Binding Interaction Analysis of the Active Site and Its Inhibitors for Neuraminidase (N1 Subtype) of Human Influenza Virus by the Integration of Molecular Docking, FMO Calculation and 3D-QSAR CoMFA Modeling

Qingye Zhang,[†,||] Jiaoyan Yang,[§,||] Kun Liang,[†] Lingling Feng,[†] Sanpin Li,[†] Jian Wan,*[,†] Xin Xu,*[,‡] Guangfu Yang,[†] Deli Liu,[§] and Shao Yang[§]

Key Laboratory of Pesticide & Chemical Biology (CCNU), Ministry of Education, College of Chemistry, Central China Normal University, Wuhan 430079, P R China, State Key Laboratory of Physical Chemistry of Solid Surfaces, College of Chemistry and Chemical Engineering, Center for Theoretical Chemistry, Xiamen University, Xiamen 361005, P R China, and College of Life Science, Central China Normal University, Wuhan 430079, P R China

Received February 8, 2008

Recently, the worldwide spread of A/H5N1 avian influenza with high virulence has highlighted the potential threat of human influenza pandemic. Tamiflu and Relenza are currently the only two anti-influenza drugs targeting the neuraminidase (NA) enzyme of human influenza virus. Reports of the emergence of drug resistance further make the development of new potent anti-influenza inhibitors a priority. The X-ray crystallographic study of A/H5N1 avian influenza NA subtypes (Russell, R. J. *Nature* 2006, *443*, 45−49) has demonstrated that there exist two genetically distinct groups, group-1 (N1, N4, N5 and N8) and group-2 (N2, N3, N6, N7 and N9), whose conformations are substantially different. The detailed comparison of their active sites has established, heretofore, the most accurate and solid molecular basis of structure and mechanism for the development of new anti-influenza drugs. In the present study, a three-dimensional structure of N1 subtype of human influenza type A virus (N1hA) has been generated by homology modeling using the X-ray crystallographic structure of N1 subtype of avian influenza virus (N1aA) as the template. Binding interaction analysis between the active site and its inhibitors has been performed by combining ab initio fragment molecular orbital (FMO) calculations and three-dimensional quantitative structure−activity relationship with comparative molecular field analysis (3D-QSAR CoMFA) modeling. Integrated with docking-based 3D-QSAR CoMFA modeling, molecular surface property (electrostatic and steric) mapping and FMO pair interaction analysis, a set of new receptor−ligand binding models and bioaffinity predictive models for rational design and virtual screening of more potent inhibitors of N1hA are established. In addition, the flexibility of the loop-150 of N1hA and N1aA has been examined by a series of molecular dynamics simulations.

INTRODUCTION

Influenza is a respiratory infection associated with significant morbidity in the general population and mortality in elderly and high-risk patients.[1] Recently, the worldwide spread of avian influenza with terrible mortality has raised concerns that the avian influenza virus might mutate and cause a pandemic of influenza in humans. Understanding the molecular mechanism and cellular biology of influenza virus, especially, the substantial differences between humans and avian species is crucial for the design and development of novel inhibitors with potent anti-influenza activity. Considerable progress, achieved currently, in elucidating the life circle of the influenza virus has provided several potential molecular targets for drug design. These include haemagglutinin,[2] neuraminidase,[3] M2 protein,[4] and endonuclease.[5]

Haemagglutinin (HA) and neuraminidase (NA) are two major surface glycoproteins expressed by both influenza A and B viruses. HA mediates binding of viruses to target cells via terminal sialic acid residue in glycoconjugates. NA is important for viral release from the infected cell and viral transport through the mucus in the respiratory tract.[6] Influenza A viruses could be classified into various subtypes in terms of the distinct antigenic properties of different HA (H1−H16) and NA (N1−N9) and represented as the explicit/concrete combinations of HA and NA.[3] For example, the current avian influenza virus belongs to influenza A virus H5N1 subtype (A/H5N1); the three historic pandemics in humans were caused by A/H1N1 (in 1918), A/H2N2 (in 1957), and A/H3N2 (in 1968), respectively.[7,8] Neuraminidases of viruses, currently circulating in humans, belong to two phylogenetically distinct groups,[9] group-1, composed of the N1, N4, N5, and N8 subtypes, and group-2, composed of N2, N3, N6, N7 and N9.

The determined crystal structure of N2 or N9 has provided opportunities for the design of broad-spectrum anti-influenza drugs.[10,11] Two classes of anti-influenza virus agents target-

* To whom correspondence should be addressed. Phone: +86-27-67862022. Fax: +86-27-67862022. E-mail: jianwan@mail.ccnu.edu.cn (J.W.); xinxu@xmu.edu.cn (X.X.)
[†] College of Chemistry, Central China Normal University.
[‡] Center for Theoretical Chemistry, Xiamen University.
[§] College of Life Science, Central China Normal University.
[||] These authors contributed equally to this study.

10.1021/ci800041k CCC: $40.75
Published on Web 08/16/2008

ing either the M2 ion channel or NA are currently available for influenza treatment and under consideration for stockpiling in the event of influenza pandemic. The weakness of two M2 blockers, amantadine and its analogue rimantadine, consists of a lack of inhibitory effects against the type B influenza viruses in which M2 protein does not exist, many side effects, and resistance problems.[12] Moreover, recent A/H5N1 avian influenza viruses are naturally resistant to these inhibitors.[13] In addition to the M2 inhibitors, there are only two NA (both A and B viruses) inhibitors, oseltamivir (commercially named as Tamiflu) and zanamivir (commercially named as Relenza), which have been approved for treatment of influenza virus infections in humans.[14−16] Structure-based inhibitor design[17] and two and three dimension quantitative structure−activity relationship (2D and 3D QSAR) studies [6,18−21] have played an important role accelerating the discovery of the NA inhibitors including Tamiflu and Relenza. The success of these drug developments has been attributed, at least partially, to the validity of the approximation that the catalytic sites of the NA enzymes are an invariant feature that can be exploited for subtype-independent therapy and to the observations that subtypes N2 and N9 are comparatively rigid, with only minor conformational changes in the sites of inhibitor binding. The crystallographic structure information of subtype N2 or N9 has been, therefore, applied to the previous structure-based NA inhibitor designs[16] and the associated 3D-QSAR analysis studies.[6,18,21]

Recently, a deliberated X-ray crystallographic structure study of N1 subtype of avian influenza A/H5N1 (PDB code 2HU0) has been reported by R. J. Russell et al.[22] This breakthrough study has clearly demonstrated that there are two genetically distinct groups, group-1 (N1) and group-2 (N2 and N9), that possess substantial conformational differences, centered on the "150-loop" and the "150-cavity" adjacent to the active site. This finding has established, heretofore, the most direct and solid molecular basis of structure and mechanism for the development of novel potent N1 subtype inhibitors. One report[23] also declared that inhibitor structure−activity relationships (SAR) did not apply across subtypes. Reports of the emergence of NA inhibitor resistance [24−27] have further made the development of new potent anti-influenza inhibitors a priority. Therefore, it is worthy to carry out a systematic comparison study of the catalytically active sites between N1 subtype of human influenza type A virus (N1hA) and that of avian influenza A/H5N1 and to do further 3D-QSAR study based upon the X-ray crystallographic structure of N1 subtype of avian influenza A/H5N1 (N1aA) instead of the N2 or N9 subtype as carried out before.

In the present work, the 3D structure of N1hA was built by homology modeling, employing the X-ray crystallographic structure of 2HU0 as the protein template. Subsequently, we analyzed and compared the receptor−ligand binding interactions in detail by performing ab initio fragment molecular orbital (FMO)[28] calculations based on docking Tamiflu into the active site of the N1hA homology model and the crystal complex structure for Tamiflu binding to the active site of 2HU0.[22] Furthermore, a molecular docking-based 3D-QSAR study with comparative molecular field analysis (CoMFA)[29] was performed on 27 well-known N1 inhibitors. Integration of the results from the 3D-QSAR CoMFA modeling, molecular surface property (electrostatic and steric) mapping of homology model, and binding interaction analysis has put forward a set of new receptor−ligand binding models and bioaffinity predictive models, which are valuable for rational design and virtual screening of more potent inhibitors of N1hA.

