J. Chem. Inf. Model., **2008**, *48* (10), pp 2074–2080

Design of Multi-Binding Site Inhibitors, Ligand Efficiency and Consensus Screening of Avian Influenza H5N1 Wild-Type Neuraminidase and of the Oseltamivir-Resistant H274Y Variant

Alfonso T. García-Sosa,[*,1] *Sulev Sild*[1] *and Uko Maran*[1]

[1] Institute of Chemistry, University of Tartu, Jakobi 2-319, Tartu 51014, Estonia.

atgs@cantab.net

RECEIVED DATE (to be automatically inserted after your manuscript is accepted if required according to the journal that you are submitting your paper to)

Tel: +372 737 5270; Fax: +372 735 5264.

Abstract The binding sites of wild-type avian influenza A H5N1 neuraminidase as well as of the Tamiflu (oseltamivir)-resistant H274Y variant were explored computationally in order to design inhibitors that target simultaneously several adjacent binding sites of the open conformation of the virus protein. The compounds with the best computed free energies of binding, in agreement by two docking methods, consensus scoring, and ligand efficiency values, suggest that mimicking a polysaccharide, β-lactam and other structures, including known drugs, could be routes for multi-binding site inhibitor

design. This new virtual screening method based on consensus scoring and ligand efficiency indices is introduced, which allows the combination of pharmacodynamic and pharmacokinetic properties into unique measures.

KEYWORDS Drug design, Antiviral agents, Influenza, H5N1, Ligand Efficiency, Efficiency Index, Drug resistant, High-performance computing, Grid computing, Virtual screening

Introduction

The avian influenza A strain H5N1 emerged in late 2003. It is a deadly disease in birds, in certain mammals, as well as in humans (236 reported deaths, 63% mortality rate).[1] Its continual global spread, as well as the not improbable event of a mutation that could confer the strain with an easier human-to-human transmission, have been described as possible factors for a human pandemic.[2] While the current treatment is focused on Tamiflu (oseltamivir, **1**)[3] and Relenza (zanamivir, **2**),[4] which were developed to treat the HxN2 group of human influenza virus, there are already reports of resistance.[5-6] The identification of oseltamivir-resistant strains of avian influenza H5N1, such as the mutant H274Y isolated from European samples, and the realization that proper counter strategies should include more than one inhibitor[7], has spurred the search for new drug candidates.

The crystal structure of the H5N1 neuraminidase (EC 3.2.1.18) has been made available,[8] as well as a structure from molecular dynamics (MD) simulations,[9] where the binding site of the protein shows additional adjacent open sites (the 150- and 430-loops) which are proposed to be amenable to drug design. Multi-binding site inhibitors may have properties and interactions which provide them with tighter affinity and better profiles than single-site inhibitors. The newly-disclosed oseltamivir-resistant H274Y neuraminidase crystal structure[7] has also been employed.

We have explored computationally these new sites with known inhibitors as well as with readily available chemicals in addition to existing, in-use drugs and natural products, and suggest compounds

with the capability of binding to several of these sites in order to develop a multi-binding site inhibitor. We also employed the calculation of molecular properties and ligand efficiency values for better characterizing the pharmacokinetic behaviour of the compounds. This is a new method for virtual screening, since it employs several protein structures in different conformations (including a drug-resistant mutant), consensus scoring, in addition to ligand efficiency indices to characterize and prioritize compounds for inhibiting a virus target.

Methods

Ligand efficiency indices can give an indication of the binding energy per heavy atom, per unit of molecular weight, or other unit, which can better identify drug candidates,[10-14] such as those having a ΔG/NHA deeper than -0.24 kcal/mol·NHA which has been described as a limit over which small molecules are able to disrupt protein-protein interfaces.[14] Molecular weight lower than 500 g/mol, and log*P* lower than 5,[15] or molecular weight between 160 and 480 g/mol, in addition to log*P* between -0.4 and 5.6, and number of heavy atoms between 20 and 70,[16] have also been proposed to be useful to predict the pharmacokinetic drug-likeness of a compound.

For the virtual screens, the libraries selected were the collection of known and approved small-molecule drugs obtained from the DrugBank,[17,18] and the N.C.I. diversity set.[19] In addition, a set of 65,000 compounds from the ZINC database[20] was screened with both Glide XP v4.5,[21] and Autodock v4.0.[22] The collection of compounds cover a diverse area of chemical space, as they are from a variety of chemical vendors, in addition to a representative set created in order to cover ample molecular structure (N.C.I. diversity set), as well as the large differences in structure observed in the drugs in use that were also included. The set of small-molecule approved drugs was treated with LigPrep v2.1[23] to generate all possible tautomers, ionization and protonation states in a target pH of 7.0 +/- 2.0, as well as generating a limited amount of stereoisomers. The total size of the libraries was 67,768 compounds. Four known inhibitors were also employed: oseltamivir (**1**), zanamivir (**2**), peramivir (**3**), and DANA (**4**, 2-deoxy-2,3-dehydro-N-acetylneuraminic acid), as well as dimers and trimers of these in order to link

several binding sites. All the Autodock calculations were carried out with the Chemomentum system,[24-26] which uses UNICORE 6 middleware[27,28] to provide a computer grid for distributed computing. This allowed achieving large parallelization of docking jobs by pooling available computer resources for distributed virtual screening.

For virtual screening, we removed from the collections of molecules those compounds containing metals, as well as salt counter-ions. Hydrogen atoms were added to the small molecule drug collection using Babel.[29] The National Cancer Institute (N.C.I.) diversity list and ZINC database already contain hydrogens. The 65,000 structures from the ZINC 7 database were selected based on the properties of the molecules, such as number of rotatable bonds between 5 and 10. However, it must be noticed that all the other sets were not constricted for number of rotatable bonds, and indeed, the compounds docked have between 3 and 23 or more rotatable bonds. The protein structures were protonated through Maestro.[30] A consensus score was generated by combining the Autodock and Glidescore values, taking the arithmetical average of both scores.

