Homology Modeling and Docking Studies Of Hemagglutinin Protein Of Influenza A (H1N1) Virus With Selected Ligands - A Computer Aided Structure Based Drug Design Approach To Find A Suitable Inhibitor

A.Bhattacharjee [1*], P.Yadav[2], B.J. Mylliemngap[3], E.Annipindi [4]

Abstract

Swine flu is an emerging infectious disease cause by the influenza A (H1N1) virus. The virus has a special protein found on the surface known as Hemagglutinin, which is an antigenic glycoprotein responsible for binding virus to the monosaccharide sialic acid of its target cell that is being infected. An *In silico* attempt was made to characterize a newly sequenced Hemagglutinin protein of influenza A virus (H1N1) to deduce its structural information and to identify the potential drug to inhibit the protein. For that an effort was taken to deduce the 3-D structure of this protein and to identify and bind the active site of Hemagglutinin protein with docking technique. Using drug data bank and NIST Standard Reference Database the 3-D structure of ligand were retrieved. QSAR studies were done for the available synthetic drugs like Oseltamivir and zanamivir and also the natural product like shikimic acid to know about the physiochemical properties using software Molinspiration. It was noted that Amantadine an antiviral drug which is not in used for the treatment of swine flu and a natural compound Shikimic acid also shows a better docking energy in conjunction to available synthetic drug oseltamivir and zanamivir with this protein.

Keywords: influenza A (H1N1) virus, Homology modelling, docking, QSAR, ligand

Introduction

Swine influenza (also called swine flu, hog flu and pig flu) is an infection of a host animal by any one of several specific types of swine influenza virus. A swine influenza virus (SIV) is any strain of the influenza family of viruses that is usually hosted by pigs. As of 2009, the known SIV strains are the influenza C virus and the subtypes of the influenza A virus known as H1N1, H1N2, H3N1, H3N2, and H2N3. Swine flu is common throughout pig populations worldwide. Transmission of swine influenza virus from pigs to humans is not common and does not always cause human influenza, often only resulting in the production of antibodies in the blood. The meat of the animal poses no risk of transmitting the virus when properly cooked. If transmission does cause human influenza, it is called zoonotic swine flu. People who work with pigs, especially people with intense exposures, are at increased risk of catching swine flu. Since then, fifty confirmed transmissions have been recorded. Rarely, these strains of swine flu can pass from human to human. In humans, the symptoms of swine flu are similar to those of influenza and of influenza-like illness in general, namely chills, fever, sore throat, muscle pains, severe headache, coughing, weakness and general discomfort (Knobler *et al.,* 2005). Pigs can also become infected with human influenza, and this appears to have happened during the 1918 flu pandemic. The 2009 swine flu outbreak in humans is due to a new strain of influenza A virus subtype H1N1 that contains genes closely related to swine influenza (V Trifonov *et al.,* 2009). The origin of this new strain is unknown. However, the World Organization for Animal Health (OIE) reports that this strain has not been isolated in pigs (Maria Zampaglione *et al.,* 2009). This strain can be transmitted from human to human and causes the normal symptoms of influenza. More than 70 countries are now reporting cases of human infection with novel H1N1 flu. This number has been increasing over the past few weeks, but many of the cases reportedly had links to travel or were localized outbreaks without community spread. The WHO designation of a pandemic alert Phase 6 reflects the fact that there are now ongoing community level outbreaks in multiple parts of world. The virus has a special type of antigenic glycoprotein found on its surface called Hemagglutinin protein. It is responsible for binding the virus to the cell that is being infected (Nelson *et al.,* 2005). Hemagglutinin (HA) has two primary functions: allowing the recognition of target vertebrate cells, accomplished through the binding of these cells' sialic acid-containing receptors, and allowing the entry of the viral genome into the target cells by causing the fusion of host endosomal membrane with the viral membrane (Nelson et al., 2005). For that HA binds to the monosaccharide sialic acid which is present on the surface of its target cells. This causes the viral particles to stick to the cell's surface. The cell membrane then engulfs the virus and the portion of the membrane that encloses it pinches off to form a new membrane-bound compartment within the cell called an endosome, which contains the engulfed virus. The cell then attempts to begin digesting the contents of the endosome by acidifying its interior and transforming it into a lysosome. However, as soon as the pH within the endosome drops to about 6.0, the original folded structure of the HA molecule becomes unstable, causing it to partially unfold, and releasing a very hydrophobic portion of its peptide chain that was previously hidden within the protein. This so-called "fusion peptide" acts like a molecular grappling hook by inserting itself into the endosomal membrane and locking on. Then, when the rest of the HA molecule refolds into a new

structure (which is more stable at the lower pH), it "retracts the grappling hook" and pulls the endosomal membrane right up next to the virus particle's own membrane, causing the two to fuse together. Once this has happened, the contents of the virus, including its RNA genome, are free to pour out into the cell's cytoplasm (David S. Goodsell *et al.,* April 2006).

