203 formed precipitation at 25 μM, whereas NSC85561 and NSC73413 at 12.5 μM. Also, NSC47715 and NSC7223 at 6.25 μM stood out from the thirteen compounds (Fig. 4b). We picked out these five compounds based on the outcome of the interaction profile and HA assay. Next, to validate

these five compounds' efficacy, we used virus plaque reduction assay in different concentrations. The MDCK cells were seeded in 6-well plates for 24 hours, mixed with or without the five compounds at 25 μM or 12.5 μM (Fig. 5a-5b). When our five compounds were at 12.5 μM, the efficacy of NSC47715, NSC85561, and NSC7223 was much higher than NSC4203 and NSC73413. Figure 5c shows that the bar plot of the relative cell viability of NSC85561, NSC47715, and NSC7223 with at least three individual experiments. The average of the values are NSC85561: 0.88, NSC47715: 0.83, and NSC7223: 0.83 at 100 μM. CC50 was used to evaluate the safety of the three compounds. The CC50 lies on NSC85561: 700 μM, NSC7223: 900 μM, and NSC47715: > 1000 μM. EC50 was taken to measure the efficacy of the three compounds. The MDCK cells infected with the virus (100 PFU) were mixed with NSC47715, NSC85561, and NSC7223, and virus plaque reduction assay was performed (fig. 5d). NSC47715 showed the lowest values, which means that NSC47715 could inhibit IAV at the lowest concentration among the three compounds. SI value, the ratio of the CC50 and EC50, was used to evaluate safety and toxicity simultaneously. The summary table is showed in table 2.

Discussion

In our study, HA was the target to block IAV infection, which was different from the recent research focusing on neuraminidase (NA)[22-24]. However, the NA-drug resistance events happened continuously, so NA might not be a reliable target to quench IAV. From a 2019 research, the authors emphasized the vulnerable HA stalk trimer domain and depicted the conserved residues over various IAV strains[25]. Besides, they found that FluA-20 (human antibody) tends to bind between the 220-loop and the 90-loop. Our models and predictions were consistent with their findings, so the anchors might play an essential role in interfering with IAV's fusion.

The interaction profile based on the post-screening analysis presents some intriguing outcomes (fig. 4a). Although six compounds from NSC85561 down to NSC73413 showed a similar pattern, NSC87862 did not demonstrate the HA assay's corresponding efficacy (fig. 4b). NSC87862 shows a lower binding affinity toward the side chain of N91 's hydrogen-bonding force and Y102's van der Waal force compared with the other five compounds. Therefore, the 2D interaction diagrams of six compounds with the pocket environment were examined by the PROTEINS PLUS server [26](fig. 6a-6c and S. 1a-1c). All compounds have two naphthalene rings except NSC87862, which might be the critical reason for weak binding affinity. Also, NSC4203 and NSC73413 possess weak efficacy toward IAVs that might lie in their orientation of one naphthalene between residue T93 and G92 (fig. 6d-6f and S. 1d-1f). The interface view of RBS with docked NSC47715, NSC85561, and NSC7223 was shown in Figure 6g-6i. The most significant difference is the additional nitrite in NSC85561 and

NSC7223 compared with NSC47715, and it might interfere with their binding to the HA. Consequently, virtual screening and bio-assay methods complement each other to see more clearly and more objective about identifying novel inhibitors.

Conclusions

Our research sheds light on the utility and feasibility of virtual screening with the outcome of SiMMap for identifying lead compounds. NSC85561 showed the highest affinity to the RBS initially. The same experimental assays were conducted on derivatives of NSC85561, and three compounds, NSC85561, NSC7223, and NSC47715 in total, were finally identified as the potential inhibitors of HA. These strategies reveal that SiMMap enables us to search for desired functional groups and analyze the connection between the RBS and the compounds. To sum up, we considered that the virtual screening methods possessed the potential to eliminate the threat of IAV shortly.

Methods and Materials

Dataset Preparation and Virtual Screening

The structure of H1N1 (PDB ID: 1RUY) was selected as the target protein for virtual screening from Protein Data Bank[27]. It belongs to the H1 class from the 1930 swine influenza strain. The method to determine the structure of 1RUY is through X-ray diffraction with a resolution of 2.7Å. The compounds employed for virtual screening were assembled from the NCI database[28]. The 208,023 compounds were finally filtered out according to Lipinski's rules of five. The iGEMDOCK software[19], a widely used and accurately predicted molecular docking method, was used to dock the compounds into the RBS of the HA1 subunit. Besides, the iGEMDOCK consists of three types of interactions: electrostatic, hydrogen-bonding, or van der Waals forces. Based on the interactions (docking energy), which consist of a simple empirical scoring function and a pharmacophore-based scoring function, the top 1,000 compounds were selected.

SiMMap Construction and Hemagglutinin Inhibitors Identification

The top 1,000 compounds ranked by docking energy with their binding poses were submitted to the SiMMap server to reevaluate[20]. The SiMMap was used to analyze a protein binding site to identify anchors (antigenic sites), often critical binding environments[24]. The moiety compositions between these compounds and the SiMMap anchors provide clues for lead optimization. Based on the protein-compound interaction profiles generated by the SiMMap, consensus interactions between compound moieties and binding site composed of consensus interacting residues were identified. Next, a site-moiety map score for each chosen compound was generated. The identified compounds were reordered

with the scores. Finally, the promising HA1 inhibitors were picked out for further rigorous examination based on their ranking, feasibility, physicochemical property, and interaction type.