The program XLOGP v2.0[31] was used for calculating the octanol/water partition coefficient (log*P*).

Marvin Calculator Plug-ins were used for the calculation of ligand molecular formulae, molecular mass (MW), and Wiener index using Marvin 4.8.1, 2007, ChemAxon. (http://www.chemaxon.com)

Computational parameters: For Blind docking using Autodock v4.0, we used the following parameters: grid spacing = 0.55 Å, number of runs = 100, npts = 70 70 70, ga_num_evals = 20000000, ga_pop_size = 250, ga_num_generations = 27000.

For fine docking with Autodock v4.0, we used parameters: grid spacing = 0.375 Å, number of runs = 50, npts = 70 70 70, ga_num_evals = 20000000, ga_pop_size = 250, ga_num_generations = 27000.

For virtual screening with Glide XP v4.5, we used the default parameters included in the virtual screening workflow.[32]

For virtual screening with Autodock v4.0, we used the following parameters: grid spacing = 0.375 Å, number of runs = 1, npts = 70 70 70, ga_num_evals = 20000000, ga_pop_size = 250, ga_num_generations = 27000.

Results and Discussion

Exploration of binding sites.

First, we employed a technique called Blind docking[33,34] where sites on the surface of the protein are explored by ligands in order to identify favorable regions on the protein surface for binding. Using 4 known inhibitors of neuraminidase: oseltamivir (**1**), zanamivir (**2**), peramivir (**3**), and DANA (**4**, 2-deoxy-2,3-dehydro-N-acetylneuraminic acid), the protein surfaces of the 'open' structures of wild-type neuraminidase from crystallography (Protein DataBank code 2HU0) and from MD simulations,[9] as well as a structure in the 'closed' conformation (2HU4) were explored using 100 docking runs per ligand per protein using the program Autodock v4.0.[22] Several complexes were obtained with the inhibitors above occupying the cavities formed by the residues of the 150-loop, as well as at the entrance of the 430-loop cavity, and also some interacting with Arg156 deep inside the protein binding region (as well as reproducing their native crystal structures). The results showed that the lowest energies corresponded to the compounds binding with the open neuraminidase conformation (structures 2HU0, and from MD simulation) and that the known drugs and inhibitors could bind in several regions (binding sites) of the protein. The molecular weight, number of heavy atoms, and Wiener index ligand efficiency indices ($\Delta G/MW$, $\Delta G/NHA$, $\Delta G/Wiener$, respectively, where ΔG is the free energy of binding) were also computed.[10-14] A new ligand efficiency index that uses *P*, the partition coefficient between octanol and water, namely $\log(-\Delta G/P)$, was also calculated.

Construction of dimers and trimers.

Once the new binding sites were identified as well as the binding modes of the inhibitors in different regions of neuraminidase, molecules were constructed that could span the regions by linking all the combinations of X-CH_2-Y, where X, Y = **1**-**4** (the known drugs and inhibitors), as well as the homodimers X-CH_2-CH_2-X, and X-CH_2-CH_2-CH_2-X. The 'snowflake', symmetrical homotrimers $(X)_3$-CH and $(X\text{-}CH_2)_3$-CH were also generated. In order to mimic a disaccharide unit, all combinations of X-

O-Y were also built. These compounds were re-docked (50 runs) into the open protein conformations, with a finer grid separation of 0.375 Å. The results showed that the dimers in the open crystal structure (2HU0) had stronger calculated binding energies. A top ranking compound, oseltamivir(*R*)-CH_2-(*R*)oseltamivir (compound **5**) is shown in Table 1 as well as in Figure 1. It is predicted to form a complex with neuraminidase in the oseltamivir site as well as in the Arg156, 150- and early 430-loop sites. It makes hydrogen bonds with residues Arg118, Glu227, Glu276, Arg292, Arg371, Asp150 and Arg151 which are crucial to the binding of the natural substrate (sialoside), as well as with residues Ser246, Asn294, and Tyr406, and van der Waals interactions with more residues.

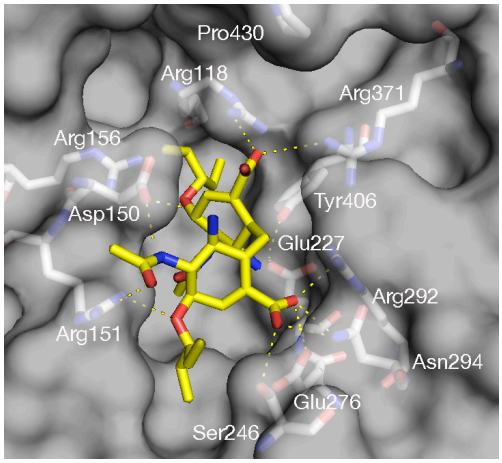
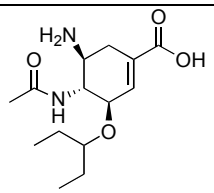
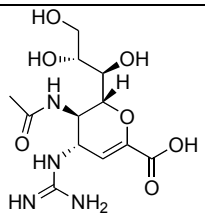
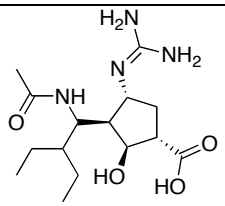
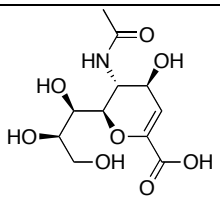
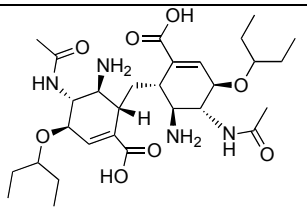
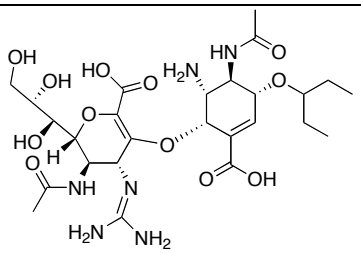
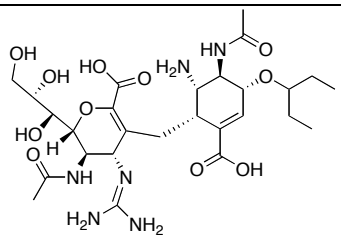
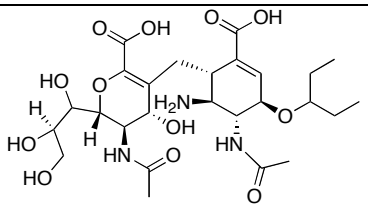