MATERIALS AND METHODS

Homology Modeling. An accurate 3D structure of the target enzyme is important for molecular docking and rational development of novel potential inhibitors. The template enzyme employed here is avian influenza A/H5N1 (PDB code 2HU0). The target enzyme is N1hA of A/PR/8/34 H1N1, whose amino acid sequence (ID P03468) can be obtained from the NCBI Web site.[30] The homology model of N1hA was built based on sequence alignment. The target sequence P03468 was submitted to the SWISS-MODEL server (Automated Comparative Protein Modeling Server, Version 3.5, GlaxoWellcome Experiment Research, Geneva, Switzerland) [31,32] for a comparative structural modeling. All hydrogen atoms were subsequently added to the unoccupied valence of heavy atoms at the corresponding neutral state using the biopolymer module of SYBYL 7.0 package.[33]

Molecular Docking. Molecular docking analysis was carried out by the FlexX module of SYBYL package to explore the interaction mechanism and to illustrate the accurate binding model for the active site of N1hA with its ligands, especially with Tamiflu. FlexX is a fast automated docking program that takes ligand's conformational flexibility into account during the docking process by an incremental fragment placing technique.[34,35] We defined the active site as follows: the model structure of N1hA and the template protein (2HU0) were superposed first, and the oseltamivir ligand of 2HU0 was then merged into the corresponding site of the N1hA model structure. All atoms located within the range of 6.5 Å from any atom of the oseltamivir ligand in the model were selected into the active site, and the residue was included into the active site if at least one of its atoms was picked out. Other default parameters in the FlexX module were adopted in the calculations of FlexX docking.

Binding Energy and Pair Interaction Analysis. We have performed ab initio fragment molecular orbital (FMO) calculations for the binding energies of active site and pair interactions of specific amino acids with Tamiflu. FMO is an approximate molecular orbital method developed for the calculations of large molecules consisting of thousands of atoms. It has been successfully applied for the quantum mechanical calculations of binding energies and pair interaction energies of receptor ligands.[36−38] In the FMO method, the whole molecular system is divided into small fragments, and ab initio MO calculations are performed on fragments and fragment pairs, which are referred to as "monomers" and "dimers", respectively.[39−41] The total electronic energy of the molecular system is calculated by the following equation, where E_I and E_{IJ} are self-consistent field (SCF) energies of a monomer (I) and a dimer (IJ), respectively

$$\mathrm{E} = \sum_I E_I + \sum_{I>J} (E_{IJ} - E_I - E_J) \qquad (1)$$

Equation 1 can be transformed as follows

$$E = \sum_I E_I + \sum_{I>J} \Delta E_{IJ}$$

where ΔE_{IJ} is the pair interaction energy

$$\Delta E_{IJ} = E_{IJ} - E_I - E_J \quad (2)$$

The computation of the receptor−ligand binding energy is performed according to

$$\Delta E_{RL} = E_{RL} - (E_R + E_L) \quad (3)$$

where E_{RL} is the total energy of complex *RL*, and E_R and E_L are the total energies of individual receptor *R* and ligand *L* with the same geometries as in the complex.

We adopted the residues located within a radius of 6.5 Å from the center of the complex of the docking ligand with N1hA. The ends of the polypeptide fragments, −NH and −CO, were capped with hydrogen atoms. The N-terminus Lys, Arg, and His residues were protonated, while the C-terminus Asp, and Glu residues were deprotonated. To make the fragmentation of the receptor protein, the peptide chain was divided at the Cα atom into blocks of each residue in a manner as shown in Figure 1. The ligand was treated as a single fragment.

All FMO calculations in the present study were performed by using GAMESS software[42] at the restricted Hartree−Fock level with the 6−31G* basis set. Other parameters adopted their default values.

Data Sets for CoMFA Modeling Analysis. As is well-known, the validation of 3D QSAR analysis is strongly dependent upon the selected training data set. Therefore, the determined reliability and comparability of bioaffinity (pIC_{50}), its dynamic range and structural diversity of compounds should be selected delicately for doing 3D QSAR such as CoMFA modeling. There is an obvious variation in biological activity measured under different experimental materials and conditions. Hence, to obtain a correct quantitative relationship between structures and inhibitors' activies, it is essential to select bioaffinities determined under identical experimental materials and conditions or carefully done by an individual research group. We selected twenty-seven compounds (see Figure 2 for specific molecular formula) from three references.[16,20,43] Their biological inhibitory activities were determined by one research group using human A/PR/8/34 (H1N1).[30,44] The structures of oseltamivir (no. 9), zanamivir (no. 12) and 4-amino-Neu5Ac2EN (no. 11) were extracted from the Brookhaven Protein Database (PDB, http://www.rcsb.org/pdb; PDB codes 2QWK, 2QWE, and 2QWD, respectively). Other structures of the oseltamivir analogues were constructed based on the oseltamivir's conformation. The 3D structures of all compounds were constructed by using the sketch molecule module of SYBYL 7.0 package. Partial atomic charges were calculated by the Gasteiger-Huckel method,[45] and energy minimizations were performed by using the Tripos force field[46] and the Powell

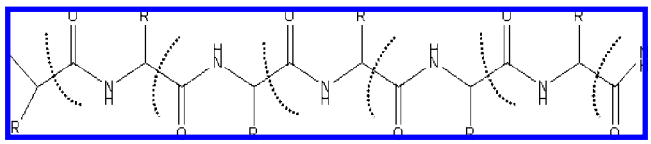

Figure 1. Schematic diagram of fragmentation of polypeptides for FMO binding energy and pair interaction energy calculations (see the context of Materials and Methods for detail).

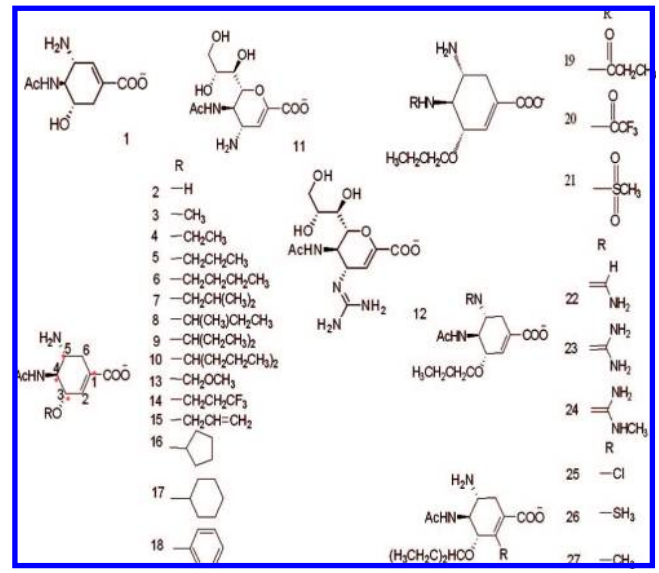

Figure 2. Molecular structural formulas of compounds collected for the 3D QSAR CoMFA modeling. The atoms used for fitting are marked with asterisks.

conjugate gradient algorithm[47] with a convergence criterion of 0.05 kcal/(mol Å). All calculations in the present study were performed on a CCNUGrid-based computational environment (CCNUGrid Web site http://grid.ccnu.edu.cn/index.htm).

Molecular Alignment Rules for CoMFA Modeling. Two different alignment rules were adopted in the CoMFA modeling: [1] The SYBYL conventional fit-atom molecular alignment rule was applied by using the module of SYBYL/Analyze/Fit-atom. The Fit-atom module adjusted the geometry of the molecule such that its steric and electrostatic fields matched the template molecule. The template molecule chosen in the present study is oseltamivir (no. 9), which is the most potent inhibitor discovered so far. The atoms used as fitting scaffold were marked with asterisks in the template molecule shown in Figure 2. [2] Molecular alignment using molecular docking active conformation was achieved by FlexX, docking the ligands into the active site of N1hA model structure. All ligands' active conformations were selected from the binding orientation in the active site, jointly evaluated by the consideration of FlexX Energy Scores (total score).

3D QSAR CoMFA Modeling. CoMFA steric and electrostatic interaction fields were calculated at each lattice intersection on a regularly spaced grid of 2 Å. The grid pattern was generated automatically by the SYBYL/CoMFA routine. An sp^3 carbon atom with a van der Waals radius of 1.52 Å and a +1.0 charge was used as the probe to calculate the steric (Lennard-Jones 6−12 potential) field and the electrostatic (Coulombic potential) field with a constant dielectric constant at each lattice point. The electrostatic energy at the point where the steric energy exceeded the steric cutoff for any molecule in the analysis was set to the mean value of the nonexcluded electrostatic field. Values of the steric and electrostatic fields were truncated at 30.0 kcal/mol. The CoMFA steric and electrostatic fields were scaled by the CoMFA-STD[48] method in SYBYL.

