POTENTIAL INFLUENZA A ANTIVIRAL NEURAMINIDASE INHIBITORS BY *IN SILICO* PHARMACOPHORE SCREENING

Massimo D. Bezoari†, Shelby Riedel*

Northwestern State University, Louisiana Scholars' College, Caspari Street, Morrison Hall, Natchitoches, LA

Abstract

Current influenza A antivirals such as Tamiflu® and Relenza® focus on inhibiting the neuraminidase (NA) enzyme but antiviral resistance mutations have been found in strains, such as the 2009 H1N1 strain. In this work, Sia(α2→6)GalMe and Sia(α2→3)GalMe–methyl derivatives of NA's host cell receptor targets which have Sia-Gal termini–were computationally docked to H1N1 and H5N1 NAs. Schrödinger's Phase program was used with the SiaGal derivatives to create various *hypotheses* that were employed to search the ZINC15 Substances database of approximately 53,000 compounds. Glide docking of the screened database gave ten compounds showing DockingScores that were strongly exergonic for the double mutated (H275Y/I223V) NA of H1N1. More than twenty compounds showed strongly exergonic Docking-Scores for the H5N1 avian flu strain without documented mutations, and more than twenty compounds were found for H5N1 with the H274Y antiviral resistant mutation. This work promotes the development of new antiviral therapies for flu strains that are showing resistance to current NA inhibitors as well as future mutations of H5N1 avian flu that may cause such flu strains to develop greater transmissibility and pandemics.

†corresponding author: bezoarim@nsula.edu.

Introduction

Seasonal influenza causes 291,000 to 646,000 deaths annually (1). More antiviral strategies are needed to mitigate future pandemics of flu strains that have gained resistance to common antivirals. This research used computational chemistry methods to verify the resistance mechanism of mutated flu strains to the common neuraminidase (NA) inhibitor, oseltamivir, and its active, acid form (AF). Sia(α2→6)Gal and Sia(α2→3)Gal are the terminal glycoprotein sites of NA hydrolysis. In this research, molecular modeling using methyl derivatives of these disaccharides as pharmacophores was used to screen the ZINC15 Substances database for compounds that might inhibit NA in mutated and non-mutated H1N1 and H5N1 strains of influenza A.

Influenza is an RNA virus from the family *Orthomyxovirus*, which consists of four genera, *Influenzavirus* A, B, C, or D, as well as the genus *Thogotovirus*. The genomes are small and change rapidly, allowing them to develop resistance to antiviral drugs and immune responses (2). Rapid mutations can result in different strains of a given genus–even with only a few amino acid residue substitutions in functional proteins–that are resistant to common antivirals.

All influenza viruses have the surface proteins hemagglutinin (HA), M2, and neuraminidase (NA), or other surface proteins that are structurally and functionally similar and are essential for the infection process (3). Influenza A is the most common seasonal flu virus strain and is of the greatest concern to public health; influenza B viruses have a limited range of hosts (humans and seals) and do not cause pandemics, possibly because they are less susceptible to antigenic variation (4), but they infect humans with regularity; influenza C viruses have been studied less as they do not cause severe disease in humans and are not deemed a public health threat (5); influenza D has received little attention as signs of acute infection in humans have not been identified.

Influenza A virus strains have 18 HA subtypes and 11 NA subtypes. Strain mutations form through antigenic drift from accumulation of mutations during replication, and antigenic shift in which HA or NA proteins in a strain are replaced by proteins from another strain, commonly, an avian one (6). New strains have the potential to cause pandemics (7, 8). Different HA and NA proteins between strains are indicated by numbered identifiers, giving influenza A strains such as H1N1, H3N2, etc. The 2009 pandemic strain was identified as a novel H1N1 strain, formed from a triple reassortment between the common human-infective H3N2 strain, a swine H1N1 strain, and an "avian-like" swine H1N1 strain. Despite the low mortality rate of that outbreak (0.15-0.25%), the appearance of antiviral resistance mutations is of great concern (9). Nearly all the 2009 pandemic strains were resistant to adamantane-like antivirals (Figure 1) which target the M2 virus protein (10), and many were resistant to the NA inhibitor, oseltamivir, due to mutations in NA. Other strains of influenza A have since appeared with resistance to common drugs including zanamivir and peramivir (9).

The surface glycoproteins of viral capsids promote binding to receptors of the host cell (11). Receptors consist of terminal sialic acids of host cell glycoproteins and glycolipids. Human-infecting influenza generally binds to Sia(α2→6)Gal (Figure 2) termini; avian-infecting influenza attaches to Sia(α2→3)Gal termini which are less common. Thus, H5N1, an influenza A strain of avian origin, generally infects the lower respiratory tract in humans whose respiratory tract cells contain more Sia(α2→3)Gal glycoproteins, many of whom are children (12).

Infection by influenza A begins with binding of HA to sialic acids of glycoproteins and glycolipids of the host cell followed by endocytosis. Various processes, which have been detailed elsewhere, result in viral protein synthesis and assembly (13). Although the mechanism of influenza assembly and budding is not well understood (14, 15), the viral structure must fuse and release from the host cell membrane. As virions bud out of the host cell, NA hydrolyzes cellular sialic acid residues bonded to viral HA,

releasing the virions to infect other cells (16).

Neuraminidase inhibitors (Table 1, Figure 1) block cleavage of new virions from host cell sialic acid receptors, preventing new virions from spreading and infecting other host cells.

Common NA inhibitors include oseltamivir (Tamiflu) and zanamivir (Relenza). Resistance to common NA inhibitors, especially oseltamivir, has become more common since the 2009 pandemic. Even though 99% of seasonal 2014-2015 strains remain susceptible to four recommended neuraminidase inhibitors, influenza A's rapid rate of mutation and propensity for antigenic shift make the spread of resistance genes a continuing possibility (17). Baloxavir marboxil (Xofluza®) (Figure 1), approved by the FDA in 2019, inhibits viral mRNA synthesis in the host cell and blocks viral proteins from forming. Other potential inhibitors also target the viral polymerase's subunits. Favipiravir is being studied as a potential treatment for COVID-19.

NA inhibitors are the most common type of antivirals because of the highly conserved nature of its active site (8). Adamantane derivatives are no longer used because rapid mutation of viral M2 have rendered them ineffective. The NA enzyme is a tetramer with four identical subunits. Although single subunits have no enzymatic activity, the active sites of each subunit seem to function independently. Each subunit is a polypeptide with four domains: in the cytoplasm there is an N-terminal cytoplasmic sequence of six amino acids, Met-Asn-Pro-Asn-Gln-Lys, the function of which is unknown; a transmembrane hydrophobic domain consisting of amino acid residues 7-29; a thin stalk on the outside of the membrane; and a globular domain that contains the active site. Even though the globular domain is highly conserved across strains, variations have produced two main groups of NA strains (18). Group 1 NA strains include N1, N4, N5, and N8, Group 2 includes N2, N3, N6, N7, and N9. Group 1 NAs show an open cavity adjacent to the active site that is not present in Group 2 strains (16). Different numbering systems for residues and mutations are used in the literature resulting in numbers that may differ slightly (usually, by 1). In the current work, the numbering used in original reports or PDB files were retained. The catalytic residues of the active site (using "N2 numbering") are Arg118, Asp151, Arg152, Arg224, Glu276, Arg292, Arg371, and Tyr406. Structural residues of the active site consist of Glu119, Arg156, Trp178, Ser179, Asp/Asn198, Ile222, Glu227, His274, Glu277, Asp293, and Glu425 (18). The adjacent cavity of Group 1 NAs is called the 150 cavity, consisting of residues 147-152. The 150 cavity remains in an open conformation which closes on binding of a target to the active site. Thus, the 150 cavity has become a target for development of antiviral compounds.

