



OPEN

Potent sialic acid inhibitors that target influenza A virus hemagglutinin

Yu-Jen Chang^{1,6}, Cheng-Yun Yeh^{1,6}, Ju-Chien Cheng², Yu-Qi Huang², Kai-Cheng Hsu³, Yu-Feng Lin⁴ & Chih-Hao Lu^{1,2,5}✉

Eradicating influenza A virus (IAV) is difficult, due to its genetic drift and reassortment ability. As the infectious cycle is initiated by the influenza glycoprotein, hemagglutinin (HA), which mediates the binding of virions to terminal sialic acids moieties, HA is a tempting target of anti-influenza inhibitors. However, the complexity of the HA structure has prevented delineation of the structural characterization of the HA protein–ligand complex. Our computational strategy efficiently analyzed > 200,000 records of compounds held in the United States National Cancer Institute (NCI) database and identified potential HA inhibitors, by modeling the sialic acid (SA) receptor binding site (RBS) for the HA structure. Our modeling revealed that compound NSC85561 showed significant antiviral activity against the IAV H1N1 strain with EC₅₀ values ranging from 2.31 to 2.53 μM and negligible cytotoxicity (CC₅₀ > 700 μM). Using the NSC85561 compound as the template to generate 12 derivatives, robust bioassay results revealed the strongest antiviral efficacies with NSC47715 and NSC7223. Virtual screening clearly identified three SA receptor binding site inhibitors that were successfully validated in experimental data. Thus, our computational strategy has identified SA receptor binding site inhibitors against HA that show IAV-associated antiviral activity.

Seasonal influenza virus poses a severe threat to human health. An estimated 56 million people worldwide were infected with influenza virus between October 1, 2019 and April 4, 2020, causing serious illness in 39 to 56 million and 62,000 deaths¹. Influenza-related deaths most commonly occur in people aged 65 years and over in industrialized countries², while among young children, a systematic review published in 2008 of the global burden of respiratory infections related to seasonal influenza among young children calculated that of all deaths in children aged younger than five years attributable to influenza-associated acute lower respiratory infections, 99% occurred in developing countries³. The effectiveness of seasonal vaccines intended to protect against infection frequently fail, because of influenza's genetic drift and reassortment ability⁴. Yearly vaccination is the primary means of preventing and controlling influenza⁵, but vaccine efficacy varies according to genetic relatedness among viruses in the vaccine and circulating strains⁶. Currently, three influenza antiviral drugs approved by the US Food and Drug Administration (FDA) are recommended by the US Centers for Disease Control and Prevention (CDC) for use against circulating influenza viruses; Rapivab (peramivir), Relenza (zanamivir), and Tamiflu (oseltamivir), all of which are neuraminidase (NA) inhibitors and interfere with the release of the progeny virion from infected host cells. This process prevents the infection of new host cells and halts infection spread in the respiratory tract⁷. However, despite the apparent treatment efficacy shown by these NA inhibitors, they have also been associated with drug resistance. In Japan, oseltamivir-resistant H1N1 viruses were isolated from 7 of 43 patients (16%) in 2000/2001⁸ and almost one-third (27%) of patients infected with H1N1 viruses between 2005 and 2007 exhibited resistance to oseltamivir in the UK⁹. We therefore sought to determine the feasibility of targeting the influenza virus glycoprotein hemagglutinin (HA) as a potential alternative therapeutic option to NA-mediated therapeutic strategies, capable of inhibiting influenza virus entry.

HA plays a key role during the influenza A virus (IAV) life cycle, particularly in the attachment and penetration of influenza virus strains to host cell molecules¹⁰. No HA inhibitors are as yet available for clinical use, mainly because of the diversity of HAs created by antigenic drift and shift^{11,12}. One report has suggested that

¹The Ph.D. Program of Biotechnology and Biomedical Industry, China Medical University, Taichung, Taiwan. ²Department of Medical Laboratory Science and Biotechnology, China Medical University, Taichung, Taiwan. ³Graduate Institute of Cancer Biology and Drug Discovery, Taipei Medical University, Taipei, Taiwan. ⁴Department of Medical Laboratory Science and Biotechnology, Asia University, Taichung, Taiwan. ⁵Graduate Institute of Biomedical Sciences, China Medical University, Taichung, Taiwan. ⁶These authors contributed equally: Yu-Jen Chang and Cheng-Yun Yeh. ✉email: chlu@mail.cmu.edu.tw

glycosylations in the globular head of the HA is an essential step for viruses to gain virulence and antigenic properties¹³. Mature HA is a homotrimer, with a globular domain containing a RBS specific for SA expressed by host cell glycoproteins and glycolipids¹⁴. 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 uncleaved HA0 subunits, which are unable to fuse the host membranes. By interrupting the correct folding of HAs in progeny influenza viruses, HA0 inhibitors elicit nonfunctional HA conformation and thus block viral entry¹⁵. For instance, the HA0 inhibitor AF4HIK1 blocks immature HA0 cleavage and inhibits the infectivity of IAV¹⁶. Another example is nafamostat, a serine protease inhibitor, which reduces cleavage of the precursor protein HA0¹⁷. By folding into a jelly-roll motif of eight stranded antiparallel β -sheets with a shallow pocket at the distal tip, the HA1 subunit acts as an RBS surrounded by antigenic sites¹⁸. The antiviral mechanism of HA1 inhibitors is mainly attributed to their ability to block receptor binding and prevent the acquisition of a viral infection. An analysis of extracts from a traditional medicinal plant used in Borneo for treating symptoms of influenza infection has revealed that the useful substances were not limited to SA-like compounds, but also included non-SA-like components that might act against other viral proteins besides HA and NA¹⁹. Other research has shown that synthetic SA-mimic peptides have a high affinity to the RBS and thus block the initial infection²⁰. Another study has described how multivalent 6'-sialyllactose-polyamidoamine (6SL-PAMAM) conjugates, which include a dendrimer scaffold against the SA receptor binding site, are capable of effectively inhibiting IAV infection²¹. Notably, recent research has detailed the isolation and characterization of 1428A33/1, 1428B5/1 and F3A1, human monoclonal antibodies (mAbs) that target residues of the RBS of HA and neutralize A(H1N1)pdm09 escape mutant viruses that very occasionally escape from these mAbs²². The HA2 subunit mediates the host endosomal membrane's fusion with the viral membrane, allowing viral ribonucleoprotein entry into the host cell²³. During the process of HA-mediated fusion, metastable HA2 subunits undergo irreversible rearrangement under acidic pH, which causes membrane fusion and the completion of viral entry¹⁵. The HA2 inhibitor Arbidol (umifenovir) interacts directly with HA2 and can prevent the acidification-induced conformation change of HA2, impeding fusion with endosome membranes²⁴.

After considering all HA subunits in the IAV structure, our study focused on the RBS of the HA1 subunit. We analyzed records of compounds held in the United States National Cancer Institute (NCI) database to identify potential inhibitors. Over 200,000 compounds were screened using the molecular docking method. Docked compounds were used to recognize interaction preferences by analyzing the RBS with interacting residues and specific physical–chemical properties. These models enabled us to identify inhibitors with novel scaffolds. Several HA inhibitors were identified and validated by a series of bioassays, including cell proliferation, plaque reduction, and hemagglutination inhibition. To refine our analysis and discover more compounds targeting HA1, we subjected the initial compound's derivatives to the same experimental tests, which revealed some attractive leads for potential HA1 inhibitors. Our study describes our use of a comprehensive framework for efficient screening of lead compounds for further drug design and development as HA1 inhibitors.