Cells and viruses

Madin-Darby Canine Kidney MDCK cells (ATCC accession no. NBL-2) for virus infection were maintained in Dulbecco's modified eagle medium (DMEM; Gibco) containing 10% fetal bovine serum (FBS; GIBCO) and 1% penicillin and streptomycin each (PS; GIBCO) at 37 °C under 5% CO_2 incubator. The influenza virus H1N1 (A/WSN/33), which were supplied from Taiwan CDC (center for disease control), were cultured in a flu-medium (DMEM with 1% PS and 2 μg/ml L-1-tosylamide-2-phenyl chloromethyl ketone (TPCK, Sigma-Aldrich)-Trypsin from bovine pancreas (Sigma-Aldrich; St. Louis; MO)) for growth. All viruses were under -80 for storage. Further, all the experiments related to viruses were conducted in the BSL2 laboratory.

Cell Proliferation Assay (MTS)

To test the cells' corresponding viability and cytotoxic compounds' analysis, we conducted a cell proliferation assay (MTS)[29], a colorimetric method for sensitive quantification. Based on the reduction of MTS tetrazolium compound by viable cells to generate a colored formazan product, the formazan dye produced by viable cells can be quantified by measuring the absorbance at 490 nm. MDCK cells (1 x 10^4 cells/well) are seeded into 96- well microtiter plates overnight. Next, the MDCK cells were mixed with our potential compounds in flu-medium for 72 hours at 37°C with 5% CO_2. The cells were washed with phosphate buffer saline (PBS; Na2HPO4: 8mM, NaCl: 137 mM, KCl: 2.68mM, KH2PO4: 1.47 mM adjusted to pH7.2) twice after removing the previous medium in the wells. 10μL MTS solutions and 90μL DMEM were added to each well, and the cells were incubated for an additional half an hour. Finally, absorbance at 490 nm was recorded using a microplate reader (SpectraMax iD3, Molecular Devices, USA). A blank control (flu-medium only) and cell control (without compounds defined as 100% cell survival) were included in every assay plate. The mean optical density (OD, absorbance) of the three wells in the indicated groups was used to calculate the percentage of cell viability as follows: Percentage of cell viability = $(A_{treatment} - A_{blank})/(A_{control} - A_{blank}) \times 100\%$ (A: absorbance)[30]. Furthermore, the cytotoxic concentration was required for the identified compounds to reduce cell viability by 50% (CC_{50}) was determined.

Plaque Assay

To determine viral titer as plaque-forming units per ml (PFU/ml), we selected plaque assay[31,32], which is a standard method to quantify the virus. In brief, MDCK cells (1×10^6 cells/well) were seeded in 6-well plates for 24 hours at 37°C with 5 % CO_2. The single layer of cells was infected with the indicated dose of IAV, followed by washing with PBS three times. Next, the cells were covered with 0.3%

agarose in the flu-medium for an additional two days at 37°C. Cells were fixed with 10% formaldehyde and then stained with 1% crystal violet. Finally, the PFU was calculated.

Plaque Reduction Assay

To quantify the antiviral activity of the compounds versus IAV, we conducted the virus plaque reduction assay. As described above, the previous seeded monolayer cells (100 PFU/well) were incubated with or without the selected compounds. The virus suspension was removed after the IAV adsorbed the compounds for an hour. The cells were then immediately washed trice with PBS and overlaid with 0.3% agarose in the flu-medium again present or absent the compounds for additional 48 hours. The 50% effective concentration (EC50) of the compounds was determined by comparing the number of plaques in the virus-infected control. Furthermore, the EC50 was defined as half maximal inhibitory concentration, and then the selectivity index (SI) was determined by the ratio: CC50/EC50.

Hemagglutination Inhibition Assay

Our experiment observed an erythrocyte agglutination test whether sialic acid was bound to a compound and whether the filtered compounds inhibited hemagglutination from causing a precipitation reaction in the sialic acid-binding and membrane fusion of HA, respectively. If there is agglutination, we can measure the lowest concentration of the virus. First, we diluted our compounds with PBS to one-thousandth. MDCK cells were also diluted in a serial twofold to gain the titer of our IAV. To reduce the interference of the antibodies in erythrocytes, we prepared human type-O serum. Next, the sera were diluted in a serial twofold in a 96-well microtiter plate.

Data availability

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials.

References

1 Thompson, W. W. e. a. in *Influenza Other Respir. Viruses* Vol. 3 37-49 (2009).

2 Nair, H. e. a. in *Lancet* Vol. 378 1917–1930 (2011).

3 *Influenza (Flu) Antiviral Drugs and Related Information*, <https://www.fda.gov/drugs/information-drug-class/influenza-flu-antiviral-drugs-and-related-information> (2018).

4 Moscona, A. Neuraminidase Inhibitors for Influenza. *New England Journal of Medicine* **353**, 1363-1373, doi:10.1056/NEJMra050740 (2005).

5 Ward, P., Small, I., Smith, J., Suter, P. & Dutkowski, R. Oseltamivir (Tamiflu (R)) and its potential for use in the event of an influenza pandemic. *J Antimicrob Chemoth* **55**, 5-21, doi:10.1093/jac/dki018 (2005).

6 Stephenson, I. *et al.* Neuraminidase Inhibitor Resistance after Oseltamivir Treatment of Acute Influenza A and B in Children. *Clinical Infectious Diseases* **48**, 389-396, doi:10.1086/596311 (2009).