H1N1 and other strains with N1 neuraminidase have shown mutations that cause resistance, including E119V, I223R, H274Y, R292K, and N294S, which are within or adjacent to the active site. One of the most important mutations is H274Y (also numbered H275Y, which replaces His274–or His275–with Tyr) because it causes the greatest decrease in antiviral effectiveness. Another oseltamivir resistance mutation (I222V, or I223V) identified in
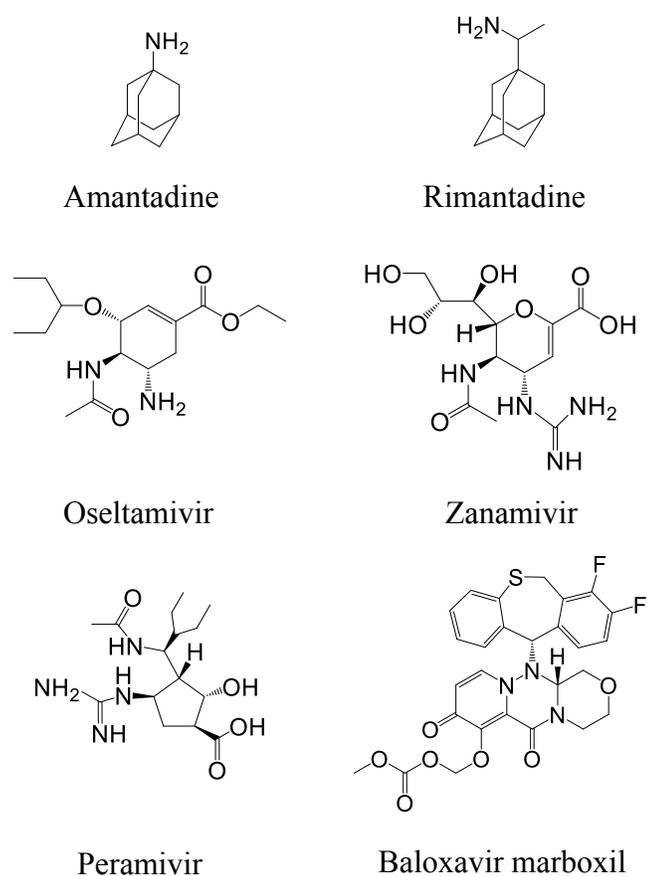

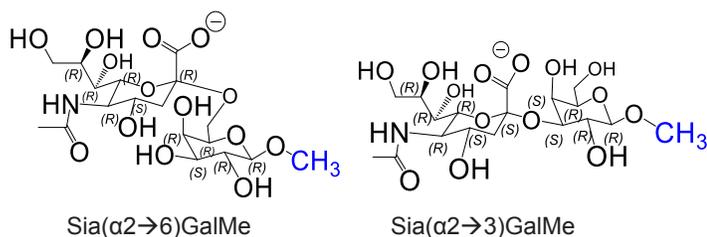**Figure 1. Influenza Antiviral Compounds.**

Amantadine · Rimantadine · Oseltamivir · Zanamivir · Peramivir · Baloxavir marboxil

Sia(α2→6)GalMe   Sia(α2→3)GalMe

Figure 2. Methyl Derivatives of Sialic Acid Structures.

Table 1. Available Influenza Antivirals.

Drug	Brand	Target	FDA Status
Amantadine	Symmetrel®	M2 protein	Approved, not recommended
Baloxavir[1]	Xofluza®	mRNA synthesis[1]	Approved, recommended
Laninamivir	Inavir®	NA protein	Not approved
Oseltamivir	Tamiflu®	NA protein	Approved, recommended
Peramivir	Rapivab®	NA protein	Approved, recommended
Rimantadine	Flumadine®	M2 protein	Approved, not recommended
Zanamivir	Relenza®	NA protein	Approved, recommended

1. Not a neuraminidase inhibitor–inhibits viral mRNA synthesis in the host cell.

H1N1 of 2009 replaces Ile222/223 with Val (19). Mutants with both H274Y and I222V have greater resistance to oseltamivir than either mutation alone (20).

H5N1 is an avian strain that is not common in humans and is not usually transmissive but is more lethal than most strains. It is being evaluated for its pandemic potential that may arise from antigenic shift (21). H5N1 may incorporate the H274Y mutation as well as N294S, in which Asn294 is replaced with Ser, that has similar oseltamivir resistance activity. H5N1 strains with antivirial resistance mutation E119A (in which Glu119 is replaced by Ala) are also of concern.

Computational Chemistry Strategy

In computational chemistry, force fields are used to calculate the stabilities of molecular structures. A force field includes the intermolecular forces within (and between) molecules using Newtonian mechanics for energy calculations. The various force fields assign different weighting factors to the contributions of given types of interaction to the overall energies being calculated or may even ignore certain interactions. Schrödinger programs currently use the OPLS3e force field, with a root mean square (RMS) error of less than 1 kcal/mol in predicting protein/ligand binding affinities that have been measured experimentally (22).

Receptor protein structures are typically obtained from crystallographic analyses and must be prepared and optimized. Preparations may involve removal of co-crystallized ligands, residual compounds from cell extracts, and salts or solvents that were part of the crystallization process. Candidate ligands must also be optimized to find lowest energy conformations (23). Crystallographic analyses and computational investigations that rely on them are subject to the possibility that the processes used to crystallize proteins and ligand-protein complexes may cause changes in their true structures. It is worth noting that PDB structure descriptions of proteins with more than one subunit may differ in the subunit chain labels; for example, a given enzyme's chain A may be named chain B in a different PDB file of the same enzyme.