Results

A framework of anchor construction and post-screening analysis. Figure 1 gives an overview of how we identified novel HA inhibitors by anchor construction and post-screening analysis. First, the structure of the RBS of the HA1 subunit (PDB ID: 1RUY)²⁵ was selected as the target protein, and 208,023 NCI compounds were collected as the screening compound database (Fig. 1a). The molecular docking and post-screening analyses were performed by iGEMDOCK²⁶ (Fig. 1b). Sialic acid served as the positive control and was docked into our prepared sialic acid RBS to validate our docking tool. SiMMap²⁷ anchors were subsequently constructed from the SiMMap server (a site-moiety map for drug discovery and mechanisms server), based on the docking scores and the types of interactions between the 1,000 top-ranked compounds and the RBS (Fig. 1c). An anchor is composed of a binding pocket with corresponding interacting residues, moiety preference, and interaction type (E: electrostatic, H: hydrogen-bonding, or V: van der Waals forces). These 1,000 compounds were re-ranked based on the SiMMap score, and the top 20 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). Figure 1 illustrates the process of how we conducted our research and validated our hypothesis.

Anchors of the receptor-binding site. Three pivotal anchors, E1, H1, and V1 (Fig. 2b), were generated by SiMMap in the RBS. Of all three anchors and their respective pockets, sulfonic acid and carboxylic acid moieties prefer to interact with R224, the basic residue that lies in the 220-loop of the E1 anchor. Notably, all compounds that have a nucleophilic tendency, including sulfonic acid and carboxylic acid moieties, have a binding affinity with 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 anchor prefers to bind to amide, ketone, amine, and hydroxyl groups (with decreasing preferences indicated in this order). Lastly, the V1 anchor has a high tendency to bind to an aromatic and heterocyclic moiety, interacting with one amide residue (N91) and one basic residue (R224), as well as the E1 anchor. Our 3-dimensional (3-D) model is shown in Fig. 2a.

Compounds derived from NSC85561. Supplementary Table S1 lists the nine compounds that were purchased according to the SiMMap scores. The compounds were subjected to bioassay examinations and safety analysis using the MTS method²⁸ (a proliferation assay that uses the tetrazolium compound, MTS, in combination with an electron coupling reagent to produce a colorimetric change). Each compound was applied as 100 μ M concentrations to epithelial Madin Darby canine kidney (MDCK) cells to measure optical density (OD) values and toxic effects of the compounds (Fig. 3a). NSC97307 was excluded due to its low relative cell viability

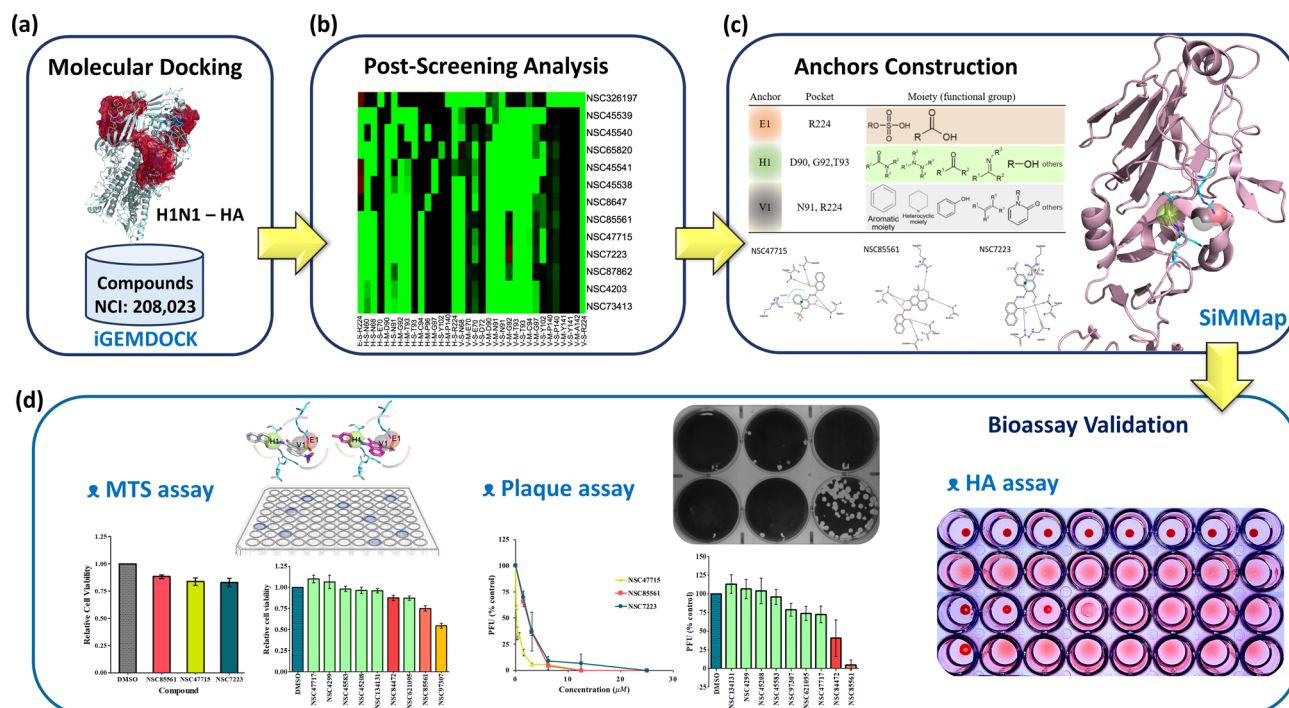


Figure 1. Framework of the study. **(a)** Virtual screening for HA inhibitors using docking compounds from the NCI and the H1N1 hemagglutinin (PDB ID: 1RUY) was based on the iGEMDOCK program. **(b)** Post-screening analysis was conducted by generating the consensus interaction profile between the docked compounds and HA residues. **(c)** Anchors were constructed using the SiMMap to detect the most favorable binding pockets and moieties. **(d)** Validation of potential compounds was performed using three bioassays: The MTS assay evaluated the toxicities of the compounds; the plaque reduction assay quantified the antiviral activity of the compounds versus IAV; and the hemagglutination inhibition assay measured influenza-specific levels in type-O human serum.

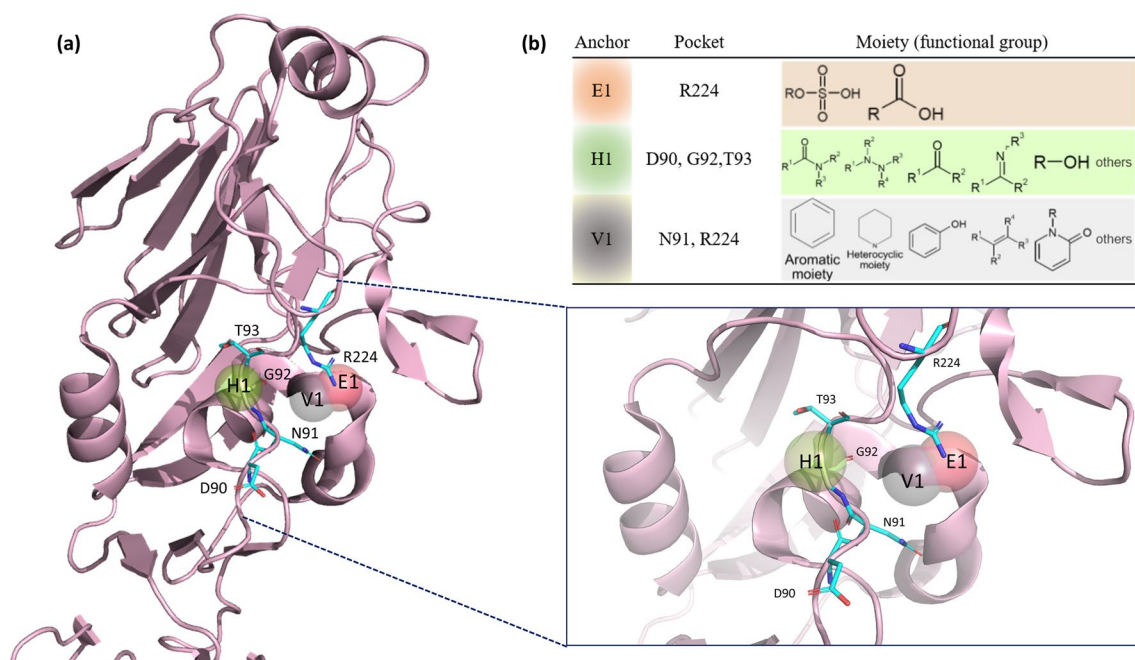


Figure 2. Receptor-binding site anchors obtained from SiMMap. **(a)** The RBS structure of the H1N1 HA (chain H of 1RUY) is depicted as a cartoon and three anchors as transparent spheres. Red E1 stands for the electrostatic force, green H1 for hydrogen bond force, and grey V1 for van der Waals forces. The corresponding binding pockets (residues) are shown in the cyan sticks. **(b)** The table presents binding pockets and moieties for each anchor. Each moiety of the anchor represents the functional group preference of the top-ranked compounds.

