

**HOMOLOGY MODELING AND STRUCTURAL STUDIES OF CWP- FMN-REDUCTASE
FROM NEUROSPORA CRASSA**S. K. Gousia¹, B. Janakiram², G. Srinivas¹ and J. Naveena Lavanya Latha^{1*}¹Department of Biotechnology, Krishna University, Machilipatnam-521001.²Department of Microbiology, NRI Medical Hospital, Guntur.

*Corresponding Author: Dr. J. Naveena Lavanya Latha

Department of Biotechnology, Krishna University, Machilipatnam-521001.

Article Received on 01/08/2017

Article Revised on 22/08/2017

Article Accepted on 13/09/2017

ABSTRACT

Background: Cell wall encompasses the cells of all fungi. Walls are composites of an assortment of parts. Commonly, the wall contains fibrillar materials bound together by sugars, proteins, lipids and an assortment of polysaccharides. While the fibrillar material of the cell wall is to great extent inactive, the structure of the cell wall changes with time. The practical segments are vital in supplement transport, digestion system of non-porous substrates, correspondence, and cell wall alterations. **Methods:** The stereo chemical quality of the protein model was checked by using insilico analysis with PROCHECK and QMEAN servers. The metal binding sites were determined by CHED. **Results:** In this paper the protein with 4 metal binding sites shows highest metal binding probability for the metal namely calcium in sites 1, 2 & 3 with the metal probability of 0.398702, **0.627722 and 0.627722**; magnesium in site 4 with the probability of 0.571190 were showed. The 97.8% residue in the core region of Ramachandran plot showing high accuracy of protein model and the QMEAN Z-score of -2.69 indicates the overall model quality of protein. **Conclusions:** The result of the study may be a guiding point for further investigations on **Flavo protein** and metal binding sites.

KEYWORDS: Flavo Protein, Cell Binding Protein, *Neurospora Crassa*, Homology Modeling, Metal Binding Sites.

INTRODUCTION

The unbending cell mass of organisms is a stratified structure comprising of chitinous microfibrils installed in a framework of little polysaccharides, proteins, lipids, inorganic salts, and shades that gives skeletal support and shape to the encased protoplast. Chitin is a (β 1-4) - connected polymer of N-acetyl-D-glucosamine (GlcNAc). It is created in the cytosol by the exchange of GlcNAc from uridine diphosphate GlcNAc into chains of chitin by chitin synthetase, which is situated in the cytosol in organelles called chitosomes. The chitin microfibrils are transported to the plasmalemma and along these lines incorporated into the new cell wall.^[1] The dynamic type of the purged catalyst has all the earmarks of being a dimer comprising of two 33-kDa subunits with noncovalently bound FMN as a cofactor. Hypotonic treatment of mitochondria uncovered that the NADH dehydrogenase is situated in the inward film/grid part confronting the lattice.^[2] FMN domain, ferredoxins or cytochrome b5 serve as the electron transport halfway between the FAD space and P450. The atomic advancement of both P450-containing frameworks and of every specific part does not take after phylogeny when all is said in done. Quality combination and flat quality exchange occasions can prompt to the presence of novel redox chains in a similar way that fake chimeric proteins

can be built by human.^[3] Mg^{2+} is crucial to all growths. It is a cofactor in enzymatic responses, settles the plasma layer, and its take-up is ATP subordinate. Fe is additionally fundamental to all fungi and it's a piece of the structure of cytochrome and cytochrome oxidase. Just hints of Zn, Mo, Mn, Cu and S were identified. Ca^{2+} can assume a roundabout part in fungi development by altering inward Ca^{2+} which controls the cytoplasmic Ca^{2+} angle, vesicle relocation to the tip and the movement of contagious proteins required in cell wall extension.^[4] The outflow of a few proteins is connected with the morphological development type of the organism and may assume a part in morphogenesis. At last, surface mannoproteins are solid immunogens that trigger and tweak the host insusceptible reaction.^[5]

METHODS

2D gel analysis of cell wall proteins: It is a proteomic examination approach, in light of two dimensional polyacrylamide gel electrophoresis was utilized to recognize and portray the cell wall related proteins.

BLAST P, multiple sequence alignment and phylogenetic tree construction: The amino acid sequence of Flavoprotein, obtained by MALDI-TOF/MS analysis of flavor protein kinase isolated from the cell

walls of *Neurospora crassa*. The protein sequences are scanned by using the BLAST P algorithm we can obtain the homologous protein sequences from the available protein sequences of various organisms. Template search with Blast and HHblits has been performed against the SWISS-MODEL template library. Overall 754 templates were found (Table 1a). The target sequence was searched with BLAST against the primary amino acid sequence contained in the SMTL.^[6] Models are built based on the target-template alignment using ProMod3. Coordinates which are conserved between the target and the template are copied from the template to the model. Insertions and deletions are remodeled using a fragment library. Side chains are then rebuilt. Finally, the geometry of the resulting model is regularized by using a force field. In case loop modelling with ProMod3 fails, an alternative model is built with PROMOD-II^[7] or MODELLER^[8] in case of reoccurring failure. Phylogenetic tree was then constructed using phylogeny.fr (<http://www.phylogeny.fr/>) to determine the evolutionary relationships.^[9-10]

Secondary structure prediction: Secondary structure of FMN_ red family protein was predicted using SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) tool in ExPasy.

