Binding Mechanism of H5N1 Influenza Virus Neuraminidase with Ligands and its Implication for Drug Design

Ke Gong[1,§], Lin Li[1,§], Jing-Fang Wang[2], Feng Cheng[3], Dong-Qing Wei[1,4,*] and Kuo-Chen Chou[4]

[1]*College of Life Science and Biotechnology, Shanghai Jiaotong University, Shanghai 200240, China;* [2]*Bioinformation Center, Key Lab of System Biology, Shanghai Institutes for Biological Sciences, Graduate School of the Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China;* [3]*Department of Biophysics, University of Virginia, Charlottesville, VA, USA;* [4]*Gordon Life Science Institute, 13784 Torrey Del Mar Drive, San Diego, California 92130, USA*

Abstract: To simulate new strategies for designing effective drugs against bird flu, we have carried out extensive studies by using various computer-aided drug design tools. Molecule AG7088 was first docked to the active site of H5N1 avian influenza neuraminidase (PBD code: 2HTY). The results thus obtained were compared with those by docking zanamivir (Relenza) and oseltamivir (Tamiflu) to the same receptor, respectively. It has been found that the compound AG7088 has better binding energy than zanamivir and oseltamivir. Thus, it was adopted as a template to perform the similarity search of 392,698 druggable compounds in order to find the leading candidates for the next step of modeling studies. Nine analogs of AG7088 were singled out through a series of docking studies. Finally, the molecular dynamics simulation technique was utilized to investigate into the binding interactions between the H5N1 receptor and the nine analogs, with a focus on the binding pocket, intermolecular surfaces and hydrogen bonds. This study may be used as a guide for mutagenesis studies for designing new inhibitors against H5N1.

Key Words: H5N1, avian influenza, bird flu, drug design, neuraminidase, structural bioinformatics, active site, AG7088.

[#]**Author Profile:** Dr. Kuo-Chen Chou is the chief scientist of Gordon Life Science Institute. He is also an Advisory Professor of several Universities. Professor Chou has published over 350 papers in the fields of computer-aided drug design, bioinformatics, protein-structural prediction, low-frequency internal motion of protein and DNA and its biological functions, graphic rules in enzyme kinetics and other biological systems, and diffusion-controlled reactions of enzymes. For more information about Professor Kuo-Chen Chou, visit http://gordonlifescience.org/members/kcchou/ or http://home.roadrunner.com/~kchou/kc_index.html, or http://www.pami.sjtu.edu.cn/people/kcchou/.

INTRODUCTION

The outbreak of H5N1 avian influenza, commonly called "bird flu", has raised concerns that this virus might acquire the ability to pass readily among humans and cause influenza pandemics [1]. The main presenting features are fever, pneumonitis, lymphopenia and diarrhea [2]. The major biological concern is that the virus may acquire the ability to mutate and develop resistance to existing drugs.

Hemagglutinin (HA) and neuraminidase (NA) are the two antigenic glycoprotein enzymes located on the surface of the influenza virus [1, 3-5]. HA mediates the cell-surface sialic acid receptor binding to initiate virus infection. After virus replication, NA removes the sialic acid from the virus and cellular glycoproteins to facilitate virus release, spreading infection to new cells. Recently, more attention are focused on NA than HA as a main target for drug design against influenza, even though HA is easier to be separated and purified [5]. This is because inhibition of NA can delay the viral release from the surface of the infected cells [6, 7].

*Address correspondence to this author at the College of Life Science and Biotechnology, Shanghai Jiaotong University, Shanghai 200240, China; E-mail: dqwei@sjtu.edu.cn

[§]Equal contribution to this work

There are two existing drugs designed based on targeting NA: one is zanamivir (commercially named "Relenza") and the other is oseltamivir (commercially named "Tamiflu"). That zanamivir is administered by inhalation has limited its usage for treating the elderly because it may induce bronchospasm [8] 1393. Oseltamivir is an orally active influenza neuraminidase inhibitor approved for treatment and prevention of influenza virus infection. But recently some oseltamivir-resistant mutant neuraminidases have been reported from influenza A virus isolated from the influenza-infected humans treated by oseltamivir [9-12]. The present study was initiated in an attempt to provide useful insights to deal with the drug-resistance problem.

METHODS

Many useful insights for drug design can be acquired through the approach of structural bioinformatics [13] and other bioinformatics tools (see, e.g., [14-17]). In the present study, the following structural bioinformatics techniques were utilized.

Molecular Docking

Molecular docking has been increasingly utilized in the course of drug design (see, e.g., [18-38]). In the present

1573-4064/09 $55.00+.00

study, the Metropolis algorithm developed by Morris *et al*. [39] was used to find the most favorable binding interaction. The ligands were flexible during the docking process. The program generates a diversified set of conformations by making random changes to the ligand coordinates. Using the simulated annealing approach [40], the search for the favorable binding configurations was conducted within a specified 3D (dimensional) docking box, when a new conformation of the ligand was generated. Both methods seek to optimize the purely spatial contacts as well as electrostatic interactions. The interaction energy was calculated using the electrostatic and van der Waals potentials. In all our computations, the CHARMM22 [41] force field parameters were used. The binding pocket of H5N1-NA for the ligand is defined by those residues that have at least one heavy atom (i.e., other than hydrogen) with a distance 5 Å from a heavy atom of the drug molecule [13, 42-46]. A similar definition has been used for the binding pocket of ATP in the Cdk5-Nck5a* complex [47] that has later proved quite useful in identifying functional domains and stimulating the relevant truncation experiments [48].

Three ligands, AG7088, zanamivir and oseltamivir, were selected for the initial docking studies. AG7088 was developed by Pfizer and is currently in clinical trials for the treatment of rhinovirus, a pathogen that can cause the common cold; while zanamivir and oseltamivir are two anti-influenza drugs on the market.