So, if we modulate the hemagglutinin protein then H1N1 virus would not be able to bind to sialic acid of target cell surface and that way it would be unable to enter in to the cell cytoplasm. So there is a urgent need to build a model structure of influenza A virus Hemagglutinin protein (IAVHA) whose only primary information in the form of sequence is available. Therefore the present work involves the extensive use of tool & graphical software for creation of protein structure and to screen the drug that will dock/bind to this site to inhibit the action of protein. The process involves the prediction of ligand conformation and orientation within the target binding site. The suitable ligand was chosen at the end of docking process. The energy value of docking and the number of hydrogen bonds between the active site residues and the ligands under investigation was taken into consideration for coming into conclusion regarding the best pose and the binding ability

Materials and Method

Various tool and software were used to analyze this protein sequence and assign its structure and to study its docking properties. The primary sequence of hemagglutinin protein (Acc.No.gi|227831759|gb|ACP41926.1|) was obtained from Gene Bank at NCBI (www.ncbi.nlm.nih.gov). Science only primary structural information was available from NCBI, and no structure in the form of X-Ray crystallographic data was available from the Protein Data Bank (www.pdb.org/pdb/), hence to identify 3-D structure of protein Homology modeling was done. For that template sequence and structure are retrieved from PDB and modeller9v7(A. Sali *et al.,*1993) it was used for modeling protein however loop modeling and energy minimization work was done by Swiss PDB Viewer(http://spdbv.vital-it.ch/). Then pymol (http://pymol.sourceforge.net) and rasmol (http://www.openrasmol.org) was used for molecular visualization. Physical and chemical parameters for protein were calculated by Protoparam (http://www.expasy.ch/tools/protparam.html) and for secondary structure prediction GOR (http://npsa-pbil.ibcp.fr) was used. Validity of 3-D structure of modeled protein was checked by procheck software (http://nihserver.mbi.ucla.edu/SAVES). Active sites of protein were retrieved by software CASTp server. The structures of ligand molecules were obtained from the Drug Data Bank while the 3-D structure of natural drug product shikimic acid were obtained by NIST Standard Reference Database(http://webbook.nist.gov/chemistry). To observe their molecular properties Molinspiration (www.molinspiration.com) was used. ArgusLab 4.0.1 (MarkA.Thompson Planaria Software LLC, Seattle, WA) was used for docking purpose and for visualization docking molecule structure Molegro Molecular Viewer was used.

Result & Discussion

Prediction of interaction energies between ligand and receptor has been a major challenge for drug docking. Influenza A virus Hemagglutinin protein (IAVHA) is an antigenic glycoprotein and its structure deduced by homology modelling shows that there are 6 possible templates for protein modelling out of which we have chosen 1RVT having 84% identity to hemagglutinin protein structure of modeled protein by modeller9v7. The Modeled protein structure (fig 1) showed 329 H-bond, 566 groups and 4450 atoms. The secondary structure is composed of 14 helices, 59 turns, and 59 strands. The physiochemical property of

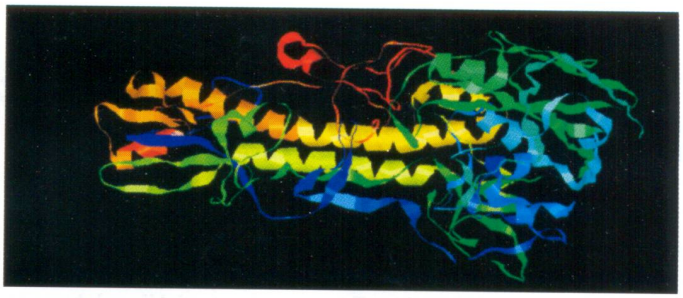


Figure 1: Three dimensional structure of modeled hemagglutinin protein showing helices and turns

hemagglutinin revealed the number of amino acid to be 566, Mol. Wt. of 63239.5 and theoretical isoelectric point at 6.93.The maximum number of amino acid sequence present in sequence was found to be that of serine (8.5) and least was that of methionine (1.4).The total number of positively charge residue (A+L) were 60 and the negatively charged residue (A+G) were 61.The grand hydropathicity was calculated to be -0.355. The secondary structure analysis of protein was done and random coil found to be most frequent (48.76%), followed by extended strand (26.15%) and alpha helix was found to be least frequent (25.09%).