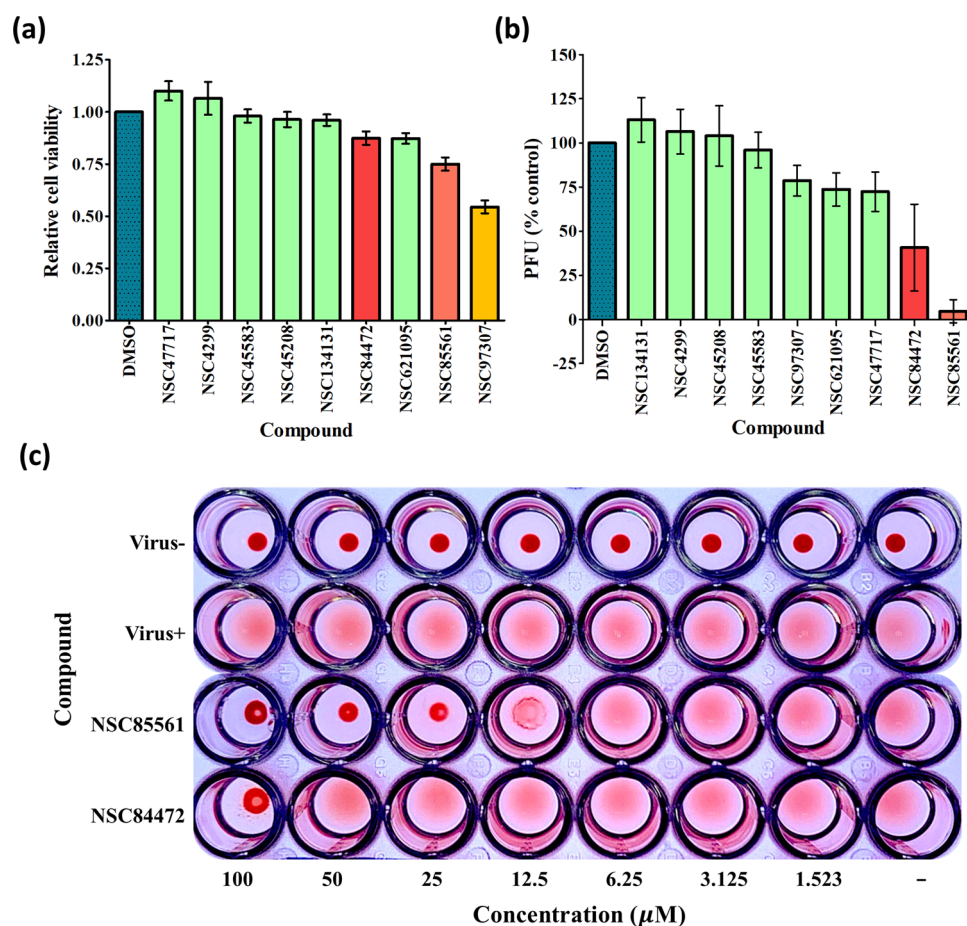


Figure 3. Identification of NSC85561 from the initial compounds. **(a)** MDCK cells were treated with the initial nine compounds at the concentration of 100 μ M and cell viability was measured using the MTS assay. **(b)** MDCK cells were infected with A/WSN/1933 (H1N1, 100 PFU) were treated with the mixture at the concentration of 25 μ M. A virus plaque reduction assay was performed. **(c)** Representation of the agglutination of NSC85561 and NSC84472 in serial two-fold dilutions started with 100 μ M and was performed in three individual experiments.

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

Table 1. Plaque reduction assay of the nine purchased compounds. All experiments were performed in triplicate. Plaque-forming unit (PFU) values are shown as the mean \pm standard deviation.

(below 0.75). When the cells were treated with 25 μ M concentrations of each compound, identification of antiviral effects by plaque reduction assay revealed dramatic decreases from baseline in plaque-forming unit (PFU) values with NSC85561 and NSC84472 (mean PFUs of 5 \pm 5% and 40 \pm 20% versus control, respectively), identifying these compounds as effective inhibitors of the RBS (Fig. 3b and Table 1). Antiviral efficacy examination by HA assay began with concentrations of 100 μ M for NSC85561 and NSC84472, followed by two-fold serial dilution. Whereas precipitation formed at 100 μ M with NSC84472, precipitation did not occur with NSC85561 until the compound was diluted to 12.5 μ M (Fig. 3c). Since the HA assay revealed much greater efficacy of NSC85561,