Molecular Dynamics

To reflect the dynamic feature of the interaction beween the H5N1 receptor and its ligand, one of the feasible approaches is to use the molecular dynamics (MD) tool [49], which can simulate the motions in a system consisting of the target protein and its ligand under specified ensembles. In the current study, the energy favorable structures derived by the aforementioned molecular docking studies were further investigated with the MD tool triggered by breaking hydrogen bonds and making events of the ligand and receptor interactions. The simulations were performed at 300K and the normal atmospheric pressure. All the backbone atoms were fixed to maintain the correct protein structure with side chains being allowed to move freely. Some fictitious degree of freedom was added to the system to represent the motion of heat in and out of the system. This would generate a series of conformations in the important phase space, providing configuration and momentum information for each relevant atom, from which the thermodynamic properties of the system could be computed. The trajectory represents an exploration of the energy landscape with ligands sitting in a specific domain of the receptor. The above approach by combining the molecular docking operation with the MD simulation would provide very useful clues for the energy landscape of binding sites of H5N1-NA with the ligands and their interaction modes.

Similarity Search

Based on the binding conformation discovered *via* molecular docking into the crystal structure of H5N1-NA (PDB code 2HTY) [50], a similarity search was conducted to virtually screen for the leading compounds: 392,698 candidates were filtered and 556 potential leading compounds were obtained. Virtual screening has the advantages of being less expensive and easier to perform than the real experiments in searching for leading compounds [19], and hence has been widely used to find initial leading structures from a large collection of compounds. The MACCS structural keys, typed atom distances, typed atom triangles, typed graph distances, typed graph triangles were used to fingerprint the molecules. The fingerprints of the template molecule and each of these compounds in the database were calculated and was stored internally as a vector of indices, where the presence of an index in the vector indicates the presence of the corresponding substructure in the molecule. Once a fingerprint is derived from a chemical structure, a metric is needed to compare the fingerprint. Overlap determines the strictness of the search. All similarity metrics return a number between 0 and 100%, where 100% means the most similar. If the similarity threshold was assigned as 75%, 566 conformers were retrieved as an outcome of the similarity search. To further narrow down the investigation scope for molecular docking, a check for druggable rules was performed. At last, 143 potential leading compounds were screening out for further study later.

Screening Criteria

Binding energy (electrostatic and van der Waals interactions) was used as the criterion for ranking the binding state. Moreover, we also considered the other factors, such as the location of ligand, hydrophobic effects, steric complementarities and the size of the ligand. Ten top binding conformers were found that all have better binding interaction with the receptor than zanamivir and oseltamivir. Meanwhile, the intermolecular hydrogen bonds and electrostatic interaction, whose effects had already been counted in computing the binding energy, were further investigated in order to single out useful information for drug design.

RESULTS AND DISCUSSIONS

Comparison between AG7088 and Existing Drugs

The results of the binding free energies are shown in Table **1**, from which we can see that the binding energy of H5N1-NA with AG7088 is stronger than those of the existing drugs (zanamivir and oseltamivir). These findings have been further supported by the MD simulation studies later. A close view of the binding interaction is given in Fig. (**1A**). The AG7088 molecule is placed in active site. The binding pocket [43, 47] consists of residues ARG98, VAL129, ASP131, ARG132, SER160, ASN202, ILE203, ARG205, THR206, GLU208, PRO226, SER227, ASN228, GLU258, ARG273, ASN275, ALA323, TYR324, GLY325, ARG348, TYR382, ARG410, PRO411, and THR418. As shown in Fig. (**1A**), the AG7088 molecule should have enough room to move around in the pocket due to its tiny volume. The AG7088 molecule is tethered to the H5N1-NA complex by six hydrogen bonds (green lines) which may provide useful information for in-depth understandings on the mechanism of the binding between the H5N1-NA and the AG7088 molecule. The detailed atoms forming the hydrogen bonds are given as follows. There are two hydrogen bonds holding the AG7088 molecule with the receptor: one is between an

Table 1. List of Interaction Energies (kcal/mol) Obtained be Docking Ligands to the Crystal Structure of H5N1 (PDB code 2HTY)

Rank	Molecules	E (Electrostatic)	E (van der Waals)	E (Binding)
*	AG7088	-11.29	-24.60	-35.89
*	zanamivir	-7.41	-23.11	-30.52
*	osteltamivir	-11.42	-20.45	-31.87
1	1-6	-11.72	-32.41	-44.13
2	2-7	-15.34	-22.58	-37.92
3	1-8	-10.61	-27.51	-37.12
4	1-7	-7.19	-29.43	-36.621
5	3-2	-12.39	-24.13	-36.52
6	1-9	-14.95	-20.98	-35.93
7	3-10	-14.31	-20.96	-35.27
8	2-12	-8.07	-25.33	-33.40
9	3-11	-7.63	-24.10	-31.73

oxygen atom on H5N1-NA and the H atom on ARG132 of the AG7088 molecule (2.63 Å), and the other is between the same oxygen atom linked to another H atom on ARG132 (2.42 Å), respectively. The other four hydrogen bonds involve the interactions of AG7088 with ARG348 and ASN202 of the receptor.

The binding conformation of H5N1-NA and zanamivir is shown in Fig. (**1B**). The corresponding binding pocket consists of ARG98, GLU99, ASP131, ARG132, ARG205, GLU208, ARG273, and ARG348. In comparison with the case of AG7088, the hydrogen bonding interaction of zanamivir with H5N1-NA is much weaker.

Likewise, we also obtained the binding interaction of the receptor with oseltamivir, as shown in Fig. (**1C**). Now the binding pocket consists of ARG98, ASP131, ASN202, ILE203, ARG205, GLY225, PRO226, SER227, ASN228, GLU257, GLU258, ARG273, and ASN275. Also, the hydrogen bonding interaction of oseltamivir with H5N1-NA is much weaker, compared with AG7088.