Figure 1. Compound **5**, oseltamivir(*R*)-CH_2-(*R*)oseltamivir, in complex with wild-type H5N1 neuraminidase in the open conformation.

The same calculations were performed with another docking program, Glide XP v4.5.[21] A consensus score was generated by combining both the Autodock Binding Energy, as well as the XP GlideScore energy, in order to have more confidence in the resulting score energy (A molecule would require both programs with their different methodologies and scoring functions, to agree in a high-ranking compound in able for it to score highly in the consensus score). Top ranking compounds were **5**, as well as oseltamivir(*R*)-O-zanamivir, **6,** compounds **7**, and **8** (see Table 1 and Table S1). 4 out of the top 10 binders were concurrently predicted by both methods. The constructed molecules present interesting stereochemistry as well as flexibility, which is characteristic of carbohydrates. Oseltamivir and zanamivir themselves were designed as sialic acid (the neuraminidase natural product) transition-state

mimics. This stereochemistry provides opportunities to develop new compounds by modifying the stereocenters. The flexibility may also help the molecules to adapt to viral protein residue mutations.

Table 1. Name, ID (in bold), and free energy of binding (kcal/mol) for two drugs for wild-type H5N1 neuraminidase (open crystal structure 2HU0) together with compounds predicted as inhibitors by their XP Glidescore, their Autodock Binding Energy (in italics) and consensus score (underlined). In brackets, MW (g/mol), ΔG/MW (kcal·g/mol^2), NHA, ΔG/NHA (kcal/mol·NHA), Wiener index, ΔG/Wiener (kcal/mol), log*P*, and log(-ΔG/*P*)(kcal/mol).[a]

oseltamivir **1** -5.47, -*9.04*, <u>-7.26</u> (284.35, -0.026, 20, -0.363, 801, -0.009, -0.21,1.071)		zanamivir **2** -6.99, -*9.02*, <u>-8.00</u> (332.31, -0.024, 23, -0.348, 1114, -0.007, -3.14, 4.043)	
peramivir **3** -6.28, -*8.23*, <u>-7.26</u> (328.41, -0.022, 23, -0.316, 1102, -0.007, 0.64, 0.221)		DANA **4** -5.09, -*8.59*, <u>-6.84</u> (291.26, -0.023, 20, -0.342, 773, -0.009, -2.7, 3.535)	
oseltamivir(*R*)-CH_2-(*R*)oseltamivir **5** -7.25, -*12.21*, <u>-9.73</u> (578.7, -0.017, 41, -0.237, 5424, -0.002, -0.53, 1.518)		oseltamivir(*R*)-O-zanamivir **6** -10.25, -*10.57*, <u>-10.41</u> (628.63, -0.017, 47, -0.221, 6262, -0.002, -3.71, 4.727)	
oseltamivir(*R*)-CH_2-zanamivir **7** -7.06, -*10.44* , <u>-8.75</u> (626.66, -0.014, 41,-0.199, 6262,-0.001, -4.01,4.952)		oseltamivir(*R*)-CH_2-DANA **8** -8.08, -*9.77* , <u>-8.92</u> (585.6, -0.015, 41, -0.218, 5354, -0.002, -3.56, 4.51)	

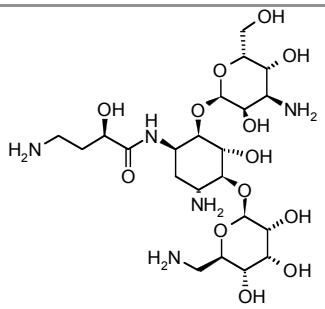
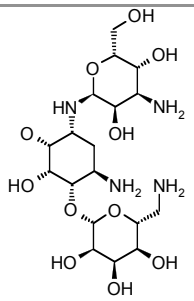
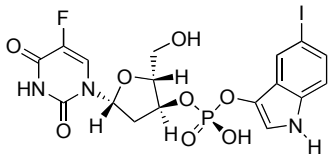
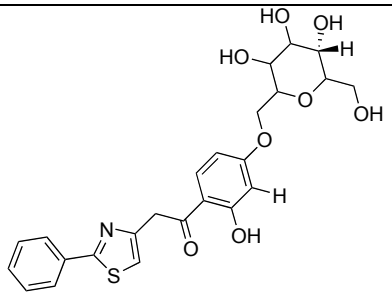
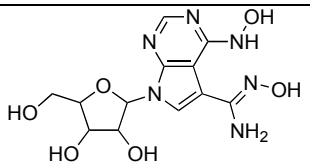
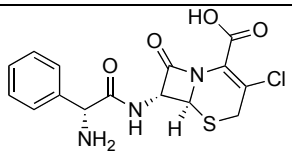
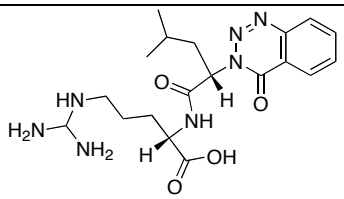
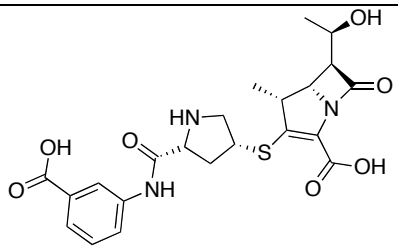
[a] MW = molecular weight, NHA = number of heavy atoms.

