articles

molecular pharmaceutics

In Silico Identification of the Potential Drug Resistance Sites over 2009 Influenza A (H1N1) Virus Neuraminidase

Huanxiang Liu,*[,†,‡] Xiaojun Yao,[‡,§] Chengqi Wang,[§] and Jian Han[‡]

School of Pharmacy, Department of Chemistry, and Key Lab of Preclinical Study for New Drugs of Gansu Province, Lanzhou University, Lanzhou 730000, China

Received February 18, 2010; Revised Manuscript Received April 20, 2010; Accepted April 26, 2010

Abstract: The outbreak and high speed global spread of the new strain of influenza A (H1N1) virus in 2009 poses a serious threat to the general population and governments. At present, the most effective drugs for the treatment of 2009 influenza A (H1N1) virus are neuraminidase inhibitors: mainly oseltamivir and zanamivir. The use of these two inhibitors will undoubtedly increase, and therefore it is more likely that drug-resistant influenza strains will arise. The identification of the potential resistance sites for these drugs in advance and the understanding of corresponding molecular basis to cause drug resistance are no doubt very important to fight against the new resistant influenza strains. In this study, first, the complexes of neuraminidase with the substrate sialic acid and two inhibitors oseltamivir and zanamivir were obtained by fitting them to the 3D structure of 2009 influenza A (H1N1) neuraminidase obtained by homology modeling. By using these complexes as the initial structures, molecular dynamics simulation and molecular mechanics generalized Born surface area (MM-GBSA) calculations were performed to identify the residues with significant contribution to the binding of substrate and inhibitors. By analyzing the difference of interaction profiles of substrate and inhibitors, the potential drug resistance sites for two inhibitors were identified. Parts of the identified sites have been verified to confer resistance to oseltamivir and zanamivir for influenza virus of the past flu epidemic. The identified potential resistance sites in this study will be useful for the development of new effective drugs against the drug resistance and avoid the situation of having no effective drugs to treat new mutant influenza strains.

Keywords: 2009 H1N1 influenza A virus; drug resistance; neuraminidase inhibitors; molecular dynamics simulation; molecular mechanics generalized Born surface area (MM-GBSA)

1. Introduction

Since March 2009, the outbreak of a new strain of influenza A (H1N1) virus infection in humans has caused a global flu epidemic. At the end of 2009, from the World Health Organization (WHO) report, more than 208 countries and overseas territories or communities worldwide have reported laboratory confirmed cases of pandemic influenza H1N1 2009, including at least 12799 deaths.[1] Its continual and high speed global spread has heightened the awareness of both the general population and governments to the threat of influenza virus.

The virus currently spreading as a new subtype A/H1N1 is a recombined virus by human, swine, and avian influenza viruses with a high transmissible ability among human beings.[2] This new strain confers resistance to M2 channel inhibitors amantadine (AMT) and rimantadine (RMT) (Figure 1).[3] Fortunately, the neuraminidase (NA) inhibitors oseltamivir and zanamivir (Figure 1) are still effective against the new virus. Among them, oseltamivir (OTV) is recommended by WHO and the Centers for Disease Control and Prevention

* Author to whom correspondence should be addressed. Current address: School of Pharmacy, Lanzhou University, Lanzhou 730000, China. E-mail: hxliu@lzu.edu.cn. Tel: +86-931-8915686. Fax: +86-931-8915686.

† School of Pharmacy.

‡ Key Lab of Preclinical Study for New Drugs of Gansu Province.

§ Department of Chemistry.

(1) http://www.who.int/csr/don/2010_01_08/en/index.html, accessed Jan 9, 2010.

10.1021/mp100041b
Published on Web 04/26/2010

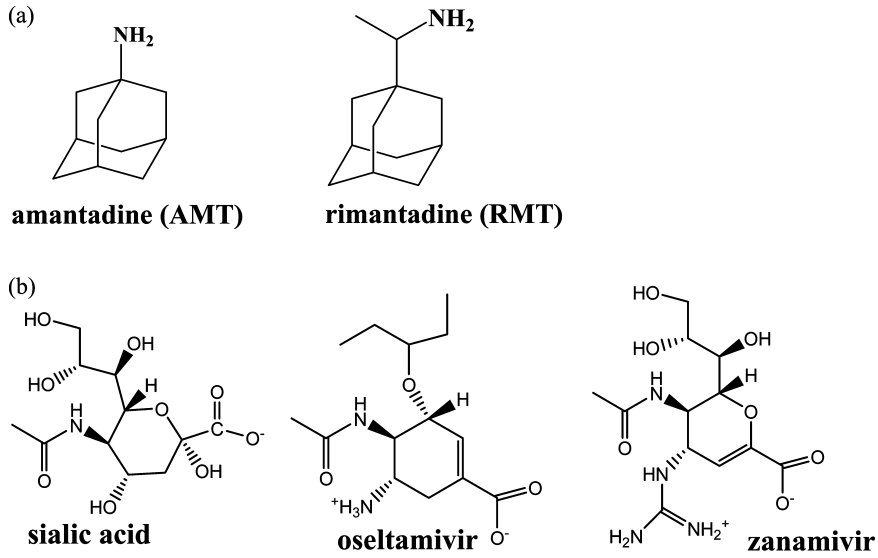

Figure 1. The molecular schemes of the ligands: (a) M2 channel inhibitors; (b) neuraminidase substrate and inhibitors.

(CDC) as the most effective antiviral agent for the treatment of the infected patients. While oseltamivir is widely used in persons infected with pandemic (H1N1) 2009 virus, resistance has been observed recently.[4] Faced with the possible serious circumstance and future pandemic caused by drug resistance, it will be very useful if we can know the potential drug resistance sites of the effective neuraminidase drugs in advance. Information about the potential drug resistance sites can provide some insights to design a new effective medicine against the new resistant influenza strains and would be very significant to avoid the situation of having no effective drugs to treat new mutant influenza strains.