7 Stevens, J. *et al.* Structure and Receptor Specificity of the Hemagglutinin from an H5N1 Influenza Virus. *Science* **312**, 404, doi:10.1126/science.1124513 (2006).
8 Caton, A. J., Brownlee, G. G., Yewdell, J. W. & Gerhard, W. The antigenic structure of the influenza virus A/PR/8/34 hemagglutinin (H1 subtype). *Cell* **31**, 417-427, doi:10.1016/0092-8674(82)90135-0 (1982).
9 Medina, R. A. *et al.* Glycosylations in the globular head of the hemagglutinin protein modulate the virulence and antigenic properties of the H1N1 influenza viruses. *Sci Transl Med* **5**, 187ra170, doi:10.1126/scitranslmed.3005996 (2013).
10 Kosik, I. & Yewdell, J. W. Influenza Hemagglutinin and Neuraminidase: Yin–Yang Proteins Coevolving to Thwart Immunity. *Viruses* **11**, 346, doi:10.3390/v11040346 (2019).
11 Shen, X., Zhang, X. & Liu, S. Novel hemagglutinin-based influenza virus inhibitors. *Journal of thoracic disease* **5 Suppl 2**, S149-S159, doi:10.3978/j.issn.2072-1439.2013.06.14 (2013).
12 Xiao, H. *et al.* Light chain modulates heavy chain conformation to change protection profile of monoclonal antibodies against influenza A viruses. *Cell Discovery* **5**, 21, doi:10.1038/s41421-019-0086-x (2019).
13 Isin, B., Doruker, P. & Bahar, I. Functional Motions of Influenza Virus Hemagglutinin: A Structure-Based Analytical Approach. *Biophysical Journal* **82**, 569-581, doi:https://doi.org/10.1016/S0006-3495(02)75422-2 (2002).
14 Rajasekaran, D. *et al.* Identification of Traditional Medicinal Plant Extracts with Novel Anti-Influenza Activity. *Plos One* **8**, doi:ARTN e7929310.1371/journal.pone.0079293 (2013).
15 Matsubara, T. *et al.* Sialic Acid-Mimic Peptides As Hemagglutinin Inhibitors for Anti-Influenza Therapy. *Journal of Medicinal Chemistry* **53**, 4441-4449, doi:10.1021/jm1002183 (2010).
16 Sriwilaijaroen, N. & Suzuki, Y. Molecular basis of the structure and function of H1 hemagglutinin of influenza virus. *Proceedings of the Japan Academy. Series B, Physical and biological sciences* **88**, 226-249, doi:10.2183/pjab.88.226 (2012).
17 Boriskin, Y. S., Leneva, I. A., Pecheur, E. I. & Polyak, S. J. Arbidol: a broad-spectrum antiviral compound that blocks viral fusion. *Curr Med Chem* **15**, 997-1005 (2008).
18 Gamblin, S. J. *et al.* The Structure and Receptor Binding Properties of the 1918 Influenza Hemagglutinin. *Science* **303**, 1838, doi:10.1126/science.1093155 (2004).
19 Yang, J.-M. & Chen, C.-C. GEMDOCK: A generic evolutionary method for molecular docking. *Proteins: Structure, Function, and Bioinformatics* **55**, 288-304, doi:10.1002/prot.20035 (2004).
20 Chen, Y. F. *et al.* SiMMap: a web server for inferring site-moiety map to recognize interaction preferences between protein pockets and compound moieties. *Nucleic Acids Res* **38**, W424-430, doi:10.1093/nar/gkq480 (2010).
21 Flocco, M. M. & Mowbray, S. L. Planar Stacking Interactions of Arginine and Aromatic Side-Chains in Proteins. *Journal of Molecular Biology* **235**, 709-717, doi:https://doi.org/10.1006/jmbi.1994.1022 (1994).
22 Zhang, L. *et al.* Virtual screening approach to identifying influenza virus neuraminidase inhibitors using molecular docking combined with machine-learning-based scoring function. *Oncotarget* **8**, 83142-83154, doi:10.18632/oncotarget.20915 (2017).
23 Abdalsalam, A. A. A. *In-silico Virtual Screening and ADMET Study to Find Novel Neuraminidase N1 Inhibitors Extended to the 150-Cavity*. (0001).
24 Hsu, K.-C. *et al.* Identification of neuraminidase inhibitors against dual H274Y/I222R mutant strains. *Scientific Reports* **7**, 12336, doi:10.1038/s41598-017-12101-3 (2017).
25 Bangaru, S. *et al.* A Site of Vulnerability on the Influenza Virus Hemagglutinin Head Domain Trimer Interface. *Cell* **177**, 1136-1152 e1118, doi:10.1016/j.cell.2019.04.011 (2019).

26 Stierand, K., Maass, P. C. & Rarey, M. Molecular complexes at a glance: automated generation of two-dimensional complex diagrams. *Bioinformatics* **22**, 1710-1716, doi:10.1093/bioinformatics/btl150 (2006).
27 ww, P. D. B. c. Protein Data Bank: the single global archive for 3D macromolecular structure data. *Nucleic Acids Research* **47**, D520-D528, doi:10.1093/nar/gky949 (2019).
28 Milne, G. W. A., Feldman, A., Miller, J. A. & Daly, G. P. The NCI Drug Information System. 3. The DIS chemistry module. *Journal of Chemical Information and Computer Sciences* **26**, 168-179, doi:10.1021/ci00052a004 (1986).
29 Barltrop, J. A., Owen, T. C., Cory, A. H. & Cory, J. G. 5-(3-carboxymethoxyphenyl)-2-(4,5-dimethylthiazolyl)-3-(4-sulfophenyl)tetrazolium, inner salt (MTS) and related analogs of 3-(4,5-dimethylthiazolyl)-2,5-diphenyltetrazolium bromide (MTT) reducing to purple water-soluble formazans As cell-viability indicators. *Bioorganic & Medicinal Chemistry Letters* **1**, 611-614, doi:https://doi.org/10.1016/S0960-894X(01)81162-8 (1991).
30 Park, C.-M. & Xian, M. in *Methods in Enzymology* Vol. 554 (eds Enrique Cadenas & Lester Packer) 127-142 (Academic Press, 2015).
31 Cooper, P. in *Advances in virus research* Vol. 8 319-378 (Elsevier, 1962).
32 Dulbecco, R. & Vogt, M. in *Cold Spring Harbor symposia on quantitative biology.* 273-279 (Cold Spring Harbor Laboratory Press).
33 Schrödinger, L. The PyMOL Molecular Graphics System, Version 1.8. (2015).

