Current Research Journal of Biological Sciences 3(3): 223-228, 2011
ISSN: 2041-0778

Received: February 10, 2011 Accepted: March 21, 2011 Published: May 05, 2011

# The Evaluation of Neuraminidase Type 2 Models

Sigit Jaya Herlambang and Rosari Saleh
Departement of Physics, Mathematics and Science Faculty, University of Indonesia,
Kampus UI Depok, F Building, Depok, 16424 Indonesia

**Abstract:** The aim of this study is to predict a model with the smallest errors by comparing built homology models using three different methods instead of using just one method. The generated models from building homology models always contain errors. The outcomes with smaller errors are the best models to be used for further study. The combination of three methods resulted in the determination of one model, the model no. 2, which is better than the other constructed models. The DOPE score of this model was -36613.3 which held the 4$^{th}$ place in DOPE score rank and the score from profiles-3D was 195.05 which was the best in the verify score rank. The online protein models assessment PROCHECK results showed that 98.8% residues (without Pro and Gly) of this model are in the most favoured and additional allowed regions. The results show that the three methods used, may help to select a better model to be simulated by considering different perspectives than if we choose a model with only one method perspective.

**Key words:** Homology modeling, influenza, model assessment, neuraminidase

## INTRODUCTION

Neuraminidase (NA) is one of the surface glycoprotein of the influenza virus that has been considered to be a suitable target to inhibit the virus spread because it has a major role in viral replication (Liu *et al*., 1995; McKimm-Breschkin *et al*., 1996; Palese and Schulman, 1977). After the budding virion in the edge of host cell mature, this sialydase catalyze the cleavage process in the junction of budding virus and host cells. The cleavage process could determine the virulence of virus. The rate of the releasing process could determine the number of viruses that may infect another host cells. Furthermore, it could aggravate the symptoms in the infected species. The NA glycoprotein has been shown by many experimental studies (in vivo and in vitro) to have a major role like haemagglutinin (HA) in viral replication (Mitnaul *et al*., 2000). Following experimental studies, researchers then diversified the research field into a complementary theoretical and computational-based research named in silico.

In silico research improved the knowledge of molecular proteins, including NA. The ability of NA binding affinity and high catalytic activity was the greatest aim in the study of this molecule. NA is also known to have the unique ability of rejecting inhibitor drug molecules after antigenic shift and antigenic drift even if it was by a single mutation (Russell *et al*., 2006; Chachra and Rizzo, 2008; McKimm-Breschkin *et al*., 1998; Mishin *et al*., 2005; Sheu *et al*., 2008; Wetherall *et al*., 2003; McKimm-Breschkin *et al*., 2003; Yen *et al*., 2007; Meijer *et al*., 2009; Monto *et al*., 2006; Tamura *et al*., 2009). The latest issue in NA research is to design and improve the structure of ligand based on the sialic acid molecule which could inhibit better than the rejected drugs using Molecular Dynamics Simulation (MDS).

To provide the accurate data for MDS, we must prepare an optimal molecule to simulate. The limitation of the crystallized structure molecule of neuraminidase provides potential growth of the technique known as homology modeling. This is based on the three dimensional (3D) structure prediction used as an alternative way to gain insight on the activity and the behavior of protein as long as that particular model posses the correct amount of sequence identity for further homology modeling (Krieger *et al*., 2003).

The objective of this study is to determine the best model by comparing structures with three different methods. The evaluations of the models generated from homology modeling are needed to create a better simulation. In this study, we try to evaluate the best model we can achieve with the methods known as DOPE (Shen and Sali, 2006) and 3D-profiles (Lüthy *et al*., 1992; Ramachandran *et al*., 1963). Both methods have its advantage. DOPE score is useful to check the reliability of the model by its stability and 3D-profiles evaluates the likelihood that a residue should be present within its current environment. After the evaluation, we make re-evaluation with PROCHECK (Fiser and Sali, 2000; Laskowski *et al*., 1993; Morris *et al*., 1992) to look at the