Generally, in the presence of inhibitors, neuraminidase variants are under selection pressure and those not binding to the inhibitors but still maintaining catalytic activity would become dominant in the virus population because of the high replication rate and low fidelity of the viral replication. Thus, it is widely accepted that drug resistance is more easily caused by mutations that significantly impair the neuraminidase's interaction with drugs but not with substrates, i.e., if one residue interacts more favorably with a drug than with the substrate, mutation of this residue may not affect the function of the neuraminidase but may impair the binding of the drug and thus cause resistance to the drug.[5,6]

Recently, molecular dynamics (MD) simulation combined with binding free energy calculation has been developed to investigate the molecular mechanism of drug resistance and evaluate the potency of an inhibitor to combat resistance before it goes through the long and costly procedure of drug development.[7−10] The methods based on protein structure are particularly useful for mechanistic study and *de novo* prediction of resistant mutants with little or no prior knowledge. Among the molecular dynamics simulation-based

(2) Garten, R. J.; Davis, C. T.; Russell, C. A.; Shu, B.; Lindstrom, S.; Balish, A.; Sessions, W. M.; Xu, X.; Skepner, E.; Deyde, V.; Okomo-Adhiambo, M.; Gubareva, L.; Barnes, J.; Smith, C. B.; Emery, S. L.; Hillman, M. J.; Rivailler, P.; Smagala, J.; de Graaf, M.; Burke, D. F.; Fouchier, R. A.; Pappas, C.; Alpuche-Aranda, C. M.; López-Gatell, H.; Olivera, H.; López, I.; Myers, C. A.; Faix, D.; Blair, P. J.; Yu, C.; Keene, K. M.; Dotson, P. D., Jr; Boxrud, D.; Sambol, A. R.; Abid, S. H.; St George, K.; Bannerman, T.; Moore, A. L.; Stringer, D. J.; Blevins, P.; Demmler-Harrison, G. J.; Ginsberg, M.; Kriner, P.; Waterman, S.; Smole, S.; Guevara, H. F.; Belongia, E. A.; Clark, P. A.; Beatrice, S. T.; Donis, R.; Katz, J.; Finelli, L.; Bridges, C. B.; Shaw, M.; Jernigan, D. B.; Uyeki, T. M.; Smith, D. J.; Klimov, A. I.; Cox, N. J. Antigenic and genetic characteristics of swine-origin 2009 A (H1N1) influenza viruses circulating in humans. *Science* **2009**, *325*, 197–201.
(3) CDC, Antiviral Drugs and H1N1 Flu (Swine Flu). Available from: http://www.cdc.gov/h1n1flu/antiviral.htm (accessed May 2009).
(4) http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5832a3.htm, accessed Jan. 9, 2010.
(5) Condra, J. H.; Schleif, W. A.; Blahy, O. M.; Gabryelski, L. J.; Graham, D. J.; Quintero, J. C.; Rhodes, A.; Robbins, H. L.; Roth, E.; Shivaprakash, M.; Titus, D.; Yang, T.; Teppler, H.; Squires, K. E.; Deutsch, P. J.; Emini, E. A. Invivo emergence of HIV-1 variants resistant to multiple protease inhibitors. *Nature* **1995**, *374*, 569–571.
(6) Kaplan, A. H.; Michael, S. F.; Wehbie, R. S.; Knigge, M. F.; Paul, D. A.; Everitt, L.; Kempf, D. J.; Norbeck, D. W.; Erickson, J. W.; Swanstrom, R. Selection of multiple human-immunodeficiency-virus type-1 variants that encode viral proteases with decreased sensitivity to an inhibitor of the viral protease. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 5597–5601.
(7) Cao, Z. W.; Han, L. Y.; Zheng, C. J.; Ji, Z. L.; Chen, X.; Lin, H. H.; Chen, Y. Z. Computer prediction of drug resistance mutations in proteins. *Drug Discovery Today* **2005**, *10*, 521–529.
(8) Chachra, R.; Rizzo, R. C. Origins of Resistance Conferred by the R292K Neuraminidase Mutation via Molecular Dynamics and Free Energy Calculations. *J. Chem. Theory Comput.* **2008**, *4*, 1526–1540.
(9) Hou, T.; McLaughlin, W. A.; Wang, W. Evaluating the potency of HIV-1 protease drugs to combat resistance. *Proteins* **2008**, *71*, 1163–1174.
(10) Zhang, J.; Hou, T.; Wang, W.; Liu, J. S. Detecting and understanding combinatorial mutation patterns responsible for HIV drug resistance. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 1321–1326.

approaches, a commonly used and tractable approach is the molecular mechanics Poisson−Boltzmann surface area (MM-PBSA) method.[11,12] In this first-principle based method, the gas-phase energy, calculated using conventional molecular mechanics force fields such as AMBER,[13,14] is combined with a continuum model of solvation that includes a surface-area based nonpolar contribution and a polar solvation free energy calculated with the Poisson−Boltzmann (PB) model. Solute entropy is computed from statistical thermodynamics with normal-mode analysis. Recently, there has been an increased interest in the faster molecular mechanics generalized Born surface area (MM-GBSA) variant of MM-PBSA, which replaces the PB electrostatics with the generalized Born (GB) approximate model of electrostatics in water.[11,15,16] Although MM-PBSA is widely used, the MM-GBSA method has been demonstrated to be sometimes more efficient than MM-PBSA, for protein−drug systems,[17] carbohydrates[18] or nucleic acids.[19] Another advantage of MM-GBSA is that it utilizes a fully pairwise potential useful for decomposing the total binding free energy into atomic or group contributions, in a structurally nonperturbing formalism.[20] By decomposing the total binding free energy to each residue, we can obtain detailed binding energies such as the contribution of each residue on the binding process. By comparing the contribution of residues for drugs and substrate, we can know which mutation will significantly impair the protein's interaction with drugs but not with substrate. Based on such information, the potential drug-resistance sites can be predicted.

In this study, the 3D structure of A/H1N1-NA was first modeled using homology modeling. Next, the substrate sialic acid and two widely used drugs, oseltamivir and zanamivir (Figure 1), were fitted into the binding site of A/H1N1-NA by superposing the H1N1-NA homology model to the crystal structure of other subtype neuraminidase with the corresponding ligands. Following this step, molecular dynamics simulation was performed to obtain the dynamic structural information of the complexes of neuraminidase with the substrate and two drugs by using the modeled structures as the initial structures. Based on the obtained MD trajectory, MM-GBSA combined with pair interaction analysis was applied to calculate the binding free energy and understand the detailed interaction profile as well as characterize the contribution of each residue for the binding of substrate and two drugs. By analyzing the differences of interaction profiles for substrate and inhibitors, the potential drug resistance sites for two inhibitors were identified.

2. Methods

2.1. Preparation of Systems. Since no crystal structure for the 2009 influenza A H1N1 neuraminidase is available, Program Prime from Schrodinger Suite 2008[21] was used to search the homologous proteins of A/H1N1-NA and construct the 3D structure. Among the searched results, the H5N1 neuraminidase (H5N1-NA) in complex with zanamivir (pdb code: 3CKZ)[22] shows the highest sequence identity (91.4%). Therefore, the 3D structure of A/H1N1-NA was modeled based on the genomic sequence data of virions isolated from infected patients in Mexico (GenBank accession no.: ACQ99623)[23] by using the crystal structure 3CKZ as a template.

To compare the interaction of the substrate and inhibitors with neuraminidase for A/H1N1, first, we aligned the built 3D structure of A/H1N1-NA and the crystal structure of complexes with sialic acid (pdb code: 2C4A), oseltamivir (pdb code: 3CL0) and zanamivir (pdb code: 3CKZ), respectively. Based on this alignment, the three ligands sialic acid, oseltamivir and zanamivir were input into the homology modeled structure of H1N1-NA to get their complexes.

