ISSN 2348-6201

In Silico Characterization of Hemagglutinin Protein of Influenza a Virus [A/Canine/Beijing/Cau9/2009(H1N1] of H1N1 Subtype

Rajneesh Prajapat[1]*, Avinash Marwal[1], Mukesh Goyal[2]
[1]Department of Science, Mody Institute of Technology & Science (Deemed University), Lakshmangarh, Sikar, Rajasthan, India.
prajapat.rajneesh@gmail.com
[1]Department of Science, Mody Institute of Technology & Science (Deemed University), Lakshmangarh, Sikar, Rajasthan, India.
marwal_avinash@yahoo.co.in
[2]Sriganganagar College of Allied Health Sciences, Tantia University, Sriganganagar, Rajasthan, India.
goyalmukesh17@gmail.com
Rajneesh Prajapat and Avinash Marwal have contributed equally to the present research

ABSTRACT

Using 3D-Jigsawn protein comparative modeling server, Verify-3D structure evaluation server, CHIMERA the homology modeling study of the mentioned protein was carried out and it holds the aim of the study. In this work, we *in silico* characterized the hemagglutinin protein of influenza A virus [A/canine/Beijing/cau9/2009(H1N1] (AEM89472). The model was validated by using protein structure checking tools RAMPAGE server for reliability. In NJ tree AEM89472 is place in a separate monophyletic and show close evolutionary relationship with ADD97678. 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.

Indexing terms/Keywords

H1N1; Hemagglutinin; 3D-JIGSAWN; Verify-3D; Hex 6.3

Academic Discipline And Sub-Disciplines

Bioinformatics .

Subject Classification

Animal virology.

Type (Method/approach)

Research Article .

ISSN 2348-6201

INTRODUCTION

In 2009 (Abhilash and Nandhini, 2010) 8 RNA fragments that enclosed in a lipid envelope was identified in Swine influenza virus A, H1N1 subtype (family *Orthomyxoviridae*) which contains segmented genome. 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. The pig infecting swine flu strain, acquired the capability for human to human transmission (Butler, 2009; Cohen and Enserink, 2009).

Therefore, we *in silico* characterize the hemagglutinin protein of influenza A virus [A/canine/Beijing/cau9/2009(H1N1] (AEM89472) 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.

MATERIALS AND METHODS

The homology modeling (Cueno *et al*., 2013) procedure can be divided into four sequential steps: template selection, target template alignment, model construction and model assessment (Marti-Renom *et al*., 2000). 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. Hemagglutinin sequence of H1N1 subtype, in FASTA format was mined from GenBank-NCBI database AEM89472 for homology modeling and docking study with Tamiflu.

Template Selection and Sequence Alignment: The homology modeling requires a query sequence with unknown 3D structure and target sequence that have known 3D structure with at least 35% similarity. 2WR0 (Influenza H2 Hemagglutinins chain A) with the highest identity of 63%, Positives 77% and gaps 0%. For this BLASTp was used (Altschul *et al*., 2005) to search against the PDB (Protein Databank) to find out the related homologues of the AEM89472 sequence (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Target Sequence: The PDB file of both proteins 2WR0 was downloaded from PDB (http://www.rcsb.org/pdb) and the FASTA format of AEM89472 sequence were mined from GenBank-NCBI.