Schrödinger's Glide program uses a scoring function to evaluate the most favorable (i.e., most exergonic) locations and orientations for ligand-protein binding (24). Thus,

$$\text{GlideScore} = 0.065\,E_{vdW} + 0.130\,E_{Coul} + E_{Lipo} + E_{Hbond} + E_{Metal} + E_{BuryP} + E_{RotB} + E_{Site}$$

in which E_{vdW} is the van der Waals energy, E_{Coul} is the Coulombic energy, E_{Lipo} rewards hydrophobic (London Forces) interactions, E_{Hbond} rewards hydrogen bonds, E_{Metal} is the metal binding term, E_{BuryP} penalizes internal polar groups in hydrophobic regions, E_{RotB} penalizes immobile rotatable bonds, and E_{Site} rewards polar interactions in the active site. The DockingScore calculation combines the GlideScore with Epik program state penalties. These penalties take account of ionizable groups and tautomers that contribute to the true 3-D isomeric structures of a compound within a specified pH range. GlideScores were employed (by default) in the current work for the initial database screening; DockingScores were used in more detailed analyses of complex formation by the best of the screened compounds. DockingScores provide relative estimates

of binding affinities of ligands to receptor enzymes, with several assumptions that relate to difficulties in accounting for solvation effects in the natural processes. The true binding energy (ΔG_{BE}) is given by

$$\Delta G_{BE} = G_{complex} - G_{protein} - G_{ligand}$$

The various terms include solvation effects, enthalpies, and entropies involved in the binding process. As the various compounds used in pharmacophore modeling are known to be similar in chemical and physical properties, such effects were assumed to be similar for their ligand-neuraminidase complexes. DockingScores were used to evaluate and rank the best poses. A "pose" may contain many complexes that are within the root mean square deviation (RMSD) parameters of the search. There is no immediate way to determine how many complexes are within a specific pose but energy differences between pose sets provides an indication of pose favorability. A simplified analysis using the relationship

$$\Delta G_{BE} = -RT\ln K$$

in which K is the ratio of one pose to a different one (for the same ligand and receptor), shows that a difference of 3 kcal/mol indicates a preference for the more exergonic structures within that pose of about 160-fold.

In evaluating where a ligand may bind to a protein receptor, the region of the protein that is searched is defined by a Gridbox, comprising a smaller inner box and a larger outside one. The inner box specifies the region in which a ligand's center point must lie while the entire ligand must lie within the boundary of the outside box. The receptor structure is assumed to be rigid, but a degree of induced fit can be included by scaling down the van der Waals radii that are calculated for atoms in the protein (24).

This research employed ligand-based virtual screening (LBVS). In this method, a ligand of known structure and activity (i.e., the pharmacophore) is used to search databases for compounds of similar 3D molecular properties that should have activity comparable to the known ligand (25). Schrödinger's Phase program conducts pharmacophore development and database screening. Pharmacophore structures have important physicochemical activities, such as hydrogen bond acceptor/donor atoms, hydrophobicity sites, negative or positive ionizable groups, aromatic rings, and other such features. The important features of the reference compound(s) can be assigned manually or in automated procedures. Phase generates multiple models with different combinations of features based on settings selected by the user. Each model is a *hypothesis* that is scored and ranked based on geometric and heuristic criteria giving a Phase Screen Score or other such evaluations. Initial *hypotheses* can be discarded, used individually, or merged. The final *hypothesis*–or set of *hypotheses*–is used to search databases to find compounds that match its features (26). Different pharmacophore models can be evaluated to identify the best one. Thus, the initial score is generated by using each model as a temporary reference for the others and evaluating the RMSDs for atom positions and the angles formed by corresponding pairs of sites for each. The one that best matches all the others is re-scored based on its similarity to the original reference compounds that were used. This *hypothesis*–there still may be more than one–is

then used for database screening (26, 27).

The question of how to evaluate a given *hypothesis* has been reviewed (27, 28). Essentially, the *hypothesis* is used to screen a database of compounds; the same *hypothesis* is then used to screen a different database of compounds that are physically but not chemically similar to known ligands (i.e., the "decoys") and which has been seeded with known ligands. Thus, the *hypothesis* can be evaluated by its ability to distinguish new ligands from a database relative to its ability to find known ligands in a database of decoys. One of the setbacks in using a database of decoy ligands consisting of known inhibitors (e.g., of neuraminidases) is that experimental evidence may indicate that these decoys inhibit the enzyme in vitro but have not shown success in therapeutic usage.

Prior to conducting the screening, each ligand of a database is converted to a 3D structure using LigPrep, to obtain different conformations that account for uncertainty about the specific conformation the ligand will adopt during binding to a receptor. The user decides how many conformers will be generated by specifying a relative energy window to filter the ligands and eliminate repeat structures with energies within 1 kcal/mol of each other and/or analogous pairs of atoms that are positionally identical (i.e., within a specified distance of each other).

Database screening may yield many compounds that match the *hypothesis*. Filtering of these can be carried out by selecting a percentage of compounds having the best Phase Screen Scores; these can be subjected to High Throughput Virtual Screening (HTVS), and again a percentage of these can be selected based on the HTVS docking method which uses approximations that allow faster analysis. When a reasonable number of potential ligands has been obtained, these can be subjected to regular Glide docking analyses to obtain more accurate DockingScores, best poses based on DockingScores, pose configurations, residue interactions, etc.

Molecular modelling research focused on examining the inhibition mechanisms of known inhibitors has been reported (29). A virtual screening study aimed at identifying new NA inhibitors differs from the present work in that structure-based virtual screening (SBVS), which is based on the receptor structure, was used (29); additionally, the potential inhibitor database employed was the National Cancer Institute (NCI) database (30). Sialic acid derivates DANA (2,3-didehydro-2-deoxy-N-acetylneuraminic acid) and FANA (2-deoxy-2,3-dehydro-N-trifluoroacetylneuraminic acid) have been used as transition state analogs of the Sia-Gal cellular receptors that mimic the natural target of neuraminidase (31, 32). However, DANA itself had limited inhibitory activity in pre-clinical testing and did not inhibit influenza in clinical testing (33). Molecular drug design was employed to identify two ligands that showed strong inhibitory activity for the H3N2 NA (34). A detailed analysis of known drugs at the active sites of NAs, based on docking studies with GOLD, has also been reported recently (35).

In this work, screened compounds were analyzed by dockings to neuraminidases and compared to results obtained by docking of the acid form of oseltamivir (i.e., oseltamivirAF). This acid form (Figure 3) is recognized as the actual active form of the inhibitor (36).

Materials and Methods

Ligand and Protein Structure Preparation

Methyl derivatives of Sia(α2→6)Gal, Sia(α2→3)Gal, and oseltamivirAF were optimized using Schrödinger's LigPrep with the following settings: Epik ligand states, pH of 7.0 +/- 1.0; desalt and generate tautomers; chiralities of ligands, from 3D structures (Figure 2); maximum number of conformers, 9. SiaGal derivatives and oseltamivirAF each returned 1 stable conformer.

The ZINC15 Substances database was downloaded as an .sdf file and uploaded into Maestro. Compound structures were optimized with LigPrep as described above, with chiralities from their 3D structures. This gave a Maestro output file that contained the prepared version of every compound in the database. A Phase database was created for all ZINC ligands and their conformers, with 50 possible conformers as the default option.