The 3D structures of A/H1N1-NA with two inhibitors and one substrate were treated further and used as the initial structures of molecular dynamics simulation. All hydrogen

(11) Srinivasan, J.; Cheatham, T. E.; Cieplak, P.; Kollman, P. A.; Case, D. A. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. *J. Am. Chem. Soc.* **1998**, *120*, 9401–9409.

(12) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating structures and free energies of complex molecules: combining molecular Molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889–897.

(13) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24*, 1999–2012.

(14) Lee, M. C.; Duan, Y. Distinguish protein decoys by using a scoring function based on a new AMBER force field, short molecular dynamics simulations, and the generalized born solvent model. *Proteins: Struct., Funct., Bioinf.* **2004**, *55*, 620–634.

(15) Tsui, V.; Case, D. A. Theory and applications of the generalized born solvation model in macromolecular simulations. *Biopolymers* **2000**, *56*, 275–291.

(16) Onufriev, A.; Bashford, D.; Case, D. A. Modification of the generalized Born model suitable for macromolecules. *J. Phys. Chem. B* **2000**, *104*, 3712–3720.

(17) Feig, M.; Brooks, C. L., III. Evaluating CASP4 predictions with physical energy functions. *Proteins* **2002**, *49*, 232–245.

(18) Gourmala, C.; Luo, Y.; Barbault, F.; Zhang, Y.; Ghalem, S.; Maurel, F.; Fan, B. Elucidation of the LewisX-LewisX carbohydrate interaction with molecular dynamics simulations: A glycosynapse model. *J. Mol. Struct. THEOCHEM* **2007**, *821*, 22–29.

(19) Afshan Shaikh, S.; Jayaram, B. A Swift All-Atom Energy-Based Computational Protocol to Predict DNA-Ligand Binding Affinity and ΔTm. *J. Med. Chem.* **2007**, *50*, 2240–2244.

(20) Gohlke, H.; Kiel, C.; Case, D. A. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *J. Mol. Biol.* **2003**, *330*, 891–913.

(21) *Prime*, version 2.0; Schrödinger, LLC: New York, NY, 2008.

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

(23) http://www.ncbi.nlm.nih.gov/protein/237511824, accessed Jan 9, 2010.

atoms of the proteins were added using the leap module in AMBER 10.0 software package.[24] There are eight disulfide bonds in the protein. They are very important to keep the 3D structure of neuraminidase. Thus, here, all of the disulfide bonds were also added using the leap module. Gaussian 98 program[25] was used to perform geometry optimization of the inhibitors and substrate by using Hartree−Fock method and 6-31+G* basis functions. As follows, partial charges of the drug and substrate atoms were determined using the RESP fitting technique[26] based on the optimized structures and the electrostatic potentials calculated using HF/6-31+G* in Gaussian 98 program.[25] The force field parameters for the ligands generated using the Antechamber program in AMBER10.0 were described by the general AMBER force field (GAFF).[27] The standard AMBER force field for bio-organic systems (ff03)[28] was used to describe the neuraminidase parameters. For each system, the counterions were added to neutralize each ligand-bound system. Then, the corresponding systems were solvated using atomistic TIP3P water[29] in an octahedron box with at least 10 Å distance around the complex. Three ligand-bound systems contain 28881, 28897, 28881 atoms for the sialic acid, oseltamivir and zanamivir bounded complexes, respectively.

2.2. Molecular Dynamics Simulation. The molecular dynamics package AMBER10[24] was used throughout the simulations as well as for the minimization and equilibration protocols. The energy minimization was first conducted using the steepest descent method in AMBER10 for 1000 iterations with a restraint potential of 5 kcal/mol·Å^2 applied to all atoms. This was followed by two rounds of energy minimization with steepest descent method switched to conjugate gradient every 500 steps totally for 2000 steps with a 2.0 kcal/mol·Å^2 restraint on all atoms for the first 1000 steps and 0.1 kcal/mol·Å^2 for the last 1000 steps. Long-range Coulombic interactions were handled using the particle mesh Ewald (PME) summation.[30] For the equilibration and subsequent production run, the SHAKE algorithm[31] was employed on all atoms covalently bonded to a hydrogen atom, allowing for an integration time step of 2 fs. The system was gently annealed from 0 to 300 K over a period of 50 ps using a Langevin thermostat with a coupling coefficient of 1.0/ps with a force constant 1.0 kcal/mol·Å^2 on the complex. Here, the meaning of a "gentle" annealing refers to the NMR restraint (nmropt = 1) being used when we heated the system. The use of this item can ensure that the system was heated gradually. An additional two rounds of MD (50 ps each at 300 K) were performed with decreasing restraint weights reduced from 0.5 to 0.1 kcal/mol·Å^2. All subsequent stages were carried out in the isothermal isobaric (NPT) ensemble using a Berendsen barostat[32] with a target pressure of 1 bar and a pressure coupling constant of 2.0 ps. The systems were again equilibrated for 500 ps while maintaining the force constants on the restrained atoms. The production phase of the simulation was run in the same condition for a total of 5−10 ns. Here, the same condition refers to the residues 95−216 and 449−467 being restrained as above with a fixed harmonic potential 0.1 kcal/mol·Å^2 during the production simulations. Coordinate trajectories were recorded every 1 ps throughout all equilibration and production runs. The reason to constrain these residues during the production phase is discussed as follows.

Generally, the biological form of NA is a tetramer; however, each monomer contains a functionally complete binding site. At the same time, it is difficult to obtain the tetramer structure of NA by homology modeling and it will be time-consuming to perform the simulation of a tetramer. Thus only a single monomer was used for the simulations. However, there are close contacts between each monomer mainly in two regions. One region consists of the residues from 95 to 216. The other contained the residues 449−467. In order to avoid the large motion of protein and keep the

(24) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668–1688.

(25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.9; Gaussian Inc.: Pittsburgh, PA, 1998.

(26) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges--the Resp model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(27) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(28) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J. M.; Kollman, P. A. point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **2003**, *24*, 1999–2012.

(29) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(30) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–9593.

(31) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(32) 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. Chem. Phys.* **1984**, *81*, 3684–3690.

backbone fold intact, the atoms in the two regions were constrained weakly through the production phase.