Searching for Potential Leading Compounds

The 143 candidates screened from the 392,698 druggable compounds by using AG7088 as a template were docked to H5N1-N1. Listed in Table **1** are the 15 energy most favorable binding conformations. It is interesting to note that ARG98, ASP131, ARG132, ILE203, ARG205, SER227, GLU257, ARG273 occur in all these binding modes as highlighted in Table **2**. These findings are very useful for understanding the binding mechanism and providing clues for further modification. Shown in Table **3** are all the possible hydrogen bonds in the binding interactions concerned.

Analysis of Useful Features of Candidates

As shown in Fig. (**2A**), seven of the nine binding interaction modes share a common feature; i.e., with a five members ring and a six members ring, indicating that such a pattern is indispensable for the ligand-receptor recognition. Their binding energies are more favorable than others, suggesting that the framework shared by them can be used to design more potent inhibitors against H5N1. Also, as shown

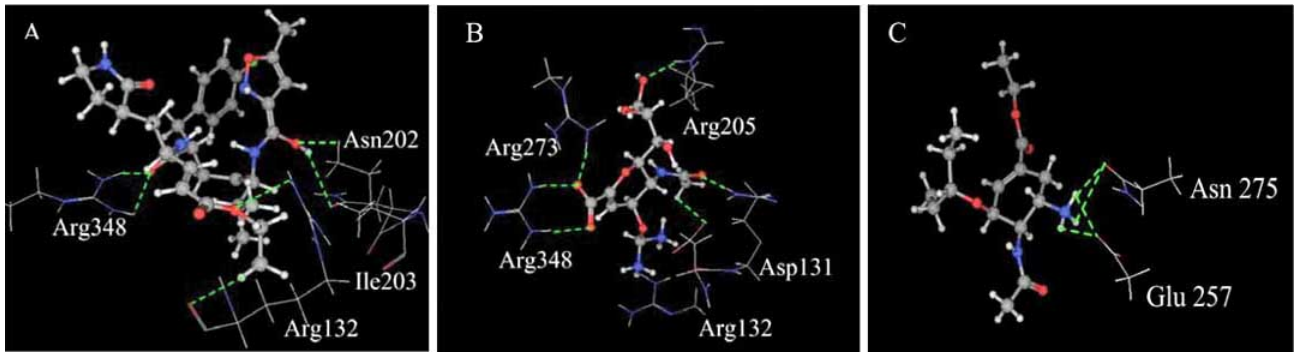

Fig. (1). A close view of the binding interaction of H5N1-NA with (**A**) AG7088, (**B**) zanamivir, and (**C**) oseltamivir: the residues of H5N1-NA are in stick drawing, the ligand is in ball and stick drawing, and the hydrogen bonds are represented by the green dotted lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper).

Table 2. Residues in Forming the Active Cavity of the Crystal Structure of H5N1-NA for the 9 Molecules Screened Out, Respectively, which is Obtained by MD Simulation

Mode	Residues[a]
1-6	ARG98, GLU99, LEU114, VAL129, ASP131, **ARG132**, ARG136, **TRP159**, SER160, ILE203, LEU204, ARG205, THR206, GLU208, SER227, **GLU257**, GLU258, ARG273, ASN275, ALA323, **TYR324**, TYR382, THR418
3-11	ARG98, **GLU99**, LEU114, **ASP131**, **ARG132**, ARG136, TRP159, **SER160**, **ILE203**, LEU204, ARG205, GLU208, SER227, GLU257, GLU258, SER227, GLU257, GLU258, ARG273, ASN275, **ALA323**, TYR324, GLY325, TYR382
1-8	ARG98, **GLU99**, LEU114, ASP131, **ARG132,** ARG136, **TRP159**, SER160, ILE203, ARG205, THR206, GLU208, SER227, GLU257, ARG273, ASN275, ALA323, **TRY324**, **TRY328**
1-9	ARG98, **GLU99**, ASP131, ARG132, ILE203, ARG205, **GLU208**, SER227, GLU257, GLU258, **ARG273**, ASN275, **ALA323**, TYR324, GLY325, ARG348, TYR382
3-10	**ARG98**, GLU99, ASP131, ARG132, TRP159, SER160, ILE203, LEU204, **ARG205**, THR206, GLU208, SER227, GLU257, GLU258, ARG273, TYR324, ARG348, TYR382
2-12	ARG98, **GLU99**, LEU114, **ASP131**, ARG132, ARG136, TRP159, SER160, ILE203, ARG205, GLU208, SER227, **GLU257**, GLU258, ARG273, ASN275, **ALA323**, TYR324, GLY325, GLY382
1-7	**ARG98**, GLU99, LEU114, ASP131, ARG132, **ARG136**, TRP159, SER160, ILE203, LEU204, ARG205, GLU208, SER227, ASN228, GLU257, GLU258, ARG273, ASN275, ALA323, TYR324, GLY325, ARG348, **TYR382**
3-2	ARG98, ASP131, **ARG132**, ASN202, ILE203, ARG205, GLY225, PRO226, **SER227**, ASN228, GLU257, GLU258, **ARG273**, ASN275, ALA323
2-7	ARG98, GLU99, VAL129, **ASP131**, ARG132, TRP159, SER160, ASN202, ILE203, ARG205, GLU208, **SER227**, GLU257, ARG273, **ASN275**, **ALA323**, TYR324, ARG348, TYR382

[a] Residue in bold-face type means having hydrogen bond interactions with the corresponding ligand. Residue in blue means the occurrence in all binding cavities.

(For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper).

in Table **2**, seven of the nine molecules share the feature with the following 11 residues occurring in their binding pockets: ARG98, GLU99, ASP131, ARG132, ILE203, ARG205, GLU208, SER227, GLU257, ARG273, and TYR324.

As shown in Table **1**, the molecule labeled as 1-6 has the most favorable binding energy. Therefore, the framework of 1-6 was adopted as the main compound template for drug design.