Acknowledgements
This work is supported by China Medical University, Taiwan, Grant Number: CMU106-S-49 and CMU108-S-34.

Author contributions statement

Y.F. Lin and K.C. Hsu performed the virtual screening. C.Y. Yeh and Y.Q. Huang performed the bioassays. Y.J. Chang analyzed the data, prepared figured, and wrote the draft. C.H. Lu, and J.C. Cheng supervised and designed this study. All authors reviewed the manuscript.

Competing interests

The author(s) declare no competing interests.

Tables

Table 1 Plaque reduction assay of the nine purchased compounds. Experiments were performed in triplicate. The value of PFU is presented as mean ± stand deviation.

Compounds	PFU (%control)
NSC85561	5 ± 5
NSC84472	40 ± 20
NSC4299	106 ± 10
NSC47717	72 ± 9
NSC97307	79 ± 7
NSC621095	74 ± 8
NSC45583	96 ± 8
NSC45208	104 ± 14
NSC134131	113 ± 10

Table 2 The EC50, CC50, and SI values of NSC47715, NSC85561, and NSC7223. The value of EC50 is presented as mean ± stand deviation.

Compounds	CC50(μM)	EC50(μM)	SI
NSC47715	>1000	0.29 ± 0.02	>3448
NSC85561	700	2.42 ± 0.11	289
NSC7223	900	2.50 ± 0.76	360

Figure legends

Figure 1 Framework of the study.

(a) Virtual screening for HA inhibitors by docking compounds from NCI and hemagglutinin of H1N1 (PDB ID: 1RUY) based on the iGEMDOCK program. **(b)** Post-screening analysis by generating consensus interaction profile between docked compounds and HA residues. **(c)** Anchors construction by using SiMMap for detecting the most favorable binding pockets and moieties. **(d)** Bioassay validation of potential compounds through three bioassays. MTS assay: to evaluate the toxicity of the compounds. Plaque reduction assay: to quantify the antiviral activity of the compounds versus IAV. Hemagglutination inhibition assay: to measure influenza-specific levels in blood serum.

Figure 2 The anchors of the receptor-binding site from SiMMap.

(a) The structure of RBS of H1N1 HA (chain H of 1RUY) is shown in cartoon, and three anchors are shown in transparent spheres. Red E1 stands for electrostatic force, green H1 for hydrogen bond force,

and gray V1 for van der Waal forces. The corresponding binding pockets (residues) are shown in the cyan sticks. **(b)** The table of binding pockets and moieties for each anchor. The moiety of anchor represents the functional group preference in the top-ranked compounds.

Figure 3 NSC85561 unveiled from the initial compounds.

(a) MDCK cells were treated with the initial nine compounds at 100 μM and cell viability was measured using the MTS assay. **(b)** MDCK cells were infected with A/WSN/1933 (H1N1, 100 PFU) and the compounds at 25 μM. The virus plaque reduction assay was performed. **(c)** Representation of agglutination of NSC85561 and NSC84472 in serial twofold dilution started with 100 μM and performed in three individual experiments.

Figure 4 Similar outcomes from post-screening analysis and HA inhibition assay.

(a) The hierarchical clustering and interaction profile of the twelve derivatives and NSC85561. The hierarchical tree represents compound similarities. The thirteen compounds are listed on the *y*-axis and the interactive residues are listed on the *x*-axis. The first code of interactive residue stands for the force between compounds and residues, E for electrostatic force, H for hydrogen bond force, V for van der Waals. The second code stands for the interaction in the main chain (M) or side chain (S). The left stands for the residue type and serial number in H1N1 HA. The interaction types color the residue with a pharmacological preference $\geq$ 0.4. E: red; H: green; V: grey. The H or E interactions are in green when the energy $\leq$ -2.5. The V interactions are in green if the energy $<$ -4. **(b)** Representation of agglutination of the thirteen compounds in serial twofold dilution starting with 25 μM via HA inhibition assay and performed in three individual experiments.

Figure 5 Identification of NSC47715, NSC85561, and NSC7223.

The MDCK cells were infected with A/WSN/1933 (H1N1, 100 PFU) and mixed with the five compounds at **(a)** 25 μM and **(b)** 12.5 μM. The virus plaque reduction assay was performed. **(c)** The value of the relative cell viability for NSC47715, NSC85561, and NSC7223 evaluated by MTS assay. **(d)** The MDCK cells infected with the H1N1 virus were mixed with the three compounds in two-fold serial diluted concentration. Plaque reduction assay performed in at least three independent experiments.

Figure 6 Docked compounds show a high affinity toward the RBS in HA.

The docked compounds **(a, d, g)** NSC47715, **(b, e, h)** NSC85561, and **(c, f, i)** NSC7223 are displayed in different visualization modes. **(a-c)** The interaction diagrams between docked compounds and protein are shown in a 2D plot by the PROTEINS PLUS server[26]. **(d-f)** Visualization of the docked compounds with anchors in RBS of HA. HA is shown in the cartoon, anchors are shown as transparent spheres, interactive residues are shown as cyan sticks, and the docked compounds are shown as grey

sticks. **(g-i)** The docked compounds in the RBS with surface mode. PyMOL software was used to draw all the 3D figures[33].

