http://dx.doi.org/10.5012/bkcs.2012.33.3.1059

A Structural Model of Glycosylated Neuraminidase Based on Molecular Dynamics Simulations for Virtual Inhibitor Screening Against Influenza Virus†

Youngjin Choi, Chanho Kwon,‡ Eunae Cho,‡ In-Cheol Kang, Hyun-Ja Jeong, Sungjun Park,§ Karpjoo Jeong,# Tak Hur,¶ and Seunho Jung‡,*

BioChip Research Center, Hoseo University, Asan 336-795, Korea
*‡Department of Bioscience and Biotechnology, Bio/Molecular Informatics Center & Center for Biotechnology Research in UBITA, Konkuk University, Seoul 143-701, Korea. *E-mail: shjung@konkuk.ac.kr*
§Department of Game Engineering, Hoseo University, Asan 336-795, Korea
#Department of Advanced Technology Fusion, Center for Biotechnology Research in UBITA, Konkuk University, Seoul 143-701, Korea
¶Department of Materials & Engineering, College of Engineering, Konkuk University, Seoul 143-701, Korea
Received November 29, 2011, Accepted December 15, 2011

Key Words : Glycosylation, Influenza virus, Neuraminidase, Virtual screening, Molecular dynamics

Swine flu is a global pandemic disease caused by influenza A-type viruses.[1] High-pathogenic flu viruses have spread to various countries, infecting not only domestic animals but also even humans.[2] Influenza A viruses can be categorized into subtypes on the basis of their main surface glycoproteins, haemagglutinin (HA) and neuraminidase (NA).[3] Both HA and NA proteins are often regarded as major drug target for the treatment of pandemic avian flu. The surface glycoprotein of the influenza virus, neuraminidase (NA) is known to be responsible for the viral attachment process by removing the terminal sialic acids bound to haemagglutinin (HA).

Oseltamivir is a commercially available drug suitable for the treatment of avian or swine flu.[4] The mode of action is known to inhibit activity of NA protein. Previous crystallographic study showed that binding cavity of the NA adapted to two different conformations, "open" and "closed" form, depending on 150-loop flexibility.[5] The open form of NA protein has been proposed as an alternative target site for the design of anti-influenza drug because of its potential cavity adjacent to active site. This multiple fixed receptor approach is a practical method that may improve virtual screening process because direct modeling of protein binding site flexibility remains challenging.[6] As yet, the molecular detail for interactions between the oseltamivir and open form of NA protein remains a matter of debate. On the drug discovery and development process of oseltamivir, open conformation concept for the NA protein has not been considered. But this open cavity form is not easily obtainable using the crystallographic method due to its flexibility. This is why a computational technique should be used in the drug screening approach based on the open NA structure as well as molecular basis of oseltamivir coupled with the open NA conformation.

†This paper is to commemorate Professor Kook Joe Shin's honourable retirement.

In this case, molecular dynamics (MD) simulations is a useful alternative to X-ray crystallography to find different binding site conformation for the NA protein, although it requires large amount of hardware resources and computing times. Recently, McCammon and coworkers in the *University of California San Diego* (UCSD)[7] introduced ensemble-based virtual screening with MD-structures of NA protein instead of X-ray conformation. They used Amber force-field to obtain wide-open structure of NA. The UCSD's NA model suggests that the 150-loop is even more pliable than observed in the X-ray structure. In this model, rotations about C_α atoms allow the 150-loop to swing outward during the MD simulations.[5] However, they modeled the unglycosylated form of NA to calculate dynamic structures of this protein. In real biological conditions, the NA is known to be glycosylated at the Asn residues,[8] implying contribution of sugar residue to the 150-loop flexibility. The glycosylation actually affects biochemical properties of protein such as activity, stability, and structure.[9,10] Therefore dynamic NA structure could be affected by attached sugar residue during the MD simulations. In the present study, series of MD simulations were performed for both glycosylated and unglycosylated model of NA proteins to obtain multiple conformation of this target protein. The oseltamivir-bound conformation of NA was compared between X-ray and MD structures. From these results, we propose alternative MD model of NA target protein with showing different binding mode for the oseltamivir to support virtual screening of antiviral compounds.