The FASTA sequences of query proteins (AEM89472) 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 software 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 protein (Karthikeyan *et al*., 2013) of [Influenza A virusz [A/canine/Beijing/cau9/2009(H1N1] (AEM89472) were built by homology modeling based on the PDB file obtained from 3D JIGSAW by using UCSF Chimera software. The secondary structure of AEM89472.pdb 4 α helix and 15 β sheets (Fig. 2).

```
>gi|344051346|gb|AEM89472.1|  hemagglutinin, partial [Influenza A virus
(A/canine/Beijing/cau9/2009(H1N1))]

ILVVLLYTFATANAVTLCIGYHANNSTDTVDTVLEKNVTVTLSVNLLEDKHNGKLCKLRGVAPLHLGKCNIAGWILGNPEC
ESLSTASSSSYIVETSSSDNGTCYPGDFIDYEELREQLSSVSSFERFEIFPKTSSWPNHDSNKGVTAACPHAGAKSFYKNL
IWLVKKGNSYPKLSKSYINDKGKDVLVLWGIHHPSTSADQPSLYQNADAYVFVGTSRYSKKFKPEIAIRPKVRDHEGRMNY
YWTLVEPGDKITLEATGNLVVPRYAFAMERNAGSGIIISDTPVHDCNTTCQTPKGAINTSLPFQNIHPITIGTCPKYVKST
KLRLATGLRNVPSIQSRGLFGAIAGFIEGGWTGMVDGWYGYHHQNEQGSGYAADLKSTQNAIDKITNKVNSVIEKMNTQFT
AVGKEFNHLEKRIENLNKKVDDGFLDIWTYNAELLVLSENERTLDYHDSNVKNLYEKVRSQLKNNAKEIGNGCFEFYHKCD
NTCMESVKNGTYDYPKYSEEAKLNREEIDGVKLESTKIYHILAIYSTVASSLVLVVSLGAISFWMCSNGSLQCRICI
```

ISSN 2348-6201

Phylogenetic analysis was done by using MEGA 4.0 showing the relationship with other closely related viruses. In Neighbor-Joining tree, hemagglutinin protein of influenza A virus (AEM89472) place in a separate monophyletic and show close evolutionary relationship with ADD97678 [hemagglutinin [Influenza A virus (A/Wisconsin/629-D01014/2009(H1N1)] (Fig. 1), therefore these results indicating that it is a new H1N1 isolate.