Neuraminidase protein structures were obtained from the Protein Data Bank: 3TI6, an NA protein with two subunits from H1N1 without documented resistance mutations, complexed with oseltamivirAF; 5NZF, a tetrameric NA protein from 2009 H1N1 with resistance mutations H275Y and I223V, complexed with oseltamivirAF; 2HTY, a structure combining 2 tetramers, non-mutated NA from H5N1 complexed with oseltamivirAF; 3CKZ, chain A of H5N1 neuraminidase with H274Y mutation, complexed with zanamivir; and 3CL0, an NA protein with one subunit from H5N1 with the H274Y resistance mutation, complexed with oseltamivirAF.

Ligands were deleted from the protein structures in Maestro. The Protein Prep Wizard selections were: remove all hydrogens in the downloaded structure and add all missing hydrogens; create zero-order bonds to metals and disulfides; fill in missing side chains using Prime; include aromatic hydrogens as hydrogen donors and halogens as acceptors in weak hydrogen bonding; delete waters that were more than 5 Å from het groups; generate het states using Epik with a pH of 7.0 +/- 1.0; optimize hydrogen bond assignments; assign protonation states at a pH of 7.0 using PROPKA; remove water molecules that contained fewer than three hydrogen bonds to non-water molecules; perform restrained minimization with any heavy atoms restricted to an RMSD of 0.30 Å. Optional orientations were found in the structure of 5NZF where some residues were found to have alternate positions–the default orientations were selected, as they comprise the majority of the possible orientations.

Pharmacophore Hypotheses

Hypotheses were created from optimized structures of Sia(α2→6)GalMe and Sia(α2→3)GalMe using Schrödinger
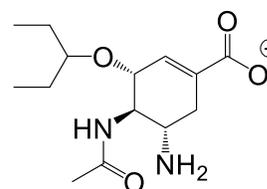

Figure 3. Oseltamivir Acid Form

Phase–Develop Pharmacophore Model. The selected settings were: Multiple ligand selection; Actives, 2–defined as the SiaGal structures; Pharmacophore method, Find best alignment and common features. *Hypothesis* should match at least 50% of actives, with 7 to 4 features including A (Acceptor), D (Donor), N (Negative Ionic); Generate conformers, target of 50.

Each *hypothesis* generated was used to screen the ZINC15 Substances database created by Phase. The settings were: Create New database; Skip duplicate ligands; Generate ligand conformers; Target number of conformers = 9, and Minimize output conformers; Prepare Ligand Structures; Generate possible states at pH = 7.0 ± 1.0 (with Epik); Determine chiralities from 3D structures; Retain at most 4 low-energy stereoisomers per ligand; Generate up to 2 low-energy 5- and 6-membered ring conformations; Remove high-energy ionization/tautomer states. The results were sorted by decreasing Phase Screen Score (PSS) with at most one hit per molecule. This gave many results that were filtered based on a PSS being within the top 25% of matches. A new Phase database was generated from the remaining compounds.

Docking Procedures with Virtual Screen Workflow

Gridboxes were created using van der Waals radius scaling at a factor of 1.0 with a partial charge cutoff of 0.25. In each case, the subunit complexed with the ligand was selected as the docking site. The specific docking coordinates were selected so that the inner box included the active site and the outer box included as much of the selected subunit of NA as possible; the inner box was defined to have the maximum possible dimensions of 40 Å x 40 Å x 40 Å; the outer box had 68-70 Å in each coordinate direction (Figure 4).

The database compounds obtained from screening the Zinc database with the hypotheses followed by retaining those with the top 25% PSS were then docked to NA structures with the following Virtual Screening Workflow selections: Input, phase database generated from PSS sorting; Prepare ligands, use Epik at pH 7.0 ± 1.0; Remove high energy ionization/tautomer states; Obtain stereochemical information = 3D geometry; Retain stereoisomers
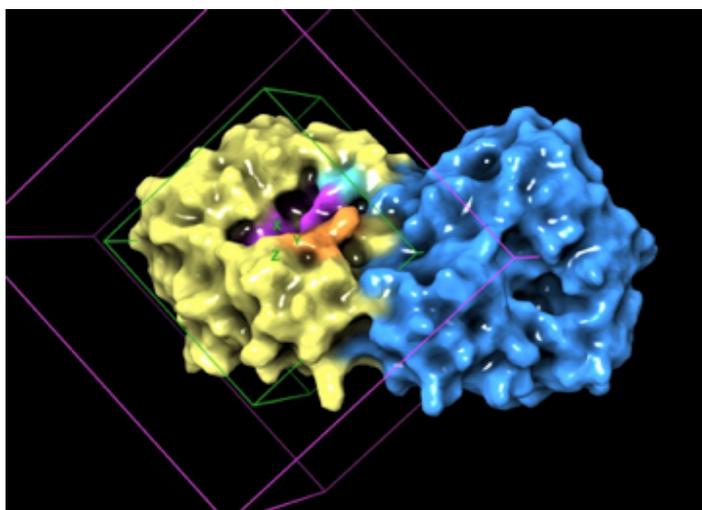
from unspecified stereocenters up to = 4; Generate low-energy ring conformations = 2; van der Waals radii of the ligands, scaled to a factor of 0.80. Dock with Glide HTVS: Dock flexibly; Perform post-docking minimization; Generate up to = 4 poses per compound state; After docking keep = 10% of best compounds; Retain = Only best scoring state. Dock with Glide SP: Dock flexibly; Perform post-docking minimization; Generate up to = 4 poses per compound state; After docking keep = 10 % of best compounds; Retain = Only best scoring state. The Complexation of ligands at the NA active sites was determined visually.

Results and Discussion

Hypotheses

Forty *hypotheses* (Figure 5) each containing 7 to 4 interactive features were obtained: 10 were AAADDDN (3 H-bond acceptors, 3 donors, and 1 negative ion), 8 AADDDN, 2 AAADDN, 4 AAADN, 3 AADDN; 3 ADDDN, 9 AAAN, 1 ADDN.

Neuraminidase Structures

The neuraminidase structures used in the present work were selected from available PDB structures. Thus, 3TI6 is a dimeric structure from 2009 H1N1 NA complexed with oseltamivir–this structure has missing residues 334, 337, 338, 393. Tetrameric structures are not always available, although they are preferred because they include surface cavities and ridges between subunits that might compete thermodynamically with the active site for complexation of ligands. PDB structures may include changes that occurred during the crystallization procedures such as chain realignments, chain breakage, and missing residues. The tetrameric NA from H1N1 with H275Y and I223V mutations, was obtained from 5NZF. This structure contained 2 sequence anomalies involving active site residues–R372E (replacing Arg372 with glutamate) and Y407V (replacing Tyr 407 with valine)–each showing that the replacement residues are located well outside the active site (Figure 6). One NA from H5N1 with no documented mutations was utilized, 2HTY, having 2 tetrameric units. Monomeric NAs from H5N1 were used, both having documented H274Y mutations: 3CKZ, complexed with zanamivir and showing chain breaks at
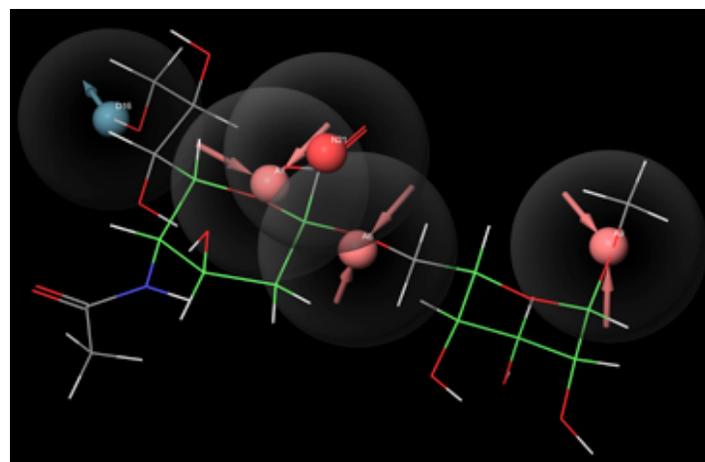