Structure Modification

Considering the aforementioned findings, we propose an inhibitor model, as shown in Fig. (**2B**). Our efforts were focused on how to make the structure modification (SM) for the inhibitor template to make its derivatives even better than the template itself in inhibiting H5N1-NA. To realize this, the problem was approached from the following five regions, as marked in Fig. (**2B**). (**A**) It can be seen from Fig. (**3A**) that

Table 3. List of the Possible Hydrogen Bonds Formed by the 9 Molecules Screened Out with the Crystal Structure of H5N1-NA (PDB code 2HTY)

Molecules	Number of H Bonds	Residues[a]
3-11	6	GLU99, ASP131, ARG132, SER160, ILE203, ALA323
1-6	5	**ARG136**, TRP159, GLU257, TYR324
1-8	5	GLU99, ARG132, TRP159, TYR324, TYR382
2-7	5	ASP131, **SER227,** ASN275, ALA323
1-9	4	GLU99, GLU208, ARG273, ALA323
2-12	4	GLU99, ASP131, GLU257, ALA323
1-7	4	ARG98, **ARG136**, TYR382
3-10	3	**ARG98**, ARG205
3-2	3	ARG132, SER227, ARG273

[a] Residue with bold-face type means that it has two possible hydrogen bonds with the corresponding ligand.

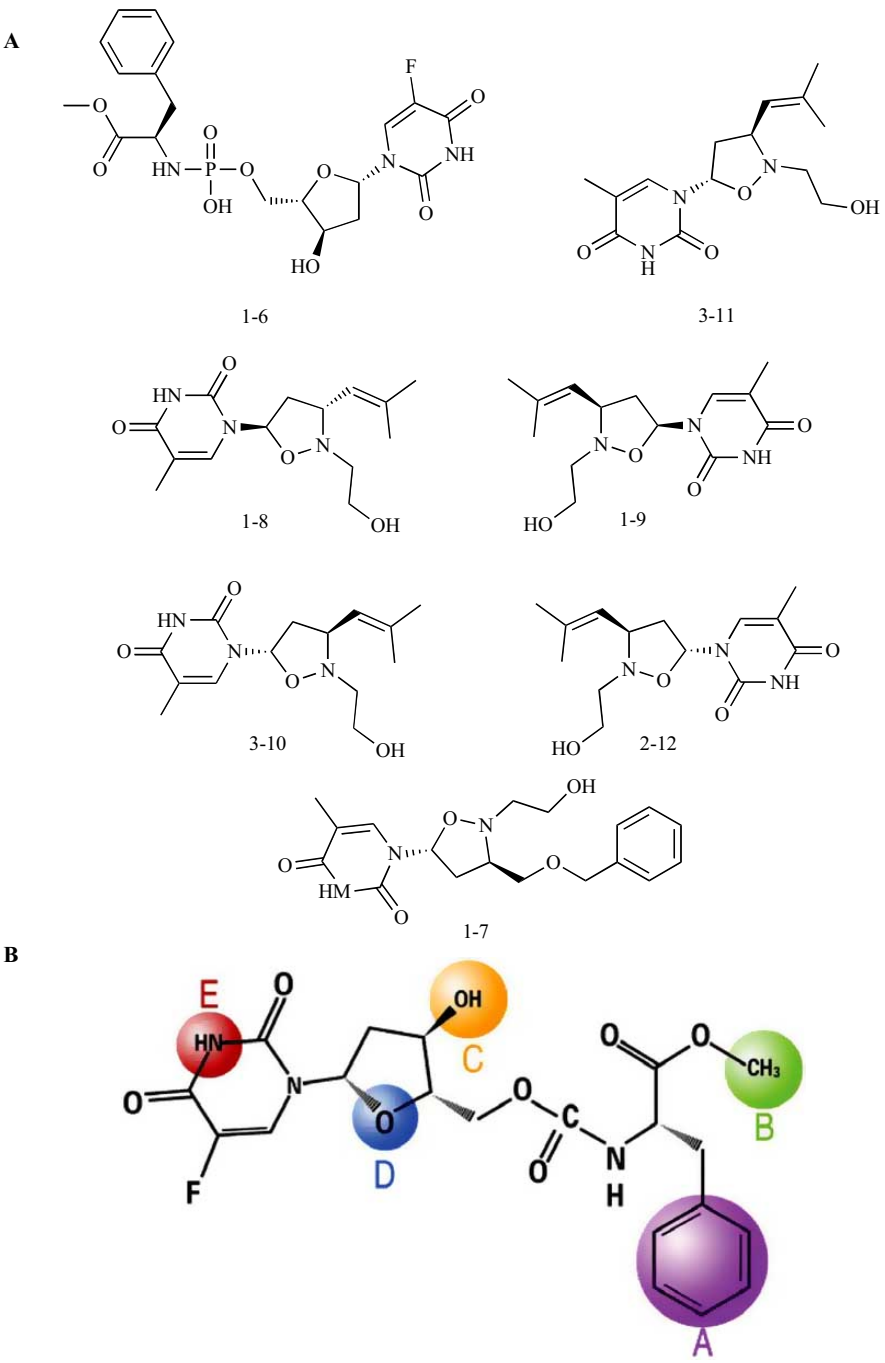

Fig. (2). Illustration to show (**A**) seven small molecule structures, and (**B**) the template model where **A,B,C,D**, and **E** are the five regions to be modified to generate new inhibitors for investigation. (For a color illustration of this figure, please see the web version of this paper).

some part of the receptor surfaces near region A is hydrophilic that can not match with the hydrophobic group. To improve their binding interaction, a logic strategy is to modify the region A by adding some hydrophilic groups, such as –OH(SM A-1), -NH2(SM A-2), -CH3(SM A-3), -SH(SM A-4), and -PH2(SM A-5) in this position. To analyze the impact by the modifications, we found that the average binding free energy for the aforementioned derivatives was more favorable than that of the original template molecule, indicating that it would improve the binding interaction by adding a hydrophilic group into the position A. **(B)** Change the chemical group -CH3 in region B in turn by -C2H5 (SM B-1), -C3H7 (SM B-2), -H (SM B-3) while keeping all the other parts constant. This modification can reduce the size of the compound and make it able to move deeper into the binding pocket so as to maximize its interaction with the receptor, as shown in Fig. (**3B**). **(C)** Change the chemical group -OH in