A partial least-squares (PLS) approach,[49−51] which is an extension of the multiple regression analysis, was used to derive the 3D QSAR models, in which the CoMFA descrip-

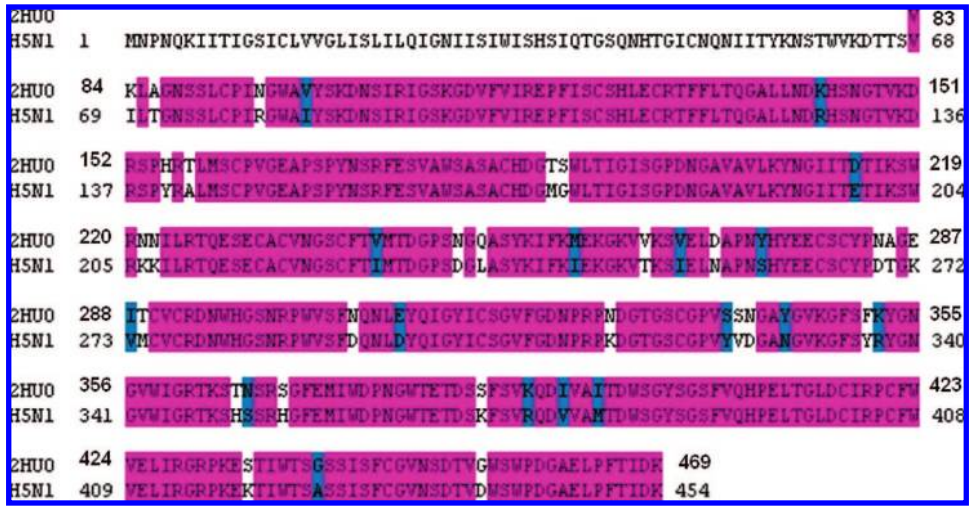

Figure 3. Alignments of the sequences of 2HU0 template and N1hA target enzymes produced by SWISS-MODEL server. Red regions denote that the amino acid residues in the individual column are identical in the sequence alignment. Blue regions denote that amino acid substitutions are observed.

tors were used as independent variables, and the experimental pIC50 values were used as dependent variables. The cross-validation with leave-one-out (LOO) option and the SAM-PLS program[52] were applied to obtain the optimal number of components to be used in the final analysis. After the optimal number of components was determined, a non-cross-validated analysis was performed without column filtering. The two best CoMFA models were obtained finally by using Region Focusing[53] with weights of StDev*Coefficient. The q^2 (cross-validated r^2), Spress (cross-validated standard error of prediction), r^2 (non-cross-validated r^2), F values, and standard error of estimate values were computed according to the definitions in SYBYL 7.0 package.

Molecular Surface Physicochemical Properties. Surface physicochemical property maps, that is, electrostatic potential, hydrophobic potential, and hydrogen bonding (donor/acceptor) potential maps of the N1hA active site were generated on the solvent assessable (Connolly) surface using the MOLCAD module[54] of SYBYL. These MOLCAD-generated property surface maps were compared to those obtained by CoMFA modeling.

Molecular Dynamics Simulation. To examine the flexibility of the loop-150 of the N1aA and N1hA, the crystallographic structure of N1aA and the initial 3D-structure of N1hA built by homology modeling were subjected to a series of MD simulations, respectively, using the SANDER module of AMBER 8 package.[55] The two systems were neutralized by addition of Na^+ and then solvated in an octahedral box of TIP3P water molecules,[56] which extended at least 8 Å from any given protein atom of N1aA and N1hA. The particle mesh Ewald method [57−59] for the long-range electrostatics, a 10 Å cutoff for nonbonding van der Waals interactions, and periodic boundary conditions were set up for the series of simulations. The eight disulfide bonds in the homology modeling of N1hA were linked handedly using Build/Edit model of SYBYL 7.0. All bonds involving hydrogen were constrained using the SHAKE algorithm.[60] Constant temperature and pressure (300 K, 1 atm) were maintained using the Berendsen coupling algorithm[61] with a time constant for heat bath coupling of 2 ps. A time step of 1 fs was used to integrate the equations of motion. Before starting the production-run phase, the following equilibration protocol was applied to the whole system: first all water of the TIP3P box water system was minimized by 3000 steepest descent steps and 3000 conjugate gradient steps with freezing the protein; then a minimization of whole system was carried out by Amber force field for 4000 steps with releasing the whole system. Finally, the system was slowly heated from 10 to 300 K over 100 ps before MD simulation. Coordinates were recorded every 1 ps during the entire simulation process. The root-mean-square deviations (rmsd) of the backbones of the whole protein and the conserved residues of the protein over the simulation time were derived using the PTRAJ module.

RESULTS AND DISCUSSION

Homology Modeling. A recent review[62] has concluded that sequence identity higher than 30% between two proteins is indicative of similar 3D structures. Figure 3 shows the sequence alignment of the 2HU0 template and the N1hA target enzymes. The sequence identity between our target and the template enzyme is 88.6%, which allows for a rather straightforward sequence alignment and guarantees the quality of homology modeling.

Recently, a crystallographic structural study of H5N1 avian influenza was reported by R. J. Russell et al.[22] They found that avian influenza neuraminidase N1 can be produced into crystallographic complexes with oseltamivir with either the "open" or "closed" conformation of the 150-loop, depending on the soaking conditions. If incubating avian influenza N1 crystals with oseltamivir in a lower concentration or for short time, the inhibitor binds to the 150-loop in the open cavity conformation. However, if it was incubated with oseltamivir in a higher concentration or for longer time, the 150-loop changes its conformation into the closed conformation, which resembles the conformation observed in group-2 NA. According to actual enzyme inhibition assay, the concentration of inhibitors is conventionally low and the time of incubation/binding interaction is not very long because of the metabolism and diffusion in vivo. Therefore, we choose the open conformation of the 150-loop of crystallographic structure 2HU0 as the template to build a 3D N1hA model from the target sequence. Figure 4a shows the superposition of 3D

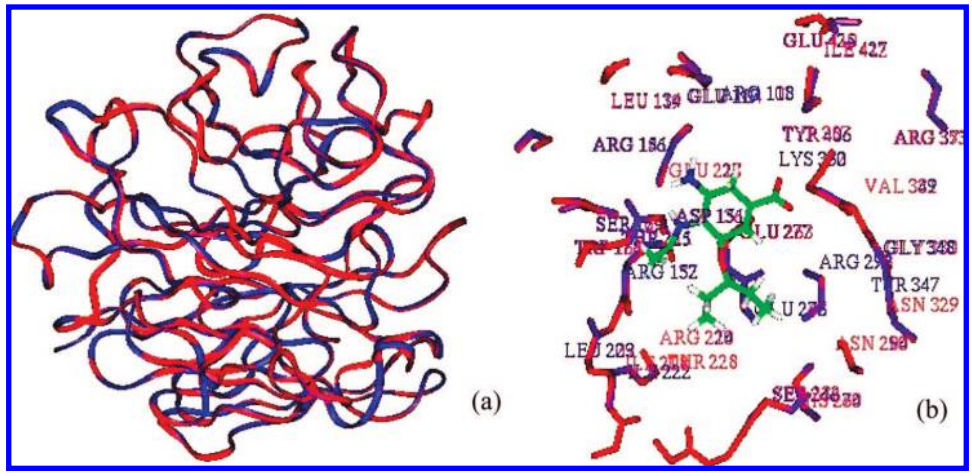

Figure 4. Schematics for the superposition of 2HU0 template and the N1hA model structure from homology modeling. Red ribbon denotes the target protein, while blue ribbon 2HU0 template in (a) entire backbone and (b) active site with ligand oseltamivir.