Screening on wild-type H5N1 neuraminidase.

The next step was to conduct virtual screening experiments on the open wild-type protein conformation (crystal structure 2HU0), targeting the new sites. These screenings would focus on determining consensus scoring energies for the free energy of binding, as well as ligand efficiency values for selecting the most promising ligands (see Table 2, and Table S2 in Supporting Information). Several interesting compounds were among the top ranking binders as they included stereoisomers of the naturally produced antibiotics amikacin (**9**), kanamycin (**10**), as well as ZINC04982827 (**11**) and ZINC03071734 (**12**), which similarly to the dimers constructed, have structures that resemble polysaccharides.

Several angiotensin converting enzyme (ACE) inhibitors, as well β-lactam antibiotics are among the top binders (Table 2). The top ranked compounds, such as **9-16,** are in Table 2 and Table S2. A stereoisomer of cefaclor (**14**) was predicted by both programs to be a top binder. The compounds NSC211332 and NSC70194 were selected as high binders (see Table S2), and have also been identified as potential inhibitors by a different study using the NCI database.[35] Both active drugs were recovered in the top 2% (top 20 molecules) of the small molecule drug screen with Glide XP, out of the total small molecule drugs of 1,053. Both active drugs were recovered in the top 5% (top 56 molecules) of the small molecule drug screen with Autodock.

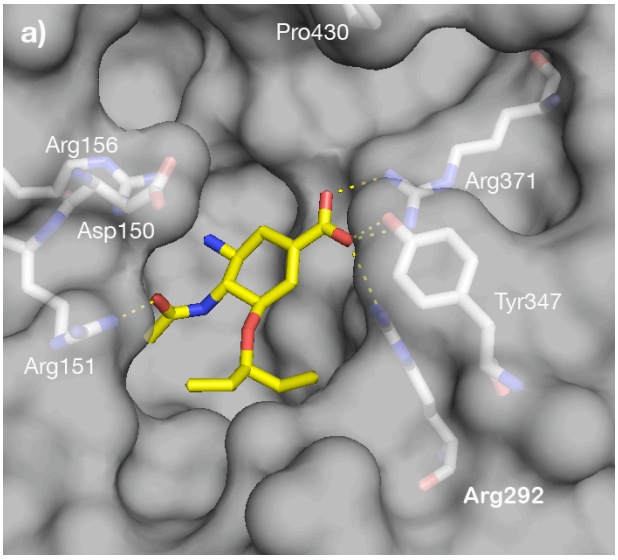
Table 2. Name, ID (in bold), and free energy of binding (kcal/mol) for wild-type H5N1 neuraminidase (open crystal structure 2HU0) of compounds predicted as inhibitors by their XP Glidescore, their Autodock Binding Energy (in italics) and consensus score (underlined). In brackets, MW (g/mol), ΔG/MW (kcal·g/mol^2), NHA, ΔG/NHA (kcal/mol·NHA), Wiener index, ΔG/Wiener (kcal/mol), log*P*, and log(-ΔG/*P*)(kcal/mol).[a]

amikacin **9** -9.63, *-11.05*, -10.34 (589.63, -0.018, 40, -0.259, 5207, -0.002, -7.63, 8.645)		kanamycin **10** -7.03, *-10.45*, -8.74 (488.53, -0.018, 32, -0.273, 3160, -0.003, -6.56, 7.502)	
ZINC04982827 **11** -10.79, *-9.28*, -10.04 (416.47, -0.018, 30, -0.324, 2595, -0.004, 0.11, 1.052)		ZINC03071734 **12** -10.31, *-9.00*, -9.66 (487.52, -0.02, 34, -0.284, 4151, -0.002, 1.57, -0.585)	
NSC154829 **13** -11.61, *-6.67*, -9.14 (340.29, -0.027, 24, -0.381, 1215, -0.008, -2.24, 3.201)		cefaclor **14** -8.39, *-10.30,* -9.34 (367.81, -0.025, 24, -0.389, 1383, -0.007, -0.11, 1.08)	
ZINC02196921 **15** -10.65, *-6.77*, -8.71 (417.46, -0.026, 30, -0.356, 2595, -0.004, -0.51, 1.539)		ertapenem **16** -11.29, *-11.37*, -11.33 (474.51, -0.024, 33, -0.343, 3464, -0.003, -1.13, 2.184)	

[a] MW = molecular weight, NHA = number of heavy atoms.

