ABSTRACT

Objective: Swine flu is an infectious disease caused by the influenza A (H1N1) virus. The virus has a particular 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. The aim of the present study is to model the 3-dimensional structure of hemagglutinin of H1N1 virus and to analyze its structural conformation.

Methods: Homology model was built using Modeller 9v8 due to non specific structure of hemagglutinin. Structural assessment was done using Procheck tool of Exome Horizon for checking quality of the protein. The interaction study of rimantadine with the modeled protein was taken place in Exome Horizon for inhibition of the active sites of the virus.

Results: The binding energy calculation showed good results and the clusters were reported with the help molecular docking study. 10 clusters were formed having minimum binding energy ranges from -1.06 to -2.99 Kcal/Mol.

Conclusion: After docking study, it was found to be the aminocids GLY 10, LEU 2, ASP 19, TRY 24, ALA 35, ALA 36, ASP 37, GLN 38, ASP 109, GLU 150, LYS 153, ASN 154 are interacting with rimantadine.

Keywords: H1N1, Modeller9v8, Exome Horizon, Homology Modelling

INTRODUCTION

In 2009 the outbreak of the swine flu epidemic gave rise to a new strain of influenza virus [1]. It appeared first in quarter of 2009 in west coastal region of North America and spread to other countries through April to June, 2009 [2]. The causative virus of swine flu is a new strain of influenza A virus subtype H1N1 which has genes closely associated to swine flu [3]. The influenza A virus is a single stranded RNA virus having eight different segments. It was found that this segments of the swine influenza A virus (H1N1) are related to swine viruses from North America and the other two [Neuraminidase (NA) and matrix protein(M)] from swine viruses isolated in Europe/Asia [3]. However, the origin of this new strain is not known [4]. But according to the World Organization for Animal Health (OIE) this strain has not been isolated from pigs, but this strain is transmitted from humans to humans [5]. Within a short span of 5 months after its emergence infection cases of more than 290,000 with deaths of 3486 were reported worldwide [6]. According to World Health Organization (WHO) there are 414000 confirmed cases and nearly 5000 deaths reported to WHO in 2009 outbreak, but as the public health system had stopped confirming individual cases so the actual number of cases is estimated to be even higher [7]. The flu viruses spreads from person to person through coughing or sneezing of infected people or by coming in contact with a surface with the flu virus and then touching the mouth or nose. The symptoms of the illness include fever, cough, sore throat, runny nose, body aches, and body chills etc. which are at times accompanied by diarrhea and vomiting [8]. The swine flu virus contains a special surface protein called hemagglutinin (HA), which is an antigenic glycoprotein. Hemagglutinin has two main functions. It recognizes the target cells and binds to the target cells sialic acid receptors present on the cell surface. Thereafter the viral genome enters into the target cells by causing fusion of the host endosomal membrane with the viral membrane [9]. The cell then tries to digest the contents of the endosome by acidifying the contents. But when the pH of the endosome drops to 6.0, the folded structure of hemagglutinin molecule becomes unstable making it partially unfold and thereafter releasing a hydrophobic portion of the peptide chain which was hidden within the protein. This causes the hemagglutinin molecule to refold to new structure and pulls the endosomal membrane to virus particles own membrane and then both fuse together. This causes the virus to pour its RNA genome to the cell cytoplasm [4]. So, if we modulate the hemagglutinin protein then H1N1 virus would not be able to bind to the target cell through sialic acid receptor, making room for the virus to enter the cell cytoplasm [10]. It is seen that the receptor binding pocket of the H1N1 HA protein was found to be smaller than other strain of influenza A virus. This allows tighter binding of virus with the receptor [11]. Another glycoprotein is there in the lipid membrane of the virus namely neuraminidase (NA) which is responsible for initiation of viral infection [12]. The NA protein initiates release and spread of progeny virions following, the intracellular viral replication cycle. Influenza A including swine influenza virus is usually treated by use of antiviral drugs. Drugs like amantidine, rimantidine which are M2 ion channel inhibitor or drugs like oseltamivir and zanamivir which are neuraminidase inhibitor are commonly used in treatment of swine influenza. Both rimantidine and amantidine are derivatives of amantidine. Amantidine inhibits the viral infection by interfering with the viral protein M2, which is an ion channel required for viral particles to become uncoated after taken up by the cell [13].