**Corresponding Author:** Sigit Jaya Herlambang, Departement of Physics, Mathematics and Science Faculty, University of Indonesia, Kampus UI Depok, F Building, Depok, 16424 Indonesia.
Tel: +62-021-7694975; Fax: +62-021-7515171

Table 1: The identity percentage of several potential templates

| Templates | Type | Identity (%) |
|---|---|---|
| 1INF | B | 30 |
| 2HU0 | N1 | 45.7 |
| 2BAT | N2 | 89.7 |
| 1V0Z | N6 | 51.2 |
| 2HT7 | N8 | 47.4 |
| 2QWK | N9 | 49 |

3D-profiles. The parameter for MODELER was set with DOPE-HR method, which is very similar to the DOPE (Discrete Optimized Protein Energy) method but is obtained at higher resolution. The parameter for 3D-profiles was set with smooth windows sized 10 and Kabsch-Sander (1983) algorithm for secondary structure stereo chemical properties of the models. This is useful to identify regions with errors in backbone geometry. Using these three evaluation methods, we can compare the results to seek a better model to conduct a simulation.

## MATERIALS AND METHODS

The NA amino acid sequence of A/chicken/ Pennsylvania/ 1370/1983 was taken from NCBI (2010) with the accession code ABI85148 and ACZ45193. This NA is obtained from the high pathogenic avian influenza virus which is mutated from the low pathogenic viruses that attacked 6 months before the pathogenic one spread in Pennsylvania in 1983 (Bean *et al*., 1985).

The NA protein templates for homology modeling were selected from crystallized structures stored in the protein data bank of RCSB (2011). The crystal structures of NA's were limited and in this study we achieved multiple sequence analysis with 6 potential templates with BLOSUM 30 scoring matrix. The sequences were 1INF, 2HU0, 2BAT, 1V0Z, 2HT7 and 2QWK.

Building homology models were produced using MODELER (Sali *et al*., 1995) which is integrated in Accelrys Discovery Studio 2.1. The parameters during this process were a medium optimization level of models and cut overhangs to remove the terminal unaligned residues in model sequence for comparison reason with the template. The homology modeling generated 10 models that were produced by known N2 structure template 2BAT after typed with CHARMm forcefield. The template 2BAT was used because by multiple sequence analysis it showed that this structure is the closest one with highest sequence identity (Table 1).

The selected model was then evaluated to verify its reliability with DOPE method using MODELER and method. The model that has a good rank in both of methods (Table 2) was sent to Swiss Expasy (Arnold *et al*., 2006) to be analyzed with PROCHECK to see the stereo chemical perspective of the model.

All procedures conducted in the study started at early December until middle January at the Physics Department, Universitas Indonesia. Software used was Accelrys Discovery Studio 2.1 ran on an amd x3, 3 GHz RAM desktop.

## RESULTS

Multiple sequence analysis was done to compare the identity of a few potential templates. By sequence identity, it showed that the best fit template molecule to the sequence is 2BAT which is a type 2 neuraminidase. This result is in accordance to Bean *et al*. (1985) whom told that the virus is neuraminidase type 2. Furthermore the results also showed that the 1370 sequences were close to such as N6 and N9 sequences which phylogeneticaly grouped in the group-2 which has a closed conformation same as N2 (Amaro *et al*., 2009; Thompson *et al*., 1994; Patel *et al*., 2009).

In Table 2 we provide three different assessment models and template. The first score column is the verify score which was obtained from profiles-3d method. The value ranges from 180.89 to 195.09. This range of value is above high expected score which only 176.92. This shows that all the models produced were valid and have good compatibility to be used as a hypothetical protein structure to be simulated. The profiles-3d verification for the known structure (the template 2BAT) was also executed to compare scores of the modeled structure with its template. From this procedure it was found that the score of the modeled structures were in the same order as