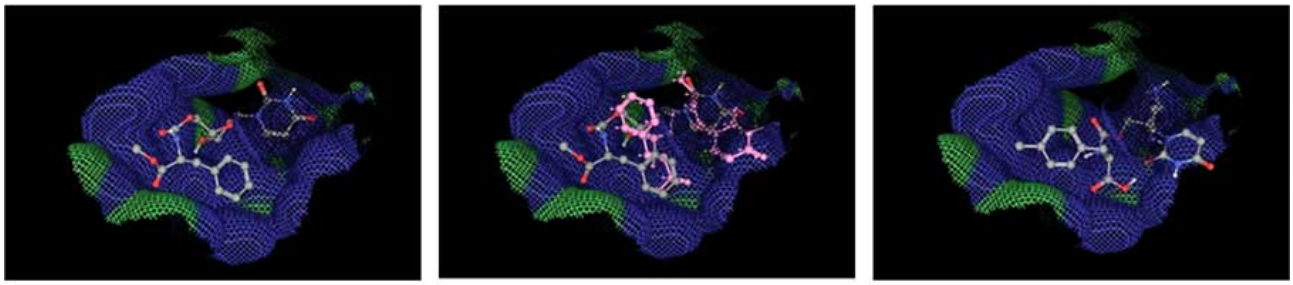

Fig. (3). A close view of the binding interactions of H5N1-NA with (**A**) the model template molecule, (**B**) both the model molecule and the modified molecule (pink) obtained by following SM B-3 scheme, and (**C**) the best modified molecule obtained by simultaneously following the SM A-3, SM B-3, SM C-1 and SM D-1 schemes, where the lipophilic and hydrophilic surfaces are colored in green and blue, respectively. The small molecules are in the ball-stick representation and the residues of the receptor are in the stick representation. Oxygen, nitrogen, carbon and hydrogen atoms are colored in red, blue, gray and white, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper).

region C in turn by –NH2 (SM C-1), -SH (SM C-2), -PH2 (SM C-3), -CH3 (SM C-4) while keeping all the other parts constant. By comparing the docking free energy, it can be observed that introducing the functional groups of SM C could make the ligand more active. The best result was achieved by following the SM C-1 scheme. **(D)** Change the chemical group -O in region D in turn by –NH (SM D-1), -CH2 (SM D-2), -C=O (SM D-3), while keeping all the other parts constant. It was found by the molecular docking studies that the van der Waals and electrostatic interactions for the original region D were 0.08 kal/mol and -0.02 kal/mol. The corresponding interactions for SM D-1, SM D-2, and SM D-3 were -0.30 and -0.10kal/mol, -0.04 and 0.01 kal/mol, -0.27 and -0.03 kal/mol, respectively, implying that SM D-1 or SM D-1 scheme would be the best in favoring the binding interaction and suggesting that this kind of modification should be taken into account for further processing in drug design. **(E)** Region E or SM-E. Change the chemical group -NH in region E in turn by –CH2 (SM E-1), -O (SM E-2), while keeping all the other parts exactly unvaried. However, this kind of group changes in region E (see Fig. (**2B**)) did not strengthen the binding interaction, implying that the original chemical group -NH was more favorable for the binding interaction and should be kept.

Finally, the molecules were assembled according to the aforementioned modifying schemes. The dockings were performed to rank their priorities as drug candidates according to the binding energy. It was observed that the analog obtained by modifying the template compound using the schemes SM A-3, SM B-3, SM C-1 and SM D-1 simultaneously could get the most competitive compound with the most favorable total docking free energy. As shown in Fig. (**3C**), such a modified molecule can very nicely match with both the hydrophilic and hydrophobic surfaces of the receptor.

ADITIONAL REMARKS

To provide the information that is often needed in targeting a particular organelle [51] or reprogramming cells for gene therapy [52, 53] but is not immediately available, the following procedures are briefly described to derive the relevant data. Shown in Fig. (**4**) is the entire sequence of the H5N1 avian influenza neuraminidase with the accession number of Q6DPL2 where the segment with X-ray PDB coordinates determined in 2HTY [50] is highlighted by an underline. Using the web server **Cell-PLoc** [54] at http://chou.med.harvard.edu/bioinf/Cell-PLoc/ or **Virus-PLoc** [55] at http://chou.med.harvard.edu/bioinf/virus/ and the input data of Fig. (**4**), we can find the H5N1 protein is located at

```
>Q6DPL2
MNPNQKIITIGSICMVTGIVSLMLQIGNMISIWVSHSIHTGNQHQSEPISNTNFLTEKAV
ASVKLAGNSSLCPINGWAVYSKDNSIRIGSKGDVFVIREPFISCSHLECRTFFLTQGALL
NDKHSNGTVKDRSPHRTLMSCPVGEAPSPYNSRFESVAWSASACHDGTSWLTIGISGPDN
GAVAVLKYNGIITDTIKSWRNNILRTQESECACVNGSCFTVMTDGPSNGQASHKIFKMEK
GKVVKSVELDAPNYHYEECSCYPNAGEITCVCRDNWHGSNRPWVSFNQNLEYQIGYICSG
VFGDNPRPNDGTGSCGPVSSNGAYGVKGFSFKYGNGVWIGRTKSTNSRSGFEMIWDPNGW
TETDSSFSVKQDIVAITDWSGYSGSFVQHPELTGLDCIRPCFWVELIRGRPKESTIWTSG
SSISFCGVNSDTVGWSWPDGAELPFTIDK
```

Fig. (4). The sequence (447 residues) of the H5N1 avian influenza neuraminidase in FASTA format, where Q6DPL2 right after the greater-than (“>”) symbol in the 1st line represents the accession number of the entire sequence, while the underscored segment (387 residues) corresponds to the sequence whose 3D structure has been determined [50] with the PDB code of 2HTY.