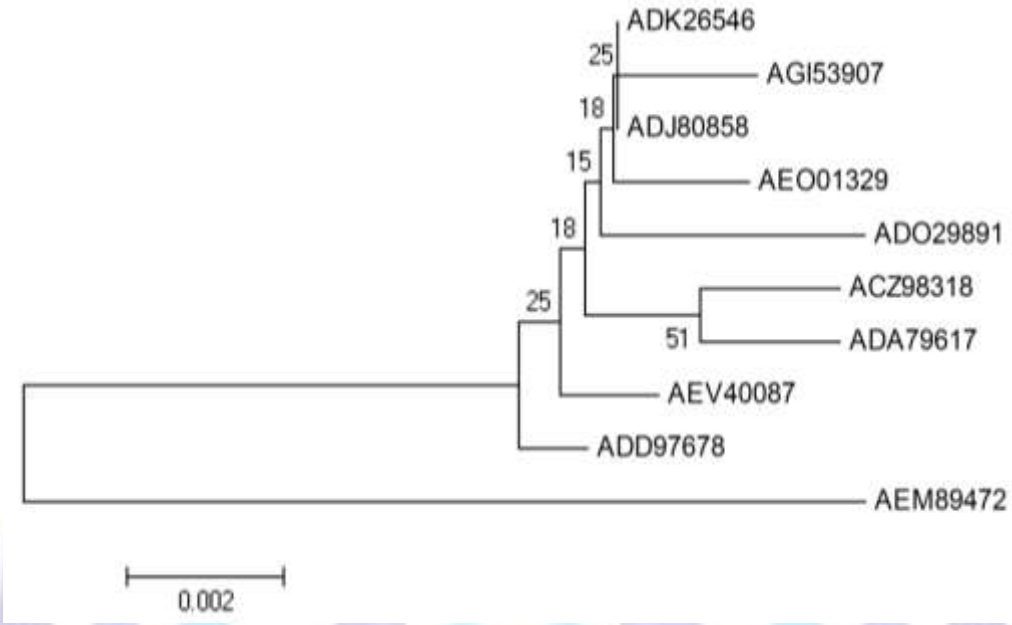


Fig. 1: Neighbor-Joining tree based on the amino acid sequence of hemagglutinin protein of Influenza A virus (AEM89472) and other sequences available in NCBI. The different proteins used in the construction of phylogenetic tree were: Influenza A virus hemagglutinin [A/Shenzhen/1430/2009(H1N1)] (AEV40087), Influenza A virus [A/Thailand/CU-MV8/2010(H1N1)] (ADK26546), Influenza A virus [A/Qingdao/1530/2009(H1N1)] (ADO29891), Influenza A virus [A/Wisconsin/629-D01014/2009(H1N1)] (ADD97678), Influenza A virus [A/Texas/44301765/2009(H1N1)] (ACZ98318), Influenza A virus [A/Russia/200/2009(H1N1)] (ADA79617), Influenza A virus [A/England/731/2009(H1N1)] (AEO01329), Influenza A virus [A/California/VRDL123/2009(H1N1)] (ADJ80858), Influenza A virus [A/Florida/03/2010(H1N1)] (AGI53907).

Fig.2: The secondary structure of AEM89472.pdb hemagglutinin, partial [Influenza A virus (A/canine/Beijing/cau9/2009(H1N1))], protein has 4 α helix and 15 β sheets.

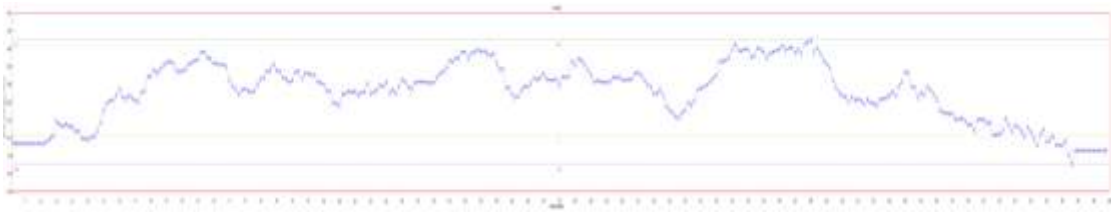


Fig. 3: Verified 3D graph of hemagglutinin, partial [Influenza A virus (A/canine/Beijing/ cau9/2009(H1N1))], protein [AEM89472].

The high score of 0.77 indicates that environment profile of the model is good (Fig. 3). The high score for homologous 2WR0 was 0.78.Profile score above zero in the Verify 3D graph (Bowie *et al.*, 1991; Luthy *et al.*, 1992) corresponds to acceptable environment of the model. By the BLAST search, we selected the closest homologue of AEM89472, was 2WR0 (Influenza H2 Hemagglutinins chain A) with the highest identity of 63%, Positives 77% and gaps 0%.

AEM89472.pdb has 85.6% of residues come in the most favoured regions, 11% residues in allowed region and 3.4 % residues in outlier regions (Table 1, Fig. 4a). The Ramachandran plot contributes to the final values of AEM89472.pdb protein. Non-proline residues, non-glycine residue regions were 97.0% and most disallowed regions were 3.0% in the plot (Fig. 4b).

The Ramachandran plot of AEM89472.pdb has only 85.6% of residues in the most favoured regions therefore it is near to good quality model (Table 1). A good quality Ramachandran plot has over 90% in the most favoured regions (Xiao *et al.*, 2004) but Homologous 2WR0 has 90.2% residues in most favoured regions therefore 2WR0 is more stable than AEM89472.pdb. Lys, Asp and Gln were identified as binding site amino acids (Table 2) that may be interact with tamiflu (ligand). The binding sites exhibit chemical specificity and ligand affinity measure strength of the chemical bond (Balakrishnan *et al.*, 2010). The binding site for AEM89472.pdb protein model was predicted using Hex 6.3. Docking used to identify possible binding modes for a ligand (Morris and Lim-Wilby, 2008).