MATERIALS AND METHODS

Sequence Retrieval: Influenza A virus protein sequences were downloaded from NCBI. The FASTA sequence of the protein was obtained from (http://www.ncbi.nlm.nih.gov). There were about 7840 sequences of protein. The sequence having accession number 1438.1 was selected for our study. The downloaded sequence is given below.
Blast was being performed in for the selected sequence. Blast tries to expand the alignments to the adjacent words without allowing gaps. If protein query sequence has been given then the output provides a list of all database sequence that is matched to the query sequence. Multiple Sequence alignment was then performed using T-Coffee tool available from EBI.

**Homology Modeling Techniques**

The Swiss-Model workspace was used to construct the hemagglutinin structure. It is a general program for homology modeling. The hemagglutinin sequence was downloaded from NCBI site (http://www.ncbi.nlm.nih.gov) in fasta format.

**Template Identification**

For building one comparative structure, one template was necessary. A template was selected with the help of template identification tool of Swiss model server. PDB ID for the selected templates was found to be 1vm7A. The sequence of 1vm7A was downloaded from NCBI.

**Alignment Of Template and Target Sequence**

Both the target and the template sequences were pasted in T-Coffee multiple sequence alignment and then these alignment results are submitted through clustal w of T-Coffee alignment tool.

**The T-Coffee Alignment**

CLUSTAL W (1.83) multiple sequence alignment

3HTO_B|2|CHAIN| GLFGAMAGIEEGGWGIDGWYGYHHQNEQSGYAADQKSTQNAIDGTN gi|1|pdb|3HTT| GLFGAMAGIEEGGWGIDGWYGYHHQNEQSGYAADQKSTQNAIDGTN  

3HTO_B|2|CHAIN| KVNNSIEKMTQNTQFTAVKGFENLNRHERENLNVKVDGFLEWVYNNAELLV gi|1|pdb|3HTT| KVNNSIEKMTQNTQFTAVKGFENLNRHERENLNVKVDGFLEWVYNNAELLV  

3HTO_B|2|CHAIN| LLENERTLDFHDSNVSNLYEKVSQQLRNNAKEINGCFFEYHKCDCEME gi|1|pdb|3HTT| LLENERTLDFHDSNVSNLYEKVSQQLRNNAKEINGCFFEYHKCDCEME  

3HTO_B|2|CHAIN| SVKNTYDYP gi|1|pdb|3HTT| SVKNTYDYP  

**Protein 3-D Structure Modeling:**

The SWISS-MODEL workspace was used for homology modeling. The alignment mode was chosen. The T-Coffee multiple sequence alignment result was pasted. Thereafter CLUSTAL W option was chosen for alignment option and then alignment file was submitted. The desired model of the protein was generated.

**Model Validation and Protein Structure Analysis**

The model was validated using Procheck, What Check, Gromas server available at Swiss Model Server. The structural analysis of the protein was carried out using the SWISS-MODEL workspace. The protein structure was evaluated by SWISS -MODEL workspace. The Ramachandran for input atoms only, Ramachandran plots for all residue types input atom only, Main – chain parameters input atom only, side chain parameters input atom only, Residue properties input atom only, Main –chain bond angles input atom only, RMS distances from planarity input atom only, Chi1-Chi2 plots input atom only and distorted geometry input atoms only were analyzed. Apart from this beta-turn input, gamma-turn input, disulphide bridges input, hairpin input, helical nets input, helical interaction input, helical geometry inputs, sheets input, strands input, summary input

**Molecular Docking Study**

Protein Ligand Docking Studies: With the help of protein ligand docking studies we could rule out the structure of the ligand and orientation of a protein when it interacts with small molecule like ligands. Protein-ligand docking module can be divided into various parts like receptor preparation, ligand preparation, binding site analysis, dock and analysis [14-15] (Parida et al.,2012; Das et al.,2013). Binding Site Analysis helps in identification and visualization of possible binding sites and the distribution of surrounding residues in the active sites. The centre of grid map was chosen as grid map values for preparation of the grids. The spacing of grids was set to 1.00 Å and no grid was taken as 60 x 60 x 60 Å and protein-ligand docking was performed using Lammackin genetic algorithm with default parameter [16] (Morris et al., 1998). Molecular Docking was performed using Exome Horizon software.

