www.ijbtjournal.com
International Journal of Biological Technology (2012) 3(1):62-66.
ISSN: 0976 – 4313

Original Article

In Silico Characterization of Hemagglutinin Protein of H1N1 Subtype

Harsh Vardhan[1], Avinash Marwal[2], Peeyush Mathur[3], Rajneesh Prajapat[2]*

[1,3] Department of Computer Science, Sobhasaria Group of Institutions, Sikar-332001, Rajasthan, India
[2]Department of Science, Mody Institute of Technology & Science (Deemed University), Lakshmangarh, Sikar, Rajasthan, India.
*To whom correspondence may be Email address: prajapat.rajneesh@gmail.com, prajapat_rajneesh@rediffmail.com.

Published: 15, April, 2012; Vol. No. 3(1):62-66; Online:www.ijbtjournal.com/documents/ijbt150412113.

Abstract

In this work, *in silico* characterized the Hemagglutinin proteins of Influenza A virus [AEN68928]. The aim of present study was to carry out the homology modeling study of the mentioned protein using 3D-Jigsawn protein comparative modeling server, Verify-3D structure evaluation server, CHIMERA. The model was validated by using protein structure checking tools RAMPAGE server for reliability. Hex 6.3 was used for protein (Hemagglutinin) - ligand (Tamiflu) docking to determine the potential ligand binding sites. These ligand (Tamiflu) binding sites identified can provide an insight to design potential inhibitors in future.

Key words: Hemagglutinin, 3D-JIGSAWN, Verify-3D, RAMPAGE, Hex 6.3

Introduction

Swine influenza virus A, H1N1 subtype (family *Orthomyxoviridae*) contains segmented genome with 8 RNA fragments that enclosed in a lipid envelope that was identified in 2009 (Abhilash and Nandhini, 2010). The pig infecting swine flu strain, acquired the capability for human to human transmission (Butler, 2009; Cohen and Enserink, 2009). The envelope glycoproteins, hemagglutinin (HA) and neuraminidase (NA) are essential for infection of host cells and also for the release of newly generated virus particles that go on to infect the other cells. According to current reports, Oseltamivir (Tamiflu®) drugs is effective for treatment (Maurer-Stroh *et al*., 2009) of mentioned infection therefore it can be used as ligand for hemagglutinin.

Therefore, *in silico* characterize the Hemagglutinin proteins of Influenza A virus Hemagglutinin proteins [AEN68928] to develop 3D model and implications on ligand identification. Docking tool Hex 6.3 was used for protein (hemagglutinin) - ligand (Tamiflu) docking to determine the binding sites.

Material and Methods

In present study different bioinformatics tools and biological database were used for homology modeling and docking studies, *e.g*., GenBank-NCBI, PDB (Protein Data Bank), UCLA-DOE and Hex, UCSF Chimera etc. The homology modeling procedure can be divided into four sequential steps: template selection, target template alignment, model construction and model assessment (Marti-Renom *et al*., 2000). Hemagglutinin sequence of H1N1 subtype, in FASTA format was mined from GenBank-NCBI database ([Influenza A virus (A/Cambodia/U306/2010(H1N1))] AEN68928]) for homology modeling and docking study with Tamiflu.

Template selection and sequence alignment: BLASTp was used (Altschul *et al*., 2005) to search against the PDB (Protein Databank) to find out the related homologues of the AEN68928 sequence (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The homology modeling requires a query sequence with unknown 3D structure and target sequence that have known 3D structure with at least 35% similarity. By the BLAST search, we selected the closest homologue of AEN68928, was 2WR3 (Influenza H2 Duck Ontario Hemagglutinin with Avian Receptor) with the highest sequence identity of 63%, Positives 79% and gaps 1%.

International Journal of Biological Technology

www.ijbtjournal.com
International Journal of Biological Technology (2012) 3(1):62-66.
ISSN: 0976 – 4313

Target sequence
The PDB file of 2WR3 was downloaded from PDB (http://www.rcsb.org/pdb) and the FASTA format of AEN68928 sequence was mined from GenBank-NCBI.