Figure 4. Neuraminidase 3TI6 with Gridbox. The inner box (green) encompasses most of subunit A (yellow). Colored residues: active site and lower boundary of the 150 cavity

Figure 5. A Representative AAADN Hypothesis. The structure shown is Sia(α2→6)GalMe; with pyranose rings shown in green. Salmon spheres, H-bond acceptor regions; pale blue spheres, H-bond donor regions; red spheres, negatively charged (carboxylate).

residues 134 and 345, and 3CL0, with chain breaks at residues 343 and 412.

Docking of Receptor-Like Pharmacophores, Sia(α2➔6)GalMe and Sia(α2➔3)GalMe vs. OseltamivirAF, to NAs

For the purposes of characterization, this paper uses "strong binding" for DockingScores in which the most favorable pose category of complexes is -6.5 kcal/mol or more exergonic; "moderate binding" denotes scores of -6.5 to -5.0 kcal/mol; and "weak binding" indicates scores that are more endergonic than -5.0 kcal/mol.

The LigPrep program returned 1 low-energy conformer each for Sia(α2➔6)GalMe and Sia(α2➔3)GalMe that were used for docking. Thus, moderate binding (most favorable DockingScores -5.738 to -5.701 kcal/mol) was observed for Sia(α2➔6)GalMe at the active sites of H1N1 NAs (non-mutated 3TI6 and double-mutated 5NZF); strong-moderate binding (DockingScores -7.715 to -5.205 kcal/mol) was seen for docking to H5N1 NA (non-mutated 2HTY, and single-mutated 3CKZ and 3CL0, Figures 7, 8). These results not only corroborate the known infectiousness of H1N1 and H5N1 but also show that antiviral-resistant mutations do not thermodynamically inhibit those abilities. Additionally, the results indicate that Sia(α2➔6)GalMe is an effective pharmacophore model for screening to find new antiviral compounds that will not be subject to known antiviral-resistant mutations.

Sia(α2➔3)GalMe demonstrated moderate binding at the active sites of H1N1 NAs (non-mutated 3TI6, and double-mutated 5NZF, -6.139 to -5.737 kcal/mol); and strong to weak binding to H5N1 NA (non-mutated 2HTY, and single-mutated 3CKZ and 3CL0, DockingScores of -8.249 to -5.045 kcal/mol). These observations corroborate the known abilities of H5N1 flu strains to infect human cells. Additionally, the docking to H5N1 NA with no documented mutations (2HTY) is strongly exergonic which would seem to corroborate a substantial thermodynamic preference for Sia(α2➔3)Gal receptors (which are common in children and in the

lower respiratory tract of adults). The use of Sia(α2➔3)GalMe as a pharmacophore model is supported.

OseltamivirAF showed strongly exergonic DockingScores for
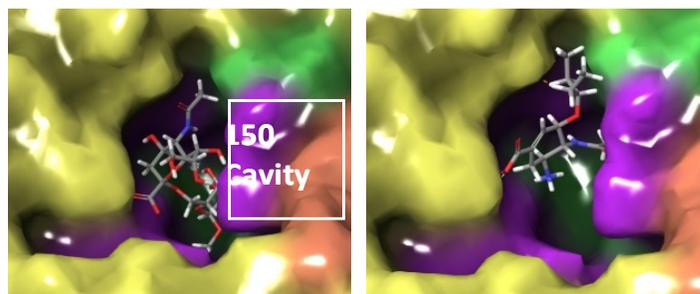
Figure 7. Docking at the active site of H5N1 (2HTY, no mutations). Left: Sia(α2➔6)GalMe. Right: oseltamivirAF. Purple: catalytic sites; green: structural sites; orange: lower boundary of 150 cavity.

DockingScores of Pharmacophore Sialic Acids and Oseltamivir to N1 Neuraminidases

Sia(α2→6)GalMe: -5.738, -5.701, -7.715, -5.205, -6.827

Sia(α2→3)GalMe: -6.139, -5.737, -8.249, -5.045, -5.800

Oseltamivir Acid Form: -8.109, -7.838, -7.724, -8.369, -8.657

DockingScore (kcal/mol)

■ H1N1, 3TI6: dimeric structure with no documented antiviral resistant mutations.
■ H1N1, 5NZF: tetrameric structure with H275Y and I223V antiviral resistant mutations.
■ H5N1, 2HTY: 2 tetrameric units, with no mutations
■ H5N1, 3 CKZ: chain A of H5N1 neuraminidase with H274Y and zanamivir. Chain breaks at 343 and 412
■ H5N1, 3CL0: monomeric, with H274Y antiviral resistant mutation. Chain breaks at 134, 345

Figure 8. Docking at the Active Sites of N1 Neuraminidases by Sialic Acid Derivatives and OseltamivirAF.
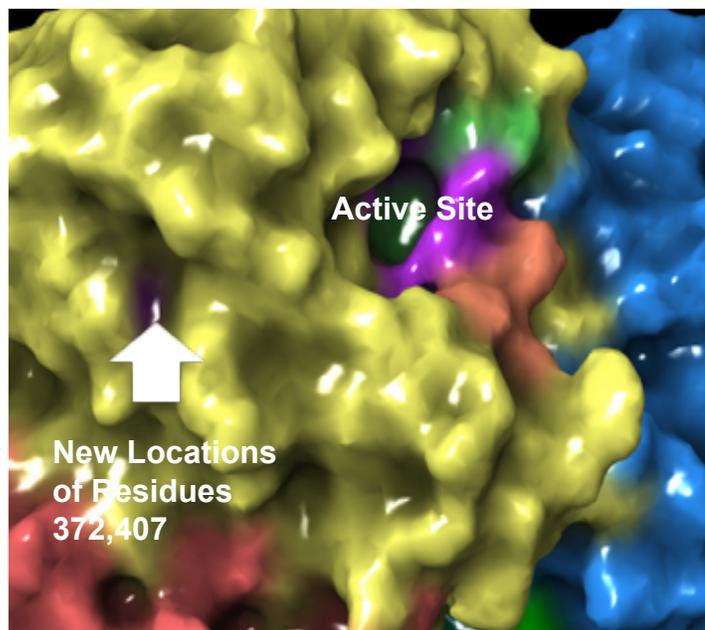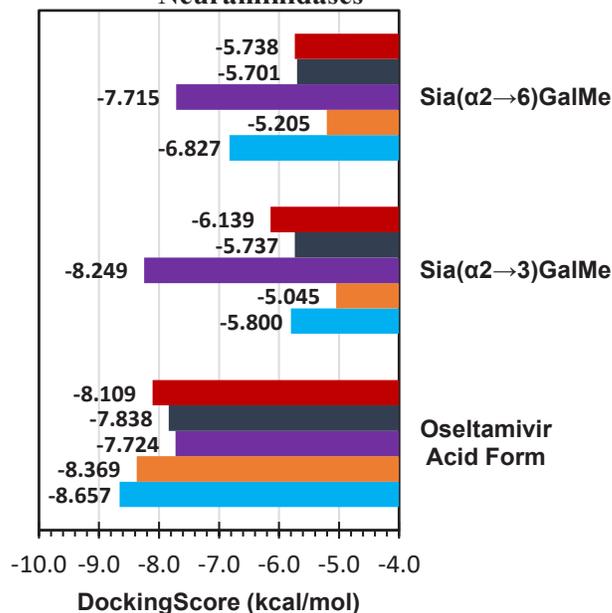