**RESULTS**

In the Ramachandran plot based on 160 residues 83.2% residues were found in most favoured region, 15.4% in additional allowed regions, 0.7% in generously allowed region and 0.7% residues were in disallowed regions. The number of non-glycine and non-proline residues was found to be 143. The number of end residues was found to be 0. The numbers of glycine and proline residues were found to be 16 and 1 respectively. In the Ramachandran plot 83.2% core, 15.4% allow, 0.7% gen, 0.7% disallow. The 158 residues were stated labeled residues. In the chi1-chi2 plots out of 112 residues 1 was labeled (Figure 2). In the main –chain parameters 5 were better, 0 inside, 0 worse (Figure 4). In the side-chain parameters 5 were better, 0 inside, 0 worse (Figure 5). The maximum deviation of the residues was found to be 4.0, the bond length/angle was found to be 3.6 (Figure 6). The dihedrals was found to be 0.05, covalent was 0.36 and overall 0.11 (Figure 7). The M/c bond length was 100% within limits and 0.0% highlighted/c bond angle was found to be 98.2% within limits and 1.8% highlighted. The planar groups were 99.5% within limits and 10.5% highlighted (Figure 8). The docking results built 10 clusters and the binding energy was given in Table 1. The interactive amino acids of the protein and the atoms of the ligand including the distances were given in the Table 2.

**DISCUSSION**

Homology modeling acts a central role in determining protein structure in the structural genomics project. The importance of homology modeling has been increasing very vast because of the large gap that exists between the overwhelming number of available protein sequences and experimentally solved protein structures, and also, because of the increasing reliability and accuracy of the method. As a fact, a protein sequence with over 30% identity to a known structure can often be predicted with an accuracy equivalent to a low-resolution X-ray structure [18]. Molecular docking studies have become more popular for study of macromolecular structure and interaction. Molecular Modeling and computer aided drug design provides most detailed possible view of receptor-ligand interaction and has built a novel rational approach to drug design [19]. The objective of this research is to conduct in silico analysis of Hemagglutinin, of H1N1 virus, which has certain influence on its high pathogenicity towards human. The general steps are the construction of homology modeling, structural assessment and the molecular docking study.

**Table 1: Binding Energy and clustering information formed by rimantadine**

<table>
<thead>
<tr>
<th>Rank</th>
<th>Energy (kcal/mol)</th>
<th>No. in cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>-2.99</td>
<td>1</td>
</tr>
<tr>
<td>2.</td>
<td>-2.46</td>
<td>1</td>
</tr>
<tr>
<td>3.</td>
<td>-2.00</td>
<td>1</td>
</tr>
<tr>
<td>4.</td>
<td>-1.95</td>
<td>1</td>
</tr>
<tr>
<td>5.</td>
<td>-1.84</td>
<td>1</td>
</tr>
<tr>
<td>6.</td>
<td>-1.72</td>
<td>1</td>
</tr>
<tr>
<td>7.</td>
<td>-1.44</td>
<td>1</td>
</tr>
<tr>
<td>8.</td>
<td>-1.40</td>
<td>1</td>
</tr>
<tr>
<td>9.</td>
<td>-1.38</td>
<td>1</td>
</tr>
<tr>
<td>10.</td>
<td>-1.06</td>
<td>1</td>
</tr>
</tbody>
</table>
**Fig. 1:** Ramachandran plot analysis of Model
Chi1-Chi2 plots
input_atom_only

Fig. 2: Chi1-Chi2 plot analysis of Model
Fig. 3: Ramachandran plot analysis of all residue types of Model
Main-chain parameters
input_atom_only