Table 2: The scores of 10 models and template 2BAT for comparison molecule

| Template and models | Verify score | DOPE | Most favoured and additional residues (%) (without Gly and Pro) |
|---|---|---|---|
| 2BAT | 195.09 | -39409.9 | 98.5 |
| Model 1 | 190.31 | -36959.3 | 98.8 |
| Model 2 | 195.05 | -36613.3 | 98.8 |
| Model 3 | 184.29 | -35980.6 | 98.5 |
| Model 4 | 184.15 | -35836.4 | 98.8 |
| Model 5 | 190.86 | -36154.7 | 98.2 |
| Model 6 | 192.80 | -35913.0 | 99.1 |
| Model 7 | 181.72 | -35566.3 | 98.5 |
| Model 8 | 182.71 | -37079.6 | 98.8 |
| Model 9 | 180.89 | -36408.1 | 98.8 |
| Model 10 | 193.56 | -36908.4 | 98.8 |

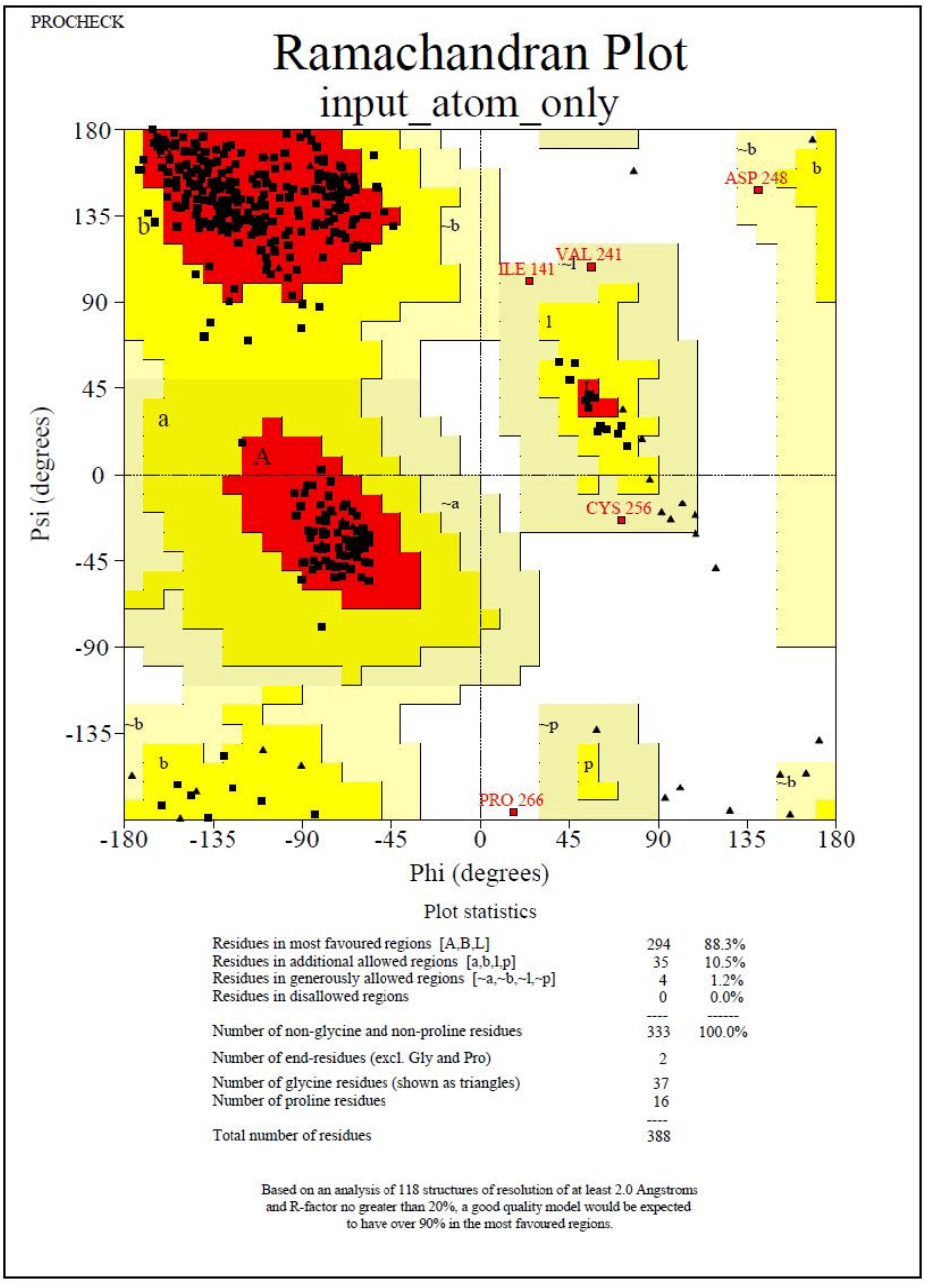


Fig. 1a: Ramachandran plots of: Model 2

the template. The residues in the structures of all models seem to be placed in appropriate environment. The best model score was the model 2 which has a value of 195.05.