Figure 6. Sequence Anomalies of Active Site Residues R372E and Y407V of H1N1 (5NZF, with known H275& and I223V mutations). Purple: catalytic sites; green: structural sites; orange: lower boundary of 150 cavity.

non-mutated H1N1 (3TI6, -8.109 kcal/mol) and double-mutated H1N1 (5NZF, -7.838 kcal/mol). The expected result was that the inhibitor would show significantly weaker binding for the mutated NA. Considering the known decrease in kinetic inhibition of mutated NA by antivirals, it seems that the resistance mutations do not affect the favorable thermodynamics of complexation at the active site.

Binding of oseltamivirAF gave strong DockingScores for the non-mutated NA of H5N1 (2HTY, -7.724 kcal/mol). Strong DockingScores of -8.369 to -8.657 kcal/mol were also obtained for H5N1 NAs with H274Y mutations, 3CKZ and 3 CL0, respectively. The initial 3CL0 PDB structure is a complex with oseltamivir, whereas the 3CKZ structure is a complex with zanamivir. This docking result addresses the possibility that, after deletion of the complexed inhibitor from the PDB structure, the remaining pro-

tein structure might be predisposed structurally to binding with oseltamivirAF. DockingScores did not show any significant (i.e., greater than or equal to 1 kcal/mol) differences between the two protein structures. The current use of oseltamivir for therapeutic treatment of H5N1 flu is supported.

Screening of ZINC15 Substances Database

DockingScores of ZINC15 compounds that passed through the screening process are listed in Table 2.

One of the primary goals of this work was to promote the discovery of new NA inhibitors for H1N1 flu strains with NA antiviral resistance mutations. Ten ZINC15 compounds were obtained that showed strong Docking Scores (-7.43 to -6.61 kcal/mol, Table2) to double-mutated N1 (5NZF). Most of these were

Table 2. Active Site DockingScores (kcal/mol) of Screened ZINC15 Compounds (in bold) to H1N1 and H5N1 Neuraminidases.
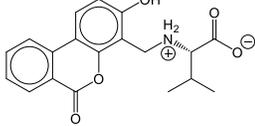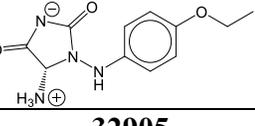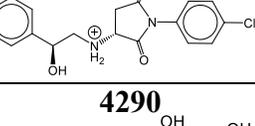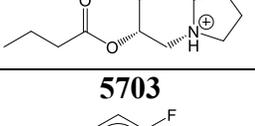
3TI6		5NZF		2HTY		3CKZ		3CL0	
34262	-7.90	44740	-7.43	28438	-8.74	44490	-8.50	44740	-8.22
44740	-7.57	33922	-7.29	36337	-8.61	44740	-8.03	44490	-7.89
36700	-7.21	32905	-7.17	31921	-8.47	5703	-7.63	5454	-7.82
5703	-7.14	4290	-7.08	38610	-8.46	36387	-7.51	42870	-7.71
843	-7.14	5703	-7.00	39683	-8.38	44660	-7.25	44082	-7.49
507	-7.10	1368	-6.98	31996	-8.35	31996	-7.25	34262	-7.49
39238	-7.08	785	-6.91	66	-8.28	1232	-7.20	25066	-7.33
33922	-7.05	24	-6.91	39856	-8.18	5454	-7.17	5266	-7.29
42870	-7.04	44490	-6.91	42675	-8.14	42870	-7.10	44660	-7.20
5454	-7.00	529	-6.61	2516	-8.11	35511	-7.07	44086	-7.20
42691	-6.92			34871	-8.09	32684	-7.00	856	-7.12
1368	-6.90			777	-8.07	44082	-6.99	5703	-7.12
				38719	-8.07	5266	-6.96	416	-7.06
				39817	-8.05	38229	-6.91	305	-7.03
				45826	-8.04	34370	-6.88	31996	-6.98
				35503	-8.02	32159	-6.87	1232	-6.88
				31928	-8.02	21406	-6.86	30342	-6.86
				31349	-8.00	35503	-6.83	38610	-6.85
				45658	-7.95	44114	-6.83	29745	-6.85
				7495	-7.90	38610	-6.77	41715	-6.81
				38601	-7.89	40263	-6.77	32684	-6.77
				36377	-7.89	25066	-6.76	38601	-6.76
				25104	-7.87	3241	-6.75	36739	-6.73
				40384	-7.86	5235	-6.75	35528	-6.73
				34858	-7.85	39317	-6.74	44114	-6.66
				36387	-7.84	44086	-6.74		
				35011	-7.82				

Yellow highlights: compounds common to 5NZF and 3TI6.
Turquoise: compounds common to 2HTY and either 3CKZ or CL0.

indistinguishable from oseltamivirAF binding favorability (-7.84 kcal/mol, Figure 8). As shown in Table 3, 4 of these also gave strongly favorable DockingScores (-7.57 to -6.98 kcal/mol) to non-mutated H1N1 (3TI6). Therefore, these 4 compounds might

provide effective leads towards therapeutic treatment of H1N1 strains that either have or do not have H275Y/I223V antiviral resistance mutations. The remaining 6 compounds might show preference for mutated H1N1 strains.

Regarding the possible future increases in the infectivity of H5N1 strains, docking to NAs without antiviral-resistant mutations (2HTY) gave 27 compounds with strong binding (DockingScores of -8.74 to -7.82 kcal/mol) with all favored poses at the NA active site. By comparison, the best DockingScore for oseltamivirAF to non-mutated NA of H5N1 (2HTY) was -7.72 kcal/mol but 5 of the 9 poses were for complexation at sites remote from the active site of the NA. Thus, the general trend for the screened compounds was that thermodynamic binding strength was better than or equivalent to oseltamivirAF.