Figure 2a shows the complex of wild-type neuraminidase with oseltamivir. Another high-ranked compound, ZINC02196921 (**15**), is shown in Figure 2b, making hydrogen bonds with residues Glu119, Glu227, Arg292, Tyr347, Arg371, Tyr406 and van der Waals interactions with more residues. The 430-loop region can be seen in the top right of the picture (to the right of Phe430), and the 150-loop region is seen on the top- and middle-left. Most high-ranking compounds make interactions with residues considered critical for binding the natural product, such as making hydrogen bonds with Glu119, Glu227, Arg292, Arg371, as well as van der Waals and other interactions with more residues. Most of

the top-ranked compounds bind in the same mode as compound **15**, with at least one aromatic ring (often also fused rings) tucked well inside the 430-loop, the region extending past Pro430, at the top-right corner of Figure 2b, also interacting (including hydrogen bonds) with residues Arg371 and Tyr406, and nearly completely occluded from solvent. There is then a `tail` or flexible region in the compounds which extends into the original oseltamivir binding site (residues Tyr347, Arg292, Arg151) and includes polar groups which realize hydrogen bonds in this area, and many times includes also a decorated aromatic ring or a heterocycle. A large proportion of the ligands also posess a branching region, which grows from the beginning of the `tail` (close to Pro430) into the binding region behind the 150 loop (the region behind and above Asp150, at the top-left corner of Figure 2b) and which include groups from methyl, hydroxyl, isopropyl, and isobutane, engaged in primarily hydrophobic contacts with the residues in this third binding site.



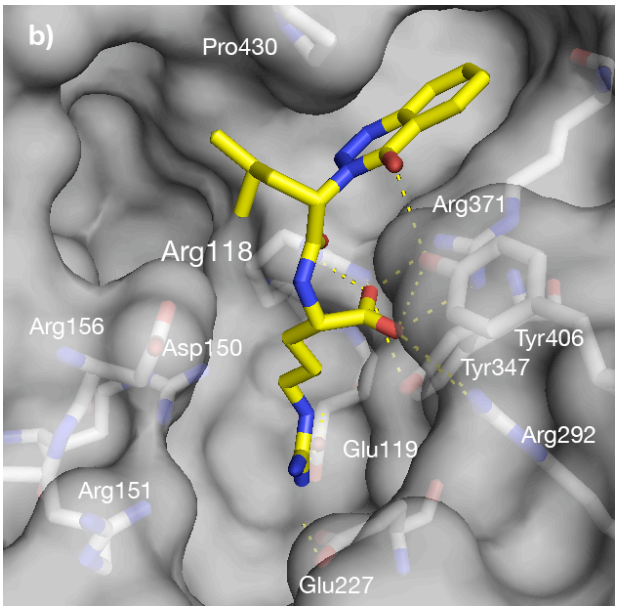


Figure 2. a) The crystal structure[8] of the complex of oseltamivir (Tamiflu, **1**) with wild-type avian influenza H5N1 neuraminidase. b) Docked compound **15** spanning multiple binding sites of wild-type H5N1 neuraminidase (structure 2HU0). Consensus score = -8.71 kcal/mol. Ligand efficiency per number of heavy atoms = -0.356 kcal/mol·NHA.

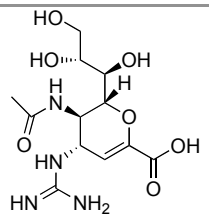
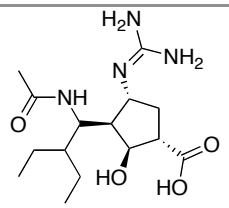
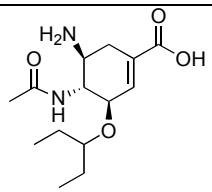
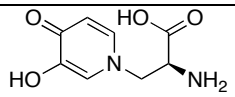
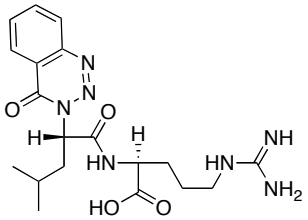
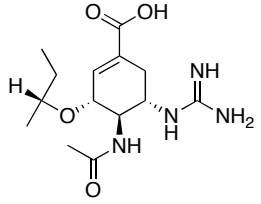
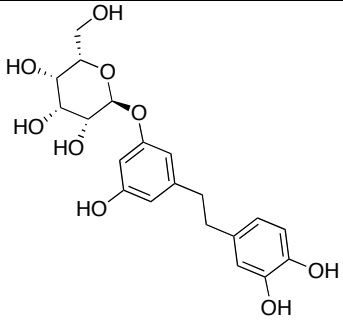
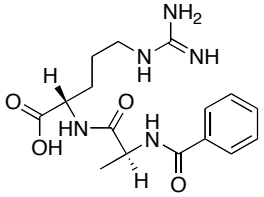
Drug-resistant mutant.

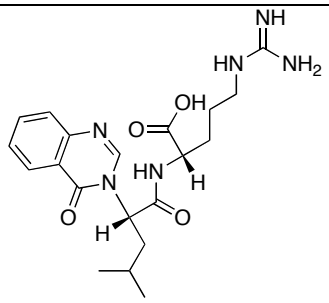
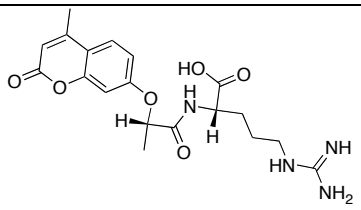
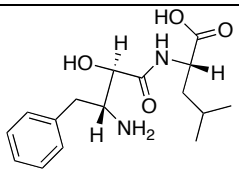
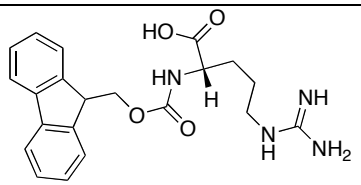
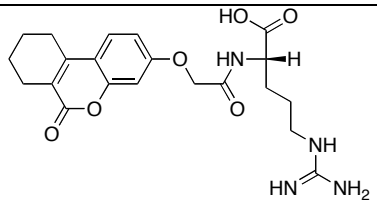
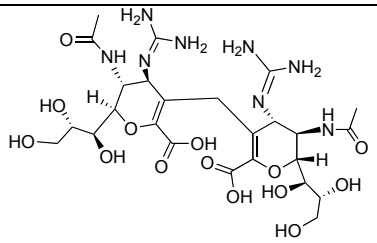
Recently, the crystal structure of a strain of oseltamivir-resistant H274Y (His264Tyr) of H5N1 avian influenza found in humans was disclosed.[7] The crystal structure of this enzyme was downloaded from the Protein Data Bank (code 3CKZ).[7] The same ligand efficiency and consensus scoring screenings were applied to this new structure.