Fig. 4: Main chain parameters of Model
Fig. 5: Side chain parameters of Model

<table>
<thead>
<tr>
<th>Stereochemical parameter</th>
<th>No. of data pts</th>
<th>Parameter value</th>
<th>Comparison values</th>
<th>No. of band widths from mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Chi-1 gauche minus st dev</td>
<td>21</td>
<td>7.5</td>
<td>22.7</td>
<td>-2.3 BEATER</td>
</tr>
<tr>
<td>b. Chi-1 trans st dev</td>
<td>44</td>
<td>13.2</td>
<td>22.7</td>
<td>-1.8 BEATER</td>
</tr>
<tr>
<td>c. Chi-1 gauche plus st dev</td>
<td>70</td>
<td>14.0</td>
<td>21.3</td>
<td>-1.5 BEATER</td>
</tr>
<tr>
<td>d. Chi-1 pooled st dev</td>
<td>135</td>
<td>12.8</td>
<td>22.0</td>
<td>-1.9 BEATER</td>
</tr>
<tr>
<td>e. Chi-2 trans st dev</td>
<td>32</td>
<td>12.5</td>
<td>23.1</td>
<td>-2.1 BEATER</td>
</tr>
</tbody>
</table>

Plot statistics
Fig. 6: Main chain bond angles of Model

Black bars > 2.0 st. devs. from mean.
Solid and dashed lines represent the mean and standard deviation values as per Engh & Huber small-molecule data.
PROCHECK

RMS distances from planarity
input_atom_only

Histograms showing RMS distances of planar atoms from best-fit plane.
Black bars indicate large deviations from planarity: RMS dist > 0.03 for rings, and > 0.02 otherwise.

▶ signifies data points off the graph in the direction shown.

Fig. 7: RMS distances from planarity of Model
CONCLUSION

Swine flu is a major infectious disease caused by the influenza A (H1N1) virus. The surface protein Hemaglutinin of influenza virus has an important role for spreading this disease. In this study, Hemaglutinin is prepared by Homology Modelling and then after docking study performed with potent anti-viral drug. This prepared protein is further used for docking study of newly discovered anti-viral drug.

Table 2: It shows Protein-Ligand interaction study

<table>
<thead>
<tr>
<th>Protein atom</th>
<th>Ligand atom</th>
<th>Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>GLY 1 O</td>
<td>31 C12 RES 1</td>
<td>3.55</td>
</tr>
<tr>
<td>LEU 2 CD1</td>
<td>31 C12 RES 1</td>
<td>3.35</td>
</tr>
<tr>
<td>ASP 19 OD2</td>
<td>28 N1 RES 1</td>
<td>3.41</td>
</tr>
<tr>
<td>TRY 24 CE 2</td>
<td>4 C4 RES 1</td>
<td>3.82</td>
</tr>
<tr>
<td>ALA 35 CB</td>
<td>28 N1 RES 1</td>
<td>3.42</td>
</tr>
<tr>
<td>ALA 36 O</td>
<td>28 N1 RES 1</td>
<td>3.03</td>
</tr>
<tr>
<td>ASP 37 OD1</td>
<td>10 C10 RES 1</td>
<td>3.38</td>
</tr>
<tr>
<td>GLN 38 CB</td>
<td>8 C8 RES 1</td>
<td>4.07</td>
</tr>
<tr>
<td>GLU 105 OE2</td>
<td>28 N1 RES 1</td>
<td>3.25</td>
</tr>
<tr>
<td>ASP 109 OD1</td>
<td>28 N1 RES 1</td>
<td>3.54</td>
</tr>
<tr>
<td>GLU 150 OE1</td>
<td>31 C12 RES 1</td>
<td>4.74</td>
</tr>
<tr>
<td>LYS 153 O</td>
<td>10 C10 RES 1</td>
<td>3.69</td>
</tr>
<tr>
<td>ASN 154 CA</td>
<td>10 C10 RES 1</td>
<td>4.49</td>
</tr>
</tbody>
</table>

ACKNOWLEDGEMENT

The authors gratefully acknowledge Bioinformatics Infrastructure Facility (BIF) funded by Department of Biotechnology, Govt. of India, at Centre for Studies in Biotechnology, Dibrugarh University.

REFERENCES