About twenty-five compounds gave strongly exergonic DockingScores (-8.50 to -6.66 kcal/mol) at the active site of H5N1 with antiviral resistant mutation H274Y (3CKZ and 3CL0). As mentioned previously, the initial PDB structures of 3CKZ and 3CL0 differ in their complexation to zanamivir and oseltamivirAF. Therefore, there is a great deal of overlap in the compounds obtained from the screening process. However, it is worth noting that the screened compounds gave DockingScores that are comparable to the best DockingScore obtained by complexation of these NAs with oseltamivirAF (about -8.5 kcal/mol, Figure 8).

Comparing non-mutated (2HTY) to mutated H5N1 (3CKZ and 3CL0), ZINC 38610 showed strong DockingScores to all 3 NAs (-8.46 to -6.76 kcal/mol), 38601 showed strong docking to 2HTY (-7.89 kcal/mol) and 3CL0 (-6.76 kcal/mol), and 36387 demonstrated strong docking to 2HTY (-7.84 kcal/mol) and 3CKZ (-7.51 kcal/mol). Therefore, most compounds obtained exhibit preference for either the non-mutated NA (2HTY) or the mutated (3CKZ and 3CL0), and 3 compounds indicated possible inhibition of both the non-mutated and mutated NAs. These observations provide a range of possible therapeutic approaches to combinatorial and selective treatments of H5N1 flu strains.

Finally, it is worth noting that some compounds showed strongly exergonic DockingScores to both sets of flu strains, H1N1 and H5N1. These include 1368, 5454, 5703, 34262, 42870, 44490, 44740. Again, such overlaps may provide many different strategies to treat flu infections and avoid future pandemics.

Table 3. Active Site DockingScores (kcal/mol) of Screened ZINC15 Compounds to H1N1 NAs.

ZINC Ligand	5NZF[1]	3TI6[2]
44740	-7.43	-7.57
33922	-7.29	-7.05
32905	-7.17	
4290	-7.08	
5703	-7.00	-7.14
1368	-6.98	-6.90
758	-6.91	
24	-6.91	-
44490	-6.91	-
529	-6.91	

1. N1 with H275Y and I223V documented mutations and active site sequence anomalies.
2. N1 with no documented mutations.

Conclusions

This research showed that sialic acid disaccharide derivatives Sia(α2→6)GalMe and Sia(α2→3)GalMe are useful pharmacophore models for *in silico* searches for novel NA inhibitors of influenza strains. The thermodynamic binding of these host-cell receptor models was strong in H1N1 and H5N1 flu strains, as revealed by exergonic DockingScores. Antiviral NA mutations gave negligible thermodynamic effects on binding of these host cell receptor models.

The current work corroborates that the known diminished efficacy of oseltamivirAF (or oseltamivir) for the treatment of H1N1 flu strains with mutated NA is kinetic and does not agree with analyses that the reported effects are of a thermodynamic nature (37).

Ten compounds were found to be suitable for further testing against H1N1 with antiviral resistance mutations H275Y and I223V, based on thermodynamic DockingScores. Twenty-seven compounds were found to be suitable candidates for non-mutated H5N1. Three of these were in common with the twenty-five to twenty-six compounds found for mutated H5N1.

The 5NZF PDB structure for H275Y/I223V mutated NA appears to contain 2 sequence anomalies involving active site residues: R372E replaces Arg372 with glutamate and Y407V replaces Tyr407 (or Tyr406) with valine–both residues, which are normally within the active site are located well outside the active site. We have not found any previous reports relating to these.

The present work promotes the development of new NA inhibitors, in anticipation of possible future viral mutations that might result in serious pandemics. The rapid spread of highly transmissible and deadly flu strains has become a greater concern with recent global events relating to the SARS-CoV-2 virus and highlights the need to develop methods and compounds to combat future diseases and avoid pandemics.

Acknowledgements

References

1. Centers for Disease Control and Prevention (2017). Seasonal Flu Death Estimate Increases Worldwide. *CDC Newsroom*. Washington D.C. Public Health Foundation. https://www.cdc.gov/media/releases/2017/p1213-flu-death-estimate.html (accessed March 14, 2020).

2. Bouvier, N.M.; Palese, P. *Vaccine* [Online] **2008**, *26*, D49-D53. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3074182/pdf/nihms71066.pdf accessed March 14, 2020).

3. Couch, Robert B. Orthomyxoviruses. In *Medical Microbiology,* 4th edition; Baron S., Ed.; University of Texas Medical Branch at Galveston: Galveston, 1996; Chapter 58.

4. Chen, R.; Holmes, E. C. *J. Mol. Evol.* [Online] **2008**, *66*(6), 655-663. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3326418/ (accessed March 13, 2020).

5. Bailey, E. S.; Choi, J. Y.; Fieldhouse, J. K.; Borkenhagen, L. K.; Zemke, J.; Zhang, D.; Gray, G. C. Evol. Med. Public Health [Online] **2018**, *2018*(1), 192-198. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6128238/ (accessed March 13, 2020).

6. Centers for Disease Control and Prevention (2017). How the Flu Virus Can Change: "Drift" and "Shift." *Influenza*. Washington D.C. Public Health Foundation. https://www.cdc.gov/flu/about/viruses/change.htm (accessed May 4, 2019).

7. Treanor, J.; *N. Engl. J. Med.* [Online] **2004**, *350*, 218-220. https://www.nejm.org/doi/10.1056/NEJMp038238?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%3dwww.ncbi.nlm.nih.gov (accessed March 14, 2020).

8. Raja, K.; Prabahar, A.; Selvakumar, S.; Raja, T. K. *Indian J. Pharm. Sci.* [Online] **2014**, *76*(1), 10-18. https://www.ncbi.

nlm.nih.gov/pmc/articles/PMC4007250/ (accessed March 13, 2020).

9. Girard, M.P.; Tam, J.S.; Assossou, O. M.; Kieny, M. P. *Vaccine* [Online] **2010**, *28*(31), 4895-4902. https://www.ncbi.nlm.nih.gov/pubmed/20553769 (accessed March 13, 2020).

10. McKimm-Breschkin, J. L. *Influenza Other Respir. Viruses* **2013**, *7*(1), 25-36. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4942987/ (accessed March 13, 2020).

11. Varki, A.; Schauer, R. Chapter 14: Sialic Acids. In *Essentials of Glycobiology*; Varki, A., Cummings, R. D., Esko, J. D., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2009, 2nd edition.

12. Nicholls, J. M.; Bourne, A. J.; Chen, H.; Guan, Y.; Peiris, J. S. M. *Respir. Res.* [Online] **2007**, *8*(73), 1-10. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2169242/pdf/1465-9921-8-73.pdf (accessed March 13, 2020).

13. Manzoor, H.; Igarashi, M.; Takada, A. *Int. J. Mol. Sci.* [Online] **2017,** *18*, 2649. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5751251/pdf/ijms-18-02649.pdf (accessed Oct 12, 2019.