The FASTA sequences of query (AEN68928) were uploaded on the 3D-Jigsawn (Protein Comparative Modeling Server) for the construction of its PDB files. 3D-Jigsawn (bmm.cancerresearchchuk.org/~3djigsaw) sends the PDB file on the e-mail address that was assigned to the modeling server. The PDB file of query and homologous target sequence were further utilized for 3D model energy validation and docking studies (Heinrichs, 2008).

Model building
Evaluation and validation: UCLA-DOE server http://nihserver.mbi.ucla.edu) provides various softwares for the study of different aspects of browsed PDB files e.g., Verify3D, Procheck etc. The Verify 3D and Procheck (Laskowski *et al*., 1993) outcomes displayed in the form of profile search and Ramachandran plots (Prajapat *et al*., 2011). In this study the model was checked with Verified-3D (Goh *et al*., 2008) server and Ramachandran plot at RAMPAGE (Lovell *et al*., 2003) server.

PDB file of both query and homologous target proteins were utilized for the structural model construction using offline bioinformatics softwares e.g., UCSF Chimera.

Docking of Hemagglutinin and Tamiflu
The binding site for tamiflu on hemagglutinin was identified by using docking program Hex 6.3 (Ritchie *et al*., 2008). The pdb file of tamiflu was retrieved from http://www-jmg.ch.cam.ac.uk/data/molecules/misc/tamiflu.html for docking study. Regularization is a procedure for fitting a protein model with the ideal covalent geometry of residues to the atomic positions of the target PDB structure (Ritchie *et al*., 2008). Based on the energy minimization the best pose of the docked complex was selected.

Results and Discussion

In this study the 3D structure of Hemagglutinin [Influenza A virus (A/Cambodia/U306/2010(H1N1))] protein [AEN68928] was built by homology modeling based on the PDB file obtained from 3D JIGSAW by using UCSF Chimera software. The secondary structure of AEN68928.pdb has 9 α helix and 17 β sheets (Fig. 1).

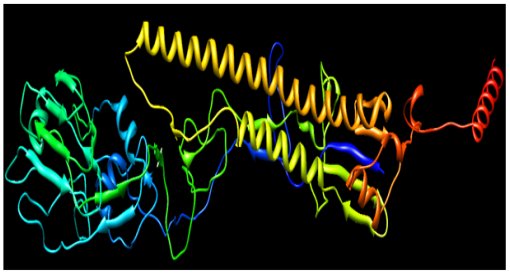


Fig.1: Ribbon diagram of Hemagglutinin [Influenza A virus (A/Cambodia/U306/2010(H1N1))], {AEN68928} designed by using UCSF Chimera.

>gi|345098955|gb|AEN68928.1| hemagglutinin [Influenza A virus (A/Cambodia/U306/2010(H1N1))]
MKAILVVLLYTFATANADTLCIGYHANNST
DTVDTVLEKNVTVTHSVNLLEDKHNGKLC
KLRGVAPLHLGKCNIAGWILGNPECESLST
ASSWSYIVETSSSDNGTCYPGDFIDYEELRE
QLSSVSSFERFEIFPKTSSWPNHDSNKGVTA
ACPHAGAKSFYKNLIWLVKKGNSYPKLSK
SYINDKGKEVLVLWGIHHPSTSADQQSLYQ
NADAYVFVGTSRYSKKFKPEIAIRPKVRDQ
EGRMNYYWTLVEPGDKITFEATGNLVVPR
YAFAMERNAGSGIIISDTPVHDCNTTCQTPK
GAINTSLPFQNIHPITIGKCPKYVKSTKLRLA
TGLRNVPSIQSRGLFGAIAGFIEGGWTGMV
DGWYGYHHQNEQGSGYAADLKSTQNAID
KITNKVNSVIEKMNTQFTAVGKEFNHLEKR
IENLNKKVDDGFLDIWTYNAELLVLLENER
TLDYHDSNVKNLYEKVRSQLKNNAKEIGN
GCFEFYHKCDNTCMESVKNGTYDYPKYSE
EAKLNREEINGVKLESTRIYQILAIYSTVASS
LVLVVSLGAISFWMCSNGSLQCRICI.