2.3. Binding Free Energy Calculations and Evaluation of Drug Resistance Based on MM-GBSA. MM-GBSA calculations were performed using AMBER10.[24] The first step of the MM-GBSA method is the generation of multiple snapshots from an MD trajectory of the protein−ligand complex, stripped of water molecules and counterions. Snapshots, equally spaced at 10 ps intervals, were extracted from the MD production runs. For each snapshot, the free energy is calculated for each molecular species (complex, protein, and ligand). The binding free energy is computed as the difference:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{protein}} - G_{\text{ligand}}$$

The free energy, G, for each species can be calculated by the following scheme using the MM-GBSA method:[11,12]

$$G = E_{\text{gas}} + G_{\text{sol}} - TS$$

$$E_{\text{gas}} = E_{\text{int}} + E_{\text{ele}} + E_{\text{vdw}}$$

$$E_{\text{int}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{torsion}}$$

$$G_{\text{sol}} = G_{\text{GB}} + G_{\text{nonpolar}}$$

$$G_{\text{nonpolar}} = \gamma\,\text{SASA}$$

Here, E_{gas} is the gas-phase energy; E_{int} is the internal energy; E_{bond}, E_{angle}, and E_{torsion} are the bond, angle, and torsion energies, respectively; E_{ele} and E_{vdw} are the Coulomb and van der Waals energies, respectively. E_{gas} was calculated using the ff03 molecular mechanics force field.[28] G_{sol} is the solvation free energy and can be decomposed into polar and nonpolar contributions. G_{GB} is the polar solvation contribution calculated by solving the GB equation.[11,12] Dielectric constants for solute and solvent were set to 1 and 80, respectively. G_{nonpolar} is the nonpolar solvation contribution and was estimated by the solvent accessible surface area (SASA) determined using a water probe radius of 1.4 Å. The surface tension constant γ was set to 0.0072 kcal/mol/Å^2.[33] T and S are the temperature and the total solute entropy, respectively. Vibrational entropy contributions are estimated by normal-mode analysis. The vibrational entropy of a substance is calculated by the following equation:[34]

$$S_{\text{vib}} = -R\ln(1 - e^{-h\nu_0/kT}) + \frac{N_A h\nu_0}{T}\frac{e^{-h\nu_0/kT}}{(1 - e^{-h\nu_0/kT})}$$

where N_A is Avogadro's number, ν_0 is the frequency of the normal mode, h is Planck's constant, and k is Boltzmann's constant. In normal-mode analysis, the frequencies of the normal modes are calculated from a molecular mechanics force field using the Hessian matrix of second energy derivatives in terms of the force constants for each type of interaction. This requires an input structure that is extracted from molecular dynamics trajectory and minimized by a Newton−Raphson protocol to obtain the Hessian matrix of second energy derivatives. Because of the high computational demand, only 20 snapshots were used in normal-mode analysis and each snapshot was optimized for 100,000 steps using a distance-dependent dielectric of $4r_{ij}$ (r_{ij} is the distance between atoms i and j) until the root-mean-square deviation of the gradient vector was less than 0.0005 kcal/mol/Å.[2]

To obtain the contribution of each residue to binding energy, MM-GBSA was used to decompose the interaction energies to each residue involved in the interaction by considering molecular mechanics and solvation energies without consideration the contribution of entropies. By analyzing the contribution difference of each residue for the substrate and inhibitors binding, if one residue interacts more favorably with a drug than with the substrate, mutations of this residue may not affect the function of the neuraminidase but may impair the binding of the drug and thus this residue is identified as potential resistant site.[35] To understand the origin of the contribution of each residue, the energy contribution of each residue was decomposed into polar and nonpolar items. The polar item composed of the electrostatic interaction in gas phase and the polar solvation contribution computed by solving the GB equation. The nonpolar interaction was obtained by summing the van der Waals energy item and nonpolar solvation contribution.

3. Results and Discussion

3.1. Homology Modeling. The sequence alignments (Figure 2) between A/H1N1-NA and the 3CKZ crystal structure were performed by PRALINE sequence alignment toolbox.[36] The crystal structure 3CKZ contained one oseltamivir-resistance mutation H274Y. From the results of sequence alignment, we found that the identity between A/H1N1-NA and 3CKZ is up to 91.4%. Here, by using 3CKZ as template, a reliable 3D structure of 2009-H1N1-NA was developed using Prime Module of Schrodinger suite 2008.[21] Figure 3 gave the overall alignment of the homology modeled 3D structure of 2009-H1N1-NA and 3CKZ with a magnifying window of the active site. The nonconserved residues were displayed in sticks. From Figure 2, it can be seen that the active site is highly conserved and only three residues (ILE149, HIS275, ASN344 numbered according to the sequence of A/H1N1-NA) are different from the template. Actually in the active site, between A/H1N1-NA and H5N1-NA, only two residues, ILE149 and ASN344, are different.

(33) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* **1994**, *98*, 1978–1988.

(34) McQuarrie, D. A. *Statistical thermodynamics*; Harper & Row: New York, 1973.

(35) Wang, W.; Kollman, P. A. Computational study of protein specificity: the molecular basis of HIV-1 protease drug resistance. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14937–14942.

(36) Simossis, V. A.; Heringa, J. PRALINE: a multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. *Nucleic Acids Res.* **2005**, *33*, W289–W294.

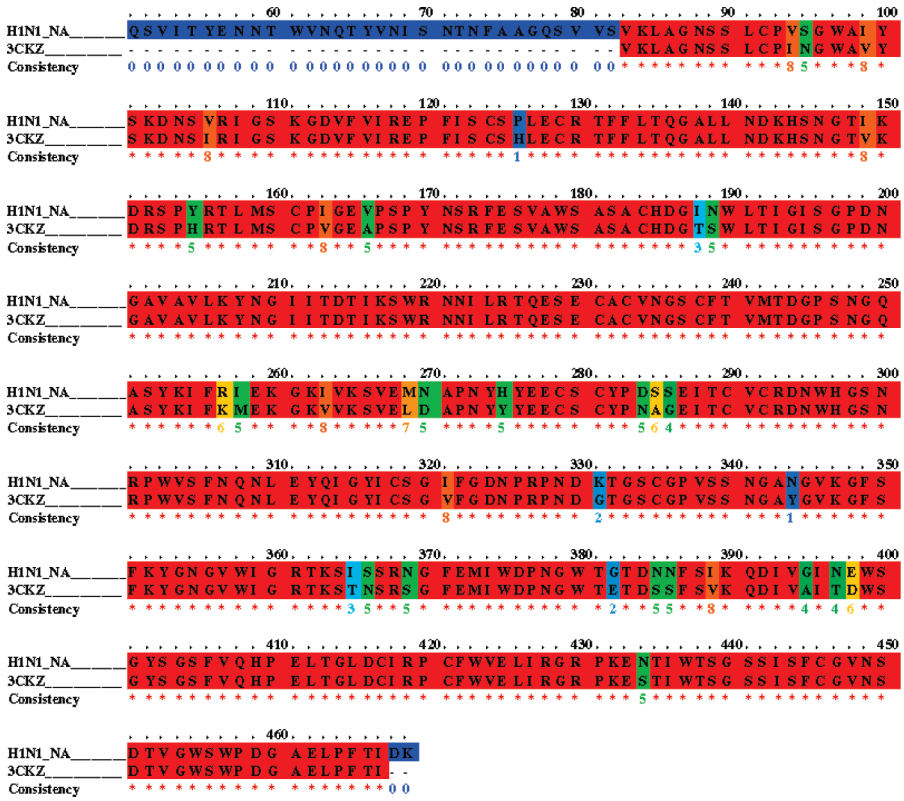

Figure 2. Sequence alignment of influenza A/H1N1-NA with 3CKZ (influenza A/H5N1-NA) used for homology modeling.