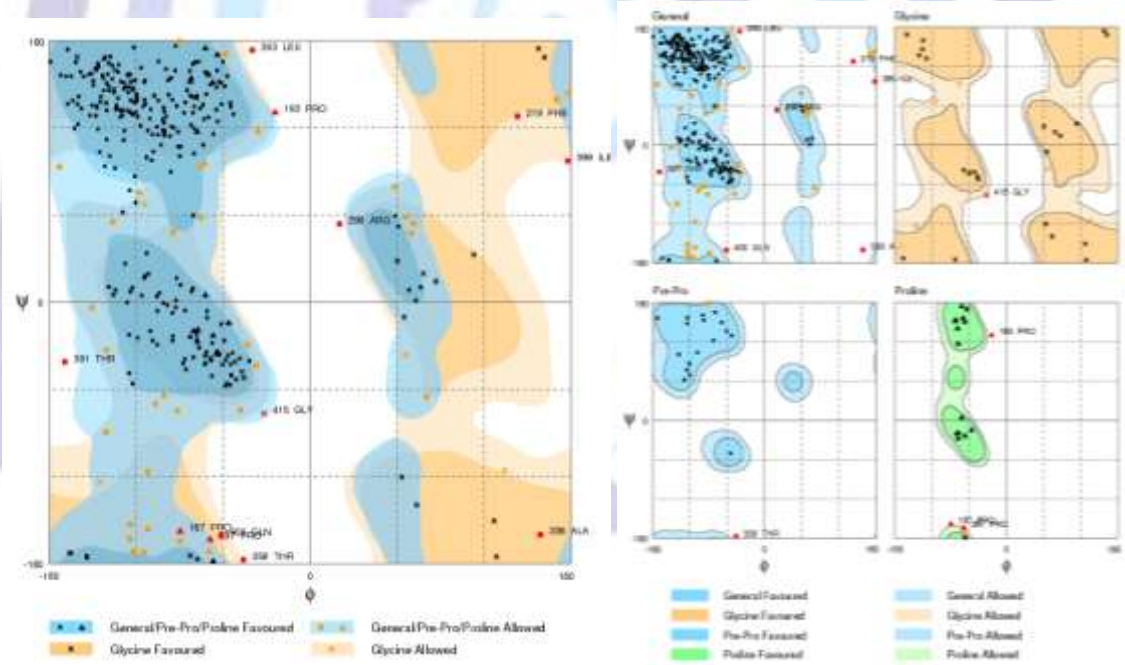


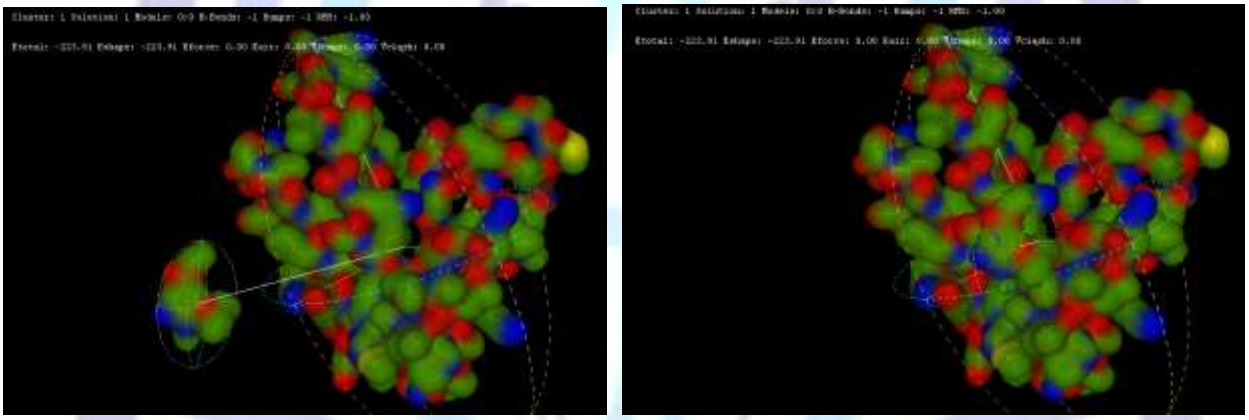
Fig. 4: (a) Ramachandran plot of 3D model of hemagglutinin, partial [Influenza A virusz(A/canine/Beijing/cau9/2009(H1N1))] , protein [AEM89472], (b) Non-proline residues and non-glycine residue regions.