Figures

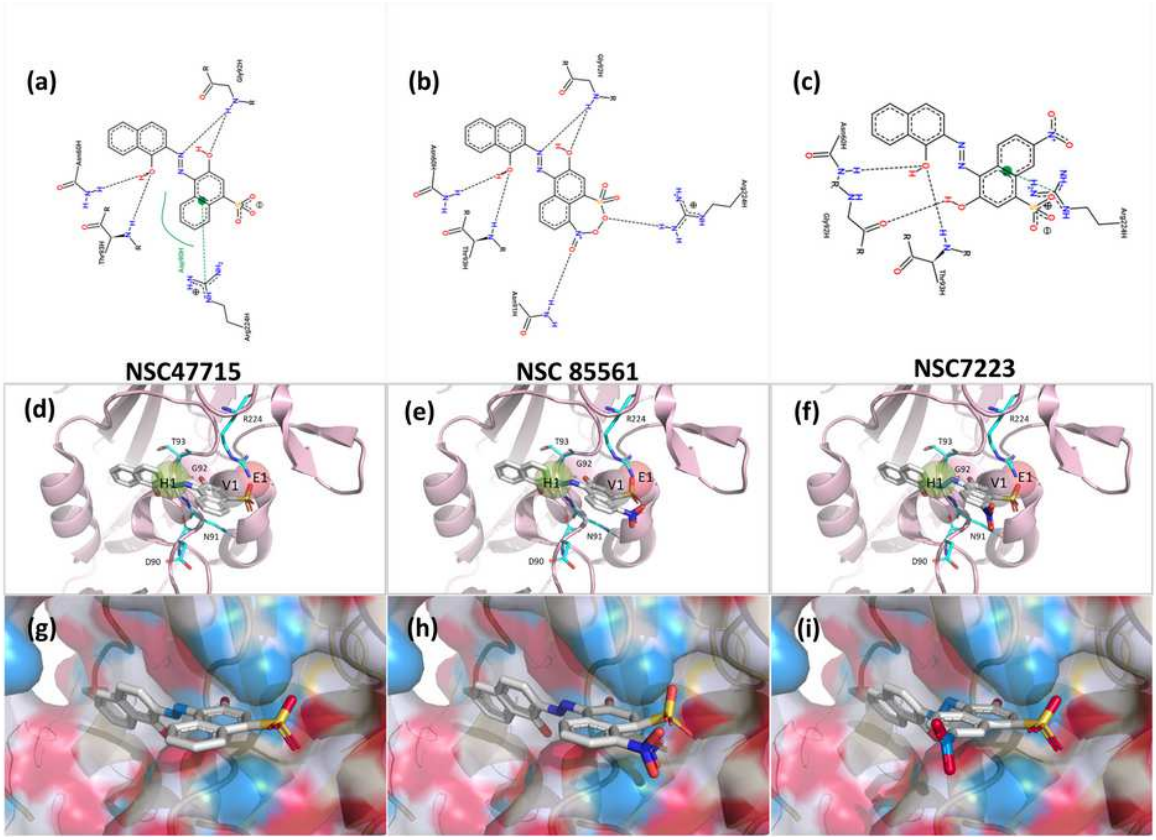

Figure 1

Docked compounds show a high affinity toward the RBS in HA. The docked compounds (a, d, g) NSC47715, (b, e, h) NSC85561, and (c, f, i) NSC7223 are displayed in different visualization modes. (a-c) The interaction diagrams between docked compounds and protein are shown in a 2D plot by the PROTEINS PLUS server26. (d-f) Visualization of the docked compounds with anchors in RBS of HA. HA is shown in the cartoon, anchors are shown as transparent spheres, interactive residues are shown as cyan sticks, and the docked compounds are shown as grey sticks. (g-i) The docked compounds in the RBS with surface mode. PyMOL software was used to draw all the 3D figures33.