To obtain initial glycosylated NA model, we used NetNGlyc 1.0 public web-service (http://www.cbs.dtu.dk/services/NetNGlyc/) which predicts *N*-glycosylation sites in a target protein. The NetNGlyc server predicts *N*-glycosylation sites in the proteins using artificial neural networks algorithm. From this prediction, three potential glycosylation sites were found at the amino acid residues Asn88, Asn146, and Asn235. The β-D-*N*-acetylglucosamine (GlcNAc)

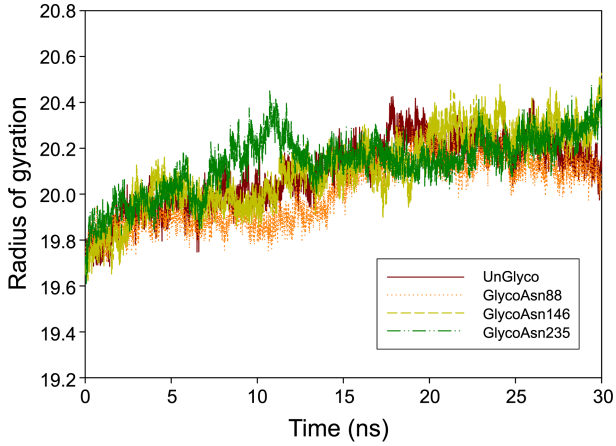

Figure 1. Radius of gyration value changes during the MD simulations.

residue was attached to X-ray structure of NA protein at each glycosylation site. The structural properties of these glycosylated NA protein models were compared with the unglycosylated model using 30-ns of MD simulations. The radius of gyration (R_G) value of each NA model was plotted in the Figure 1. The entire NA model showed similar 20-nm^2 of value after 15-ns of MD simulation time. Therefore, global conformation or size of the NA protein was not heavily changed by the glycosylation during MD simulations. The solvent-accessible surface area (SASA) of the NA protein was also reached *ca.* 160 nm^2 without significant disparity between different four NA models (data not shown).

Local conformational changes also traced by MD simulations for a region around 150-loop which is already known to be widely opened binding site. The Figure 2(a) shows RMSD changes of each NA model during MD simulations. It can be found that significant fluctuations in RMSD value of 150-loop region for between the unglycosylated and GlycoAsn146 NA models were observed. The GlycoAsn88 and GlycoAsn235 NA models favor steady increase rather than dramatic changes in the RMSD. The glycosylation far beyond 150-loop of NA seems to fortify the conformational stability of the NA protein, which is general role of sugar residue upon protein.[11,12] However, the GlcNAc sugar residue near 150-loop actually increases local flexibility around the glycosylated region of the NA protein. Therefore, the GlycoAsn146 NA model is suitable to reflect wide open conformation of the NA binding site compared to other glycosylated models. The wide open conformation of GlycoAsn146 NA model was further confirmed in terms of interatomic distance between 150- and 450-loop (Fig. 2(b)). The center of mass distance between two loops of GlycoAsn146 NA model was gradually increased and reached to 30 nm after 20-ns of MD simulation time. It was comparable to other models that showed distance below 20 nm without structural fluctuation. Therefore, the GlycoAsn146 NA model could be more suitable for reproducing a wide-open conformation of the NA protein compared to different MD structures not only unglycosylated but also glycosylated model sugar attached to distinct sites.

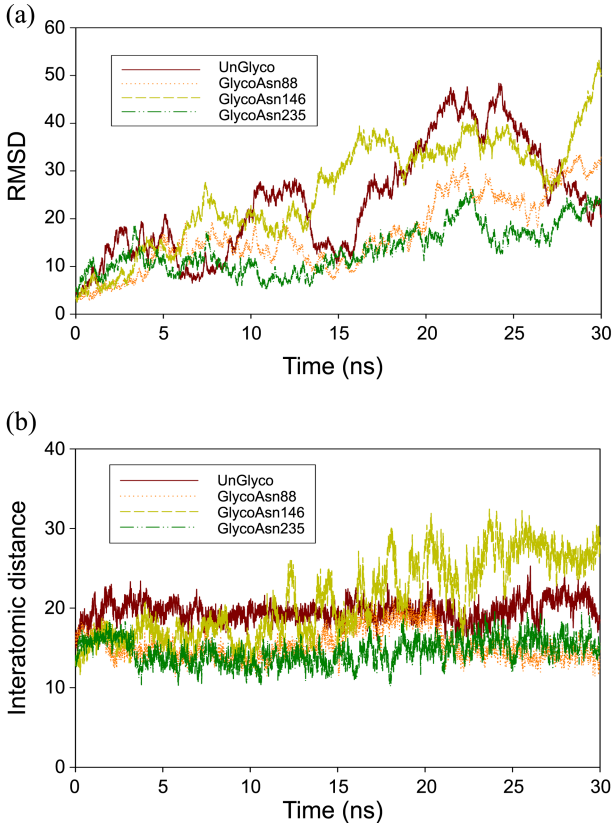

Figure 2. The RMSD (upper panel) and interatomic distance changes (bottom panel) during the MD simulations.