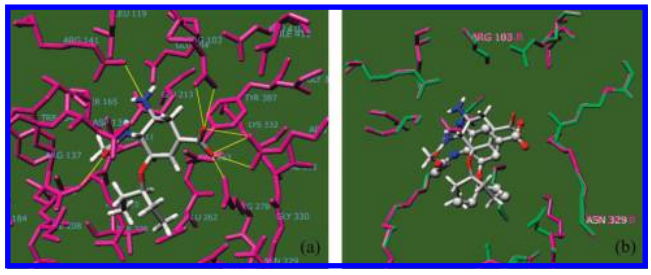

Figure 5. (a) The binding model of oseltamivir into the active site of the N1hA model. The yellow lines represent the hydrogen bonding between the ligand and the amino acid residues. (b) Schematics for the superposition of the active sites of 2HU0 template of the crystallographic complex structure (with green sticks) and the N1hA model (with red sticks) with oseltamivir located in a.

structure from the homology modeling of N1hA with the 2HU0 template: the red ribbon is the homology modeling target and the blue one represents the 2HU0 template. As expected, the overall conformation of the modeling target is very similar to the template with the rmsd of 0.089 Å. Even the amino acids involved in the oseltamivir binding site are well conserved between them as shown in Figure 4b.

Binding Energy and Pair Interaction Energy Analysis Based on Molecular Docking. We used the FlexX module to dock the ligands into the active site of the N1hA model. The structural parameters of enzyme were set rigid, while those of ligand were set soft during the present molecular docking process by FlexX. Under physiological conditions, oseltamivir was binding into the active site with the active form of carboxyl anion group. This was observed in the crystallographic structure of 2HU0.[22] We presumed that inhibitors would bind with a similar manner into both N1hA and avian A/H5N1 NA because they are well conserved in both sequence and structure. Therefore, the carboxyl anion form of oseltamivir was selected in molecular docking to simulate the interaction between the binding site of N1hA and oseltamivir.

FlexX docking results show that the binding positions of oseltamivir into the active site of N1hA are indeed similar to those observed in the crystallographic complex structure of 2HU0 with oseltamivir (see Figure 5). The carboxyl anion group of oseltamivir is held strongly by three arginine residues Arg103, Arg278, and Arg353 with positive charges in the binding site of N1hA. This suggests that strong electrostatic or hydrogen-bonding interactions play an important role in this region. There exist homologous binding interactions of oseltamivir to residues of Arg118, Arg292, and Arg371 in the case of 2HU0. In addition, electrostatic or hydrogen-bonding interactions from the oxygen atom of the C4 acetamide group interacting with the guanidyl group of residue Arg137 of N1hA (Arg152 in the case of 2HU0) also make contribution to the binding energy. The methyl group of the C4 acetamide in oseltamivir occupies a hydrophobic pocket generated by Trp164 and Ile208 of N1hA (Trp178 and Ile222 in the case of 2HU0). Meanwhile, the alkyl side chains of the C3 pentyloxy group are situated inside of a larger hydrophobic pocket created by the hydrocarbon chains of Glu262, Ser232, Arg210, and Ile208 of N1hA (Glu276, Ser246, Arg224, and Ile222 in the case of 2HU0). It was shown that, in the 150-loop region of the open conformation of 2HU0,[22] Asp151 and Glu119 are located far from the C5 amino group of oseltamivir. In contrast, in the 150-loop region of the closed conformation of 2HU0[22] and the N9 crystallographic complex with oseltamivir,[16] it was shown that the C5 amino group of oseltamivir forms strong charge−charge-type hydrogen bonding interactions with residues Asp151 and Glu119. Since our N1hA model was built by using 2HU0 with open conformation as the template, the docking results show that the C5 amino group of oseltamivir is far from Asp151 and Glu119. The distinguishing difference between 2HU0 template and N1hA model enzyme lies in that Tyr347 in 2HU0 is replaced by Asn329 in N1hA, such that the former makes an additional hydrogen bonding with the carboxyl anion group of oseltamivir while the latter does not.

To validate numerically the results obtained from molecular docking with FlexX, FMO calculations were performed for a comparative study of binding energies and pair interaction energies of oseltamivir with the binding sites of N1hA model and 2HU0 template, respectively. The calculated binding energies of oseltamivir with the selected active site of N1hA model and open conformational 2HU0 are −753.12 and −1014.09 kcal/mol, respectively. Large negative values demonstrate high stabilization contributions coming from the interactions of oseltamivir with the specific amino acid residues located in the active sites of both N1hA

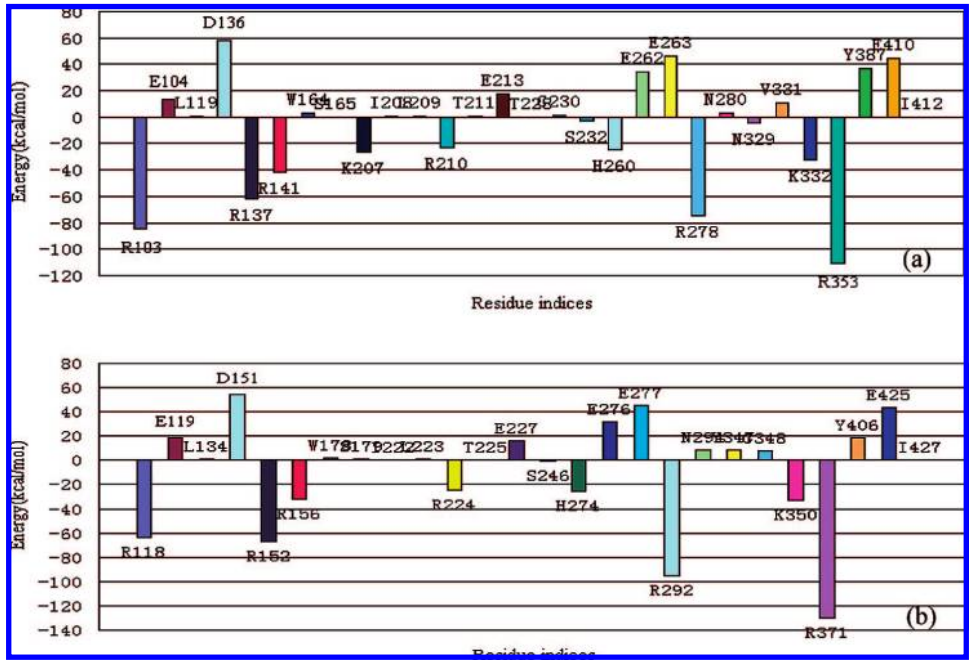

Figure 6. Pair interaction energies calculated at the FMO-RHF/6−31G* level. (a) Docking oseltamivir into the active site of the N1hA model. (b) Oseltamivir binding with the active site in the crystallographic structure of 2HU0.

model and 2HU0 template enzymes and therefore suggest that oseltamivir is most likely a potent inhibitor of N1hA as well.

Pair interaction energy analysis was performed simultaneously to detect important interactions between ligand and specific amino acid residues. The calculated pair interaction energies of oseltamivir with individual residue fragments in the selected active sites of N1hA model and 2HU0 template are illustrated in the Figure 6a and 6b, respectively. Of all amino acid residues located in the active site of N1hA model, the Arg103, Arg137, Arg141, Arg278, and Arg353 (corresponding Arg118, Arg152, Arg156, Arg292, and Arg371 in the case of 2HU0) fragments are responsible for the largest contributions to the stabilization because of the formation of hydrogen bonding of electrostatic interaction type between the carboxyl anion group of oseltamivir and these residues with positive charge centers (see also Figure 5). The second largest contributors are Lys207, Arg210, His260, and Lys332 residues (corresponding Arg224, His274, and Lys350 in the case of 2HU0), which contribute to the electrostatic and nonpolar interactions. The largest destabilization arise from Asp136, Glu262, Glu263, Tyr387, and Glu410 (corresponding Asp151, Glu276, Glu277, Tyr406, and Glu425 in the case of 2HU0) residues because of the electrostatic interactions between the electronegative C4 carbonyl group of oseltamivir and the electronegative residues given above. Generally speaking, most of the binding sites and the trends of pair interaction energies are similar to each other. Such a similarity lends credit to the reliability of the homology modeling and shows the reasonableness of the selected (default) parameters in FlexX docking in the present work.

3D QSAR CoMFA Modeling. Table 1 summarizes the PLS analysis results for two types of 3D QSAR CoMFA modeling, Model-Fit-RF and Model-Dock-RF. Model-Fit has cross-validated coefficients (q^2) up to 0.451 and regression coefficients (r^2) up to 0.789 with the optimal number of components (ONC) two. Model-Dock has cross-validated coefficients (q^2) up to 0.663 and regression coefficients (r^2) up to 0.981 with ONC four. We note that ONC in a previous CoMSIA study,[18] based on N2 and N9 structural information, was up from nine to fourteen with sample number of thirty-two. The present 3D QSAR CoMFA modeling leads to two best models based on the region focusing (RF) method. Table 2 presents the predicted activity values and their residues from the experimentally measured pIC_{50} for the twenty-seven inhibitors selected. The numerical results in Table 2 show that our present CoMFA models have good predictive power. The predicted pIC_{50} values are generally in good agreement with the experiment data with small residuals. Compared with the results of Model-Fit, Model-Dock has higher cross-validated coefficients, leading credits to the reliability of active conformations obtained by FlexX.