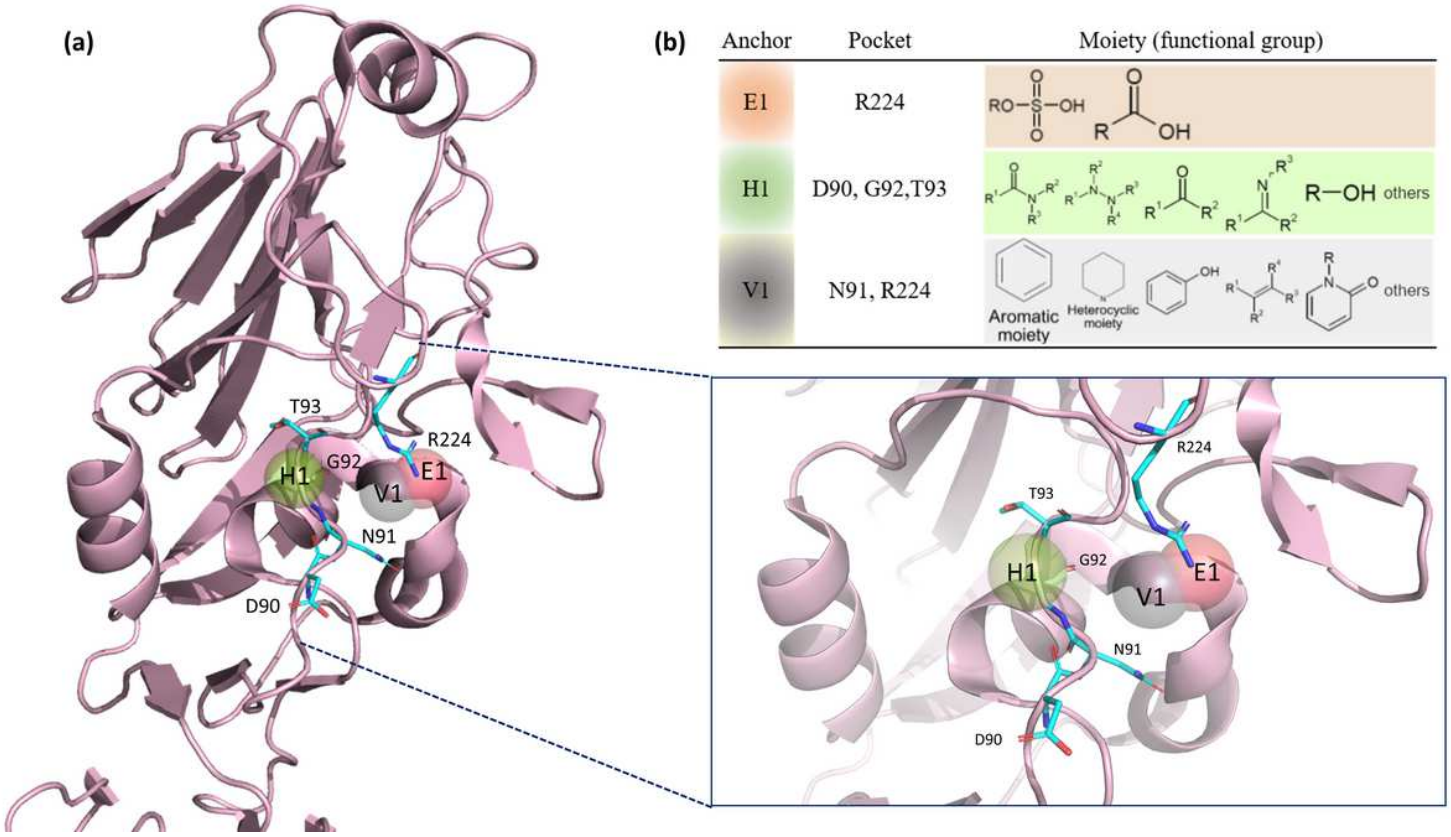

Figure 1

The anchors of the receptor-binding site from SiMMap. (a) The structure of RBS of H1N1 HA (chain H of 1RUY) is shown in cartoon, and three anchors are shown in transparent spheres. Red E1 stands for electrostatic force, green H1 for hydrogen bond force, and gray V1 for van der Waal forces. The corresponding binding pockets (residues) are shown in the cyan sticks. (b) The table of binding pockets and moieties for each anchor. The moiety of anchor represents the functional group preference in the top-ranked compounds.

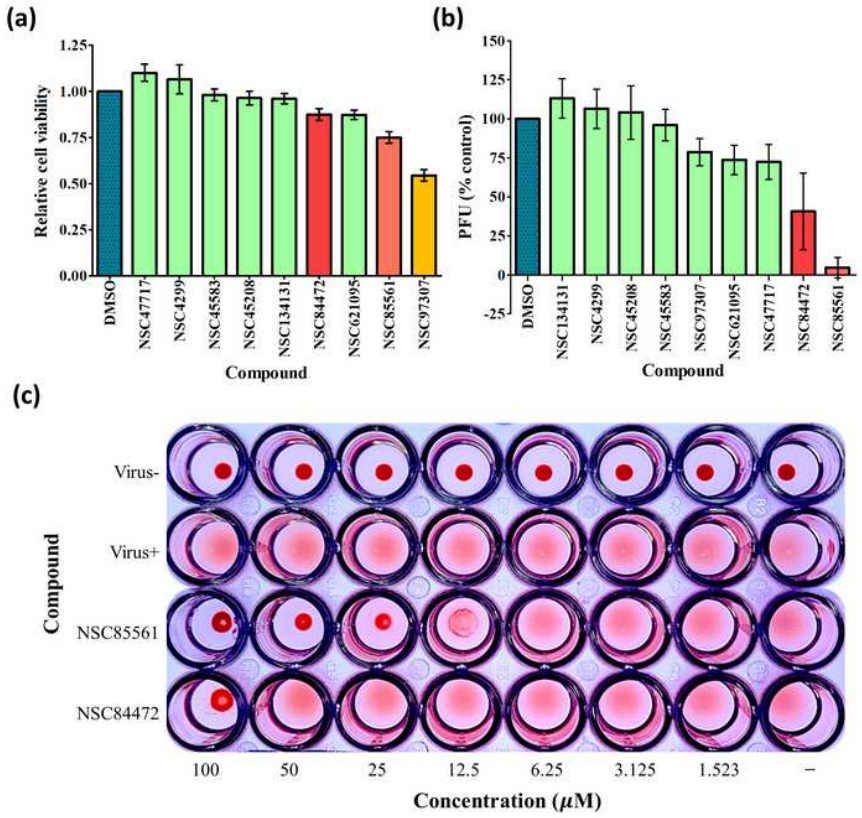

Figure 1

NSC85561 unveiled from the initial compounds. (a) MDCK cells were treated with the initial nine compounds at 100 μM and cell viability was measured using the MTS assay. (b) MDCK cells were infected with A/WSN/1933 (H1N1, 100 PFU) and the compounds at 25 μM. The virus plaque reduction assay was performed. (c) Representation of agglutination of NSC85561 and NSC84472 in serial twofold dilution started with 100 ▯▯ and performed in three individual experiments.