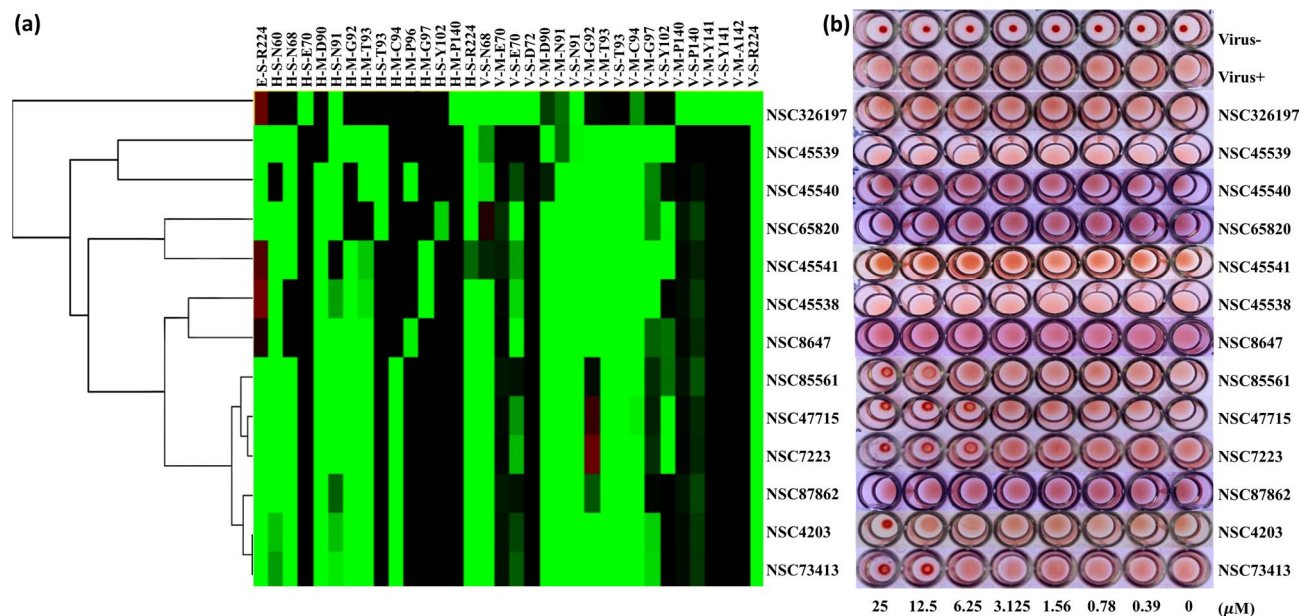


Figure 4. Similar outcomes from the post-screening analysis and HA inhibition assay. **(a)** The hierarchical clustering and interaction profiles of the 12 derivatives and NSC85561. The hierarchical tree represents compound similarities. The 13 compounds are listed on the y-axis and the interactive residues are listed on the x-axis. The first code of the interactive residue stands for the force between compounds and residues, E for electrostatic force, H for hydrogen bond force, and V for van der Waals forces. The second code stands for the interaction in the main chain (M) or side chain (S). The third code represents the residue type and serial number of the H1N1 HA. Residues with a pharmacological preference value of ≥ 0.4 are colored according to the interaction type: E = red; H = green; V = grey. The H or E interactions are represented in green when the energy ≤ -2.5 . The V interactions are in green when the energy is < -4 . **(b)** Representation of agglutination of the 13 compounds in serial two-fold dilutions, starting with 25 μM using the HA inhibition assay and performed in three individual experiments.

we used this compound as a template to generate 12 further compounds based on the AtomPair fingerprints generated by RDKit²⁹ Fingerprint in KNIME³⁰. The AtomPair approach was used to generate 825 topological features^{31,32} for all of the compounds in our database, and these topological features were then used to measure the similarities between any two compounds by calculating the Pearson's correlation coefficient (PCC). Any compound with a correlation coefficient greater than 0.8 was considered to be a derivative of the active compound NSC85561. The 2-D and PCC values of NSC85561 and the 12 derivative compounds are shown in Supplementary Table S2.

Bioassay data matched the post-screening results. The HA assay outcomes matched the iGEM-DOCK post-screening results. Figure 4a depicts the hierarchical clustering results of 13 compounds according to their interaction profiles. Patterns were similar for 6 of the compounds (from NSC85561 to NSC73413), with evidence of early clustering (Fig. 4a). The HA experiments subjected all 12 derivative compounds and the initial compound to the two-fold serial dilution process, starting with a 25 μM concentration. Precipitation with NSC4203 formed at 25 μM , whereas this phenomenon occurred at 12.5 μM and over for both NSC85561 and NSC73413, and at ≥ 6.25 μM for both NSC47715 and NSC7223 (Fig. 4b). To validate the efficacies of these five compounds, we used different concentrations of the antiviral plaque reduction assay. The MDCK cells were seeded in 6-well plates for 24 h, then incubated with the five compounds at 25 μM or 12.5 μM (Fig. 5a,b). At the concentration of 12.5 μM , the efficacies of NSC47715, NSC85561, and NSC7223 were much higher than those of NSC4203 and NSC73413. Figure 5c illustrates the bar plot of the relative cell viabilities of NSC85561, NSC47715, and NSC7223. Average values were 0.88 for NSC85561, 0.83 for NSC47715, and 0.83 for NSC7223 at a concentration of 100 μM . The compound concentration required to reduce cell viability by 50% (CC_{50}) was used to evaluate the safety of the three compounds. CC_{50} values were 700 μM for NSC85561, 900 μM for NSC7223, and > 1000 μM for NSC47715. The half-maximal concentration (EC_{50}) was used to measure the efficacy of the three compounds. MDCK cells infected with the virus (100 PFU) were mixed with NSC47715, NSC85561, and NSC7223, and subjected to the virus plaque reduction assay (Fig. 5d). Among these three compounds, NSC47715 exhibited the lowest values, so was the most efficient at the lowest concentration at inhibiting IAV. Selectivity Index (SI) values (the CC_{50} - EC_{50} ratio) evaluated cytotoxicity and efficacy. SI values were 289 for NSC85561, 360 for NSC7223, and $> 3,000$ for NSC47715. The above results are summarized in Table 2.

Our previous data demonstrated that NSC47715 had the best potential in anti-IAV activity in the early stage of virus entry. Based on the results of the time-of-addition assay, NSC47715 was as expected to block the virus binding to the cells. However, the virucidal effect obviously contribute to the anti-IAV activity of the NSC47715.

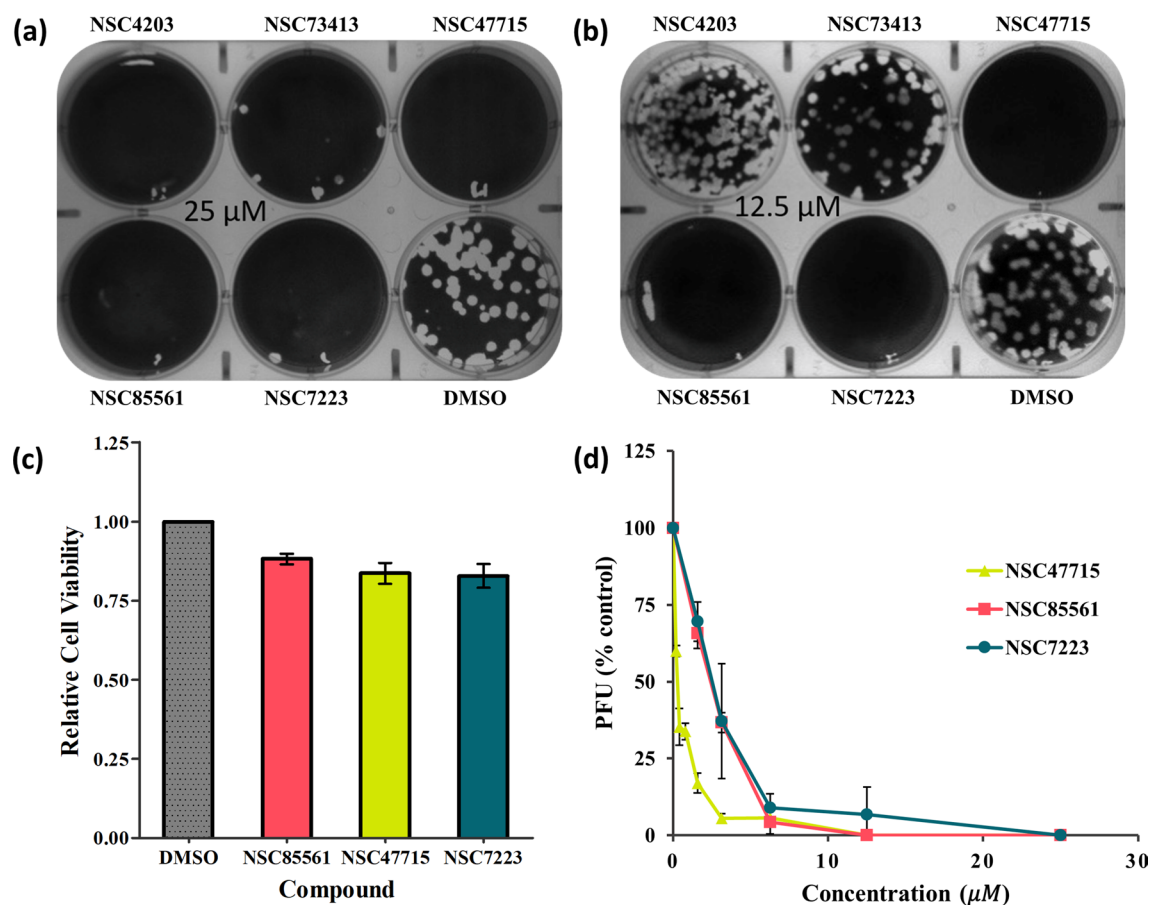


Figure 5. Identification of NSC47715, NSC85561 and NSC7223. MDCK cells were infected with A/WSN/1933 (H1N1, 100 PFU) and mixed with the five compounds at (a) 25 μM and (b) 12.5 μM . A virus plaque reduction assay was performed. (c) Relative cell viability value for NSC47715, NSC85561 and NSC7223 were evaluated by the MTS assay. (d) MDCK cells infected with the H1N1 virus were mixed with the three compounds in two-fold serial diluted concentrations. The plaque reduction assay was performed in at least three independent experiments.