Table 1. Summary of Results for the CoMFA Binding Models with Two Types of Molecular Alignments[a]

		q^2	n	r^2	F	SE	$E\%$	$S\%$
Model-Fit	STD	0.355	2	0.787	44.294	0.644	9.3	90.7
	RF	0.451	2	0.789	44.962	0.640	15.2	84.8
Model-Dock	STD	0.435	6	0.987	250.777	0.175	14.3	85.7
	RF	0.663	4	0.981	278.004	0.203	15.3	84.7

[a] STD, standard scaling. RF: region focusing (StDev × coefficient). n = the optimal number of components (ONC) to be used in the final analysis. F = the ratio of r^2 to 1.0 - r^2 (explained to unexplained), weighted so that the fewer the explanatory properties and the more the values of the target property, the higher the F ratio; SE, standard error; $E\%$, contribution of electrostatic field. $S\%$, contribution of steric field.

Qualitatively, the steric and electrostatic field derived from the conventional atom-fit molecular alignment is only ligand-centered; however, that derived from the docking-based active conformation alignment is a joint product taking both the ligands and the active site into account, which are predominantly responsible for the difference of two set of CoMFA modeling. A recent study[63] of HMG-CoA reductase showed a similar result as well. Figure 7 shows the two sets of alignments. In the following CoMFA analysis, we will

Table 2. Experimental, Predicted Activities and Residual Values by the Conventional Fit-Atom Molecular Alignment and Molecular Alignment Using Active Conformation Achieved by Molecular Docking

	pIc_{50}				
	Model-Fit-RF			Model-Dock-RF	
>compd	actual	predicted	residues	predicted	residues
1	3.70	4.249	−0.549	3.714	−0.014
2	5.20	4.804	0.396	5.217	−0.017
3	5.43	5.271	0.159	5.354	0.076
4	5.70	5.466	0.234	5.785	−0.085
5	6.74	6.452	0.288	6.725	0.015
6	6.52	6.593	−0.073	6.821	−0.301
7	6.70	6.462	0.238	6.847	−0.147
8	8.00	7.242	0.758	7.606	0.394
9	9.00	7.351	1.649	8.591	0.409
10	7.80	8.013	−0.213	8.109	−0.309
11	6.82	7.286	−0.466	7.024	−0.204
12	9.00	8.658	0.342	9.174	−0.174
13	5.70	6.345	−0.645	5.897	−0.197
14	6.65	6.613	0.037	6.483	0.167
15	5.66	5.88	−0.22	5.593	0.067
16	7.66	7.219	0.441	7.751	−0.091
17	7.22	7.20	0.02	7.241	−0.021
18	6.28	6.438	−0.158	6.215	0.065
19	5.82	6.101	−0.281	5.738	0.082
20	7.00	6.238	0.762	6.859	0.141
21	4.60	5.683	−1.083	4.663	−0.063
22	6.85	8.058	−1.208	6.577	0.273
23	8.74	8.112	0.628	8.726	0.014
24	8.34	9.239	−0.899	8.401	−0.061
25	5.51	5.126	0.384	5.347	0.163
26	5.47	5.782	−0.312	5.725	−0.255
27	5.64	5.868	−0.228	5.568	0.072

focus only on the results from Model-Dock. In addition, with this model, the bioaffinity of compound RWJ-270201 discussed in a recent experimental study[64] is predicted to be 8.01, which compares favorably with the experimental bioaffinity 8.47 testing by an essay of A/duck/Alberta/35/76 (H1N1).

Figure 8 shows the 3D steric and electrostatic maps derived from the CoMFA model using docking-based active conformation alignment. Oseltamivir (compound no. 9), one of the most potent N1 inhibitors at present, is shown to be inside the fields. The contribution of the steric field to the activity is 84.7% and that of the electrostatic field is 15.3%. The 3D contour maps show that the changes of molecular fields are

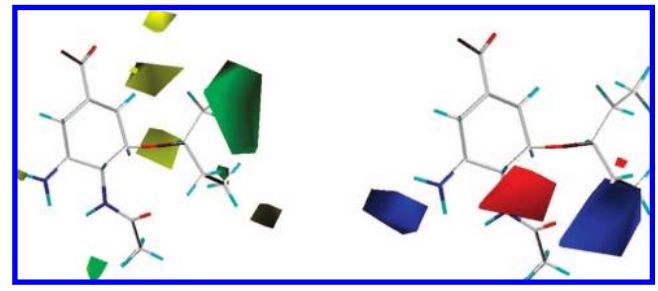

Figure 8. Steric and electrostatic maps from the CoMFA model using docking-based active conformation alignment. Oseltamivir (compound no. 9) is shown inside the field. Sterically favored areas (80% contribution) are represented by green polyhedra. Sterically disfavored areas (20% contribution) are represented by yellow polyhedra. Blue contours (80% contribution) encompass regions where an increase of positive charge will enhance affinity; whereas red contours (20% contribution) encompass regions where an increase of negative charge will enhance affinity.

associated with the differences of biological activity. The steric fields are in green and yellow. The region of green contour suggests that more bulky substituents in these positions will improve the biological activity, while the yellow region indicates that an increased steric bulk is unfavorable for the inhibitory activity. The *n*-amyl in the R group of compound no. 9 in Figure 2 is located in the green region resulting in a higher pIC_{50} value of 9.0. The *n*-heptyl in the R group of compound no. 10 is extended to the yellow region, and therefore, its pIC_{50} value is reduced to 7.8.

Figure 9A shows the steric contours projected over the solvent accessible (Connolly) topological surface (MOLCAD generated) of the active site of the N1hA model. The yellow region (a) flanks the Arg278, Glu216, Tyr387, and Arg353 amino acid residues, while the yellow regions (b) the ILE208 amino acid residue and (c) the Arg141 amino acid residue. The green region (d) around the C3 pentyloxy group of oseltamivir in the CoMFA contour map is located on the surface of residues toward outside of the active site pocket. The green regions (e and f) are on the corner with large room of the active site cavity. These steric field distributions of CoMFA model (Model-Dock) guarantee a large percentage of the ligand volume buried inside the binding pocket.[31,65]

Figure 9B shows the electrostatic contours projected over the electrostatic potential surface (MOLCAD generated) of the active site of the N1hA model. The CoMFA electrostatic fields in blue suggest that the positively charged substitutions

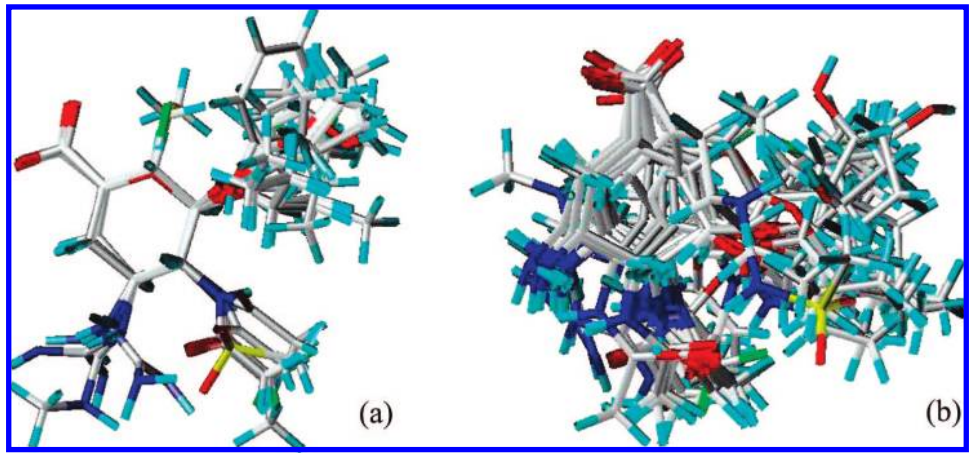

Figure 7. Schematic diagram of two sets of molecular alignments: (a) conventional atom-fit molecular alignment, and (b) docking-based active conformation alignment.