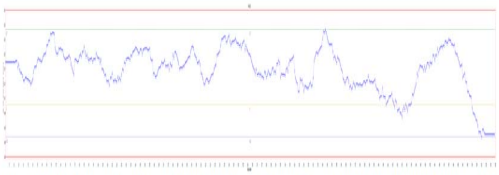


Fig. 2: Verified 3D graph of Hemagglutinin [Influenza A virus A/Cambodia/ 306/2010(H1N1))], protein [AEN68928].

Profile score above zero in the Verify 3D graph (Bowie *et al*., 1991; Luthy *et al*., 1992) corresponds to acceptable environment of the model. The high score of 0.77 indicates that environment profile of the model is good (Fig. 2). The high score for homologous 2WR3 was 0.78. By the BLAST search, we selected the closest homologue of AEN68928, was 2WR3 (Influenza

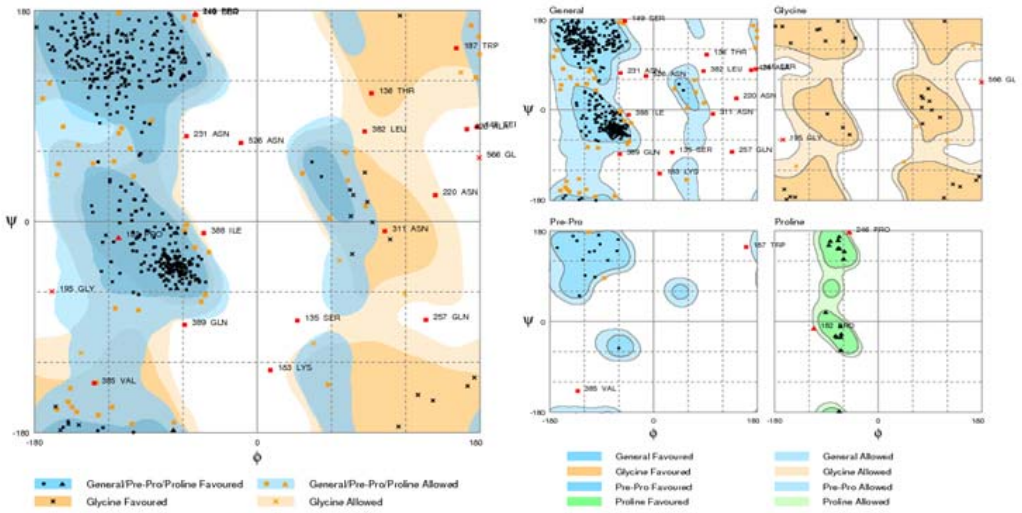
www.ijbtjournal.com
International Journal of Biological Technology (2012) 3(1):62-66.
ISSN: 0976 – 4313

H2 Duck Ontario Hemagglutinin with Avian Receptor) with the highest sequence identity of 63%, Positives 79% and gaps 1%.

The Ramachandran plot contributes to the final values of AEN68928.pdb protein e.g., 86.7% of residues comes in the most favoured regions, 9.4% residues in allowed region and 3.8 % residues in outlier regions (Table 1, Fig. 3a). Non-proline residues, non-glycine residue regions were 98.0% and most disallowed regions were 2.0% in the plot (Fig. 3b).

Table -1: Results summary of the Ramachandran plot

Accession No.	Protein	Virus Strain	Residues in favoured regions %	Residues in allowed regions %	Residues in outlier regions %
AEN68928.pdb	Hemagglutinin	[Influenza A virus A/Cambodia/ U306/2010(H1N1)]	86.7	9.4	3.8
2WR3.pdb	Hemagglutinin	Influenza H2 Duck Ontario	91.1	6.8	2.1



[a] [b]

Fig.3: (a) Ramachandran plot of 3D model of Hemagglutinin [Influenza A virus A/Cambodia/ U306/2010(H1N1))], protein [AEN68928], (b) Non-proline residues and non-glycine residue regions