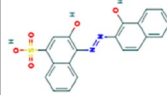
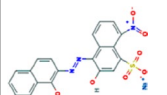
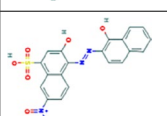
Compounds	CC ₅₀	EC ₅₀	SI	Score ^a	Energy ^b	VDW ^b	HBond ^b	Elec ^b	2-D structure
NSC47715	> 1000	0.29 ± 0.02	> 3448	3.233	-126.11	-72.98	-49.83	-3.30	
NSC85561	700	2.42 ± 0.11	289	3.228	-123.49	-69.96	-50.27	-3.25	
NSC7223	900	2.50 ± 0.76	360	3.227	-127.09	-71.48	-53.31	-2.30	

Table 2. CC₅₀ (μM), EC₅₀ (μM) and SI values for NSC47715, NSC85561 and NSC7223. The EC₅₀ value is presented as the mean ± standard deviation. CC₅₀ cytotoxic concentration 50%, EC₅₀ half-maximal effective concentration, VDW van der Waals force, Elec electrostatic force. ^aFrom SiMMap. ^bFrom iGEMDOCK.

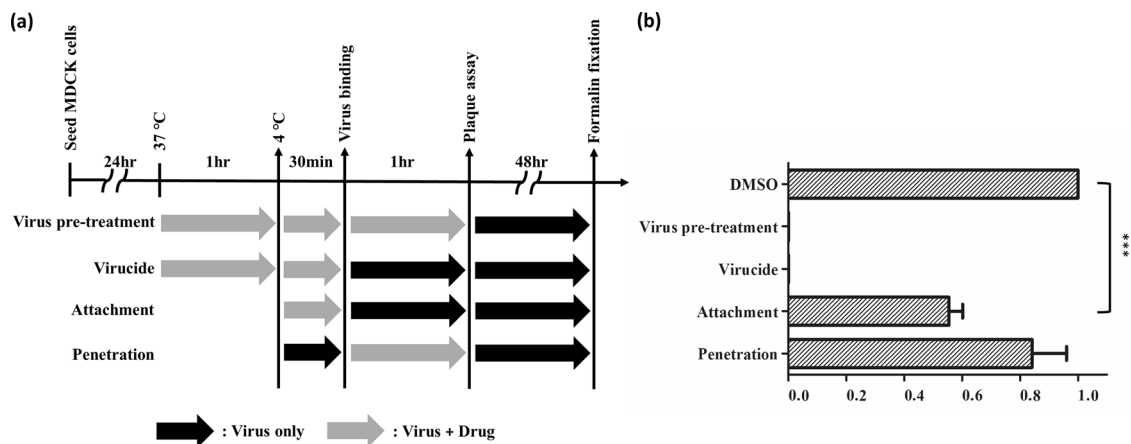


Figure 6. Time-of-addition assay. (a) Schematic representation of the antiviral activity of NSC47715 compound in different entry stages. (b) MDCK cells were infected with H1N1 (A/WSN/1933) and treated with NSC47715 compound at the concentration of $EC_{90}(\mu M)$ at the indicated time according to the different strategy presented above. The values were presented as mean \pm standard error of three independent experiments and were normalized to the values of DMSO solvent control. *** ($p < 0.0001$) shows there is a statistically different between NSC47715 and DMSO solvent control in the attachment stage stage.

The virucidal effect represents the activity by which the compound interacts with virus and consequently prevents the binding of viral particles to the cells. Though the docking strategy is based on the blockage of the interaction of HA and sialic acid receptor, the compound may interact with HA causing the instability of virus structure leading to the failure binding to the cells. Taken together, NSC47715 displayed anti-IAV activity by blockage of virus binding to cells (Fig. 6b).

Discussion

The genetic diversity of IAV poses an ongoing threat to global public health and much research has concentrated on the evolution of IAV and sought to clarify the mechanisms caused by the rapid mutations of the virus and genetic reassortment of gene segments³³. The animal-human interface has also aroused the attention of scientists³⁴ and encouraged the development of rigorous clinical guidelines for the prevention and treatment of influenza³⁵. Few US FDA-licensed anti-influenza A virus drugs are available for clinical despite use, massive screening for new agents against IAV. The NA inhibitor oseltamivir is the main agent for use against IAV, but pathogen mutations have conferred resistance to oseltamivir in influenza A (H7N9) viruses^{36,37}. Thus, NA inhibitors have limited efficacy against seasonal influenza viruses and emphasize the need to develop alternative anti-IAV therapies. Due to its critical role in the early stage of IAV, HA is a potential target protein for anti-influenza agents^{38–40}. The HA sequence structure has been extensively explored and roughly divided into two parts: the SA receptor binding site and the fusion domain²⁵.

In our study, the SA receptor binding site served as the target to block IAV infection, which differs from the recent research focusing on NA^{41–43}. We first modeled the sialic acid RBS and built a compound dataset using NCI records. Using the process of molecule docking, we identified the top 1,000 compounds based on the iGEM-DOCK energy values (Fig. 1a,b). To narrow down the list of compounds, we undertook post-virtual screening analysis on the SiMMap server to organize the orientation of the compounds and the HA residues, namely the anchors (Fig. 1c, Fig. 2). We examined the environment of anchors based on research published in 2019, in which the study researchers emphasized the vulnerable HA stalk trimer domain and depicted the conserved residues over various IAV strains⁴⁴. They also found that FluA-20 (human antibody) tends to bind between the 220-loop and the 90-loop. Our models and predictions are consistent with their findings, so the anchors may play an essential role in interfering with the binding of IAV to the SA on the host cells.

Next, we used a bioassay to validate our prediction and hypothesis (Fig. 1d). NSC85561 was initially identified and served as the template to generate the derivatives from our compound database (Fig. 3). Supplementary Table S2 lists the PCC values lying between NSC85561 and the 12 derivatives generated from the atom-pair program. The PCC values range from 0.83 to 0.98, indicating similarity in the structure of the derivatives and the atom arrangement to those of NSC85561.

The interaction profile based on the post-screening analysis presents some intriguing outcomes (Fig. 4a). Although the patterns were similar for the six compounds from NSC85561 through to NSC73413, the anti-IAV efficacy of the HA assay was not as efficient with NSC87862 (Fig. 4b). Compared with the other five compounds, NSC87862 showed a lower binding affinity towards the side chain of the hydrogen-bonding force of N91 and the van der Waals force of Y102. Subsequently, the 2-D interaction diagrams of six compounds with the pocket environment were examined by the PROTEINS PLUS server⁴⁵ (Fig. 7a–c and Fig. S1a–c), which revealed that all compounds except for NSC87862 had two naphthalene rings, which might explain their weak binding affinity. Moreover, the NSC4203 and NSC73413 naphthalene situated between residue T93 and G92 inhibits these compounds from fully binding with IAVs in that orientation, so NSC4203 and NSC73413 exhibit weak efficacy towards those IAVs (Fig. 7d–f and Fig. S1d–f). The RBS interface views of docked NSC47715, NSC85561 and