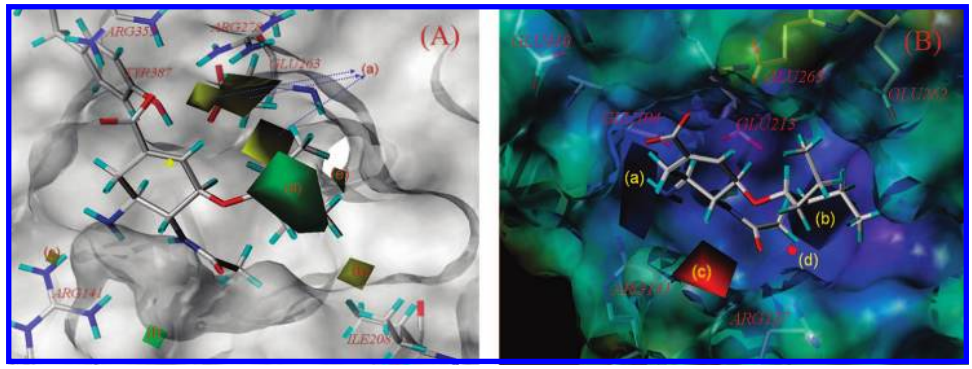

Figure 9. (A) Steric contours projected over the solvent accessible (Connolly) topological surface (MOLCAD generated) of the active site of the N1hA model; (B) Electrostatic contours projected over the electrostatic potential surface (blue, negative potential; red/brown, positive potential) of the active site of the N1hA model.

may increase the inhibitory activity, while the red region indicates that a high electronic density may play a favorable role in inhibitory potency. The blue regions of the MOLCAD generated contour maps represent a negative electrostatic potential, the red regions represent a positive electrostatic potential. The large blue regions (a and b) of CoMFA contour match well with the blue (highly electronegative) surface of the binding sites provided by the carboxyl anion groups of the Glu104, Glu213, Glu262 and Glu263, and Glu410 amino acid residues. The red regions (c and d) of the CoMFA contour match well with the red (electropositive) surface of the binding sites provided by the guanidine groups of the Arg141 and Arg137 amino acid residues. The comparison of the two types of contour maps shows that the present docking-based CoMFA model generally agrees well with the active sites of the N1hA model built in the present work.

Molecular Dynamics Simulations. A series of molecular dynamics simulations has been performed to examine the flexibility of the loop-150 of the N1aA and N1hA. We focus on the dynamic behaviors of the conserved amino acid residues and the loop-150 residues in the active site. The conserved amino acid residues consist of Arg103, Trp164, Ile208, Gln212, Glu213, Arg278, Arg353, and Tyr387 in N1hA (corresponding Arg118, Trp179, Ile223, Gln227, Glu228, Arg293, Arg368, and Tyr402 in the case of N1aA). The loop-150 residues consist of Glu104, Gln121, Gly132, Thr133, Val134, Lys135, Asp136, Arg137, and Arg141 (corresponding Glu119, Gln136, Gly147, Thr148, Val149, Lys150, Asp151, Arg152, and Arg156 in the case of N1aA).

The rmsd from the starting structure for all backbone Cα atoms of the entire enzyme, the conserved residues and the loop-150 residues were calculated, respectively, over the simulation time by using the PTRAJ module for the both N1hA and N1aA. The plots of the evolution of rmsd with simulation time were illustrated in Figure 10 (a) for the case of N1aA and (b) for the case of N1hA. The black lines represent the rmsd of backbone Cα atoms of the entire enzyme between the simulated trajectories and the initial structure of the N1aA and N1hA. The red lines represent the rmsd of backbone Cα atoms of the conserved amino acid residues between the simulated trajectories and the initial structure of the N1aA and N1hA. The green lines represent the rmsd of backbone Cα atoms of the loop-150 residues between the simulated trajectories and the initial structure of the N1aA and N1hA.

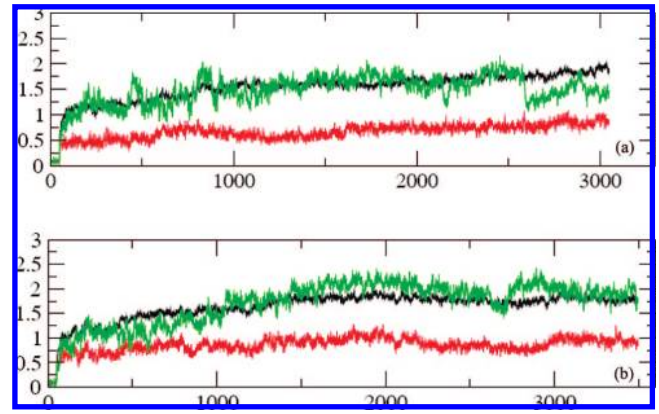

Figure 10. Plots of the root-mean-square deviation (rmsd, in Å) versus the MD simulation time in the MD-simulated structures. (a) The black line represents the rmsd of backbone Cα atoms of the entire enzyme, the red line represents the corresponding rmsd of the conserved amino acid residues, and the green line represents the corresponding rmsd of the loop-150 residues, respectively, between the simulated trajectories and the crystallographic structure of the N1aA. (b) The black line represents the rmsd of backbone Cα atoms of the entire enzyme, the red line represents the corresponding rmsd of the conserved amino acid residues, and the green line represents the corresponding rmsd of the loop-150 residues, respectively, between the simulated trajectories and the homology modeling of the N1hA.

It is obvious from Figure 10a and b that the black lines have almost achieved a dynamic convergence at around 3000 ps in the both N1aA and N1hA systems. The red lines have rapidly arrived at a dynamic convergence with the rmsd around 1 Å. However, the green lines fluctuate strongly from 1.5 to 2.25 Å without an obvious converged trend within 3000 ps. Which implies that the loop-150 region is much more fluctuant than that of conserved amino acid residues, and therefore, the flexibility of the loop-150 coincides with the observation of open and close conformations in the loop-150 region.[22] Figure 11a and b shows the schematic fluctuations behaved by the conserved amino acid residues and the loop-150 residues, respectively, in the both N1aA and N1hA.

CONCLUSION

The structural information of avian influenza N2 or N9 has been used as template in the previous homology modeling and 3D-QSAR studies. In the present study, the recent X-ray crystallographic structure of N1 subtype of avian influenza A/H5N1 (2HU0 with open conformation)

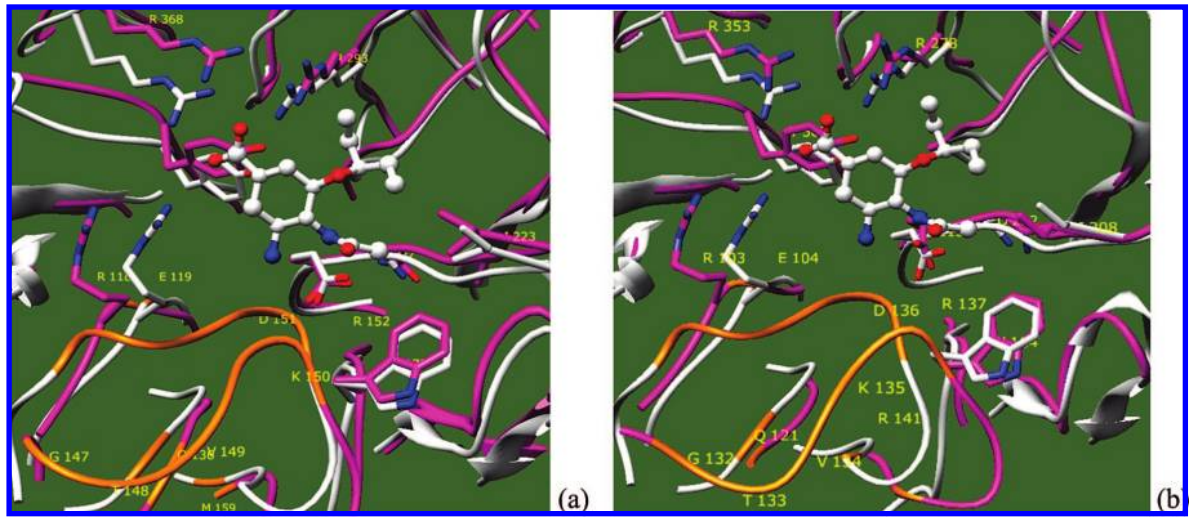

Figure 11. (a) Superposition between the crystallographic structure and the simulated structure after MD simulation in the case of N1aA. (b) The superposition between the initial homology modeling structure and the simulated structure after MD simulation in the case of N1hA. White ribbons represent the structures before MD; magenta ribbons represent the structure after MD, and orange ribbons represent the loop-150 regions.