Binding characteristics of inhibitor for the alternative MD models of NA protein by glycosylation was analyzed using molecular docking simulations of oseltamivir onto the different NA structures. A Glide docking score was varied as the structural model of NA protein changes. The docking score of oseltamivir was ranked highest-score in a genuine crystal NA structure as expected. For the MD-based models, the docking score was decreased because of the local structural changes of NA protein around oseltamivir-binding site. As wide-open binding site of MD models of the NA proteins, binding quality of the oseltamivir was deteriorated depending on structural arrangement of the models. Among the alternative MD models, GlycoAsn146 NA structure showed the highest docking score compared with other MD models including both UCSD[5] and Glyco-NA models. For example, docking scores of oseltamivir for crystal and MD-based NA model would be −7.98 versus −4.72, respectively. The docking score was reduced by ~41% when the MD-based NA model was used. Similar undesirable changes in the docking scores were also observed for UCSD's NA model. The docking score using unglycosylated UCSD's model for the NA protein was −5.85. However, the docking score of oseltamivir for GlycoAsn146 NA model was increased to −6.72, which is only ~21% of reduction compared to the score for crystal NA structure. This is indicative of usability for the GlycoAsn146 NA model to adopt advantage of both genuine and wide-open binding site conformation. The GlycoAsn146 NA model is glycoylated

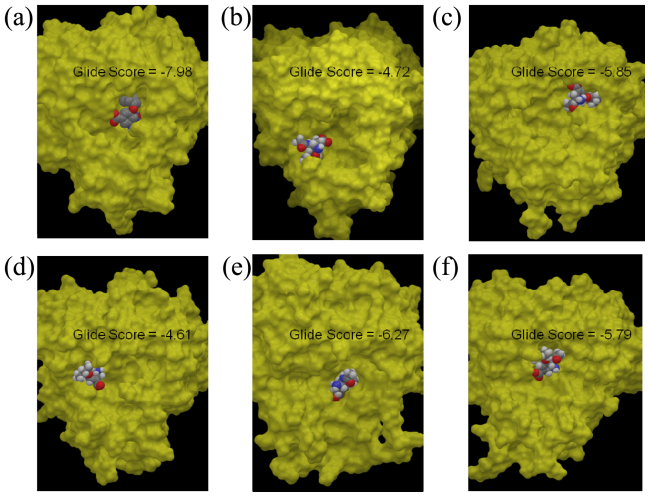

Figure 3. Representative snapshots for the docked pose of oseltamivir upon different NA models. Each panel denotes Crystal (a), UCSD's (b), CHARMM's (c) GlycoAsn88 (d), GlycoAsn146 (e), and GlycoAsn235 (f), respectively.

at the Asn146 residue which is localized in the 150-loop region of the NA protein. The 150-loop flexibility is critical to open a binding cavity of the NA temporally.

Understanding of how drug candidate interacts comes only from the 3D structure for the protein-drug complex, as it provides critical atomic details about binding. By having specific binding site for the ligand, target proteins become limited in their range of interaction partners including antagonist. Figure 3 shows binding mode of oseltamivir onto the corresponding binding site of each NA model. In the crystal structure of NA model, the oseltamivir was tightly bound into the original binding site (Fig. 3(a)). The hydroxyl groups of the oseltamivir made hydrogen bonding with three Arg residues in this NA model. These binding characteristics of oseltamivir were significantly changed when the UCSD's NA model was used. The oseltamivir was moved onto the expanded binding region rather than authentic binding site of NA protein (Fig. 3(b)). This may be due to local structural changes of UCSD's NA protein during MD simulations. Some positional transitions were also observed when the oselatmivir was docked onto our MD-based unglycosylated NA model (Fig. 3(c)). In the GlycoAsn146 or GlycoAsn234 NA model, unlike unglycosylated UCSD's or CHARMM-based model, the oseltamivir showed similar binding mode as compared with crystal NA structure. The GlycoAsn146 NA model especially adopted original-like binding character although this model had a wide open conformation (Fig. 3(e)). Accordingly, the most suitable NA model for the virtual inhibitor screening is glycosylated MD model sugar attached to the 150-loop region of NA protein. Detailed binding features between oseltamivir and each NA model were proposed as ligand-receptor interaction diagram in the supporting information (Fig. S1).