Table - 2: Predicted binding site amino acids for AEN68928.pdb

Residues	Amino acid	Contact	Av distance	JS Divergence
236	Lys	6	0.14	0.79
239	Asp	10	0.40	0.69
240	Gln	12	0.03	0.89

A good quality Ramachandran plot has over 90% in the most favoured regions (Xiao *et al*., 2004) but the Ramachandran plot of AEN68928.pdb has only 86.7% of residues in the most favoured regions therefore it is near to good quality model (Table -1).Homologous 2WR3 has 91.1% residues in most favoured regions therefore 2WR3 is more stable than AEN68928.pdb. Lys, Asp and Gln were identified as binding site amino acids (Table-2) that may be interact with tamiflu (ligand).

Docking used to identify possible binding modes for a ligand (Morris and Lim-Wilby, 2008). The binding sites exhibit chemical specificity and ligand affinity measure strength of the chemical

www.ijbtjournal.com
International Journal of Biological Technology (2012) 3(1):62-66.
ISSN: 0976 – 4313

bond (Balakrishnan *et al*., 2010). The binding site for AEN68928.pdb protein model was predicted using Hex 6.3.

The Etotal, Eshape and Eforce values for the model were -223.7, -223.7 and 0.0. Best start orientation was alpha 26 (E = -837.42) was at 17007/1312704 (Emin = -223.90, Emax = -187.21). On the basis of the RMS and energy values the best docking orientation was selected. The better RMS value of docking was -1.00.

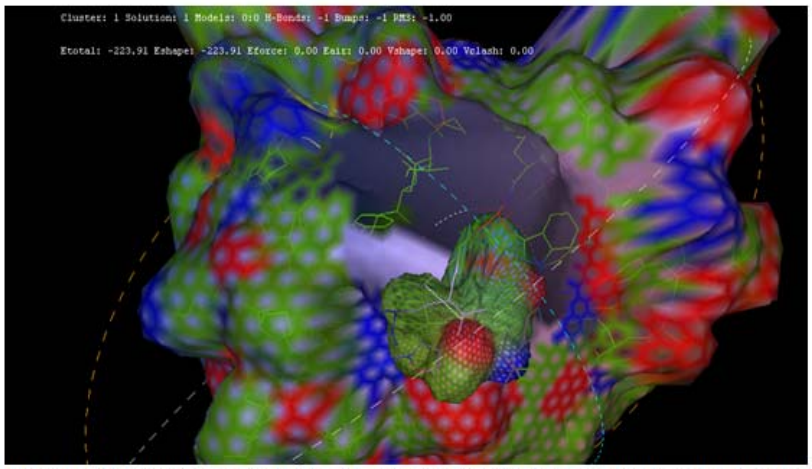


Fig. 4: Illustration of spherical harmonic surfaces to order L = 12 for the AEN68928.pdb protein domain and Tamiflu. Illustration of the AEN68928.pdb protein (Receptor) and Tamiflu (Ligand) complex shown as contoured Gaussian density surfaces and coloured by chain colour.

Table -3: Binding site model Hemagglutinin protein

Pocket/ Contouring surface for:	Polar probe	Apolar probe	Primary surface: Area	Primary surface: Volume	Typical edge arc	Typical edge length (Å)	Average radius (Å)	Surface area (Å)	Triangles		
									Min	Max	Ave
ACZ06249	0.00A	0.00A	9375.81	10326.29	4.62 °	0.31	3.91	211.56	0.21	24.65	2.17
Temiflu	0.00A	0.00A	297.15	298.23	4.62 °	0.42	4.11	351.00	0.27	32.33	1.63

Fig. 4 illustrate spherical harmonic surfaces to order L=12 for the AEN68928.pdb and Tamiflu.pdb. Illustration of the AEN68928.pdb and Tamiflu.pdb complex shown as contoured Gaussian density surfaces and background modes (Fig. 4). These docking results suggest that the tamiflu interact with the AEN68928.pdb of H1N1 and inhibit the infection in host cells. The binding pocket values for AEN68928.pdb protein model were predicted by using Hex 6.3. The predicted two pockets by the software with different primary surface area and volume showed (Table -3).