was selected, for the first time, as the template to build a 3D model of N1hA of A/PR/8/34 H1N1 from the target sequence (ID P03468). Binding energies and pair interaction energies of Tamiflu with the active sites of the N1hA model and the 2HU0 template were calculated at the FMO-RHF/6-31G* level. Most of the binding sites and the trends of pair interaction energies are similar in both cases, which lends credit to the reliability of the homology modeling and illustrates the reasonableness of the selected parameters in molecular docking in the present work. Twenty-seven N1 inhibitors, whose biological inhibitory activities have been carefully determined by one research group with human A/PR/8/34 (H1N1), were selected as a training data set for the subsequent 3D QSAR CoMFA modeling. Two 3D QSAR CoMFA predictive models, that is, Model-Fit and Model-Dock, were established for the N1hA inhibitors by using two different molecular alignment strategies. The latter was found to perform better than the former in the present study with excellent cross-validated q^2 and non-cross-validated r^2 values. Finally, the comparison of two types of contour maps from ligands-based and receptor-based, respectively, shows that the present docking-based CoMFA model generally matches well with the active site of the N1hA model. All the above information concerning the binding sites is valuable for future virtual screening of the N1hA inhibitors. The 3D QSAR CoMFA predictive model (Model-Dock) can be used jointly as a complementary score function in the process of virtual screening as was proposed in our previous study.[63] In addition, the flexibility of the loop-150 in the N1aA and N1hA has been examined by a series of MD simulations. The evolution of trajectories of the amino acids residues located in the loop-150 with the time shows that the loop-150 region is indeed much more fluctuant than that of conserved amino acid residues. The flexibility of the loop-150 is most like responsible for the observation of open and close conformations in the loop-150 region,[22] and therefore, a long-time MD simulation is needed to simulate the conformation change observed in the loop-150 in a further study.

ACKNOWLEDGMENT

This work was supported by the National Basic Research Program of China (Nos. 2007CB116302, 2004CB719902), the Natural Science Foundation of China (Nos. 20672041, 20525311, and 200423002), the Program for New Century Excellent Talents in University of China (NCET-06-0673), the Science and Technology Research Project of Ministry of Education (No.106116), and China National Technology Platform (No.2005DKA64001).

REFERENCES AND NOTES

(1) Zambon, M. C. Epidemiology and pathogenesis of influenza. *J. Antimicrob. Chemother.* **1999**, *44*, 3–9.

(2) (a) Sauter, N. K.; Hanson, J. E.; Glick, G. D.; Brown, J. H.; Crowther, R. L.; Park, S. J.; Skehel, J. J.; Wiley, D. C. Binding of influenza virus hemagglutinin to analogs of its cell-surface receptor, sialic acid: Analysis by proton nuclear magnetic resonance spectroscopy and x-ray crystallography. *Biochemistry* **1991**, *31*, 9609–9621. (b) Mammen, M.; Dahmann, G.; Whitesides, G. Effective inhibitors of hemagglutination by influenza virus synthesized from polymers having active ester groups. Insight into mechanism of inhibition. *J. Med. Chem.* **1995**, *38*, 4179–4190.

(3) (a) Colman, P. M. In *The Influenza Viruses: Influenza Virus Neuraminidase, Enzyme and Antigen*; Krug, R. M., Ed.; Plenum Press: New York, 1989; pp 175−218. (b) Colman, P. M. Influenza virus neuraminidase: Structure, antibodies, and inhibitors. *Protein Sci.* **1994**, *3*, 1687–1696.

(4) Hay, A. J.; Wolstenholme, A. J.; Skehel, J. J.; Smith, M. H. The molecular basis of the specific anti-influenza action of amantadine. *EMBO J.* **1985**, *4*, 3021–3024.

(5) Hastings, J. C.; Selnick, H.; Wolanski, B.; Tomassini, J. E. *Antimicrob. Agents Chemother.* **1996**, *40*, 1304–1307.

(6) Wang, T.; Wade, R. C. Comparative binding energy (COMBINE) analysis of influenza neruaminidase inhibitor complexes. *J. Med. Chem.* **2001**, *44*, 961–967.

(7) Bender, C.; Hall, H.; Huang, J.; Klimov, A.; Cox, N.; Hay, A.; Gregory, V.; Cameron, K.; Lim, W.; Subbarao, K. Characterization of the surface proteins of influenza A (H5N1) viruses isolated from humans in 1997−1998. *Virology* **1999**, *254*, 115–123.

(8) World Health Organization Global Influenza Program Surveillance Network. Evolution of H5N1 avian influenza viruses in Asia. *Emerg. Infect. Dis.* **2005**, *11*, 1515−1521.

(9) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. Improved sensitivity of profile searches through the use of sequence weights and gap excision. *Comput. Appl. Biosci.* **1994**, *10*, 19–29.

(10) Colman, P. M.; Varghese, J. N.; Laver, W. G. Structure of the catalytic and antigenic sites in influenza virus neuraminidase. *Nature* **1983**, *303*, 41–44.