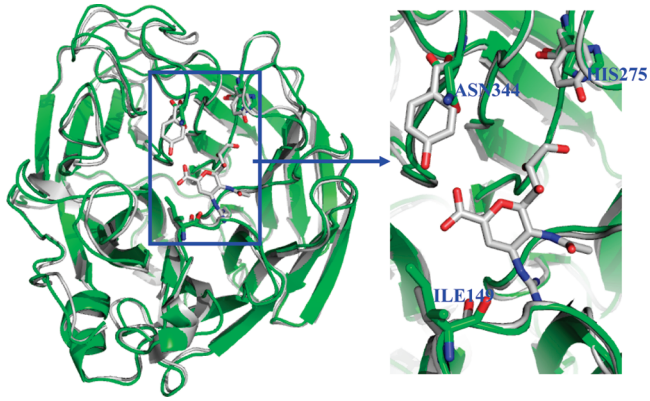

Figure 3. The alignment of the 3D structure of 2009-H1N1-NA (green) and 3CKZ (gray) (oseltamivir and nonconserved residues around active site (numbered according to H1N1 NA sequence) were given in stick style).

Here, our template 3CKZ included one oseltamivir-resistance mutation H274Y numbered according to the sequence of H5N1-NA. Among the other two residues, ILE149 is a little far from active site and has no contact with substrate or inhibitors. So, generally thinking, ASN344 should be responsible for the oseltamivir sensitivity to H1N1 and H5N1.

By fitting the substrate sialic acid and two inhibitors oseltamivir and zanamivir into the homology modeled 3D structure of A/H1N1-NA, three complexes were obtained. The fitting process is as follows: first, we aligned the crystal complexes of neuraminidase with three different ligands to the homology modeled structure of H1N1-NA, respectively. Based on this alignment, the three ligands sialic acid, oseltamivir and zanamivir were input into the homology modeled structure of H1N1-NA to get their complexes. These complexes were treated further by solvation and the addition of counterions etc. The detailed treatment process can be found in section 2.1. The resulted structures were used as the inputs of molecular dynamics simulation.

3.2. Time-Series Convergence and Binding Free Energy Calculations. Long time molecular dynamics simulations of three systems were performed to obtain the equilibrated and stable systems. Here, the equilibration of the MD trajectories was monitored from the convergence of the root-mean-square deviation (rmsd) of Cα atoms from the starting coordinates (Figure 4a). First, for each system, 5 ns production simulation was performed. By monitoring the rmsd of Cα atoms, two systems composed of the complex of sialic acid and zanamivir with the A/H1N1-NA are up to equilibration in the last two nanoseconds. However, for the complex of oseltamivir with A/H1N1-NA, it seems that the rmsd is still shifting up after 5 ns simulation. Thus for this system, we continued another 5 ns simulation. From Figure 4a, it can be seen that this system is stable in the last two nanoseconds. Since our aim is to study the interaction between the receptor and ligand, we care more about the status of binding pocket and ligand. Here, we also monitored the time evolution of rmsd of Cα atoms for the residues in 5 Å around ligand and rmsd of heavy atoms for ligand for

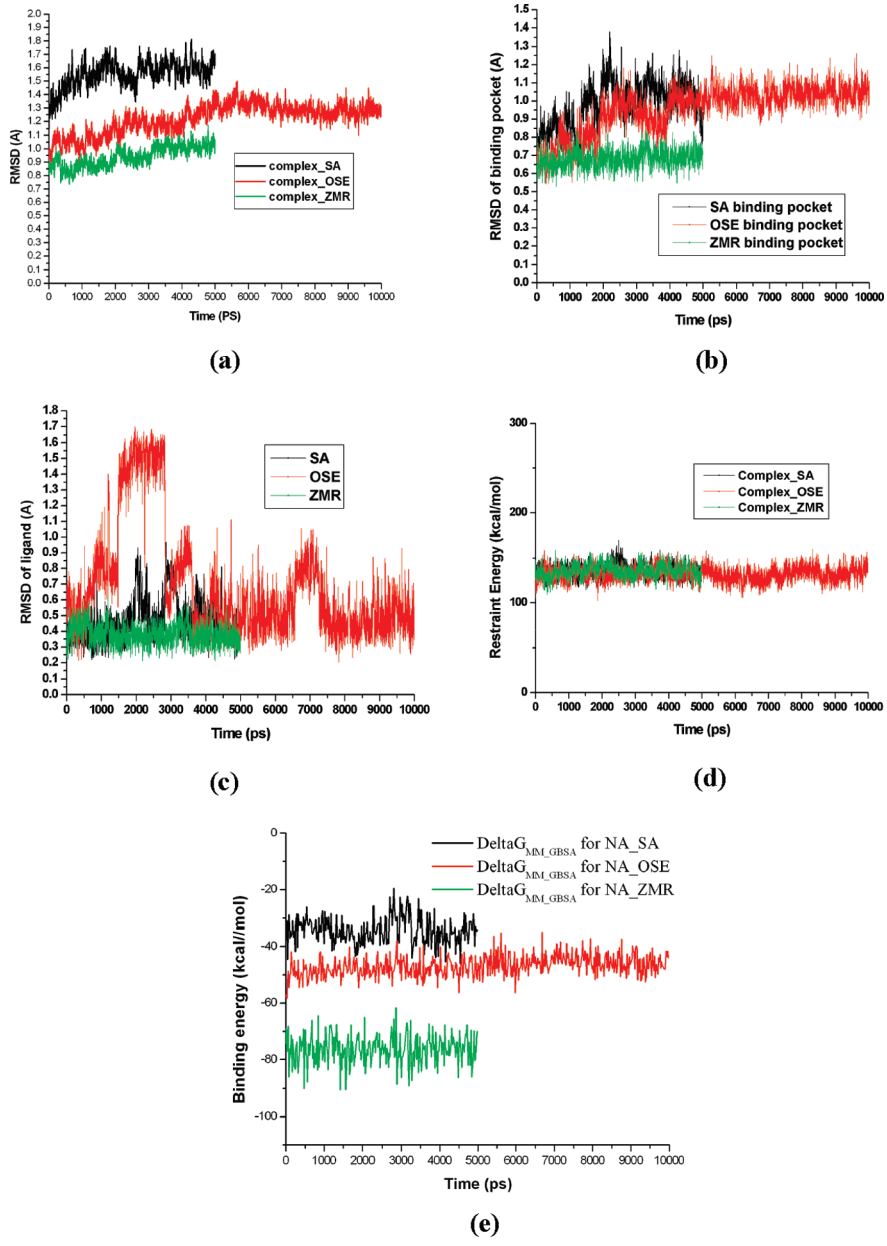

Figure 4. The monitoring for the equilibration of the MD trajectories: (a) the time series of the rmsd of Cα atoms from the initial structure; (b) time evolution of rmsd of Cα atoms for the residues around 5 Å of ligand; (c) time evolution of rmsd of heavy atoms for the ligands; (d) time evolution of the restraint potential; (e) time evolution of binding free energy components for the complex.