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 %
AEM89472.pdb	Hemagglutinin	Influenza A virus (A/canine/Beijing/cau9/2009(H1N1)	85.6%	11%	3.4 %
2WR0.pdb	Hemagglutinin	Influenza H2 Duck Ontario	90.2	7.4	2.4

Table 2: Predicted binding site amino acids for AEM89472.pdb

Residues	Amino acid	Contact	Av distance	JS Divergence
236	Lys	7	0.16	0.79
239	Asp	11	0.40	0.68
240	Gln	12	0.05	0.89



[a] **[b]**

Fig. 5: A Hex scene showing the AEM89472.pdb protein (right) domain and Tamiflu.pdb (left) in van der waals mode and with the intermolecular axis drawn in blue [a] and solid surfaces [b]

The binding site for AEM89472.pdb protein model was predicted using Hex 6.3. The Etotal, Eshape and Eforce values for the model were -223.9, -223.9 and 0.0 (Table 2). Best start orientation was alpha 26 (E = -837.44) was at 17007/1312704 (Emin = -223.91, Emax = -187.22). On the basis of the RMS and energy values the best docking orientation was selected. The better RMS value of docking was -1.00. The binding sites exhibit chemical specificity, a measure of the types of ligand that bond and the affinity that measure strength of the chemical bond (Balakrishnan *et al.*, 2010).

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
AEM89472	0.00 A	0.00A	1075.88	12526.29	4.62 °	0.33	4.92	251.55	0.03	42.47	2.42
Temiflu	0.00 A	0.00A	398.18	401.23	4.62 °	0.45	5.02	392.00	0.13	17.05	1.67

Fig. 5 illustrate of the AEM89472.pdb and Tamiflu.pdb complex shown van der waals mode. These docking results suggest that the tamiflu interact with the AEM89472.pdb of H1N1 and inhibit the infection in host cells. The binding pocket values for AEM89472.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).

Interaction of AEM89472.pdb with Tamiflu, stop different function that carryout by this protein in infected host cell and this leads to inhibit H1N1 infection (Liu et al., 2010., Tse *et al*., 2013). A few amino acids were found to be conserved in AEM89472.pdb, that forming the binding cavity for the Tamiflu (Table 2).

CONCLUSION

In NJ tree AEM89472 is place in a separate monophyletic and show close evolutionary relationship with ADD97678. Homology modeling results suggest that AEM89472.pdb is a stable protein and protein [AEM89472] - ligand [Tamiflu] docking results may allow expanding the number of other H1N1 protein analysis. The summary, provide sequence analysis and structural modeling, docking of the Hemagglutinin of the H1N1 swine flu outbreak. 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.

ACKNOWLEDGMENTS

Our thanks to the experts who have contributed towards development of the template.

REFERENCES

1. Briddon, R. W., and Stanley, J. 2006. Sub-viral agents associated with plant singlestranded DNA viruses. Virology, 344,198-210.
2. 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.
3. 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.
4. 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.
5. 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.
6. Butler, D. 2009. Swine flu goes global. Nature, 458,1082-1083.
7. Cohen, J., and Enserink, M. 2009. Infectious diseases. As swine flu circles globe, scientists grapple with basic questions. Science, 324, 572-573.
8. Cueno, ME., Imai, K., Shimizu, K., and Ochiai, K. 2013. Homology modeling study toward identifying structural properties in the HA2 B-loop that would influence the HA1 receptor-binding site. J Mol Graph Model. 44, 161-7. doi: 10.1016/j.jmgm.2013.05.011.
9. 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.
10. Heinrichs, A. 2008. Proteomics: Solving a 3D jigsaw puzzle. Nat. Rev. Mol. Cell Biol., 9, 3-3.