The binding mode of oseltamivir upon flexible open conformation of glycosylated NA protein has not been described up to now. Realization of the structural details enables the more rational design of drug candidate to perturb viral attachment process mediated by NA target protein. Therefore, the target conformation of protein should be carefully selected to reflect structural nature accurately. In the present work, glycosylated MD model for the NA protein was proposed as an alternative for the crystal NA structure.

Computational Methods

Molecular Models and Molecular Dynamics Simulations. Initial three-dimensional coordinates for the neuraminidase (NA) were obtained from the Protein Data Bank (PDB id 2HTY) as an apo-form. Single chain was selected from tetrameric crystal structure of the NA. Glycosylated NAs were built by attaching a β-D-*N*-acetylglucosamine (GlcNAc) sugar residue to each Asn88, Asn146, and Asn235, respectively. Molecular dynamics (MD) simulations were performed using a general molecular modeling package, CHARMM[13] (version 30b1), with a parm22 all-atom force field. Each unglycosylated or glycosylated NA model was solvated with explicit TIP3P[14] water model. Water molecules were removed if they were closer than 2.8 Å to heavy atoms of the NA protein. The NA system was energy-minimized by 3,000 steps of conjugate gradient, followed by Adopted Basis Newton-Raphson until the root-mean-square gradient was less than 0.01 kcal/mol. The MD simulations were performed in the isothermal-isobaric ensemble at 300 K. Bond lengths of the molecules were constrained using a SHAKE[15] algorithm and 2.0 fs of time step was used. The long-range electrostatic interactions were treated using a particle mesh Ewald summation[16] method. Lennard-Jones interactions were truncated from 13- to 8-Å with a switching function in the CHARMM software. The pressure and temperature of the NA system was regulated by Langevin piston combined with Hoover's thermostat.[17] The system was gradually heat to targeted temperature for 100 ps. Production MD runs were performed for 30 ns of time scale.

Molecular Docking Simulations. Flexible molecular docking jobs were performed by Glide5.6 module[18] in Maestro 9.2 (Schrodinger Inc.). The NA model was prepared using Protein Preparation Wizard tool in the Maestro program from the last trajectory of each MD runs. Cubical molecular grid was generated for the NA model using Receptor Grid Generation Tool in the Glide. Oseltamivir-carboxylate molecule was docked upon each NA model under XP mode with Glide score. The highest-scoring pose for the oseltamivir with the NA protein was ranked according to the Glide score.

Acknowledgments. This work was supported by Academic Research Grant of Hoseo University in 2008 (2008-0126).

References

1. Schnitzler, S. U.; Schnitzler, P. *Virus Genes* **2009**, *39*, 279.
2. Peiris, J. S.; de Jong, M. D.; Guan, Y. *Clin. Microb. Rev.* **2007**, *20*,

243.
3. Subbarao, K.; Joseph, T. *Nat. Rev. Immunol.* **2007**, *7*, 267.
4. 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. *Nature* **2006**, *443*, 45.
5. Amaro, R. E.; Minh, D. D.; Cheng, L. S.; Lindstrom, W. M.; Olson, A. J.; Lin, J. H.; Li, W. W.; McCammon, A. *J. Am. Chem. Soc.* **2007**, *129*, 7764.
6. Totrov, M.; Abagyan, R. *Curr. Opin. Struct. Biol.* **2008**, *18*, 178.
7. Cheng, L. S.; Amaro, R. E.; Xu, D.; Li, W. W.; Arzberger, P. W.; McCammon, A. *J. Med. Chem.* **2008**, *51*, 3878.
8. Li, S.; Schulman, J.; Itamura, S.; Palese, P. *J. Virol.* **1993**, *67*, 6667.
9. Wang, C.; Eufemi, M.; Turano, C.; Giartosio, A. *Biochemistry* **1996**, *35*, 7299.
10. Liang, F.; Chen, R.; Lin, C.; Huang, K.; Chan, S. *Biochem. Biophys. Res. Commun.* **2006**, *342*, 482.
11. Choi, Y.; Kim, H.; Hwang, S.; Jeong, K.; Jung, S. *Bull. Korean Chem. Soc.* **2011**, *32*, 731.
12. Kaushik, S.; Mohanty, D.; Surolia, A. *Protein Sci.* **2011**, *20*, 465.
13. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
14. Jorgensen, W. L. *J. Chem. Phys.* **1983**, *79*, 926.
15. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
16. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
17. Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613.
18. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* **2004**, *47*, 1739.