Potential Active site of protein calculated by CASTp (Fig.2) showed there are several pockets which fit in the role of active site. Active site target range from residue Leu417-Tyr534. For docking, already available market ligand zanamivir and oseltamivir were selected because it would bind to the enzyme as substrate molecule.

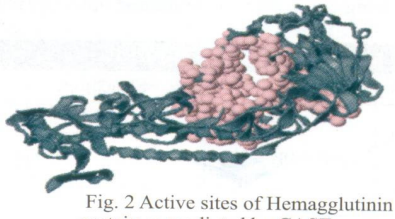


Fig. 2 Active sites of Hemagglutinin protein as predicted by CASTp server

Shikimic acid an important biochemical intermediate in plant and microorganism and used as a base material for production Tamiflu was also selected since it is a natural product and it will not cause any side effect. Other antiviral drugs were also taken in consideration for docking but out of which only one drug rimantadin gave some significant result. QSAR studies of the drugs shows that all of the drugs fall within the normal range if we calculate the molecular weight and Lipophilicity (clogP), but

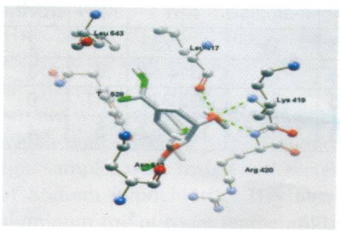


Fig 3(a)

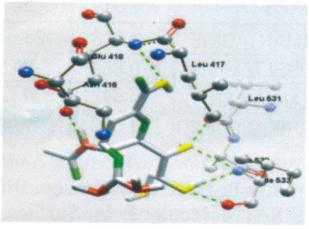


Fig 3(b)

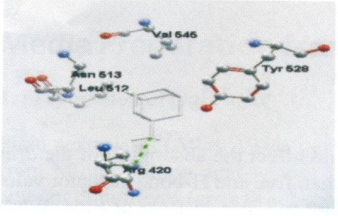


Fig 3(c)

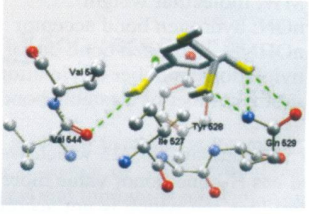


Fig 3(d)

Fig. 3 Hydrogen bond interaction of the protein with ligands 3(a) Oseltamivir and 3(b) Zanamivir 3(c) Amantadine 3(d) Shikimic acid.

if it comes to other values like H-bond Donars, H-bond Acceptor and TPSA according to the Lipinski rule of five, it was noted that oseltamivir, zanamivir and shikimic acid have high TPSA value and zanamivir has H-bond Donor more than five, H-bond Acceptor more than ten which is not according to Lipinski Rule of Five. (Lipinski et al., 1997). The QSAR analyses done for following ligands are tabulate below.

Scoring algorithm is used to calculate the energy value of the docked complex and its stability increase with decrease in energy value. The energy value obtained by docking IVAHA with potential ligand shown in Table -3 shows that Oseltamivir which is the essential drug for the treatment of Swine flu on docking shows energy of -9.33 kcal/mol and a hydrogen bond interaction of 3 (Fig 3 (a)). Though Oseltamivir is a primary drug for Swine flu it high values of TPSA (90.66) which would affect the absorption of the drug. Also Zanamivir, another drug used for the treatment of Swine flu in conjunction with oseltamivir also has TPSA value of 200.725 however has a docking energy score of -7.82kcal/mol and 5hydrogen interactions (Fig 3b). Besides these drugs, an antiviral drug has been used for docking studies was Amantadine has a Low docking energy score (-9.04) with one hydrogen bond interaction (Fig 3(c)) however a natural product shikimic acid has docking energy score of -7.37kcal/mol and four hydrogen interactions (Fig 3(d)).