A few amino acids were found to be conserved in AEN68928.pdb, that forming the binding cavity for the Tamiflu (Table-2). Interaction of AEN68928.pdb with Tamiflu, stop different function that carryout by this protein in infected host cell and this leads to inhibit H1N1 infection.

Conclusion

The summary, provide sequence analysis and structural modeling, docking of the Hemagglutinin of the H1N1 swine flu outbreak. Homology modeling results suggest that AEN68928.pdb is a stable protein and protein [AEN68928] - ligand [Tamiflu] docking results may allows expanding the number of other H1N1 protein analysis. Information obtain by this study will be used in screening of other inhibitors of the H1N1 protein and can be further applied for *in silico* drug design.

References

Abhilash M. and Nandhini, K. 2010. Homology modelling and *In silico* analysis of Hemagglutinin protein From H1N1 influenza A virus. *International Journal of Pharma Sciences and Research*, 1: 40-50.

Altschul, S.F., Wootton, J.C., Gertz, E.M., Agarwala, R., Morgulis, A., Schaffer A.A. and Yu, Y.K. 2005. Protein database searches using compositionally adjusted substitution matrices. *FEBS J*., 272: 5101-5109.

Balakrishnan, M., Srivastava, R.C. and Pokhriyal, M. 2010. Homology modeling and docking studies between HIV-1 protease and carbamic acid. *IJBT*, 9: 96-100.

Bowie, J.U., Luthy R. and Eisenberg, D. 1991. A method to identify protein sequences that fold into a known three-dimensional structure. *Science*, 253: 164-170.

Butler, D. 2009. Swine flu goes global. *Nature*, 458:1082-1083.

Cohen, J. and Enserink, M. 2009. Infectious diseases. As swine flu circles globe, scientists grapple with basic questions. *Science*, 324:572-573.

Goh, K.M., Mahadi, N.M., Hassan, O., Zaliha, R.N., Rahman. R.A. and Illias, R.M. 2008. Molecular modeling of a predominant β-CGTase G1 and analysis of ionic interaction in CGTase. *Biotechnology*, 7: 418-429.

Heinrichs, A. 2008. Proteomics: Solving a 3D jigsaw puzzle. *Nat. Rev. Mol. Cell Biol*., 9: 3-3.

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

Lovell, S.C., Davis, I.W., Arendall W.B. 3rd, P.I. de Bakker and Word, J.M.*et al*. 2003. Structure validation by Calpha geometry: Phi, Psi and Cbeta deviation. *Proteins*, 50: 437-450.

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

Maurer-Stroh, S., Jianmin, M., Raphael, T.C.L., Fernanda, L.S. and Eisenhaber, F. 2009. Mapping the sequence mutations of the 2009 H1N1 influenza A virus neuraminidase relative to drug and antibody binding sites. *Biology Direct,* 4:18. doi:10.1186/1745-6150-4-18

Morris, G.M. and Lim-Wilby, M. 2008. Molecular Docking, Methods in Molecular Biology. In: Molecular Modeling of Proteins, Kukol, A. (Ed.). Humana Press, Totowa, New Jersey, pp: 365-382.

Prajapat, R., Gaur, R.K. and Marwal, A. 2011. Homology Modeling and Docking Studies Between AC1 Rep Protein of Begomovirus and Whey α-lactalbumin. *Asian Journal of Biological Sciences*, 4: 352-361.

Ritchie, D.W., Kozakov D. and Vajda, S. 2008. Accelerating protein-protein docking correlations using a six-dimensional analytic FFT generating function. *Bioinformatics*, 24: 1865-1873.

Xiao, J.F., Li, Z.S., Sun, M., Zhang, Y. and Sun, C.C. 2004. Homology modeling and molecular dynamics study of GSK3/SHAGGY-like kinase. Comput. *Biol. Chem*., 28: 179-188.

Manuscript Progress Date

Received : 11 .02.2012
Revised : 27.03.2012
Accepted : 30.03.2012