The second column contained DOPE-HR scores which were obtained from the DOPE method. The value varies in range from -35566.3 to -37079.6. They are all higher than the DOPE-HR score of template molecule 2BAT which is -39409.9. The minimum score reached by model 8 with -37079.6 is also the closest one with the template molecule and could be said as the most relatively stable and reliable model to be simulated.

The third column shows the percentage of residues which are placed in most favoured and additionally allowed region. The results were divided into 4 ranks because some of them had the same value. The first rank was filled by model 6 with 99.1%, second rank are models 1, 2, 4, 8, 9 and 10 with 98.8%, the third rank are models 3 and 7 with 98.5%, and fourth is model 5 with

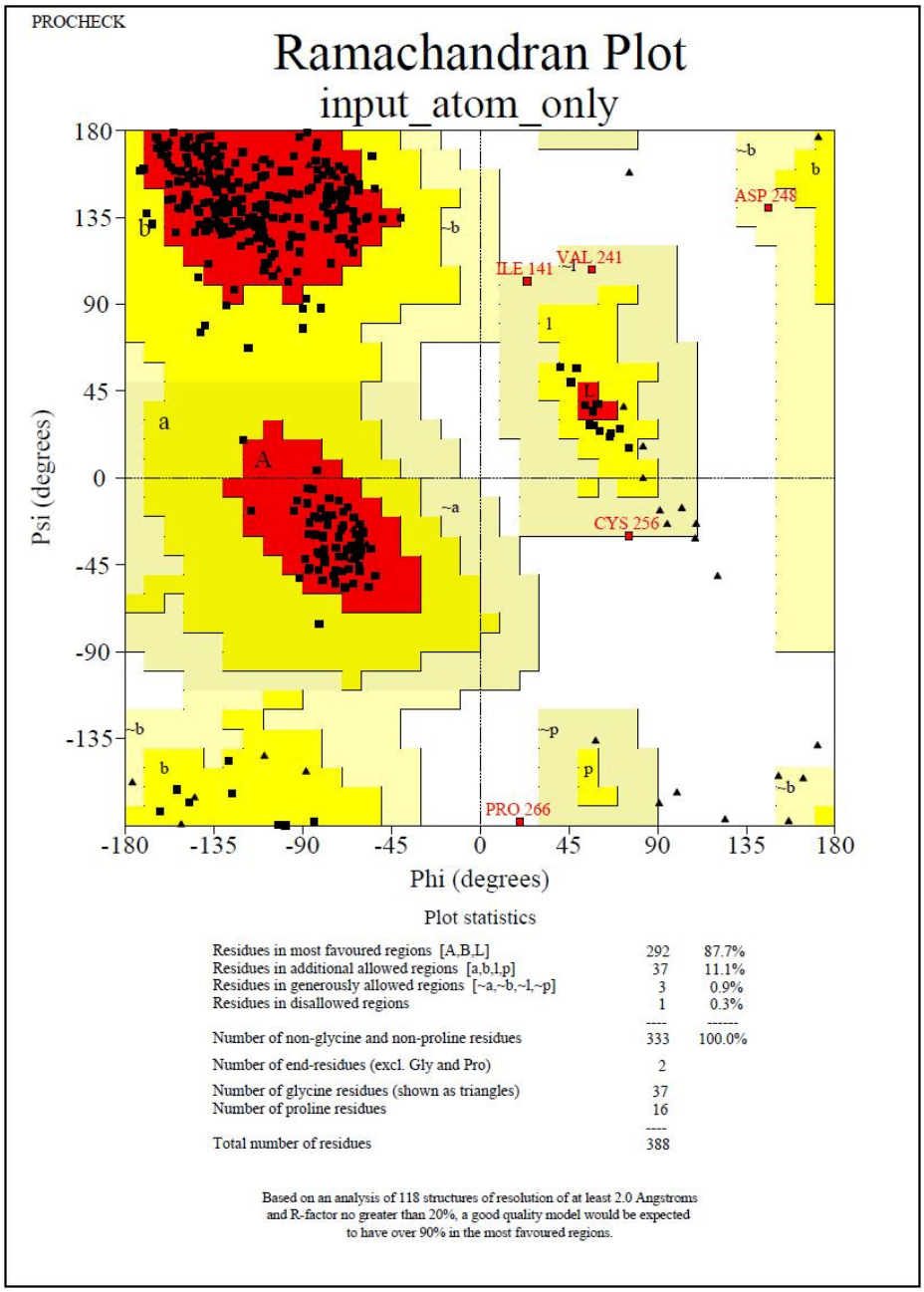


Fig. 1b: Ramachandran plots of Model 10

98.2%. The template 2BAT has the same percentage with the third rank.

The result from PROCHECK did not contain Gly and Pro residues in the same Ramachandran plots because they have special properties. It was provided in specific residue Ramachandran plots. There is no residue in Gly Ramachandran plots which placed in disallowed region but in Pro Ramachandran plots there is Pro266 residue which placed in disallowed region. It occurred in all models that were generated.

## DISCUSSION

The aim of this experiment is attempting to determine the best model by comparing three methods. Not to legitimate one or more methods by scrutinizing the other methods. We select the better model which had good scores in all the three methods. The reason for doing this is because all the methods have advantages in their own respect and by applying these three methods we could

produce an improved approach in determining the most accurate model to simulate in further studies.

From the first method, we have all models as a compatible model but cannot determine the model which has greatest the verify score before checking with another method since we need to see the stability of the model which we could see through DOPE score results. From the second method we have another three models which is better than model 2. They are model 8, 1 and 10 which placed the model 2 as the 4$^{th}$ rank. After closely looking back at the verify-3d score results, only model 10 has a potential to replace model 2 for the best selected model.