The comparative study of all these drugs shows that two drugs Oseltamivir and Amantadine have lowest docking energy compare to zanamivir but if we considered H-bond then it was reported that zanamivir has maximum number of H-bond interaction that is showing zanamivir will give better docking compare to Amantadine however has

Table 1: Linear structures of the ligands

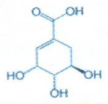
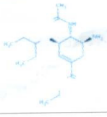
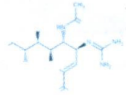
Drugs	Structures	Drugs	Structures
AMANTADINE		SHIKIMIC ACID	
OSELTAMIVIR		ZANAMIVIR	

Table 2: Validation of drugs by Molinspiration.

Drugs	miLogP	tpsa	n atoms	MW	nON	nOHNH	nviolations	nrotb	Volume
AMANTADINE	2.648	26.023	11	151.253	1	2	0	0	159.196
OSELTAMIVIR	0.852	90.66	22	314.41	6	3	0	8	309.599
ZANAMIVIR	-3.642	200.725	23	332.313	11	9	2	6	283.974
SHIKIMIC ACID	-1.569	97.983	12	174.152	5	4	0	1	147.55

Where,
miLogP: LogP (partition coefficient)
tPSA: Molecular Polar Surface Area
natoms: number of atoms
MW: molecular weight
nON: hydrogen bond acceptor
nOHNH: hydrogen bond donor
nviolations: number of violations
nrotb: number of rotatable bonds

TPSA value of 200.725 which would affect the absorption of the drug and has H-bond donor value more than five and H-bond acceptor value more than ten which is not according to Lipinski Rule of Five. While shikimic acid has highest docking energy compares to all ligands but shows 4 H-bond interaction however a High TPSA (97.983) is a problem. The energy value of the docked complex shows its increase

Table 3. Energy value after docking with ligands with H-bond interaction

S.NO.	LIGAND	ENERGY VALUE	NO. OF HYDROGEN BOND
1.	OSELTAMIVIR	-9.33	3
2.	ZANAMIVIR	-7.82	5
3.	SHIKIMIC ACID	-7.37	4
4.	AMANTADINE	-9.04	1

stability with decrease in energy value.

Conclusion:

The future insight would be to search for other possible modification in shikimic acid compound to get better drug with high efficiency. This can be studied in wet lab experimentally to identify their efficiency in binding and inhibiting the protein.

Refrences:

1. ArgusLab 4.0, Mark A. Thompson, Planaria Software LLC, Seattle, http://www.ArgusLab.com.
2. A. Sali and T. L. Blundell. Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234,779815, 1993.
3. David S. Goodsell, April 2006 PDB molecule of the month: Hemagglutinin
4. Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis 18, 2714-2723.
5. International Committee on Taxonomy of Viruses. "The Universal Virus Database, version 4: Influenza A".
6. Joe Dundas, Zheng Ouyang, Jeffery Tseng, Andrew Binkowski, Yaron Turpaz, and Jie Liang 2006. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated resiudes. Nucleic Acid Research, 34:W116-W118.
7. Knobler S, Mack A, Mahmoud A, Lemon S, ed. "1: The Story of Influenza". The Threat of Pandemic Influenza: Are We Ready? Workshop Summary (2005). Washington, D.C.: The National Academies Press. p. 75
8. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J.Experimental and computational approaches to estmate solubility and permeability in drug discovery and development settings.Adv. Drug Delivery Rev. 1997, 23, 4-25.
9. Maria Zampaglione (April 29, 2009). "Press Release: A/H1N1 influenza like human illness in Mexico and the USA: OIE statement". World Organisation for Animal Health. Http://www.oie.int/eng/press/en_090427.htm. Retrieved on April 29, 2009.
10. Nelson DL and Cox MM, 2005. Lehninger's Principles of Biochemistry, 4th edition, WH Freeman, New York, NY
11. V Trifonov, H Khiabanian, B Greenbaum, R Rabadan (30 April 2009). "The origin of the recent swine influenza A(H1N1) virus infecting humans". Eurosurveillance 4 (17). Http://www.eurosurveillance.org/images/dynamic/EE/V14N17/art19193.pdf.
12. White JM, Hoffman LR, Arevalo JH, et al., 1997. Attachment and entry of influenza virus into host cells. Pivotal roles of hemagglutinin. In Structural Biology of Viruses. Chiu W, Burnett RM, and Garcea RL, editors. Oxford

1, 2, 3, 4 Bioinformatics Laboratory,
Department of Biotechnology and Bioinformatics,
Permanent Campus, North Eastern Hill University,
Umshing Shillong-22, India
For correspondence: Email: atanubioinfo@gmail.com