Several interesting results were produced. Zanamivir and peramivir (but not oseltamivir) were the strongest binders of the known inhibitors, in agreement with experimental results. The small drug L-mimosine (**17**), a natural product used to block cell cycle progresion in breast cancer cells by chelating iron, was a strong binder with good properties. A number of coumarins (8 out of the top 89 with Glide) and indoles featured prominently among the top binders. In addition, a compound very similar to zanamivir, oseltamivir, and peramivir, the molecule ZINC04134496 (**18**) was among the strongest binders. Another top-binder, ZINC05811097 (**19**) also has a saccharide-like motif as identified in our

previous screens (see above). Importantly, there were also a number of compounds among the top binders against the H274Y strain which had also scored highly against the wild-type protein, such as **8**, **15** and **20**-**26**, shown in Table 3 (see also Table S3).

Table 3. Name, ID (in bold), and free energy of binding (kcal/mol) for oseltamivir-resistant H274Y variant of H5N1 neuraminidase together with compounds predicted as inhibitors by their XP Glidescore, their Autodock Binding Energy (in italics) and consensus score (underlined). In brackets, MW (g/mol), ΔG/MW (kcal·g/mol^2), NHA, ΔG/NHA (kcal/mol·NHA), Wiener index, ΔG/Wiener (kcal/mol), log*P*, and log(-ΔG/*P*)(kcal/mol).[a]

zanamivir **2** -9.44, *-9.41*, <u>-9.42</u> (332.31, -0.028, 21, -0.449, 1114, -0.008, -3.48, 4.454)		peramivir **3** -7.04, *-8.81*, <u>-7.92</u> (328.407, -0.024, 23, -0.345, 1102, -0.007, 0.64, 0.259)	
oseltamivir **1** -6.09, *-9.12*, <u>-7.60</u> (284.35, -0.027, 20, -0.380, 801, -0.009, -0.21, 1.091)		L-mimosine **17** -11.40, *-7.04*, <u>-9.22</u> (199.184, -0.046, 14 , -0.659, 321, -0.029, -2.21, 3.175)	
ZINC02196921 **15** -10.76, *-8.91*, <u>-9.84</u> (417.462, -0.024, 30, -0.328, 2595, -0.004, -0.51,1.503)		ZINC04134496 **18** -9.42, *-8.66*, <u>-9.04</u> (312.365, -0.029, 22, -0.411, 1017, -0.009, -0.76, 1.716)	
ZINC05811097 **19** -9.12, *-8.39*, <u>-8.76</u> (408.399, -0.021, 29, -0.302, 2478, -0.004, 0.59, 0.352)		ZINC02560874 **20** -11.35, *-8.61*, <u>-9.98</u> (349.385, -0.029, 25, -0.399, 1780, -0.006, -0.19, 1.189)	

ZINC05789694 **21** -10.97, *-9.67*, -10.32 (416.474, -0.025, 30, -0.344, 2595, -0.004, 0.11, 0.904)		ZINC06008495 **22** -10.65, *-9.12*, -9.88 (404.417, -0.024, 29, -0.341, 2634, -0.004, 0.42, 0.575)	
ZINC02545165 **23** -10.03, *-10.03*, -10.03 (308.373, -0.033, 22, -0.456, 1172, -0.009, 0.77, 0.231)		ZINC02510125 **24** -8.84, *-9.87*, -9.36 (396.44, -0.024, 29, -0.323, 2537, -0.004, 1.84, -0.869)	
ZINC02091879 **25** -8.33, *-9.23*, -8.78 (430.454, -0.020, 31, -0.283, 3226, -0.003, 0.63, 0.313)		zanamivir-CH_2-zanamivir **26** -8.62, *-7.95*, -8.23 (676.631, -0.012, 47, -0.176, 7172, -0.001, -8.07, 8.988)	

[a] MW = molecular weight, NHA = number of heavy atoms.

From the results in the tables, it can be seen that there are a variety of compounds with good predicted binding free energy. Also, it can be seen that for the wild-type neuraminidase, the drugs in use, oseltamivir and zanamivir have good (deep) ligand efficiency values such as ΔG/NHA values deeper than -0.24 kcal/mol·NHA: namely for the wild-type neuraminidase, -0.363 kcal/mol·NHA for oseltamivir and -0.348 kcal/mol·NHA for zanamivir. Their strong binding energy per unit of heavy atom may explain why they are efficient drugs against the wild-type virus. There are also a number of predicted compounds that possess good ligand efficiency values and molecular properties, such as **11**, **13-16** and others (see Table S2), which would be interesting for inhibiting the wild-type neuraminidase.

For the oseltamivir-resistant H274Y variant, zanamivir shows a very strong ligand efficiency of -0.449 kcal/mol·NHA which may explain why this variant is still susceptible to this drug. Also interesting are compounds **15, 17, 18, 20-24** (and others, see Table S3) which have good molecular properties and ligand efficiency values against the drug-resistant variant. L-mimosine, compound **17**, is

a drug which is already in use against breast cancer and could find a new use against influenza. This compound is small and binds in the same region as zanamivir. Most compounds, if small, bind in this area. The larger compounds bind in a manner similar to those presented in Table 2 (screening against wild-type protein), with strong binding energies. However, in contrast to the wild-type protein case, the oseltamivir-resistant mutant has the 150-loop in a very narrow, almost closed conformation, and so the high-ranked compounds interact mostly with neuraminidase in the oseltamivir binding site and in the 430-loop, although there are compounds with a methyl, hydroxyl, amine or amide group at the opening of the 150-loop (in the case of a hydroxyl, amine or amide group, making hydrogen bonds with Val148 and Asp150).