After selecting model 2 and 10 as potentially the best model, we take a look at the PROCHECK assessment method results. Model 10 are in the same rank with model 2 for percentage of residues in the most favourable and additionally allowed regions. Then we compared the Ramachandran plots of both models and there is one residue which placed in the disallowed region for model 10 which did not occur in model 2. By this, comparing with model 10 in Ramachandran plots, model 2 is better (Fig. 1a, b).

The value of DOPE score was not compared with another research in model assessment. It was because the DOPE score is depends on the structural probability of molecule. The pairwise statistical potential function of a protein with numbers of N atoms can be represented by the pairwise probability density function. The molecule which has more atoms will have a more negative DOPE score.

It also occurred in 3D-profiles score. Since we did not find the similar molecule to compare, we are not able to accomplish the comparison of verify score. We could not compare it with another molecule because the residues that constitute the structure are different.

The values of PROCHECK results were compared with other molecules model assessment and show that our result model is in the same order. The amylase (Patel *et al*., 2009), ninjurin (Khatri *et al*., 2010) and lycopene (Satpathy *et al*., 2010) best model evaluation showed about 87.3, 87.3 and 88.2%, respectively, residues were in the most favoured region while our best model has 88.3%.

## CONCLUSION

All neuraminidase models were generated from homology modeling may produce errors that we could not see with only one method. The use of a few methods which each of them could act on behalf of unique perspective may provide advantages to selecting the best model to simulate. Some of models were produced could have an error in stereo chemistry even its stability is the better than the other. This is why we try to choose the best model by comparing with several methods. The model chosen was model 2 which has preferably good rank in every method we carried out.

## ACKNOWLEDGMENT

We would like to express gratitude towards Ding Ming Chee of Accelrys Singapore for the Accelrys Discovery Studio 2.1 trial sent to us.

