Available online at www.scholarsresearchlibrary.com

Scholars Research Library

Archives of Applied Science Research, 2010, 2 (6): 93-97

(http://scholarsresearchlibrary.com/archive.html)

ISSN 0975-508X
CODEN (USA) AASRC9

Homology modeling of neuraminidase protein of the novel H1N1 subtype of swine influenza virus of A/California/04/2009/H1N1

[1]Vijay Tripathi; [2]Shambhavi Sankrit, [1]Dwijendra Kumar Gupta

[1]Center of Bioinformatics, University of Allahabad, Allahabad, U.P, India
[2]Department of Biotechnology, Dr. D. Y. Patil University, Pune, Maharastra, India

---

ABSTRACT

Influenza A virus H1N1 is the causative agent of the recent 2009 Swine flu or Swine influenza. The virus shows high rate of antibiotic resistance due to rapid mutation which leads to the production of new viral strains and hence, new vaccines are made to treat the new H1N1 variants. Neuraminidase plays an important biological role in this viral infection. It has been found that it has a major function at the final stage of infection when it cleaves sialic acid from cell surface and progeny virions facilitating virus release from the infected cells.The recent sequence of influenza A virus, A/California/04/2009(H1N1) for the protein, neuraminidase was selected for the study. Homology modelling was used to model the three-dimensional structure of the neuraminidase protein using Modeller 9v6, Swiss Model and GENO3D. Energy minimization was performed by GROMACS and functional motifs were identified by the ProFunc server. CASTp server was used to determine the potential ligand binding sites. These ligand binding sites identified can thus provide an insight to design potential inhibitors in future.

Keywords: Influenza A virus H1N1, Neuraminidase, Homology modeling, ligand binding sites.

INTRODUCTION

Swine influenza virus A, subtype H1N1 belongs to the family of *Orthomyxoviridae.* It contains RNA as genetic material enclosed in a lipid envelope derived from the host cell membrane during budding. The genome is segmented, with 8 RNA fragments .These can either be spherical or filamentous in form and are obligate parasites. The viral envelope is studded with two glycoproteins namely, hemagglutinin (HA) and neuraminidase (NA) which are essential for infection of host cells and also for the release of newly generated virus particles that go on to infect the neighbouring cells. There are 16 known HA proteins and 9 known NA proteins.

The neuraminidase (NA) protein is expressed on the viral surface and its structure consists of a globular head, a thin stalk region and a small hydrophobic region that anchors the protein in the virus membrane. It cleaves the terminal sialic acid residues which act as receptors for the hemagglutinin (HA) protein. It facilitates the penetration of the virus into the mucus and release of newly formed virions [1]. In the 2009 flu pandemic, the viruses that were isolated from the United States were found to be made up of genetic elements from four different flu viruses – North American Mexican influenza, North American avian influenza, human influenza, and swine influenza virus typically found in Asia and Europe and are referred to as an unusually mongrelised mix of genetic sequences[2]. The reassortment between human influenza and swine influenza viruses has resulted in the appearance of this new strain, in all four different strains of subtype H1N1.

Many variations of H1N1 occur because the influenza virus changes constantly and subtly. Every year, the influenza vaccine for people contains a modified human strain of H1N1. Although the vaccine protects against that strain, but due to the rapid mutation which the virus undergoes, the antibodies from the vaccine do not work against a newly mutated strain. Hence, a new influenza vaccine is made every year to keep up with the new H1N1 variations.

MATERIALS AND METHODS

The FASTA sequence of the recent influenza A virus, A/California/04/2009 (H1N1) for neuraminidase was taken from NCBI (National Center for Biotechnology Information).

pdb-BLAST was performed to determine the template for the above query. The sequence of the query and template were then submitted to ClustalW [4] for the determination of conserved regions. Modeller 9v6 [5], Swiss Model [6] and GENO3D [7] were used for homology modeling of the protein structure being used as template. SAVS server was used to predict the validity and to analyze this structure using Verfy3D [3] and PROCHECK [8] tools. The functional motifs were identified by the ProFunc [9] server by analyzing the sequence and structure of the neuraminidase protein. The structural motifs were identified from the results of Nest analysis and Cleft analysis. CASTp [10] server analytically measured the area and volume of the pockets and cavities present in the model. Energy minimization was performed using GROMACS [11].