each system (Figure 4b and 4c). From Figure 4b, it can be seen that the binding pocket of zanamivir has the least fluctuation and that of oseltamivir is up to stable after 4 ns. Although the complex of sialic acid is up to equilibration after 3 ns, the binding pocket still has a large fluctuation. For zanamivir, Figure 4c gives consistent information with Figure 4b that this compound always has a very stable binding mode. However for oseltamivir, as can be seen from Figure 4c, after the equilibration of the whole system, it seems that the oseltamivir binding position is flexible. The reason oseltamivir has a flexible binding mode compared with zanamivir and sialic acid should come from the 3-penta 3-yl group. This group mainly formed van der Waals interaction with binding pocket. The corresponding 1,2,3-trihydroxy propyl for zanamivir and sialic acid can form two stable hydrogen bonds. As we know, hydrogen bond interaction has a certain direction and can orient the position of the corresponding group. However, the van der Waals interaction is a nonspecific interaction and then the involved group can be in different direction.

In addition, during the MD process, we incorporated restraints on the parts with the close contact with other monomer in the biological tetramer form to avoid the large motion of protein and keep the backbone fold intact. Here, we also analyzed the evolution of the restraint potential versus time (Figure 4d). From Figure 4d, it can be seen that the restraint potential is more or less constant. Furthermore,

Table 1. The Experimental IC_{50} for the Substrate and Inhibitors

ref	sialic acid (nM)	oseltamivir (nM)	zanamivir (nM)
37		53.20 ± 10.30	5.80 ± 2.10
38		17.80	4.53
39		1.53	0.65
40	6100	0.01−2.2	0.3−2.3

Table 2. Calculated Binding Free Energy and Its Components (kcal/mol)

	ΔE_{int}^{vdw}	ΔE_{int}^{ele}	$\Delta G_{sol}^{nonpolar}$	ΔG_{sol}^{ele}	ΔH_{bind}	$-T\Delta S$	ΔG_{bind}
NA-SA	−19.18	−209.00	−4.59	198.36	−34.41	22.88	−11.53
NA-OSE	−26.01	−225.87	−4.98	210.07	−46.79	23.05	−23.74
NA-ZMR	−20.98	−327.41	−4.96	277.03	−76.33	27.83	−48.50

for three systems, the ratio of the restraint potential in comparison to the global energy is evaluated by calculating the absolute value of the average of restraint potential divided by the average of total energy. The corresponding ratio value for sialic acid, oseltamivir and zanamivir systems is 0.217%, 0.211% and 0.215%, respectively. From this result, we can see that the energetic effect from the restraint on the whole system is very small.

To learn the energy changing information during the MD process, we also monitored the time evolution of binding free energy components for three complexes. The value of the binding energy for the substrate and inhibitors from Figure 4e is well consistent with the experimental value.[37−40] Although the experimental binding energy data from different laboratories are different (Table 1), most of them have the same trends for three ligands except ref 40. The general trend of inhibition constant is zanamivir > oseltamivir > sialic acid. Two inhibitors have a much stronger binding affinity than the substrate sialic acid. Zanamivir is more potent to inhibit H1N1-NA than oseltamivir. This is in good agreement with our calculated results shown in Figure 4e. Along the whole trajectory, each system is stable at the last 1.5 ns both from rmsd and from energy. So, our following pair interaction analysis is based on the snapshots extracted from the last 1.5 ns trajectory for the corresponding system.

The binding energy value from Figure 4e does not include the entropy contribution. In order to evaluate the binding free energy more accurately, we calculated the contribution of entropy by using normal-mode analysis. Because of the high computational demand, only 20 snapshots for every system extracted from the last 1.5 ns were used in normal-mode analysis. The calculated results including the enthalpy and entropy are given in Table 2. From Table 2, two conclusions can be obtained. One is that the inhibitor oseltamivir has much stronger van der Waals interactions with neuraminidase than the substrate and zanamivir. Such difference in the contribution of van der Waals interactions comes from the difference of their structure. Oseltamivir has a more hydrophobic group in the meta-position of carboxyl group and then forms the stronger van der Waals interaction. Besides, compared with sialic acid and oseltamivir, polar interactions have more contribution to the binding of zanamivir because of the guanidino group in another meta-position of carboxyl group.

3.3. The Identification of Potential Drug Resistance Sites from Pair Interaction Analysis. First, pair interaction energy analysis was performed to identify key residues of the NA interface to the binding of substrate. The residue energy decomposition results for substrate were given in Figure 5. From Figure 5a, it can be seen several residues of NA surface such as Glu119, Arg225, Glu277, Glu278, Arg293 and Arg368 give a large contribution to substrate binding. Figure 5b gives the interactions between sialic acid and NA. By analyzing the detailed interaction between these residues and substrate, we can see that each residue forms an important hydrogen bond or salt bridge with sialic acid. This indicates that polar interaction plays an important role during the binding of substrate. From Table 2, it can be seen that both polar interaction and nonpolar interaction have obvious contribution to the binding of sialic acid. The polar interaction mainly comes from the formation of hydrogen bond and salt bridge. Where does the large nonpolar interaction come from? To answer this question, we further decompose the contribution of each residue to polar and nonpolar interaction (Figure 5c). The items of polar and nonpolar interaction components have been described in Methods. By comparing Figure 5a and 5c, it can be seen that almost each involved residue with obvious contribution except Arg225 has a large polar interaction with sialic acid. It is not surprising since these residues either form the important hydrogen bond or salt bridge. For the polar residue Arg225, its main contribution to the binding process does not come from the formation of polar interaction, but comes from the obvious van der Waals interaction by the packing of its side chain with the cyclic center of sialic acid. From Figure 5c, we can see that there are other residues besides Arg225 with nonpolar contribution, such as Glu278, Arg293, Arg295 and Tyr402. These resides like Arg225 have a close van der Waals contact and form the van der Waals interaction.