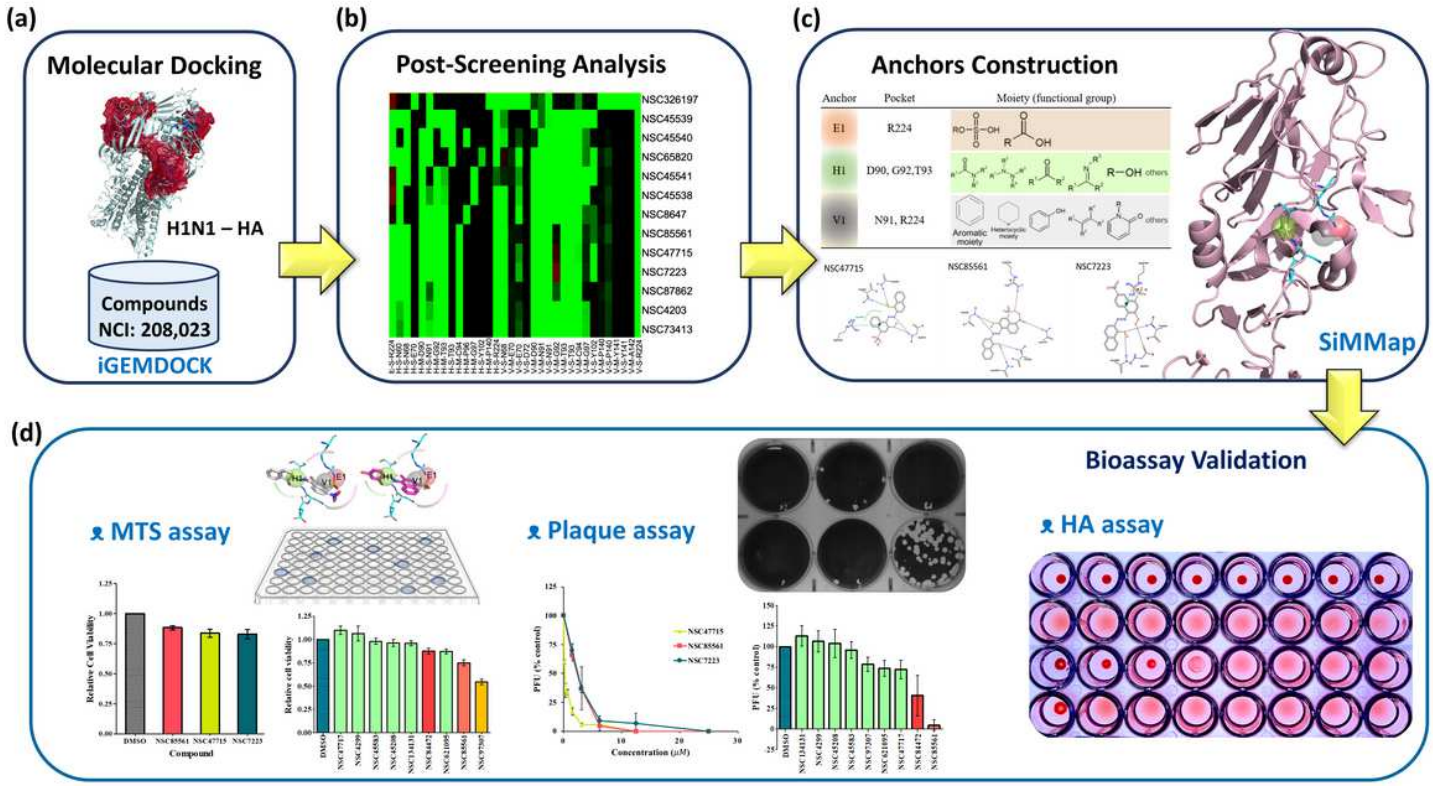

Figure 1

Framework of the study. (a) Virtual screening for HA inhibitors by docking compounds from NCI and hemagglutinin of H1N1 (PDB ID: 1RUY) based on the iGEMDOCK program. (b) Post-screening analysis by generating consensus interaction profile between docked compounds and HA residues. (c) Anchors construction by using SiMMap for detecting the most favorable binding pockets and moieties. (d) Bioassay validation of potential compounds through three bioassays. MTS assay: to evaluate the toxicity of the compounds. Plaque reduction assay: to quantify the antiviral activity of the compounds versus IAV. Hemagglutination inhibition assay: to measure influenza-specific levels in blood serum.