RESULTS AND DISCUSSION

The sequence A/California/04/2009(H1N1) of the recent influenza A virus H1N1 for the protein neuraminidase was taken in the FASTA format from NCBI. BLAST was performed and 2hty|A was selected as the template for the protein. Multiple sequence alignment between the query and template sequences was performed using Crustal W (Fig. 1).

Homology modeling was used for model building of the template structure using Modeller 9v6, Swiss Model and GENO3D tools. The structural validity of the predicted model was analyzed using SAVS server.Verify3D gave 100.00% of the residues an averaged 3D-1D score > 0.2, for the model generated by Modeller. PROCHECK predicted the validity of the models generated by the three modeling tools using Ramachandran plot (Table 1).

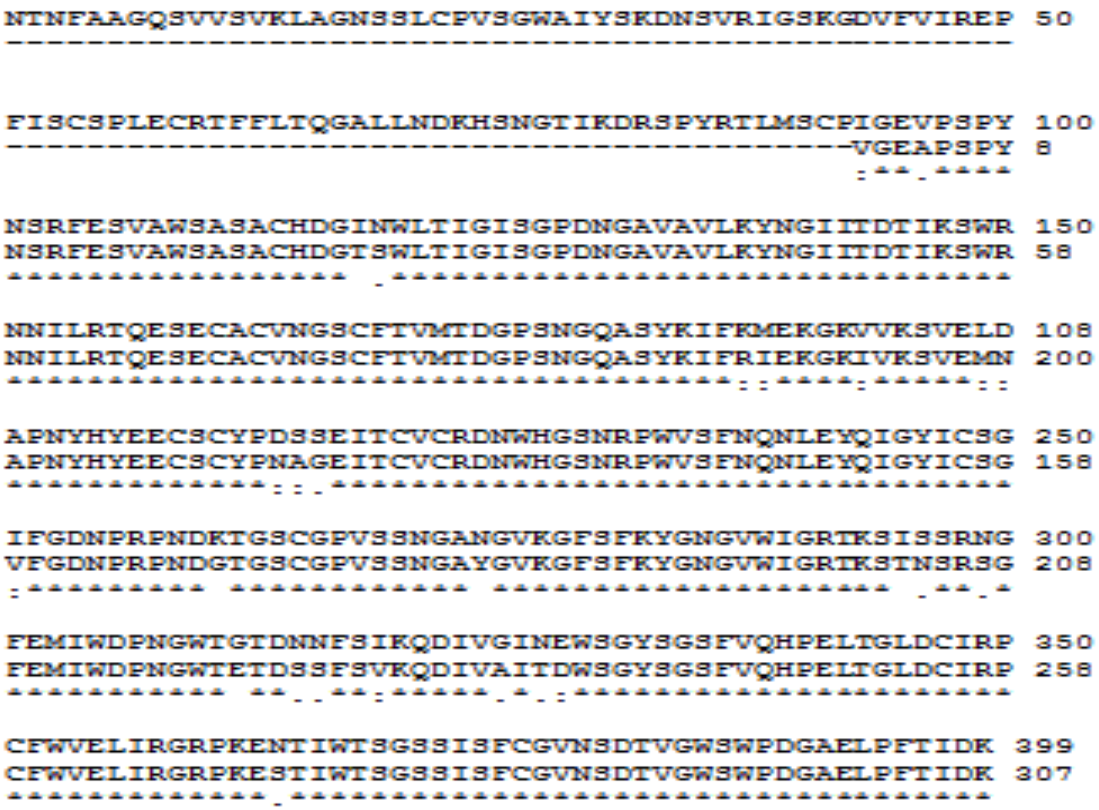

Fig. 1: Alignment between query and template

Table 1: Comparative analysis of the model generated by the Homology Modeling software tools

	MODELLER	SWISS MODEL	GENO3D
Residues in the most favoured region	83.5%	77.0%	68.9%
Residues in the additionally allowed region	15.2%	21.4%	27.6%
Residues in the generously allowed region	0.9%	1.2%	2.9%
Residues in the disallowed region	0.5%3%	0.3%	0.6%

The energy for the models were also generated by the three modelling tools (Table 2).