(37) 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.

(38) Bantia, S.; Parker, C. D.; Ananth, S. L.; Horn, L. L.; Andries, K.; Chand, P.; Kotian, P. L.; Dehghani, A.; El-Kattan, Y.; Lin, T.; Hutchison, T. L.; Montgomery, J. A.; Kellog, D. L.; Babu, Y. S. Comparison of the anti-influenza virus activity of RWJ-270201 with those of oseltamivir and zanamivir. *Antimicrob. Agents Chemother.* **2001**, *45*, 1162–1167.

(39) Boivin, G; Goyette, N. Susceptibility of recent Canadian influenza A and B virus isolates to different neuraminidase inhibitors. *Antiviral Res.* **2002**, *54*, 143–147.

(40) Masukawa, K.; Kollman, P. A.; Kuntz, I. D. Investigation of Neuraminidase-Substrate Recognition Using Molecular Dynamics and Free Energy Calculations. *J. Med. Chem.* **2003**, *46*, 5628–5637.

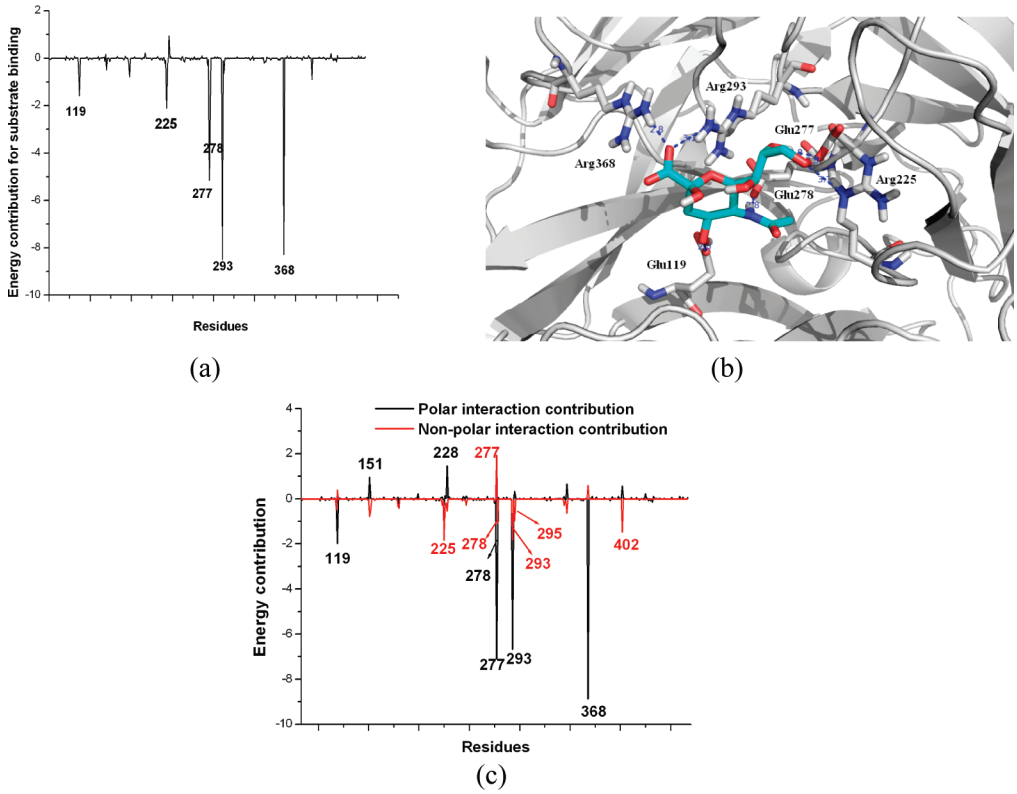

Figure 5. (a) The residue contributions for substrate binding; (b) the distribution of the critical residues in NA; (c) the polar and nonpolar interaction contribution of each residue for substrate binding.

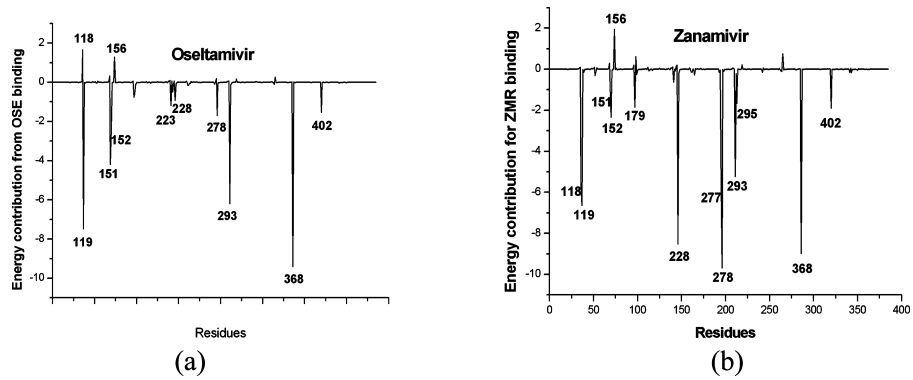

Figure 6. The residue contributions for two inhibitors binding: (a) Oseltamivir; (b) Zanamivir.

We also analyzed the detailed interaction between two inhibitors with neuraminidase. Figure 6 gives the residue contribution to the binding process of two inhibitors. From the comparison of Figure 5a and Figure 6, the interaction profile for substrate and inhibitors has similar features but also with some differences. For example, either for oseltamivir or zanamivir, the new interaction with Asp151 and Asp 152 occurs compared with substrate. As for zanamivir, there are many more new interactions with the NA. In detail, residues 118, 228, and 295 have a large contribution for zanamivir binding.

Since the difference of residue contribution between substrate and inhibitors determines the potential drug resistance sites, we plotted the energy difference of each residue contribution to substrate and two inhibitors, respectively (Figure 7). From Figure 7, we can see that several residues have much stronger interactions with drugs than substrate. This indicates that these residues are much easier to mutate to confer resistance of drugs. As a result, for oseltamivir, the most potential resistant sites will be the residues Glu119, Asp151 and Asp152, which have more than 2 kcal/mol energy difference of contribution for oseltamivir and sialic acid binding. In addition, although Glu277 has an obvious contribution for sialic acid, it does not affect the binding of oseltamivir. Thus generally, drug resistance should not happen at this site. As for zanamivir, from the Figure 7c, there are more residues with the more favorable interaction to drug. That is, it will have more potential drug resistant sites. Most obviously, Arg118, Glu119, Glu228, Glu278 have more than 4 kcal/mol energy contribution difference for the