## REFERENCES

Amaro, R.E., X. Cheng, I. Ivanov, X. Dong and J.A. McCammon, 2009. Characterizing loop dynamics and ligand recognition in human-and avian-type influenza neuraminidases via generaslized born molecular dynamics and end-point free energy calculations. J. Am. Chem. Soc., 131(13): 4702-4709.

Arnold, K., L. Bordoli, J. Kopp and T. Schwede, 2006. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics, 22: 195-201.

Bean, W.J., Y. Kawaoka, J.M. Wood, J.E. Pearson and R.G. Webster, 1985. Characterization of virulent and avirulent A/Chicken/Pennsylvania/83 influenza A viruses: Potential role of defective interfering RNAs in nature. J. Virol., 54: 151-160.

Chachra, R. and R.C. Rizzo, 2008. Origins of resistance conferred by the R292K neuraminidase mutation via molecular dynamics and free energy calculations. J. Chem. Theory Comput., 4: 1526-1540.

Fiser, R.K. and A. Sali, 2000. Modeling of loops in protein structures. Protein Sci., 9: 1753-1773.

Kabsch, W. and C. Sander, 1983. Biopolymers Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers, 22: 2577-2637.

Khatri, S., S. Patel, J. Choubey, S.K. Gupta and M.K. Verma, 2010. Insilico 3D structure prediction of cell membrane associated protein ninjurin (Homosapiens). Curr. Res. J. Biol. Sci., 2(1): 1-5.

Krieger, E., S.B. Nabuurs and G. Vriend, 2003. Homology Modeling. Structural Bioinformatics. Wiley-Liss Inc., pp: 507-508.

Laskowski, R.A., M.W. MacArthur, D.S. Moss and J.M. Thornton, 1993. PROCHECK - a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr., 26: 283-291.

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

Lüthy, R., J.U. Bowie and D. Eisenberg, 1992. Assessment of protein models with three-dimensional profiles. Nature, 356: 83-85.

McKimm-Breschkin, J.L., M. McDonald, T.J. Blick and P.M. Colman, 1996. Mutation in the influenza virus neuraminidase gene resulting in decreased sensitivity to the inhibitor 4-guanidino-Neu5Ac2en leads to instability of the enzyme. Virology, 225: 240-242.

McKimm-Breschkin, J.L., A. Sahasrabudhe, T.J. Blick, M. McDonald, P.M. Colman, G.J. Hart, R.C. Bethell and J.N. Varghese, 1998. Mutations in a conserved residue in the influenza virus neuraminidase active site decreases sensitivity to Neu5Ac2en-derived inhibitors. J. Virol., 72: 2456-2462.

McKimm-Breschkin, J., T. Trivedi, A. Hampson, A. Hay, A. Klimov, M. Tashiro, F. Hayden and M. Zambon, 2003. Neuraminidase sequence analysis and susceptibilities of influenza virus clinical isolates to zanamivir and oseltamivir. Antimicrob. Agents Ch., 47: 2264-2272.

Meijer, A., A. Lackenby, O. Hungnes, B. Lina, S. van der Werf, B. Schweiger, M. Opp, J. Paget, J. van de Kassteele, J. Hay and M. Zambon, 2009. European Influenza Surveillance Scheme: Oseltamivir-resistant influenza virus A (H1N1), Europe, 2007-2008 season. Emerging Infectious Diseases Vol. 15: 552-560Mishin, V.P., F.G. Hayden and L.V. Gubareva, 2005. Susceptibilities of antiviral-resistant influenza viruses to novel neuraminidase inhibitors. Antimicrob. Agents Ch., 49: 4515-4520.

Mitnaul, L.J., M.N. Matrosovich, M.R. Castrucci, A.B. Tuzikov, N.V. Bovin, D. Kobasa and Y. Kawaoka, 2000. Balanced hemagglutinin and neuraminidase activities are critical for efficient replication of influenza A virus. J. Virol., 74: 6015-6020

Monto, A.S., J.L. McKimm-Breschkin, C. Macken, A.W. Hampson, A. Hay, A. Klimov, M. Tashiro, R.G. Webster, M. Aymard, F.G. Hayden and M. Zambon, 2006. Detection of influenza viruses resistant to neuraminidase inhibitors in global surveillance during the first 3-years of their use. Antimicrob. Agents Ch., 50: 2395-2402.