14. Nayak, D. P.; Hui, E. K.; Barman, S. *Virus Res.* [Online] **2004**, *106*(2), 147-165. https://www.sciencedirect.com/science/article/abs/pii/S0168170204003235 (accessed March 13, 2020).

15. Samji, T. *Yale J. Biol. Med.* [Online] **2009**, *82*(4), 153-159. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2794490/ (accessed March 13, 2020).

16. Air, G. M. *Influenza Other Respir. Viruses* [Online] **2012**, *6*(4), 245-256. https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1750-2659.2011.00304.x (accessed March 13, 2020).

17. Hurt, A. C.; Besselaar, T. G.; Daniels, R. S.; Ermetal, B.; Fry, A.; Gubareva, L.; Huang, W.; Lackenby, A.; Lee, R. T. C.; Lo, J.; et al. *Antiviral Res.* [Online] **2016**, *132*(2016), 178-185. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5357725/ (accessed March 13, 2020).

18. Hussain, M.; Galvin, H. D.; Haw, T. Y.; Nutsford, A.N.; Husain, M. *Infect. Drug Resist.* [Online] **2017**, *10*, 121-134. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5404498/ (accessed March 13, 2020).

19. Pizzorno, A.; Abed, Y.; Bouhy, X.; Beaulieu, E.; Mallett, C.; Russell, R.; Bovin, G. *Antimicrob. Agents Chemother.* [Online] **2012**, *56*(3), 1208-1214. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3294894/ (accessed March 13, 2020).

20. van der Vries, E.; Collins, P. J.; Vachieri, S. G.; Xiong, X.; Liu, J.; Walker, P. A.; Haire, L. F.; Hay, A. J.; Schutten, M.; Osterhaus, A. D. M. E.; Martin, S. R.; Boucher, C. A. B.; Skehel, J .J.; Gamblin, S. J. *PLoS Pathog.* [Online] **2012**, *8*(9), 1-8. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3447749/ (accessed March 13, 2020).

21. Govorkova, E. A.; Baranovich, T.; Seiler, P.; Armstrong, J.; Burnham, A.; Peiris, M.; Webby, R. J.; Webster, R. G. *Antiviral Res.* [Online] **2013**, *98*(2), 297-304. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3648604/ (accessed March 13, 2020).

22. Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wnag, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A. J. Chem. Theory Comput. [Online] **2016**, *12*(1) 281-296. https://pubs.acs.org/doi/pdf/10.1021/acs.jctc.5b00864 (accessed May 19, 2020).

23. Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. *J. Comput. Aided Mol. Des.* [Online] **2013**, *27*, 221-234. https://link.springer.com/arti-

cle/10.1007%2Fs10822-013-9644-8 (accessed March 13, 2020).

24. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T .A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* [Online] **2004**, *47*, 1739-1749. https://pubs.acs.org/doi/10.1021/jm0306430 (accessed March 13, 2020).

25. Shoichet, B. K. *Nature* [Online] **2004**, *432*, 862-865. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1360234/ (accessed March 13, 2020).

26. Dixon, S. L.; Smondyrev, A. M.; Knoll, E. H.; Rao, S. N.; Shaw, D. E.; Friesner, R. N. *J. Comput. Aided Mol. Des.* [Online] **2006,** *20*, 647-671. https://cyberleninka.org/article/n/1304769 (accessed March 13, 2020).

27. Gimeno, A.; Ojeda-Montes, M. J.; Tomás-Hernández, S.; Cereto-Massagué, A;, Beltrán-Debón, R; Mulero, M.; Pujadas, G.; Garcia-Vallvé, S. Int, J, Mol. Sci. [Online] 2019, 20(6), 1375-1398. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6470506/pdf/ijms-20-01375.pdf (accessed April 19, 2020).

28. Réau, M.; Langenfeld, F.; Zagury, J.-F.; Lagarde, N.; Montes, M.; Front Pharmacol. 2018, 9, 11. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5787549/ (accessed March 15, 2020).

29. Vavricka, C. J.; Li, Q.;Wu, Y; Qi, J.; Wang, M.; Liu, Y.; Gao, F.; Liu, J.; Feng, E.; He, J.; Wang, J.; Liu, H.; Jiang, H.; Gao, G. F. *PLOS Pathogens* **2011**, *7*(10), e1002249. https://journals.plos.org/plospathogens/article/file?id=10.1371/journal.ppat.1002249&type=printable (accessed March 15, 2020).

30. Hsu, K.-C.; Hung, H.-C.; HuangFu, W.-C.; Sung, T.-Y.; Lin, T. E.; Fang, M.-Y.; Chen, I.-J.; Pathak, N.; Hsu, J. T.-A.; Yang, J.-M. *Scientific Reports* **2017**, *7*, 12336 https://www.nature.com/articles/s41598-017-12101-3 (accessed February 22, 2020).

31. Greengard, O.; Poltoraskaia, N.; Leikina, E.; Zimmerberg, J.; Moscona, A. *J. Virol.* [Online] **2000**, *74*(23), 11108-11114. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC113191/ (accessed March 13, 2020).

32. Palese, P.; Compans, R. W. *J. Gen. Virol.* [Online] **1976**, *33*, 159-163. https://www.microbiologyresearch.org/docserver/fulltext/jgv/33/1/JV0330010159.pdf?expires=1584121825&id=id&accname=guest&checksum=32934D791FDAB647204A1F-4247BE2F3E (accessed March 13, 2020).

33. Wu, X.; Wu, X.; Sun, Q.; Zhang, C.; Yang, S.; Li, L.; Jia, Z. *Theranostics* [Online] **2017**, *7*(4), 826-845. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5381247/pdf/thnov07p0826.pdf (accessed March 13, 2020).

34. Ye, J.; Yang, X.; Xu, M.; Chan, P. K-s.; Cong, M.; *Eur. J. Med. Chem.* [Online] **2019**, *182*(111635). https://www.sciencedirect.com/science/article/abs/pii/S022352341930769X?via%3Dihub#abs0010 (accessed March 14, 2020).

35. Selvaraj, G. F.; Piramanayagam, S.; Devadasan, V.; Hassan, S.; Krishnasamy, K.; Srinivasan, S. *Informatics in Medicine Unlocked* [Online] **2020**, *18*(100284). https://reader.elsevier.com/reader/sd/pii/S2352914819303016?token=3529EFBB-D0E37EDC2DA07F878C93D8F3B0B726EE443763833B-877FEC9DCD96090211EE21E6D6728EE123036C21EBCA-CA (accessed March 14, 2020).

36. Amaro, R. E.; Li, W. W. *Methods Mol. Biol*. [Online] **2012**, *819*, 575-594. https://europepmc.org/article/PMC/3352029#FN21 (accessed March 13, 2020).

37. Pokorná, J.; Pachl, P.; Karlukova, E. ; Hejdánek, J.; Řezáčová,

P.; Machara, A.; Hudlický, J.; Konvalinka, J.; Kožíšek, M. *Viruses* [Online] **2018**, *10*(339), 1-15. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6071225/pdf/viruses-10-00339.pdf (accessed March 14, 2020).