Table 2:Analysis of the energy of the models predicted by the modelling tools

	MODELLER	SWISS MODEL	GENO3D
Energy(KJ/mol)	-16530.457	-19266.50	-16659.40

Table 3: Energy with its respective values

Energy	Average	RMSD	Fluctuation	Drift	Total Drift
Potential	-315787	1245.54	008.905	-376.32	-3763.95
Kinetic	59533.8	295.47	295.469	0.180898	1.80934
Total Energy	-256254	1246.75	612.298	-376.139	-3762.14
Temperature Heat Capacity, Cv	303.335 12.4722 J/mol K (factor=2.4632e-05)	1.50547	1.50547	0.000921036	0.0092122
Pressure	134.295	188.949	177.919	-22.0318	-220.362

The model generated by Modeller 9v6 was predicted to have maximum validity and was therefore selected for further study. GROMACS performed energy minimization of the model and the result has been shown in

Table 3. The functional motifs were determined by the ProFunc server. Nest analysis predicted the structural motifs present in the predicted model (Fig. 2).

Nest	Score	Residue range	View in RasMol	Residue	Ramachandran region	Solvent accessibility	Cleft	Depth in cleft	Residue conservation
1.	4.67	Gln54-Ala56		Gln54	RIGHT	0.22%	1	-	1.00
				Gly55	LEFT	0.00%	6	8.96	1.00
				Ala56	-	0.00%	-	-	1.00
2.	3.33	Asp377-Ala379		Asp377	RIGHT	0.24%	-	-	1.00
				Gly378	LEFT	0.23%	5	7.28	1.00
				Ala379	-	0.00%	-	-	1.00
3.	2.50	Asn64-Ile67		Asn64	RIGHT	0.00%	6	7.68	1.00
				Gly65	LEFT	0.87%	6	7.92	1.00
				Thr66	RIGHT	0.00%	-	-	1.00
				Ile67	-	0.00%	-	-	1.00
4.	2.33	Asn296-Trp298		Asn296	LEFT	0.80%	-	-	1.00
				Gly297	RIGHT	0.00%	-	-	1.00
				Trp298	RIGHT	0.00%	-	-	1.00
5.	2.33	Thr331-Leu333		Thr331	RIGHT	0.00%	-	-	1.00
				Gly332	LEFT	0.16%	-	-	1.00
				Leu333	-	0.00%	-	-	1.00
6.	1.95	Asp21-Val24		Asp21	RIGHT	0.55%	-	-	1.00
				Asn22	LEFT	0.00%	-	-	1.00
				Ser23	RIGHT	0.00%	-	-	1.00
				Val24	RIGHT	0.00%	-	-	0.78
7.	1.00	Arg138-Asn140		Arg138	RIGHT	0.00%	-	-	1.00
				Asn139	LEFT	0.00%	-	-	1.00
				Asn140	-	0.00%	-	-	1.00
8.	1.00	Gln145-Ser147		Gln145	RIGHT	0.00%	-	-	1.00
				Glu146	LEFT	0.00%	-	-	1.00
				Ser147	-	0.00%	-	-	1.00