There exist similarities in some of the compounds in Table 2 and Table 3, namely, the presence of an aromatic or cyclic group, with added functionality such as hydroxy and amine groups which resemble a saccharide unit (see for example compounds **1-4**, **9-13**, **18**, **19**, **26**). However, the β-lactam structures that appeared during screening against the wild-type protein (for example structures **14** and **16**), no longer are privileged structures in the screening against the drug-resistant mutant protein.

Particularly interesting are compounds **15, 20-24** which have good molecular properties and ligand efficiency values for inhibiting both wild-type and oseltamivir-resistant H274Y neuraminidases, since they may be able to inhibit either one or both forms of the virus protein. It is worthy to note that these last compounds have a common substructure: a central amide region, in addition to an aromatic group of 1-3 fused rings on one side of the amide, and (excluding compound **23**) on the other side an aliphatic chain of 3 carbons length plus a terminal guanidinium group.

The results include compounds from all of the initial sets and therefore, the diversity in the compound sets is also present in the top ranked compounds.

Conclusions

Certain structures appear repeatedly among the most interesting ligands, such as polysaccharide-like molecules, β-lactam structures, carboxy-quinazolines or carboxy-cinnolines, coumarins, and others,

which can be pursued as new directions for treatments against H5N1 influenza. These molecules could provide additional drugs necessary for the now-recognized[7] multi-drug strategy required for effective treatment against avian influenza H5N1.

We have shown that the method of including several protein structures with several binding sites, as well as considering consensus scoring and ligand efficiency indices to characterize and select promising compounds for new inhibitors, works well, provides increased confidence and important information for distinguishing those compounds with the best free energy of binding (pharmacodynamic properties) at the same time as distinguishing those with the best pharmacokinetic properties, such as molecular weight, number of heavy atoms, or log*P*, which can ultimately be related to phenomena such as permeability, absorption, distribution, metabolism, excretion, toxicity as well as others. New virtual screening procedures using different molecular properties can be developed in a manner similar to that presented in this work.

Our results show that the design of inhibitors targeting simultaneously several binding sites of wild-type and drug-resistant avian influenza H5N1 neuraminidase has good potential for developing new and potent drugs against the disease. The lowest binding energy dimer and trimer compounds in agreement by both docking methods, as well as those compounds obtained through virtual screening, and especially those with good ligand efficiency values and molecular properties, can be tested for inhibition of the avian influenza A H5N1 wild-type as well as the drug-resistant H274Y strain. Since they are proposed to bind specifically to the (neuraminidase 1) N1 form of the disease they might be useful for targeting this more dangerous variety of influenza if a larger outbreak in humans occurs. In order to combat a possible human pandemic, information should be revealed as soon as it is available, which may allow other scientists to further possible leads that can serve to provide new therapeutics in a fast manner.

ACKNOWLEDGMENT We thank the Estonian Science Foundation grant JD80 and the EU FP6

Chemomentum project IST-033437.

Supporting Information Available. Full Tables S1-S3 with compound names, consensus scores, ligand efficiencies, and molecular structures for the drugs, monomers, dimers, trimers and other compounds with the wild-type and oseltamivir-resistant H274Y mutant avian influenza neuraminidase. This material is available free of charge via the Internet at http://pubs.acs.org .

REFERENCES

1) World Health Organization. Cumulative Number of Confirmed Human Cases of Avian Influenza A/(H5N1) Reported to WHO; 18 Mar 2008, **2008.**

2) World Health Organization. Influenza A (H5N1) in Hong Kong Special Administrative Region of China; 19 Feb 2003, **2003**.

3) Kim, C.U.; Lew, W.; Williams, M.A.; Liu, H.T.; Zhang, L.J.; Swaminathan, S.; Bischofberger, N.; Chen, M.S.; Mendel, D.B.; Tai, C.Y.; Laver, W.G.; Stevens, R.C. Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: Design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. *J. Am. Chem. Soc.* **1997**, *119*, 681-690.

4) von Itzstein, M.; Wu, W.Y.; Kok, G.B.; Pegg, M.S.; Dyason, J.C.; Jun, B.; Van Phan, T.; Smythe, M.L.; White, H.F.; Oliver, S.W.; Colma, P.M.; Varghese, J.N.; Ryan, D.M.; Woods, J.M.; Bethell, R.C.; Hotham, V.J.; Cameron, J.M.; Penn, C.R. Rational design of potent sialidase-based inhibitors of influenza-virus replication. *Nature* **1993**, *363*, 418-423.

5) Gubareva, L.V.; Kaiser, L.; Matrosovich, M.N.; Soo-Hoo, Y.; Hayden, F.G. Selection of influenza virus mutants in experimentally infected volunteers treated with oseltamivir *J. Infect. Dis.* **2001**, *183*, 523-531.

6) European Union. European Center for Disease Prevention and Control. Resistance to oseltamivir

(Tamiflu) found in some European Influenza samples; http://ecdc.europa.eu (accessed Apr 16, 2008).

7) Collins, P.J.; Haire, L.F.; Lin, Y.P.; Liu, J.; Russell, R.J.; Walker, P.A.; Skehel, J.J.; Martin, S.R., Hay, A.J.; Gamblin, S.J. Crystal structures of oseltamivir-resistant influenza virus neuraminidase mutants. *Nature* **2008**, *453*, 1258-1261.