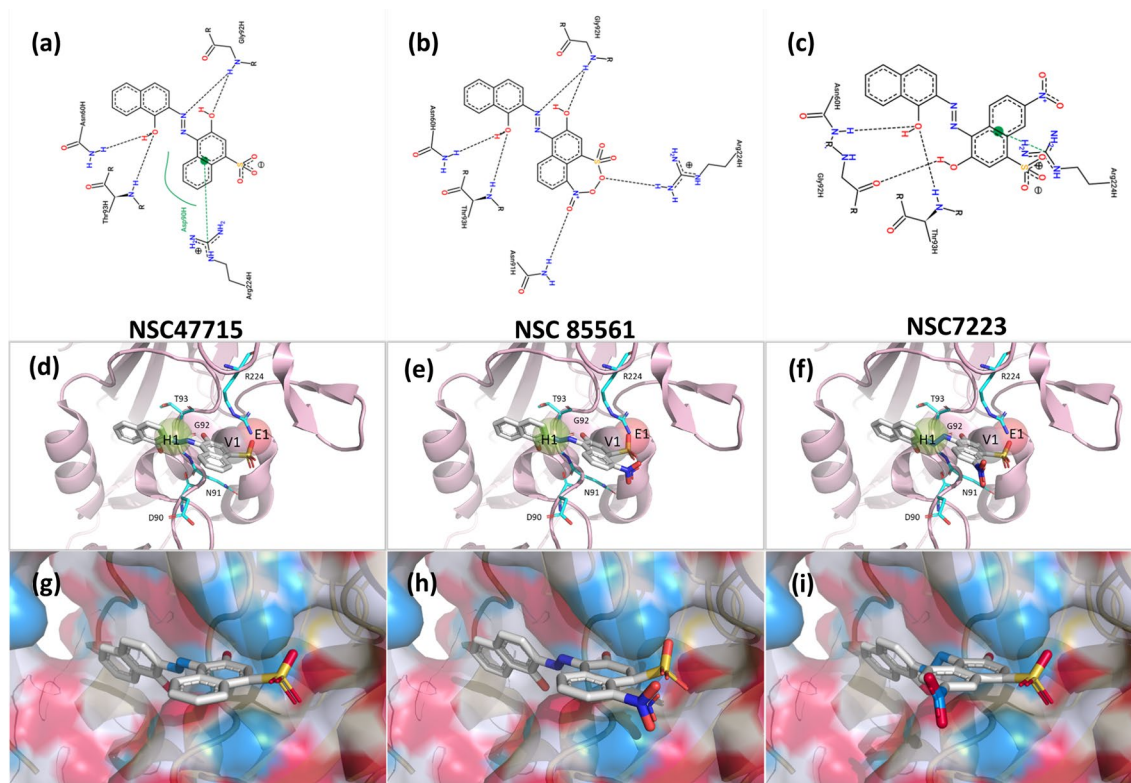


Figure 7. Docked compounds show a high affinity towards the RBS of the HA. The docked compounds (**a,d,g**) NSC47715, (**b,e,h**) NSC85561, and (**c,f,i**) NSC7223 are displayed in different visualization modes. (**a–c**) The interaction diagrams between docked compounds and proteins are shown in a 2-D plot that were examined by the PROTEINS PLUS server⁴⁵. (**d–f**) Visualization of the docked compounds with anchors in the RBS. The HA structure 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 are represented in surface mode. PyMOL software was used to draw all of the 3-D figures⁵⁵.

NSC7223 is shown in Fig. 7g–i. Compared with NSC47715, the additional nitrite in NSC85561 and NSC7223 might interfere with their interaction with the HA binding site. As our investigations have shown, combining virtual screening and bioassay methods allows scientists to clearly and objectively identify novel SA binding site inhibitors.

Conclusions

Our research sheds light on the feasibility of virtual screening generating SiMMap scores for identifying lead anti-IAV compounds. NSC85561 initially showed the highest affinity to the RBS, according to bioassay results. After conducting the same experimental assays on 12 derivatives of NSC85561, we identified three compounds, NSC85561, NSC7223, and NSC47715, as strong potential inhibitors of HA. The SiMMap server enables efficient searching for desired compound functional groups and elucidates the connection between the RBS and compounds NSC85561, NSC7223, and NSC47715. Thus, our data support the use of virtual screening methods to easily identify potential anti-IAV inhibitors that appear to exhibit much greater antiviral efficacy than existing strategies.

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 the Protein Data Bank⁴⁶. H1N1 belongs to the H1 class from the 1930 swine influenza strain. Our preparation of the SA receptor binding site was as follows: Any residues located within 10 Å of the critical residues, K58, E70, S89, P140 and R224⁴⁷, was selected as the SA receptor binding site/pocket in the 1RUY HA1 structure. To confirm whether or not our binding pocket is located in the conserved region of the HA1 subunit, the different hemagglutinin strains (all belonging to the H1N1 clade) were aligned using the multiple sequence alignment tool, Clustal Omega. The alignment results show that our prepared binding pocket is highly conserved in the HA1 subunit (Fig. S2). The 208,023 compounds employed for virtual screening were selected from the NCI database⁴⁸ and were filtered out according to Lipinski's Rule of Five. iGEMDOCK software⁴⁶, a widely used docking program that provides accurate predictions, was used to dock the compounds into the RBS of the HA1 subunit. iGEMDOCK contains three types of interactions: electrostatic, hydrogen-bonding, and 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 identification of hemagglutinin inhibitors. The top 1000 compounds ranked by docking energy with their binding poses were submitted to the SiMMap server for reevaluation²⁷. The SiMMap was used to analyze a protein binding site to identify anchors (antigenic sites), in often critical binding environments⁴³. 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 sites composed of consensus interacting residues were identified. Next, a site-moiety map score was generated for each chosen compound. The identified compounds were reordered with the scores. Finally, potential HA1 inhibitors were selected for rigorous examination based on their ranking, feasibility, physicochemical properties and interaction types.

Cells and viruses. MDCK cells (ATCC accession no. NBL-2) were maintained in Dulbecco's modified eagle medium (DMEM; Gibco) containing 10% fetal bovine serum (FBS; Gibco), 1% penicillin and 1% streptomycin (PS; Gibco) at 37 °C in a 5% CO₂ incubator. Influenza virus H1N1 (A/WSN/33) cells were supplied by Taiwan's Centers for Disease Control and cultured in an influenza 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, USA]) for growth. All viruses were stored at -80 °C. All virus-related experiments were conducted in a BSL2 laboratory.

Cell proliferation assay (MTS). To test for cellular viability and cytotoxicity, we conducted a cell proliferation assay (MTS)⁴⁹, a colorimetric method for sensitive quantification. Based on the reduction of MTS tetrazolium compound by viable cells, the MTS assay generates a colored formazan dye that can be quantified by measuring the absorbance at 490 nm. MDCK cells (1×10^4 cells/well) were seeded into 96-well microtiter plates overnight then mixed with our potential compounds in influenza medium for 72 h at 37 °C under 5% CO₂. The cells were washed twice with phosphate buffer saline (PBS; Na₂HPO₄ 8 mM, NaCl 137 mM, KCl 2.68 mM, KH₂PO₄ 1.47 mM, adjusted to pH 7.2), after the previous medium was removed from the wells. We then added 10 µL of MTS solution and 90 µL DMEM to each well, and the cells were incubated for 30 min. Finally, absorbance at 490 nm was recorded using a microplate reader (SpectraMax iD3, Molecular Devices, USA). A blank control (influenza medium only) and cell controls (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 = $(A_{\text{treatment}} - A_{\text{blank}}) / (A_{\text{control}} - A_{\text{blank}}) \times 100\%$ (A: absorbance)⁵⁰. The cytotoxic concentration required for the identified compounds to reduce cell viability by 50% (CC₅₀) was determined.