Fig. 2: Analysis of the structural motifs

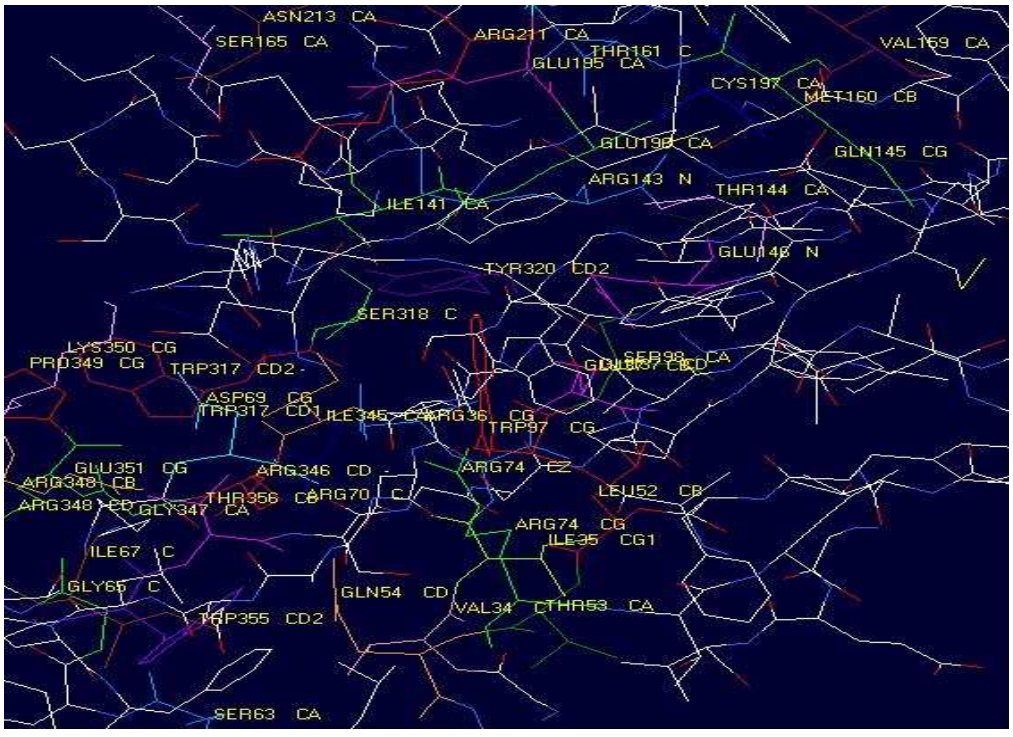

Figure 3: The potential ligand binding site with its residues

From Cleft analysis the maximum volume was found to be 1979.02 with R1 ratio 2.13. The predicted model was then submitted to CASTp server to determine the potential ligand binding pockets. A total of 44 ligand binding sites were determined of which the one with maximum area (1027.2) and volume (1325.3) was selected. Swiss Pdb Viewer 4.0 was used to visualize the structure of this pocket (Fig. 3).

The result also provided information about the residues alongwith their positions in the pocket (Table 4).

Table 4: Residues with their respective positions

Residue:	V	I	R	E	L	T	Q	H	S	G	I	D	R	R	W	S	I	R	T	Q	E	V	M	T	S
Position:	34	35	36	37	52	53	54	62	63	65	67	69	70	74	97	98	141	143	144	145	146	159	160	161	165

Residue:	Y	E	E	C	R	N	R	W	S	Y	I	R	G	R	P	K	E	W	T
Position:	194	195	196	197	211	213	286	317	318	320	345	346	347	348	349	350	351	355	356

CONCLUSION

This work was undertaken to perform the homology modeling of the neuraminidase protein of the recent influenza A virus H1N1. Modeling tools were used to predict the structural models and the potential ligand binding sites were determined that can serve as key to design the inhibitors.

Acknowledgement
The work has been supported by BBT-BIF Grant to DKG under its BTISNet scheme.

REFERENCES

[1] A. H. Reid, G. T. Fanning et al., *Proc. Nat.l Acad Sci.*, **2000**, 97(12), 6785–6790.
[2] W. R. Gallaher., *Journal of Virology*, **2009**.6:51.
[3] Bowie et al., *Science*. **1991**, 253(5016):164-70.
[4] J. D. Thompson, D. G. Higgins, & T. J. Gibson., *Nucleic Acids Research*, **1999,** 22:4673-4680.
[5] N. Eswar et al., *John Wiley & Sons, Inc., Supplement*, **2006**, 15, 5.6.1-5.6.30, 200.
[6] K. Arnold et al., *Bioinformatics*, **2006**, 22,195-201.
[7] Combet et al., *Bioinformatics*, **2002**, 18:213
[8] R. A. Laskowski et al., *J. Appl. Cryst.*, **1993,** 26, 283-291.
[9] R. A. Laskowski et al., *Nucleic Acids Res.*, **2005**, 33, 89-93.
[10] J. Dundas et al., *Nucleic Acid Research*, **2006**, 34:116-118.
[11] H. J. C. Berendsen et al., *Comp. Phys. Comm*. **1995**, 91:43–56.