"membrane" or "plasma membrane". It was identified as a single-pass type II membrane protein by using the web server **MemType-2L** [56] at http://chou.med.harvard.edu/bioinf/MemType/. Furthermore, it was found its main enzyme functional class belonging to hydrolase family and sub functional class to glycosylase family by using the web server EzyPred [57] at http://chou.med.harvard.edu/bioinf/EzyPred/. If needed, its signal peptide can also be predicted by the web server Signal-CF [58] or Signal-3L [59].

It is instructive to note that in recent several years M2 mutations caused most strains of flu virus type A to become amantadine- and rimantadine-resistant, which is probably a response to overuse of the drugs [60]. To find out how and why flu virus has learned to so effectively evade amantadine and rimantadine, scientists have been seeking the structure of their molecular target, the ion-channel protein M2 on the flu virus surface. M2 is a membrane-spanning ion channel that conducts only protons and is required for the viral replication and infectivity. However, for decades the 3D structure of M2, the "holy grail" for structural biologists and virologists, has proved very difficult to analyze [60]. Fortunately, the M2 protein channel of influenza A virus was recently successfully determined by NMR [61]. This represents the first proton channel that has ever been analyzed structurally. Such a breakthrough will certainly provide a solid base, shedding new light on this area and stimulating new strategies to deal with endemic outbreaks of influenza and future pandemics.

CONCLUSION

The binding interactions of H5N1-NA with the drug molecule were investigated by molecular docking and MD simulation. Nine compounds were singled out from a series of molecular docking operations. The details of the interactions between the receptor and ligands were analyzed, and their binding pockets were explicitly defined. These findings may provide useful insights or at least a footing for designing new inhibitors against H5N1. It has not escaped our notice that the recent milestone work in determining the 3D structure (2RLF.pdb) of the flu virus proton channel (M2 channel) and revealing its subtle mechanism by the NMR studies [61] has provided very useful information for finding drugs against bird flu [60] from a different strategy. Further studies along such a line are currently under way in our lab.

ACKNOWLEDGEMENTS

This work was supported by the grants from the national 863 Bioinformatics Projects under the contract No. 2007AA02Z333, the national 973 program under the contract No. 2005CB724303 and the Chinese National Science Foundation under the Contract No.20773085 and 30870476, as well as the Major Chinese National Funding of New Drug Discovery for the Integrated Platform and the Virtual Laboratory for Computational Chemistry of CNIC, and the Supercomputing Center of CNIC, Chinese Academy of Sciences.

REFERENCES

[1] Horimoto, T.; Kawaoka, Y. Pandemic threat posed by avian influenza A viruses. *Clin. Mircobiol. Rev.*, **2001**, *14*, 129-149.

[2] Trampuz, A.; Prabhu, R. M.; Smith, T. F.; Baddour, L. M. Avian influenza: a new pandemic threat? *Mayo Clin. Proc.*, **2004**, *79*, 523-530.

[3] Colman, P. M.; Tulip, W. R.; Varghese, J. N.; Tulloch, P. A.; Baker, A. T.; Laver, W. G.; Air, G. M.; Webster, R. G. Three-dimensional structures of influenza virus neuraminidase-antibody complexes. *Philos. Trans. R Soc. Lond. B Biol. Sci.*, **1989**, *323*, 511-518.

[4] Smith, B. J.; Colman, P. M.; von Itzstein, M.; Danylec, B.; Varghese, J. N. Analysis of inhibitor binding in influenza virus neuraminidase. *Protein Sci.*, **2001**, *10*, 689-696.

[5] Oxford, J. S.; Novelli, P.; Sefton, A.; Lambkin, R. New millennium antivirals against pandemic and epidemic influenza: the neuraminidase inhibitors. *Antivir. Chem. Chemoth.*, **2002**, *13*, 205-217.

[6] Liu, C.; Eichelberger, M. C.; Compans, R. W.; Air, G. M. Influenza type A virus neuraminidase does not play a role in viral entry, replication, assembly, or budding. *J. Virol.*, **1995**, *69*, 1099-1106.

[7] McKimm-Breschkin, J. L. Resistance of influenza viruses to neuraminidase inhibitors, a review. *Antivir. Res.*, **2000**, *47*, 1-17.

[8] Hayden, F. G. Perspectivies on antriviral use during pandemic influenza. *Phios. Trans. R Soc. Lond. B Biol. Sci.*, **2001**, *356*, 1877-1884.

[9] Ward, P.; Small, I.; Smith, J.; Suter, P.; Dutkowski, R. Oseltamivir (Tamiflu) and its potential for use in the event of an influenza pandemic. *J. Antimicrob. Chemoth.*, **2005**, *55* (Suppl 1), 5-21.

[10] Mishin, V. P.; Hayden, F. G.; Gubareva, L. V. Susceptibilities of antiviral-resistant influenza viruses to novel neuraminidase inhibitors. *Antimicrob. Agents Chemoth.*, **2005**, *49*, 4515-4520.

[11] Le, Q. M.; Kiso, M.; Someya, K.; Sakai, Y. T.; Nguyen, T. H.; Nguyen, K. H.; Pham, N. D.; Ngyen, H. H.; Yamada, S.; Muramoto, Y.; Kawaoka, Y. Avain flu: isolation of drug-resistant H5N1 virus. *Nature*, **2005**, *437*, 1108.

[12] de Jong, M. D.; Tran, T. T.; Truong, H. K.; Vo, M. H.; Smith, G. J.; Nguyen, V. C.; Bach, V. C.; Phan, T. Q.; Do, Q. H.; Guan, Y.; Peiris, J. S.; Tran, T. H.; Farrar J. Oseltamivir resistance during treatment of influenza A (H5N1) infection. *New Engl. J. Med.*, **2005**, *353*, 2667-2672.