Morris, A.L., M.W. MacArthur, E.G. Hutchinson and J.M. Thornton, 1992. Stereochemical quality of protein structure coordinates. Proteins, 12: 345-364

NCBI, 2010. Influenza Virus Sequence Database. Retrieved from: http://www.ncbi.nlm.nih.gov/genomes/FLU/Database/nph-select.cgi?go=database

Palese, P. and J.L. Schulman, 1977. Inhibitors of Viral Neuraminidase as Potential Antiviral Drugs. In: Oxford, J.S. (Ed.), Chemoprophylaxis and Virus Infections of the Respiratory Tract. CRC Press, Inc., Cleveland, pp: 189-205.

Patel, A., R. Dewangan, S. Khatri, J. Choubey, S.K. Gupta and M.K. Verma, 2009. Identification of insilico 3D structure of amylase (Drosophila melanogaster) and comparative computational studies. J. Eng. Technol. Res., 1(2): 039-045.

Ramachandran, G.N., C. Ramakrishnan and V. Sasisekharan, 1963. Stereochemistry of polypeptide chain configurations. V. Mol. Biol., 7: 95.

RCSB Protien Data Bank, 2011. A Resource for Studying Biological Macromolecules. Retrieved from: http://www.pdb.org/pdb/home/home.do

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

Sali, A., L. Pottertone, F. Yuan, H. van Vlijmen and M. Karplus, 1995. Evaluation of comparative protein modeling by MODELLER. Proteins, 23: 318-326.

Satpathy, R., R.K. Guru, R. Behera and A. Priyadarshini, 2010. Homology modelling of lycopene cleavage oxygenase: The key enzyme of bixin production. J. Comput. Sci. Syst. Biol., 3(3): 059-061.

Shen, M.Y. and A. Sali, 2006. Statistical potential for assessment and prediction of protein structures. Protein Sci., 15: 2507-2524.

Sheu, T.G., V.M. Deyde, M. Okomo-Adhiambo, R.J. Garten, X. Xu, R.A. Bright, E.N. Butler, T.R. Wallis, A.I. Klimov and L.V. Gubareva, 2008. Surveillance for neuraminidase inhibitor resistance among human influenza A and B viruses circulating worldwide from 2004 to 2008. Antimicrob. Agents Ch., 52: 3284-3292.

Tamura, D., K. Mitamura, M. Yamazaki, M. Fujino, M. Nirasawa, K. Kimura, M. Kiso, H. Shimizu, C. Kawakami, S. Hiroi, S. Takahashi, M. Hata, H. Minagawa, Y. Kimura, S. Kaneda, S. Sugita, T. Horimoto, N. Sugaya and Y. Kawaoka, 2009. Oseltamivir-resistant influenza a viruses circulating in Japan. J. Clin. Microbiol., 47: 1424-1427.

Thompson, J.D., D.G. Higgins and T.J. Gibson, 1994. Improved sensitivity of profile searches through the use of sequence weights and gap excision. Comput. Appl. Biosci., 10: 19-29.

Wetherall, N.T., T. Trivedi, J. Zeller, C. Hodges-Savola, J.L. McKimm-Breschkin, M. Zambon and F.G. Hayden, 2003. Evaluation of neuraminidase enzyme assays using different substrates to measure susceptibility of influenza virus clinical isolates to neuraminidase inhibitors: Report of the Neuraminidase Inhibitor Susceptibility Network. J. Clin. Microbiol., 41: 742-750.

Yen, H., N.A. Ilyushina, R. Salomon, E. Hoffmann, R.G. Webster and E.A. Govorkova, 2007. Neuraminidase inhibitor-resistant recombinant A/Vietnam/1203/04 (H5N1) influenza viruses retain their replication efficiency and pathogenicity *in vitro* and *in vivo*. J. Virol., Vol 81: 12418-12426.