11. Karthikeyan, M., Kirubakaran, P., Singh, K.D., Sampath, B., and Krishnasamy, G. 2013. Understanding the evolutionary relationship of hemagglutinin protein from influenza viruses using phylogenetic and molecular modeling studies. J Biomol Struct Dyn. DOI:10.1080/07391102.2013.793211
12. 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.
13. Liu, H., Yao, X., Wang, C., and Han, J. 2010. *In silico* identification of the potential drug resistance sites over 2009 influenza A (H1N1) virus neuraminidase. Mol Pharm., 7, 894-904. doi: 10.1021/mp100041b.
14. Lovell, S.C., Davis, I.W., Arendall, W.B. 3rd, De Bakker, P.I. and Word, J.M. et al. 2003. Structure validation by Calpha geometry: Phi, Psi and Cbeta deviation. Proteins, 50, 437-450.
15. Luthy, R., Bowie, J.U., and Eisenberg, D. 1992. Assessment of protein models with three-dimensional profiles. Nature, 356, 83-85.
16. Marti-Renom, M.A., A.C. Stuart, A. Fiser, R. Sanchez, F. Melo and A. Sali, 2000. Comparative protein structure modeling of genes and genomes. Ann. Rev. Biophys. Biomol. Struct., 29, 291-325.
17. 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
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.
19. 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.
20. 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.
21. Tse, H., Kao, R.Y., Wu, W.L., Lim, W.W., Chen, H., Yeung, M.Y., Woo, P.C., Sze, K.H., and Yuen, K.Y. 2011. Structural basis and sequence co-evolution analysis of the hemagglutinin protein of pandemic influenza A/H1N1 (2009) virus. Exp Biol Med (Maywood). 236, 915-25. doi: 10.1258/ebm.2011.010264.
22. 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.

Author' biography with Photo

Dr. Rajneesh Prajapat

Dr Rajneesh Prajapat has completed his Ph.D under the supervision of Dr R K Gaur entitled "Molecular and *In Silico* Characterization of Begomovirus Components Infecting Weeds of North India. He graduated from M.D.S University, Ajmer, Rajasthan (India) in 2003 and did M.Sc in Biotechnology from Universiry of Rajasthan, Jaipur, Rajasthan (India) in 2005 and obtains M.Phil in Environmental Biotechnology from Department of Environmental studies, M.D.S University, Ajmer, Rajasthan (India) in 2006. He has 30 publications and published 3 books on Bioinformatics and Molecular Virology. Dr. Rajneesh Prajapat working as HOD-Biotechnology, BBV, BITS campus, Pilani, Rajasthan (India).

Avinash Marwal

Avinash Marwal has submitted the thesis and completed his PhD from a DBT funded project (DBT project No. BT/PR13129/GBD/27/197/2009), under the supervision of Dr R. K. Gaur, Principal Investigator. He has 35 international publications in reputed journals and has command over various molecular, biotechnological and bioinformatics tools and techniques. His core interests are Homology Modeling, Molecular Docking, Biological Database development, *In Silico* drug designing, Plant Molecular Biology and Development of Transgenic Plants.

ISSN 2348-6201

Dr Mukesh Goyal

Dr Mukesh Goyal has completed his Ph.D in physiotherapy 2012, PG in physiotherapy from H.N.B. Garhwal University, Garhwal, Utrakhand in 2007 and UG in physiotherapy from Utkal University, Bhubanaswar, Orissa in 2005. In B.P.T: Efficacy of Spinal Flexion and Extension Exercises on the Basis of Movement Bias in Patients with Chronic Low Back Pain. Under the guidance of Dr P.P. Mohanty Asst. Professor & H.O.D. of Physiotherapy Department, Swami Vivekanand National Institute of Rehabilitation Training and Rech. (SVNIRTAR), (An Autonomous Institute under the Ministry of Social Justice & Empowerment, Govt. Of India) Cuttck, Orissa. In M.P.T: “A study to compare the effect of sacroiliac joint manipulation and pelvic tilt exercise in Low Back Pain patients with sacroiliac joint dysfunction” under the guidance of Dr. Rambabu, MPT, Principal of D.A.V. Institute of physiotherapy& rehab ,Yamunanagar. In Ph.D: “Significance of physiotherapy treatment in case of total knee replacement for cronic osteoarthritis” under the guidance of Dr. Ravi Sankar Singh.