[13] Chou, K. C. Review: Structural bioinformatics and its impact to biomedical science. *Curr. Med. Chem.*, **2004**, *11*, 2105-2134.

[14] Lubec, G.; Afjehi-Sadat, L.; Yang, J. W.; John, J. P. Searching for hypothetical proteins: theory and practice based upon original data and literature. *Prog. Neurobiol.*, **2005**, *77*, 90-127.

[15] Chou, K. C. Review: Prediction of HIV protease cleavage sites in proteins. *Anal. Biochem.*, **1996**, *233*, 1-14.

[16] Gonzalez-Diaz, H.; Gonzalez-Diaz, Y.; Santana, L.; Ubeira, F. M.; Uriarte, E. Proteomics, networks, and connectivity indices. *Proteomics*, **2008**, *8*, 750-778.

[17] Shen, H. B.; Chou, K. C. HIVcleave: a web-server for predicting HIV protease cleavage sites in proteins. *Anal. Biochem.*, **2008**, *375*, 388-390.

[18] Chou, K. C.; Wei, D. Q.; Zhong, W. Z. Binding mechanism of coronavirus main proteinase with ligands and its implication to drug design aganist SARS. *Biochem. Biophys. Res. Commun.*, **2003**, *308*, 148-151. (Erratum: ibid., **2003**, *310*, 675)

[19] Sirois, S.; Wei, D. Q.; Du, Q. S.; Chou, K. C. Virtual screening for SARS-CoV protease based on KZ7088 pharmacophore points. *J. Chem. Inf. Comput. Sci.*, **2004**, *44*, 1111-1122.

[20] Wei, D. Q.; Sirois, S.; Du, Q. S.; Arias, H. R.; Chou, K. C. Theoretical studies of Alzheimer's disease drug candidate [(2,4-dimethoxy)benzylidene]-anabaseine dihydrochloride (GTS-21) and its derivatives. *Biochem. Biophys. Res. Commun.*, **2005**, *338*, 1059-1064.

[21] Du, Q. S.; Wang, S.; Wei, D. Q.; Sirois, S.; Chou, K. C. Molecular modeling and chemical modification for finding peptide inhibitor against SARS CoV Mpro. *Anal. Biochem.*, **2005**, *337*, 262-270.

[22] Du, Q. S.; Wang, S. Q.; Jiang, Z. Q.; Gao, W. N.; Li, Y. D.; Wei, D. Q.; Chou, K. C. Application of bioinforamtics in search for cleavable peptides of SARS-CoV Mpro and chemical modification of octapeptides. *Med. Chem.*, **2005**, *1*, 209-213.

[23] Wei, D. Q.; Zhang, R.; Du, Q. S.; Gao, W. N.; Li, Y.; Gao, H.; Wang, S. Q.; Zhang, X.; Li, A. X.; Sirois, S.; Chou, K. C. Anti-SARS drug screening by molecular docking. *Amino Acids*, **2006**, *31*, 73-80.

[24] Wei, D. Q.; Du, Q. S.; Sun, H.; Chou, K. C. Insights from modeling the 3D structure of H5N1 influenza virus neuraminidase and its

binding interactions with ligands. *Biochem. Biophys. Res. Commun.*, **2006**, *344*, 1048-1055.

[25] Chou, K. C.; Wei, D. Q.; Du, Q. S.; Sirois, S.; Zhong, W. Z. Review: Progress in computational approach to drug development against SARS. *Curr. Med. Chem.*, **2006**, *13*, 3263-3270.

[26] Zhang R.; Wei, D. Q.; Du, Q. S.; Chou, K. C. Molecular modeling studies of peptide drug candidates against SARS. *Med. Chem.*, **2006**, *2*, 309-314.

[27] Gan, Y. R.; Huang, H.; Huang, Y. D.; Rao, C. M.; Zhao, Y.; Liu, J. S.; Wu, L.; Wei, D. Q. Synthesis and activity assess of an octapeptide inhibitor designed for SARS coronavirus main proteinase. *Peptides*, **2006**, *27*, 622-625.

[28] Wang, S. Q.; Du, Q. S.; Chou, K. C. Study of drug resistance of chicken influenza A virus (H5N1) from homology-modeled 3D structures of neuraminidases. *Biochem. Biophys. Res. Commun.*, **2007**, *354*, 634-640.

[29] Wang, J. F.; Wei, D. Q.; Li, L.; Zheng, S. Y.; Li, Y. X.; Chou, K. C. 3D structure modeling of cytochrome P450 2C19 and its implication for personalized drug design. *Biochem. Biophys. Res. Commun.*, **2007**, *355*, 513-519. (Corrigendum: ibid, **2007**, *357*, 330)

[30] Wang, J. F.; Wei, D. Q.; Lin, Y.; Wang, Y. H.; Du, H. L.; Li, Y. X.; Chou, K. C. Insights from modeling the 3D structure of NAD(P)H-dependent D-xylose reductase of *Pichia stipitis* and its binding interactions with NAD and NADP. *Biochem. Biophys. Res. Commun.*, **2007**, *359*, 323-329.

[31] Wang, S. Q.; Du, Q. S.; Zhao, K.; Li, A. X.; Wei, D. Q.; Chou, K. C. Virtual screening for finding natural inhibitor against cathepsin-L for SARS therapy. *Amino Acids*, **2007**, *33*, 129-135.

[32] Du, Q. S.; Sun, H.; Chou, K. C. Inhibitor design for SARS coronavirus main protease based on "distorted key theory". *Med. Chem.*, **2007**, *3*, 1-6.

[33] Du, Q. S.; Wang, S. Q.; Chou, K. C. Analogue inhibitors by modifying oseltamivir based on the crystal neuraminidase structure for treating drug-resistant H5N1 virus. *Biochem. Biophys. Res. Commun.*, **2007**, *362*, 525-531.