8) Russell, R.J.; Haire, L.F.; Stevens, D.J.; Collins, P.J.; Lin, Y.P.; Blackburn, G.M.; Hay, A.J.; Gamblin, S.J.; Skehel, J.J. The structure of H5N1 avian influenza neuraminidase suggests new opportunities for drug design. *Nature* **2006**, *443* (7107), 45-49.

9) Amaro, R.E.; Minh, D.D.L.; Cheng, L.S.; Lindstrom, W.M.; Olson, A.J.; Lin, J.-H.; McCammon, J.A. Remarkable loop flexibility in avian influenza N1 and its implications for antiviral drug design *J. Am. Chem. Soc*. **2007**, *129*, 7764-7765.

10) Kuntz, I.D.; Chen, K.; Sharp, K.A.; Kollman, P.A. The maximal affinity of ligands. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9997-10002.

11) Hopkins, A.L.; Groom, C.R.; Alex, A. Ligand efficiency; a useful metric for lead system. *Drug Discovery Today* **2004**, *9*, 430-431.

12) Abad-Zapatero, C.; Metz, J.T. Ligand efficiency indices as guideposts for drug discovery. *Drug Discovery Today* **2005**, *10*, 464-469.

13) Hetényi, C.; Maran, U.; García-Sosa, A.T.; Karelson, M. Structure-based calculation of drug efficiency indices. *Bioinformatics* **2007**, *23*, 2678-2685.

14) Wells, J.A.; McClendon, C.L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **2007**, *450*, 1001-1009.

15) Lipinski, C.A.; Lombardo, F.; Dominy, B.W.; Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv*.

Drug Del. Rev. **1997**, *23*, 3-25.

16) Ghose, A.K.; Vellarkad, V.N.; Wendoloski, J.J. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *J. Comb. Chem*. **1998**, *1*, 55-68.

17) University of Alberta, Canada. DrugBank www.drugbank.ca (accessed Mar 16, 2008).

18) Wishart, D.S.; Knox, C.; Guo, A.C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J. DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res*. **2006**, *34*, D668-D672 Sp. Iss. SI.

19) National Cancer Institute/National Institutes of Health USA. DTP – Diversity Set Information. http://dtp.nci.nih.gov/branches/dscb/diversity_explanation.html (accessed Mar 16, 2008).

20) Irwin, J.J.; Shoichet, B.K. ZINC - A free database of commercially available compounds for virtual screening . *J. Chem. Inf. Model*. **2005**, *45*, 177-182.

21) *Glide* v.4.5, Schrödinger, LLC, New York, NY, **2007**.

22) Morris, G.M.; Goodsell, D.S.; Halliday, R.S.; Huey, R.; Hart, W.E.; Belew, R.K.; Olson, A.J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem*. **1998**, *19*, 1639-1662.

23) *LigPrep* v.2.1, Schrödinger, LLC, New York, NY, **2007**.

24) Chemomentum consortium. Chemomentum. http://www.chemomentum.org . (accessed Jan 15, 2008).

25) Maran, U.; Sild, S.; Mazzatorta, P.; Casalegno, M.; Benfenati, E.; Romberg, M. Grid computing for the estimation of toxicity: acute toxicity on fathead minnow (Pimephales promelas), *In Distributed,*

High-Performance and Grid Computing in Computational Biology, Dubitsky, W.; Schuster, A.; Sloot, P.M.A.; Schroeder, M.; Romberg, M., Eds.; Springer-Verlag; Berlin Heidelberg, LNCS 4360, 2007.

26) Schuller, B.; Demuth, B.; Mix, H.; Rasch, K.; Romberg, M.; Sild, S.; Maran, U.; Bała, P.; del Grosso, E.; Casalegno, M.; Piclin, N.; Pintore, M.; Sudholt, W.; Baldridge K. K. Chemomentum - UNICORE 6 based infrastructure for complex applications in science and technology. *In Theoretical Computer Science and General Issues*, Bougé, L.; Forsell, M.; Larsson Träff, J.; Streit, A.; Ziegler, W.; Alexander, M.; Childs, S., Eds.; Springer-Verlag: Berlin Heidelberg, LNCS 4854, **2008**.

27) Distributed Systems and Grid Computing, Juelich Supercomputing Centre, Institute for Advanced Simulation, Research Centre Juelich, Germany. UNICORE – Distributed computing and data resources. http://www.unicore.eu (accessed Jan 15, 2008).

28) Almond, J.; Snelling, D. UNICORE: uniform access to supercomputing as an element of electronic commerce. *Future Generation Computer Systems* **1999**, *613*, 539-548.

29) Open Babel. http://openbabel.org (accessed 15 Mar, 2007).

30) *Maestro* v8.0, Schrödinger, LLC, New York, NY, **2007**.

31) Wang, R.; Gao, Y.; Lai, L. Calculating partition coefficient by atom-additive method. *Perspectives in Drug Discovery and Design* **2000**, *19*, 47-66.

32) *Virtual Screening Workflow*, Schrödinger, LLC, New York, NY, **2007**.

33) Hetényi, C.; van der Spoel, D. Efficient docking of peptides to proteins without prior knowledge of the binding site. *Protein Sci.* **2002**, *11*, 1729-2737.

34) Hetényi, C.; van der Spoel, D. Blind docking of drug-sized compounds to proteins with up to a thousand residues. *FEBS Lett.* **2006**, *580*, 1447-1450.

35) Cheng, L.S.; Amaro, R.E.; Xu, D.; Li, W.W.; Arzberger, P.W.; McCammon, A. Ensemble-based virtual screening reveals potential novel anti-viral compounds for avian influenza neuraminidase. J. Med. Chem. published online, doi: 10.1021/jm8001197

SYNOPSIS TOC