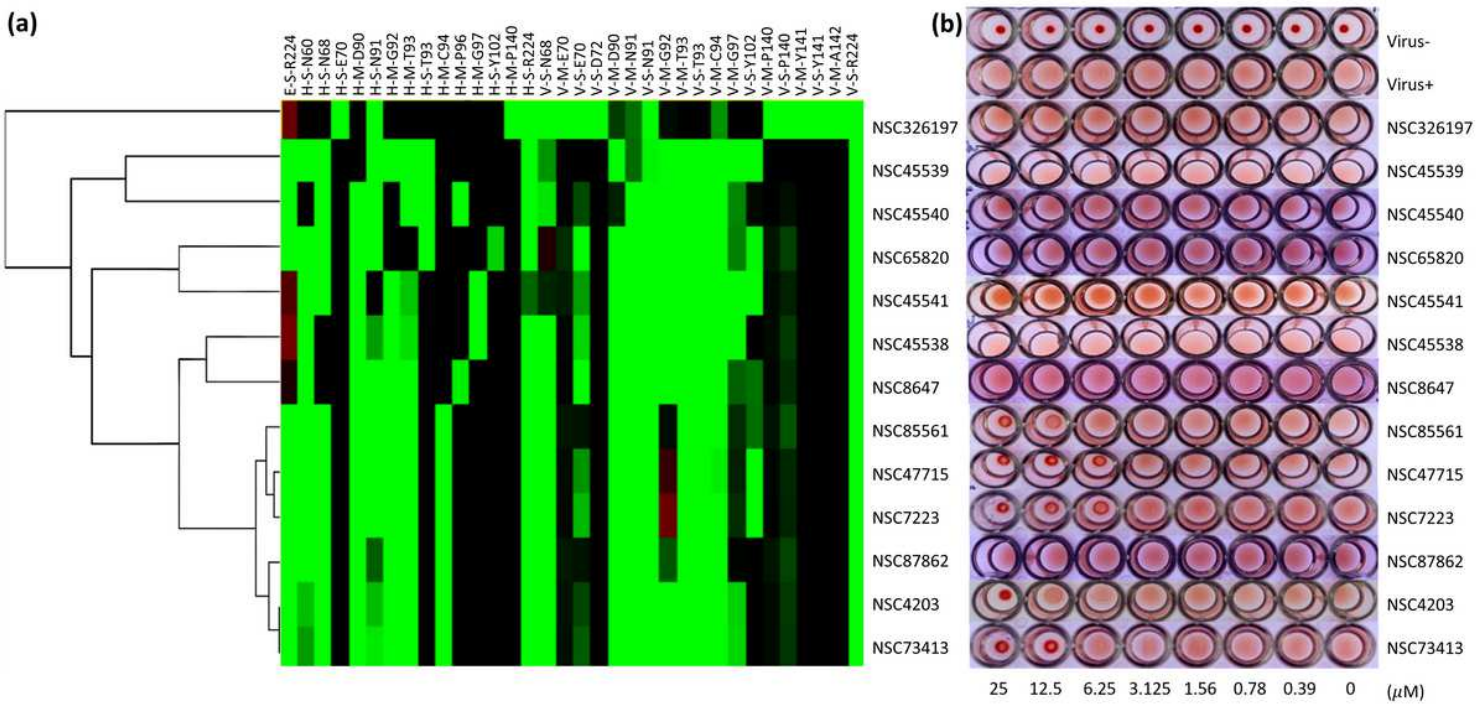

Figure 1

Similar outcomes from post-screening analysis and HA inhibition assay. (a) The hierarchical clustering and interaction profile of the twelve derivatives and NSC85561. The hierarchical tree represents compound similarities. The thirteen compounds are listed on the y-axis and the interactive residues are listed on the x-axis. The first code of interactive residue stands for the force between compounds and residues, E for electrostatic force, H for hydrogen bond force, V for van der Waals. The second code stands for the interaction in the main chain (M) or side chain (S). The left stands for the residue type and serial number in H1N1 HA. The interaction types color the residue with a pharmacological preference ≥ 0.4. E: red; H: green; V: grey. The H or E interactions are in green when the energy ≤ -2.5. The V interactions are in green if the energy < -4. (b) Representation of agglutination of the thirteen compounds in serial twofold dilution starting with 25 μM via HA inhibition assay and performed in three individual experiments.

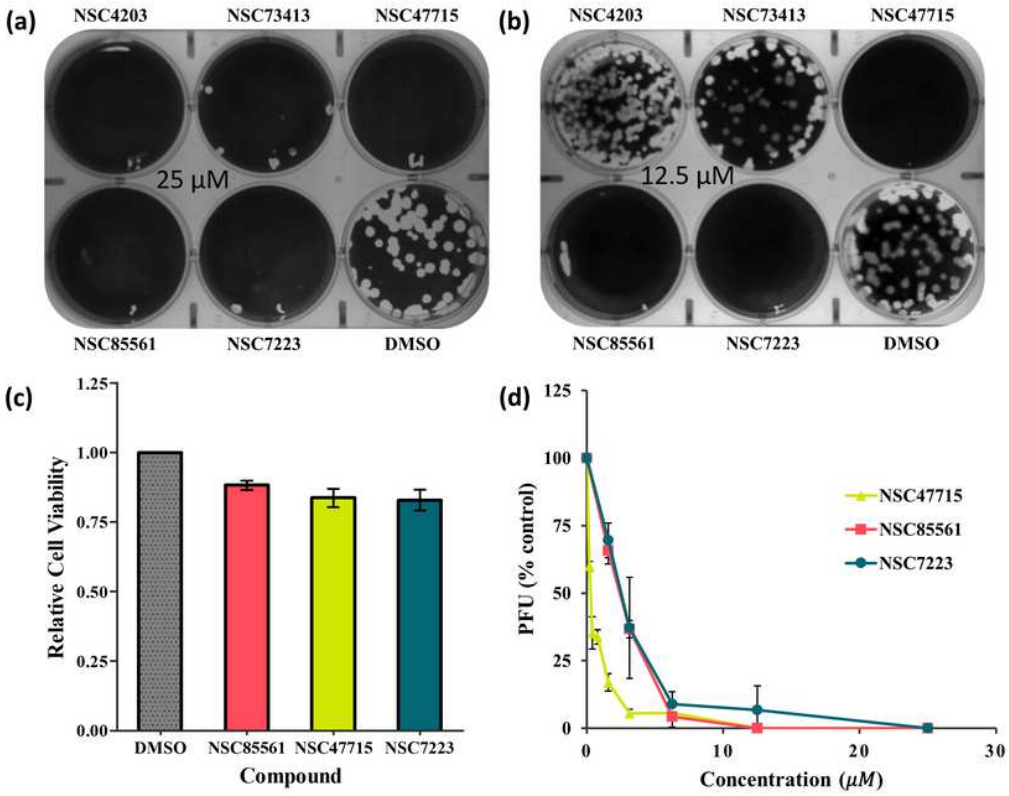

Figure 1

Identification of NSC47715, NSC85561, and NSC7223. The MDCK cells were infected with A/WSN/1933 (H1N1, 100 PFU) and mixed with the five compounds at (a) 25 μM and (b) 12.5 μM. The virus plaque reduction assay was performed. (c) The value of the relative cell viability for NSC47715, NSC85561, and NSC7223 evaluated by MTS assay. (d) The MDCK cells infected with the H1N1 virus were mixed with the three compounds in two-fold serial diluted concentration. Plaque reduction assay performed in at least three independent experiments.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryData.docx