Plaque assay. To determine viral titers as plaque-forming units per mL (PFU/mL), we selected the plaque assay^{51,52}, which is a standard method for virus quantification. In brief, MDCK cells (1×10^6 cells/well) were seeded in 6-well plates for 24 h at 37 °C under 5% CO₂. A single layer of cells was infected with the indicated dose of IAV, followed by three washes with PBS. Next, the cells were covered with 0.3% agarose in the influenza 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 activities of the compounds versus IAV, we conducted the virus plaque reduction assay. As described above, the previously seeded cellular monolayer (100 PFU/well) was incubated with or without the selected compounds. The virus suspension was removed following 1 h of IAV adsorption. The cells were immediately washed three times with PBS and overlaid with 0.3% agarose in the influenza medium, with or without the compounds, for 48 h. The EC₅₀ values for all compounds were determined by comparing the number of plaques with those in the virus-infected control. The EC₅₀ was defined as the half-maximal inhibitory concentration and the selectivity index (SI) was determined by the ratio of CC₅₀/EC₅₀.

Time-of-addition assay. The time-of-addition (TOA) protocol determines how long the addition of a newly identified agent can be temporized before losing its antiviral ability during cell culture⁵³. To understand the timing and stage at which the identified compounds inhibit IAV, the virus pretreatment protocol was divided into three stages (virucide, attachment, and penetration) to present the steps that virus may be blocked in the viral entry process⁵⁴. The virucide stage: After hybridizing the drug and incubating the H1N1 (A/WSN/33) for an hour in a 6-well cell culture plate, we subjected the mixture to a 30-min rest at 4 °C before removing the mixture from the plate and subjecting it to incubation at 37 °C, followed by the plaque reduction assay. In the attachment stage, the drug and virus were directly subjected to a 30-min rest at 4 °C in a cell culture plate, and then the mixture was removed, followed by the plaque reduction assay. In the penetration stage, the cell plate mixed with IAV was subjected to a 30-min rest at 4 °C, then was added with the drug for another hour of incubation at 37 °C, and then the mixture was removed, followed by the plaque reduction assay. This TOA approach is illustrated in Fig. 6.

Hemagglutination inhibition assay. We used an erythrocyte agglutination test to determine whether SA bound to a compound and we examined whether the filtered compounds inhibited hemagglutination from causing a precipitation reaction in the SA binding and membrane fusion of HA. We diluted our compounds with

PBS to 1:1,000. MDCK cells were diluted in a series of two-fold dilutions to achieve the IAV titer. To reduce the interference of the antibodies in erythrocytes, we prepared human type-O serum, which was diluted by two-fold serial dilutions 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.

Received: 10 December 2020; Accepted: 6 April 2021

Published online: 21 April 2021

References

- Cameron, C. A. *Flu Season Numbers for 2020/21*. <https://hive.rochesterregional.org/2020/01/flu-season-2020>. Accessed 27 Jan 2021 (2021).
- Thompson, W. W. *et al.* Estimates of US influenza-associated deaths made using four different methods. *Influenza Other Respir. Viruses* **3**, 37–49. <https://doi.org/10.1111/j.1750-2659.2009.00073.x> (2009).
- Nair, H. *et al.* Global burden of respiratory infections due to seasonal influenza in young children: A systematic review and meta-analysis. *Lancet* **378**, 1917–1930. [https://doi.org/10.1016/S0140-6736\(11\)61051-9](https://doi.org/10.1016/S0140-6736(11)61051-9) (2011).
- Hay, A. J., Gregory, V., Douglas, A. R. & Lin, Y. P. The evolution of human influenza viruses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **356**, 1861–1870. <https://doi.org/10.1098/rstb.2001.0999> (2001).
- FDA, U. S. *Influenza (Flu) Antiviral Drugs and Related Information*. <https://www.fda.gov/drugs/information-drug-class/influenza-flu-antiviral-drugs-and-related-information> (2020).
- Basu, A. *et al.* New small molecule entry inhibitors targeting hemagglutinin-mediated influenza A virus fusion. *J. Virol.* **88**, 1447–1460. <https://doi.org/10.1128/JVI.01225-13> (2014).
- Moscona, A. Neuraminidase inhibitors for influenza. *N. Engl. J. Med.* **353**, 1363–1373. <https://doi.org/10.1056/NEJMr050740> (2005).
- Ward, P., Small, I., Smith, J., Suter, P. & Dutkowski, R. Oseltamivir (Tamiflu) and its potential for use in the event of an influenza pandemic. *J. Antimicrob. Chemother.* **55**(Suppl 1), i5–i21. <https://doi.org/10.1093/jac/dki018> (2005).
- Stephenson, I. *et al.* Neuraminidase inhibitor resistance after oseltamivir treatment of acute influenza A and B in children. *Clin. Infect. Dis.* **48**, 389–396. <https://doi.org/10.1086/596311> (2009).
- Luo, M. Influenza virus entry. *Adv. Exp. Med. Biol.* **726**, 201–221. https://doi.org/10.1007/978-1-4614-0980-9_9 (2012).
- Stevens, J. *et al.* Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science* **312**, 404. <https://doi.org/10.1126/science.1124513> (2006).
- 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. [https://doi.org/10.1016/0092-8674\(82\)90135-0](https://doi.org/10.1016/0092-8674(82)90135-0) (1982).
- 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. <https://doi.org/10.1126/scitranslmed.3005996> (2013).
- Kosik, I. & Yewdell, J. W. Influenza hemagglutinin and neuraminidase: Yin-Yang proteins coevolving to thwart immunity. *Viruses* **11**, 346. <https://doi.org/10.3390/v11040346> (2019).
- Shen, X., Zhang, X. & Liu, S. Novel hemagglutinin-based influenza virus inhibitors. *J. Thorac. Dis.* **5**(Suppl 2), S149–S159. <https://doi.org/10.3978/j.issn.2072-1439.2013.06.14> (2013).
- Xiao, H. *et al.* Light chain modulates heavy chain conformation to change protection profile of monoclonal antibodies against influenza A viruses. *Cell Discov.* **5**, 21. <https://doi.org/10.1038/s41421-019-0086-x> (2019).
- Yamaya, M. *et al.* The clinically used serine protease inhibitor nafamostat reduces influenza virus replication and cytokine production in human airway epithelial cells and viral replication in mice. *J. Med. Virol.* <https://doi.org/10.1002/jmv.26700> (2020).
- Isin, B., Doruker, P. & Bahar, I. Functional motions of influenza virus hemagglutinin: A structure-based analytical approach. *Biophys. J.* **82**, 569–581. [https://doi.org/10.1016/S0006-3495\(02\)75422-2](https://doi.org/10.1016/S0006-3495(02)75422-2) (2002).
- Rajasekaran, D. *et al.* Identification of traditional medicinal plant extracts with novel anti-influenza activity. *PLoS ONE* **8**, e79293. <https://doi.org/10.1371/journal.pone.0079293> (2013).
- Matsubara, T. *et al.* Sialic acid-mimic peptides as hemagglutinin inhibitors for anti-influenza therapy. *J. Med. Chem.* **53**, 4441–4449. <https://doi.org/10.1021/jm1002183> (2010).
- Kwon, S. J. *et al.* Nanostructured glycan architecture is important in the inhibition of influenza A virus infection. *Nat. Nanotechnol.* **12**, 48–54. <https://doi.org/10.1038/nnano.2016.181> (2017).
- Yasuhara, A. *et al.* Isolation and characterization of human monoclonal antibodies that recognize the influenza A(H1N1)pdm09 virus hemagglutinin receptor-binding site and rarely yield escape mutant viruses. *Front. Microbiol.* **9**, 2660. <https://doi.org/10.3389/fmicb.2018.02660> (2018).
- Sriwilajaroen, N. & Suzuki, Y. Molecular basis of the structure and function of H1 hemagglutinin of influenza virus. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.* **88**, 226–249. <https://doi.org/10.2183/pjab.88.226> (2012).
- 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).
- Gamblin, S. J. *et al.* The structure and receptor binding properties of the 1918 influenza hemagglutinin. *Science* **303**, 1838–1842. <https://doi.org/10.1126/science.1093155> (2004).
- Yang, J.-M. & Chen, C.-C. GEMDOCK: A generic evolutionary method for molecular docking. *Proteins Struct. Funct. Bioinform.* **55**, 288–304. <https://doi.org/10.1002/prot.20035> (2004).
- 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. <https://doi.org/10.1093/nar/gkq480> (2010).
- Cory, A. H., Owen, T. C., Barltrop, J. A. & Cory, J. G. Use of an aqueous soluble tetrazolium/formazan assay for cell growth assays in culture. *Cancer Commun.* **3**, 207–212. <https://doi.org/10.3727/095535491820873191> (1991).
- RDKit: Open-Source Cheminformatics* (2020).
- Berthold, M. R. *et al.* KNIME: The Konstanz Information Miner. In: Preisach C., Burkhardt H., Schmidt-Thieme L., Decker R. (eds) *Data Analysis, Machine Learning and Applications. Studies in Classification, Data Analysis, and Knowledge Organization*. Springer, Berlin, Heidelberg. 319–326. https://doi.org/10.1007/978-3-540-78246-9_38 (2008).
- Carhart, R. E., Smith, D. H. & Venkataraghavan, R. Atom pairs as molecular features in structure-activity studies: definition and applications. *J. Chem. Inf. Comput. Sci.* **25**, 64–73. <https://doi.org/10.1021/ci00046a002> (1985).
- Willett, P., Barnard, J. M. & Downs, G. M. Chemical similarity searching. *J. Chem. Inf. Comput. Sci.* **38**, 983–996. <https://doi.org/10.1021/ci9800211> (1998).
- Shao, W., Li, X., Goraya, M. U., Wang, S. & Chen, J. L. Evolution of influenza A virus by mutation and re-assortment. *Int. J. Mol. Sci.* **18**. <https://doi.org/10.3390/ijms18081650> (2017).