[34] Li, Y.; Wei, D. Q.; Gao, W. N.; Gao, H.; Liu, B. N.; Huang, C. J.; Xu, W. R.; Liu, D. K.; Chen, H. F.; Chou, K. C. Computational approach to drug design for oxazolidinones as antibacterial agents. *Med. Chem.*, **2007**, *3*, 576-582.

[35] Li, L.; Wei. D. Q.; Wang, J. F.; Chou, K. C. Computational studies of the binding mechanism of calmodulin with chrysin. *Biochem. Biophys. Res. Commun.*, **2007**, *358*, 1102-1107.

[36] Zheng, H.; Wei, D. Q.; Zhang, R.; Wang, C.; Wei, H.; Chou, K. C. Screening for new agonists against Alzheimer's disease. *Med. Chem.*, **2007**, *3*, 488-493.

[37] Wei, H.; Zhang, R.; Wang, C.; Zhang, H.; Chou, K. C.; Wei, D. Q. Molecular insights of SAH enzyme catalysis and their implication for inhibitor design. *J. Theor. Biol.*, **2007**, *244*, 692-702.

[38] Gao, W. N.; Wei, D. Q.; Li, Y.; Gao, H.; Xu, W. R.; Li, A. X.; Chou, K. C. Agaritine and its derivatives are potential inhibitors against HIV proteases. *Med. Chem.*, **2007**, *3*, 221-226.

[39] Morris, G. M.; Goodsell, D. S.; Huey, R.; Olson, A. J. Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. *J. Comput. Aided Mol. Des.*, **1996**, *10*, 293-304.

[40] Chou, K. C.; Carlacci, L. Simulated annealing approach to the study of protein structures. *Protein Eng.*, **1991**, *4*, 661-667.

[41] Buck, M.; Bouguet-Bonnet, S.; Pastor, R. W.; MacKerell, A. D. Importance of the CMAP correction to the CHARMM22 protein force field: dynamics of hen lysozyme. *Biophys. J.*, **2006**, *90*, 36-38.

[42] Chou, K. C. Insights from modeling the tertiary structure of BACE2. *J. Proteome Res.*, **2004**, *3*, 1069-1072.

[43] Chou, K. C. Insights from modeling the 3D structure of the extracellular domain of alpha 7 nicotinic acetylcholine receptor. *Biochem. Biophys. Res. Commun.*, **2004**, *319*, 433-438.

[44] Chou, K. C. Molecular therapeutic target for type-2 diabetes. *J. Proteome Res.*, **2004**, *3*, 1284-1288.

[45] Chou, K. C. Modeling the tertiary structure of human cathepsin-E. *Biochem. Biophys. Res. Commun.*, **2005**, *331*, 56-60.

[46] Chou, K. C. Coupling interaction between thromboxane A2 receptor and alpha-13 subunit of guanine nucleotide-binding protein. *J. Proteome Res.*, **2005**, *4*, 1681-1686.

[47] Chou, K. C.; Watenpaugh, K. D.; Heinrikson, R. L. A model of the complex domain of neuronal Cdk5 activator. *Biochem. Biophys. Res. Commun.*, **1999**, *259*, 420-428.

[48] Zhang, J.; Luan, C. H.; Chou, K. C.; Johnson, G. V. W. Identification of the N-terminal functional domains of Cdk5 by molecular truncation and computer modeling. *Proteins.*, **2002**, *48*, 447-453.

[49] Allen, M. P.; Wilson, M. R. Computer simulation of liquid crystals. *J. Comput. Aided Mol. Des.*, **1989**, *3*, 335-353.

[50] Russell, R. J.; Haire, L. F.; Stevens, D. J.; Collins, P. J.; Lin, Y. P.; Blcakburn, 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*, 45-49.

[51] Chou, K. C.; Shen, H. B. Review: Recent progresses in protein subcellular location prediction. *Anal. Biochem.*, **2007**, *370*, 1-16.

[52] Hagmann, M. Colleagues say "Amen" to this year's (Nobel Prizes) choice. *Science*, **1999**, *286*, 666.

[53] Chou, K. C. Review: Prediction of protein signal sequences. *Curr. Protein Pept. Sci.*, **2002**, *3*, 615-622.

[54] Chou, K. C.; Shen, H. B. Cell-PLoc: A package of web-servers for predicting subcellular localization of proteins in various organisms. *Nat. Protoc.*, **2008**, *3*, 153-162.

[55] Shen, H. B.; Chou, K. C. Virus-PLoc: A fusion classifier for predicting the subcellular localization of viral proteins within host and virus-infected cells. *Biopolymers*, **2007**, *85*, 233-240.

[56] Chou, K. C.; Shen, H. B. MemType-2L: A web server for predicting membrane proteins and their types by incorporating evolution information through Pse-PSSM. *Biochem. Biophys. Res. Commun.*, **2007**, *360*, 339-345.

[57] Shen, H. B.; Chou, K. C. EzyPred: A top-down approach for predicting enzyme functional classes and subclasses. *Biochem. Biophys. Res. Commun.*, **2007**, *364*, 53-59.

[58] Chou, K. C.; Shen, H. B. Signal-CF: a subsite-coupled and window-fusing approach for predicting signal peptides. *Biochem. Biophys. Res. Commun.*, **2007**, *357*, 633-640.

[59] Shen, H. B.; Chou, K. C. Signal-3L: a 3-layer approach for predicting signal peptides. *Biochem. Biophys. Res. Commun.*, **2007**, *363*, 297-303.

[60] Borman, S. Flu virus proton channel analyzed: structures of key surface protein suggest different drug mechanisms. *Chem. Eng. News*, **2008**, *86*, 53-54.

[61] Schnell, J. R.; Chou, J. J. Structure and mechanism of the M2 proton channel of influenza A virus. *Nature*, **2008**, *451*, 591-595.

Received: 22 December, 2008 **Revised: 20 March, 2009** **Accepted: 24 March, 2009**