(11) Wilson, I. A.; Skehel, J. J.; Wiley, D. C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **1981**, *289*, 366–373.
(12) Hayden, F. G. Amantadine and rimantadine-clincal aspects In *Antiviral Drug Resistance*; Richman, D. D., Ed; John Wiley, Sons Ltd.: New York, 1996; Vol. 5, pp 9−77.
(13) Trampuz, A.; Prabhu, R. M.; Smith, T. F.; Baddour, L. M. Avian influenza: A new pandemic threat. *Mayo Clin. Proc.* **2004**, *79*, 523–530.
(14) von Itzstein, M.; Wu, W.-Y.; Kok, G. B.; Pegg, M. S.; Dyason, J. C.; Phan, T. V.; Smythe, M. L.; White, H. F.; Oliver, S. W.; Colman, 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.
(15) 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.
(16) Chand, P.; Babu, Y. S.; Bantia, S.; Rowland, S.; Dehghani, A.; Kotian, P. L.; Hutchison, T. L.; Ali, S.; Brouillette, W.; El-Kattan, Y.; Lin, T.-H. Syntheses and neuraminidase inhibitory activity of multisubstituted cyclopentane amide derivatives. *J. Med. Chem.* **2004**, *47*, 1919–1929.
(17) Chand, P.; Kotian, P. L.; Dehghani, A.; El-Kattan, Y.; Lin, T.-H.; Hutchison, T. L.; Babu, Y. S.; Bantia, S.; Elliott, A. J.; Montgomery, J. A. Systermatic structure-based design and stereoselective synthesis of novel multisubstituted cyclopentane derivatives with potent anti-influenza activity. *J. Med. Chem.* **2001**, *44*, 4379–4392.
(18) Yi, X.; Guo, Z. R.; Chu, F. M. Study on molecular mechanism and 3D-QSAR of influenza neuraminidase inhibitors. *Bioorg. Med. Chem.* **2002**, *11*, 1465–1474.
(19) Williams, M. A.; Lew, W.; Mendel, D. B.; Tai, C. Y.; Escarpe, P. A.; Laver, W. G.; Stevens, R. C.; Kim, C. U. Structure−activity relationship of carbocyclic influenza neuraminidase inhibitors. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1837–1842.
(20) 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.
(21) Verma, R. P.; Hansch, C. A QSAR study on influenza neuraminidase inhibitors. *Bioorg. Med. Chem.* **2006**, *14*, 982–99.
(22) 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*, 45–49.
(23) Brouillette, W. J.; Bajpai, S. N.; Ali, S. M.; Velu, S. E.; Atigadda, V. R.; Lommer, B. S.; Finley, J. B.; Luo, M.; Air, G. M. Pyrrolidinobenzoic acid inhibitors of influenza virus neuraminidase: modifications of essential pyrrolidinone ring substituents. *Bioorg. Med. Chem.* **2003**, *11*, 2739–2749.
(24) Ives, J.; Carr, J.; Roberts, N. A.; Tai, C. Y.; Mendel, D. B.; Kelly, L.; Lambkin, R.; Oxford, J. An oseltamivir treatment selected influenza A/Wuhan/359/95 virus with a E119V mutation in the neuraminidase gene has reduced infectivity in vivo. *J. Clin. Virol.* **2000**, *18*, 251–269.
(25) 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.
(26) Carr, J.; Ives, J.; Kelly, L.; Lambkin, R.; Oxford, J.; Mendel, D.; Tai, L.; Roberts, N. Influenza virus carrying neuraminidase with reduced sensitivity to oseltamivir carboxylate has altered properties in vitro and is compromised for infectivity and replicative ability in vivo. *Antiviral Res.* **2002**, *54*, 79–88.
(27) Ward, P.; Small, I.; Smith, J.; Suter, P.; Dutkowski, R. Oseltamivir (Tamiflu) and its potential for use in the event of an influenza pandemic. *J. Antimicrob. Chemother.* **2005**, *55*, 5–21.
(28) Bernardi, F.; Bottoni, A.; Garavelli, M. Exploring organic chemistry with DFT: radical, organo-metallic, and bio-organic applications. *Quant. Struct.−Act. Relat.* **2002**, *21*, 128–148.
(29) Cramer, R. D., III; Bunce, J. D. Comparative molecular field analysis (CoMFA) 1.Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* **1988**, *110*, 5959–5967.
(30) National Centerfor Biotechnology Information. http://www.ncbi.nlm-.nih.gov (accessed May 2005).
(31) Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res.* **2003**, *31*, 3381–3385.
(32) Guex, N.; Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modelling. *Electrophoresis* **1997**, *18*, 2714–2723.
(33) *Sybyl 7.0*; Tripos Inc.: St. Louis, MO, 2003; http://www.tripos.com.
(34) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470–489.
(35) Kramer, B.; Rarey, M.; Lengauer, T. Evaluation of the FlexX incremental construction algorithm for protein-ligand docking. *Proteins: Struct. Funct. Genet.* **1999**, *37*, 228–234.
(36) Fukuzawa, K.; Kitaura, K.; Nakata, K.; Kaminuma, T.; Nakano, T. Workshop 1.5 Fragment molecular orbital study of the binding energy of ligands to the estrogen receptor. *Pure Appl. Chem.* **2003**, *75*, 2405–2410.
(37) Nemoto, T.; Fedorov, D. G.; Uebayasi, M.; Kanazawa, K.; Kitaura, K.; Komeiji, Y. Ab initio fragment molecular orbital (FMO) method applied to analysis of the ligand−protein interaction in a pheromone-binding protein. *Comput. Biol. Chem.* **2005**, *29*, 434–439.
(38) Yamagishi, K.; Yamamoto, K.; Yamada, S.; Tokiwa, H. Functions of key residues in the ligand-binding pocket of vitamin D receptor: Fragment molecular orbital−interfragment interaction energy analysis. *Chem. Phys. Lett.* **2006**, *420*, 465–468.
(39) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. Fragment molecular orbital method: an approximate computational method for large molecules. *Chem. Phys. Lett.* **1999**, *313*, 701–706.
(40) Komeiji, Y.; Inadomi, Y.; Nakano, T. PEACH 4 with ABINIT-MP: A general platform for classical and quantum simulations of biological molecules. *Comput. Biol. Chem.* **2004**, *28*, 155–161.
(41) Nakano, T.; Kaminuma, T.; Sato, T.; Fukuzawa, K.; Akiyama, Y.; Uebayasi, M.; Kitaura, K. Fragment molecular orbital method: use of approximate electrostatic potential. *Chem. Phys. Lett.* **2002**, *351*, 475–480.
(42) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. General atomic and molecular electronic-structure system. *J. Comput. Chem.* **1993**, *14*, 1347–1363.
(43) Zhang, L. J.; Williams, M. A.; Mendel, D. B.; Escarpe, P. A.; Kim, C. U. Synthesis and activity of C_2-substituted analogs of influenza neuraminidase inhibitor gs4071. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1847–1850.
(44) Taubenberger, J. K.; Reid, A. H.; Krafft, A. E.; Bijwaard, K. E.; Fanning, T. G. Initial genetic characterization of the 1918 "Spanish" influenza virus. *Science* **1997**, *275*, 1793–1796. (b) Baez, M.; Taussig, R.; Zazra, J. J.; Young, J. F.; Palese, P.; Reisfeld, A.; Skalka, A. M. Complete nucleotide sequence of the influenza A/PR/8/34 virus NS gene and comparison with the NS genes of the A/Udorn/72 and A/FPV/Rostock/34 strains. *Nucleic Acids Res.* **1980**, *8*, 5845–5858.
(45) Gasteiger, J.; Marsili, M. Iterative partial equalization of orbital electronegativity−A rapid access to atomic charges. *Tetrahedron* **1980**, *36*, 3219–3228.
(46) Clark, M.; Cramer, R. D., III; van Opdenbosch, N. The Tripos force field. *J. Comput. Chem.* **1989**, *10*, 982–1012.
(47) Powell, M. J. D. Restart procedures for the conjugate gradient method. *Math. Programming* **1977**, *12*, 241–254.
(48) Nilsson, J. Multiway Calibration in 3D QSAR: Applications to dopamine receptor ligands. 1998. http://www.ub.rug.nl/eldoc/dis/science/j.nilsson.
(49) Wold, S.; Rhue, A.; Wold, H.; Dunn, W. J. I. The covariance problem in linear regression. The partial least squares (PLS) approach to generalized inverses. SIAM. *J. Sci. Stat. Comput.* **1984**, *5*, 735–743.
(50) Wold, S.; Albano, C.; Dunn, W. J., III; Edlund, U.; Esbensen, K.; Geladi, P.; Hellberg, S.; Johanson, E.; Lindberg, W.; Sjostrom, M. Multivariate data analysis in chemistry. *NATO ASI Ser., Ser. C* **1984**, *138*, 17–95.
(51) Clark, M.; Cramer, R. D., III. The probability of chance correlation using partial least squares (PLS). *Quant. Struct.−Act. Relat.* **1993**, *12*, 137–145.
(52) Bush, B. L.; Nachbar, R. B. Sample-distance partial least-squares PLS optimized for many variables, with application to CoMFA. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 587–619.
(53) Lindgren, F.; Geladi, P.; Rä nnar, S.; Wold, S. Interactive variable selection (IVS) for PLS. Part 1:Theory and algorithms. *J. Chemom.* **1994**, *8*, 349–363.
(54) Waldherr-Teschner, M.; Goetze, T.; Heiden, W.; Knoblauch, M.; Vollhardt, H.; Brickmann, J. MOLCAD−Computer aided visualization and manipulation of models in molecular science In *Advances in Scientific Visualization*; Post, F. H., Hin, A. J. S., Eds.; Springer: Heidelberg, Germany,1992, 2005; Vol. 5, pp 8−67.
(55) *AMBER*, version 8; University of California: San Francisco, CA, 2004.
(56) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.
(57) Darden, T.; York, D. L. Pedersen, Particle mesh Ewald: An N log-(N) method for Ewald sums in large system. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(58) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(59) Sagui, C.; Darden, T. A. P3M and PME: A comparison of the two methods In *Simulation and Theory of Electrostatic Interactions in Solution*; AIP Conference Proceedings, Melville, 1999; Pratt, L. R.; Hummer, G., Eds; American Institute of Physics: New York, 1999; pp 104−113.

(60) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical intergration of the Cartesian equations of motion of a system with constraints: Molecular synamics of *n*-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(61) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Comput. Phys.* **1984**, *81*, 3684–3690.

(62) Marti-Renom, M. A.; Stuart, A. C.; Fiser, A.; Sanchez, R.; Melo, F.; Sali, A. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291–325.

(63) Zhang, Q. Y.; Wan, J.; Xu, X.; Yang, G. F.; Ren, Y. L.; Liu, J. J.; Wang, H.; Guo, Y. Structure-based rational quest for potential novel inhibitors of human HMG-CoA reductase by combining CoMFA 3D QSAR modeling and virtual screening. *J. Comb. Chem.* **2007**, *9*, 131–138.

(64) Govorkova, E. A.; Leneva, I. A.; Goloubeva, O. G.; Bush, K.; Webster, R. G. Comparison of efficacies of RWJ-270201, zanamivir, and oseltamivir against H5N1, H9N2, and other avian influenza viruses. *Antimicrob. Agents Chemother.* **2001**, *45*, 2723–2732.

(65) Sulpizi, M.; Folkers, G.; Rothlisberger, U.; Carloni, P.; Scapozza, L. Applications of density functional theory-based methods in medicinal chemistry. *Quant. Struct.−Act. Relat.* **2002**, *21*, 173–181.

CI800041K