34. Mostafa, A., Abdelwhab, E. M., Mettenleiter, T. C. & Pleschka, S. Zoonotic potential of influenza A viruses: A comprehensive overview. *Viruses* **10**, <https://doi.org/10.3390/v10090497> (2018).
35. Paules, C. & Subbarao, K. Influenza. *Lancet* **390**, 697–708. [https://doi.org/10.1016/S0140-6736\(17\)30129-0](https://doi.org/10.1016/S0140-6736(17)30129-0) (2017).
36. Kyaw Win, S. M. *et al.* Epidemic of influenza A(H1N1)pdm09 analyzed by full genome sequences and the first case of oseltamivir-resistant strain in Myanmar 2017. *PLoS ONE* **15**, e0229601. <https://doi.org/10.1371/journal.pone.0229601> (2020).
37. Hussain, M., Galvin, H. D., Haw, T. Y., Nutsford, A. N. & Husain, M. Drug resistance in influenza A virus: The epidemiology and management. *Infect. Drug Resist.* **10**, 121–134. <https://doi.org/10.2147/IDR.S105473> (2017).
38. Wu, G. *et al.* Chemoreactive-inspired discovery of influenza A virus dual inhibitor to block hemagglutinin-mediated adsorption and membrane fusion. *J. Med. Chem.* **63**, 6924–6940. <https://doi.org/10.1021/acs.jmedchem.0c00312> (2020).
39. van Dongen, M. J. P. *et al.* A small-molecule fusion inhibitor of influenza virus is orally active in mice. *Science* **363**, <https://doi.org/10.1126/science.aar6221> (2019).
40. Yao, Y. *et al.* An influenza A hemagglutinin small-molecule fusion inhibitor identified by a new high-throughput fluorescence polarization screen. *Proc. Natl. Acad. Sci. U S A* **117**, 18431–18438. <https://doi.org/10.1073/pnas.2006893117> (2020).
41. 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. <https://doi.org/10.18632/oncotarget.20915> (2017).
42. Abdusalam, A. A. A. In-silico virtual screening and ADMET study to find novel neuraminidase N1 inhibitors extended to the 150-cavity. *J. Appl. Pharmaceut. Sci.* **7**, 24–33. <https://doi.org/10.7324/JAPS.2017.70505> (2017).
43. Hsu, K.-C. *et al.* Identification of neuraminidase inhibitors against dual H274Y/I222R mutant strains. *Sci. Rep.* **7**, 12336. <https://doi.org/10.1038/s41598-017-12101-3> (2017).
44. Bangaru, S. *et al.* A site of vulnerability on the influenza virus hemagglutinin head domain trimer interface. *Cell* **177**, 1136–1152 e1118, <https://doi.org/10.1016/j.cell.2019.04.011> (2019).
45. Stierand, K., Maass, P. C. & Rarey, M. Molecular complexes at a glance: Automated generation of two-dimensional complex diagrams. *Bioinformatics* **22**, 1710–1716. <https://doi.org/10.1093/bioinformatics/btl150> (2006).
46. ww, P. D. B. c. Protein Data Bank: The single global archive for 3D macromolecular structure data. *Nucleic Acids Res.* **47**, D520–D528, <https://doi.org/10.1093/nar/gky949> (2019).
47. Li, X. B., Wang, S. Q., Xu, W. R., Wang, R. L. & Chou, K. C. Novel inhibitor design for hemagglutinin against H1N1 influenza virus by core hopping method. *PLoS ONE* **6**, e28111. <https://doi.org/10.1371/journal.pone.0028111> (2011).
48. Milne, G. W. A., Feldman, A., Miller, J. A. & Daly, G. P. The NCI drug information system. 3. The DIS chemistry module. *J. Chem. Inf. Comput. Sci.* **26**, 168–179, <https://doi.org/10.1021/ci00052a004> (1986).
49. 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. *Bioorg. Med. Chem. Lett.* **1**, 611–614. [https://doi.org/10.1016/S0960-894X\(01\)81162-8](https://doi.org/10.1016/S0960-894X(01)81162-8) (1991).
50. Park, C.-M. & Xian, M. In *Methods in Enzymology* Vol. 554 (eds Cadenas, E. & Packer, L.) 127–142 (Academic Press, 2015).
51. Cooper, P. In *Advances in Virus Research* Vol. 8 319–378 (Elsevier, 1962).
52. Dulbecco, R. & Vogt, M. In *Cold Spring Harbor Symposia on Quantitative Biology*. 273–279 (Cold Spring Harbor Laboratory Press).
53. Daelemans, D., Pauwels, R., De Clercq, E. & Pannecouque, C. A time-of-drug addition approach to target identification of antiviral compounds. *Nat. Protoc.* **6**, 925–933. <https://doi.org/10.1038/nprot.2011.330> (2011).
54. Li, Y. H., Lai, C. Y., Su, M. C., Cheng, J. C. & Chang, Y. S. Antiviral activity of *Portulaca oleracea* L. against influenza A viruses. *J. Ethnopharmacol.* **241**, 112013, <https://doi.org/10.1016/j.jep.2019.112013> (2019).
55. Schrödinger, L. *The PyMOL Molecular Graphics System, Version 1.8.* (2015).

Acknowledgements

This work was financially supported by China Medical University, Taiwan (Grant Numbers CMU106-S-49 and CMU108-S-34).

Author contributions

Y.F.L. and K.C.H. performed the virtual screening. C.Y.Y. and Y.Q.H. performed the bioassays. Y.J.C. analyzed the data, prepared the figures, and wrote the draft. C.H.L. and J.C.C. supervised and designed this study. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-87845-0>.

Correspondence and requests for materials should be addressed to C.-H.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021