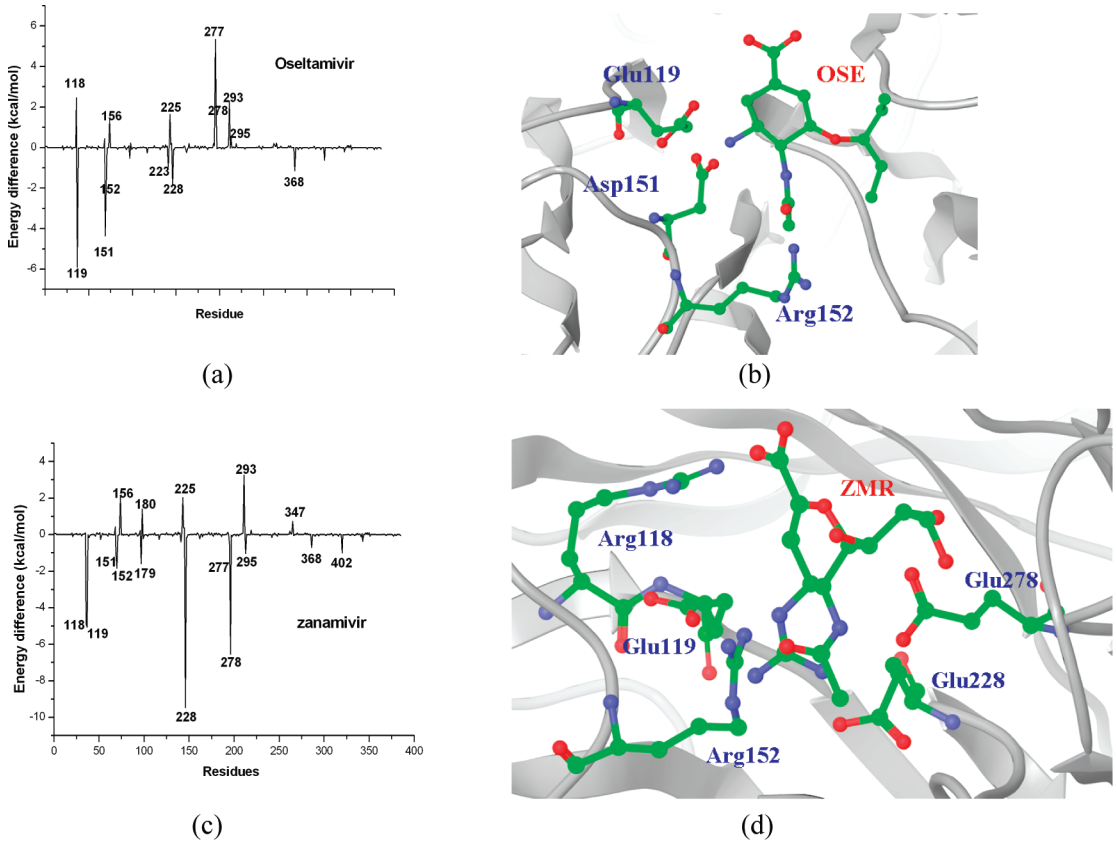

Figure 7. (a) The energy difference of each residue contribution to the binding of oseltamivir and substrate; (b) the distribution of the potential drug resistant sites in NA for oseltamivir; (c) the energy difference of the contribution of each residue to the binding of zanamivir and substrate; (d) the distribution of the potential drug resistant sites in NA for zanamivir.

binding of zanamivir and substrate. Compared with oseltamivir, three residues Arg118, Glu228 and Glu278 are new potential drug resistance sites. The potential drug resistant site distribution for oseltamivir and zanamivir from structure is shown in Figure 7b and 7d, respectively. Among the potential drug resistance sites, Glu119 and Asp152 are common sites for both drugs.

Although our identified potential drug resistant sites have not been isolated from the A/H1N1 influenza virus population due to the limited time that this virus spreads and the corresponding drugs were applied, some of them have been detected from other types of influenza. For example, the Glu119 and Asp152 mutations are known to confer resistance to oseltamivir and zanamivir for influenza A virus H3N2.[41,42] Along with the long time use of oseltamivir and zanamivir, more potential drug resistance sites will be confirmed from the experiments and clinical results.

In addition, although MD/MM-GBSA method based on single trajectory can be used to identify potential drug resistance sites fast and accurately, it still has some limitations. For example, it is only effective to find the potential drug resistance sites around the active site or with close contact to the drugs and cannot find the potential drug resistance sites far from the active site. For residues far away from the drug binding site, if their mutation can confer drug resistance, generally, they will occur from different mechanism, such as those mutations changing the stability of protein or changing the conformation of some residues could change the interaction of drug with the residues in active site. For such a circumstance, it is difficult to obtain effective results with the single trajectory method. As for this study, one typical drug resistance site H275Y, which is an oseltamivir-resistant variant of seasonal influenza A virus (H1N1),[4] cannot be identified by the single trajectory method since the residue H275 has no direct contact with ligand. Of course, it is possible to predict accurately if such residues' mutation causes drug resistance by studying the dynamics behavior and energy profile of wild-type and the mutant type, respectively. Certainly, this method will be more time-consuming and it is impossible to study the behavior of each residue mutant type. So this method is only effective when enough information of potential drug-resistance sites is available. Here, since our aim is to predict potential drug-

(41) Gubareva, L. V.; Kaiser, L.; Hayden, F. G. Influenza virus neuraminidase inhibitors. *Lancet* **2000**, *355*, 827–35.
(42) Kiso, M.; Mitamura, K.; Sakai-Tagawa, Y.; Shiraishi, K.; Kawakami, C.; Kimura, K.; Hayden, F. G.; Sugaya, N.; Kawaoka, Y. Resistant influenza A viruses in children treated with oseltamivir: descriptive study. *Lancet* **2004**, *364*, 759–65.

resistance sites, the method MD/MM-GBSA based on single trajectory is much more suitable.

4. Conclusions

In this study, molecular dynamics simulation combined with molecular mechanics generalized Born surface area calculations were used to identify potential drug resistance sites of the most widely used drugs oseltamivir and zanamivir for the therapy of new strain 2009 influenza A (H1N1) virus. From the obtained results, for oseltamivir, resistant mutations may occur in the sites Glu119, Asp151 and Asp152. Compared with oseltamivir, zanamivir should have more potential drug resistant sites. Besides the common potential sites Glu119 and Asp152, the additional three residues Arg118, Glu228 and Glu278 are also easy to mutate to confer drug resistance. The information about the potential drug resistance sites can provide some useful insights for the rational design of a new effective medicine against the new resistant influenza strains. The results of this study would be also very significant to avoid the situation of having no effective drugs to treat the new mutant influenza strain.

Acknowledgment. This work was supported by the National Natural Science Foundation of China (Grant No: 20905033) and the Fundamental Research Funds for the Central Universities (Grant No: lzujbky-2009-97). The authors also would like to thank the Supercomputing Center of Gansu Province for providing the computing resources.

MP100041B