Homology modeling: The sequence of Flavoprotein was downloaded from the universal protein resource^[11] (Uniprot KB) (<http://www.uniprot.org/>) (entry ID: F7VLA6). The suitable template for homology modeling was identified through searching Flavoprotein on PDB using the BLAST P algorithm.^[12] The 3D structure of Flavoprotein was downloaded from PDB (PDB ID: 3b6k.1.A) as the template structure.

Model validation: The quality of the homology model was validated by assessing the stereo chemical quality of the model using Ramachandran plot obtained from the RAMPAGE (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) server.^[13] Verify 3D^[14] and ERRAT^[15] were used to assess the amino acid environment from the UCLA-DOE server (<http://www.doe-mbi.ucla.edu/services>).

Metal binding sites: Metalloproteins account for almost half of all proteins in biology. Protein metallic-binding sites are liable for catalyzing probably the most difficult and but important functions, including photosynthesis, respiration, water oxidation, molecular oxygen reduction, and nitrogen fixation. In this paper the protein shows highest metal binding probability for metal namely calcium & magnesium (Table 3).

Model Quality Estimation: The global and per-residue model quality has been assessed using the QMEAN scoring function.^[16] For improved performance, weights of the individual QMEAN terms have been trained specifically for SWISS-MODEL.

RESULTS AND DISCUSSION

2D gel analysis of cell wall proteins

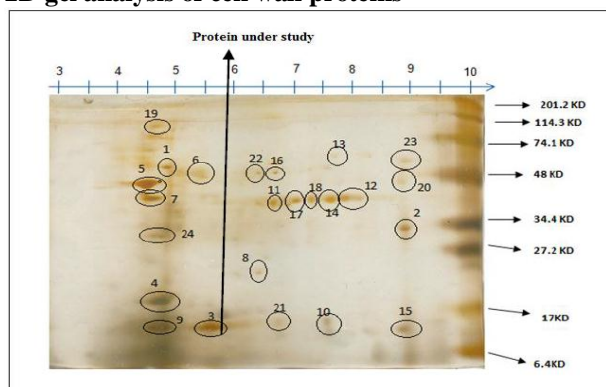


Fig. 1: 2D gel analysis showing protein under study.

Scanning of protein sequence databases using BLAST P with the sequence obtained by MALDI-TOF/MS analysis of the purified lipase revealed that the protein is a hypothetical protein from Flavoprotein with an entry ID: F7VLA6 was shown in (Fig. 2). A phylogram constructed based on multiple sequence alignment using phylogeny.fr revealed that FMN_reductase was closely related to a conserved hypothetical protein from *Sordaria macrospora* was shown in (Fig.3).

```
>tr|F7VLA6|F7VLA6_SORMK WGS project CABT00000000 data, contig 2.1
OS=Sordaria macrospora (strain ATCC MYA-333 / DSM 997 / K(L3346) / K-hell)
GN=SMAC_00501 PE=4 SV=1
MAPRIATIIYYSVMYGHIRQLADAAKAGIEKAGGTADLYQVPEITLSDEVLAKMHPKPTDV
PVIEDPAILKQYDGFLLGIPTRYGNFPAQWKTFFWDTQQLWATGGLYKKAAGLPISTAGL
GGGQESTAIAMSTFAHHGIIYVPLGYAKVFAELSDLSQVHGGSPWGAGTLAGGDSRQP
SENELKVGAIQGEFFYNVLSKLTGA
```

Fig. 2: Search for the Flavoprotein sequence in UNIPROT KB revealed that the sequence is FMN_reductase protein from *Sordaria macrospora*.

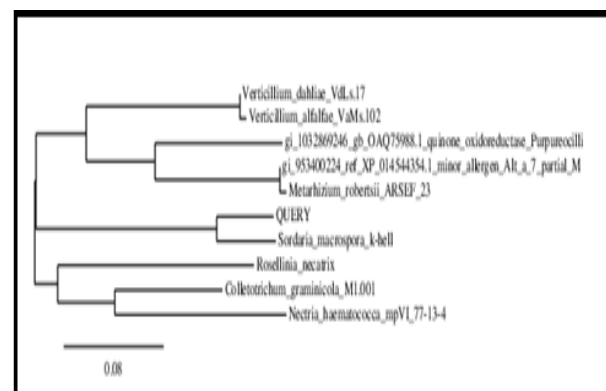


Fig. 3: A phylogenetic tree of Flavoprotein, constructed using phylogeny.fr showing that Flavoprotein is closely related to *Sordaria macrospora*.

4laf.1.D	44.00	homo-tetramer	BLAST	X-ray	1.76Å	0.41	0.98	NAD(P)H dehydrogenase (quinone)
5f4b.1.A	39.70	homo-tetramer	HHblits	X-ray	2.50Å	0.40	0.97	NAD(P)H dehydrogenase (quinone)
5f51.1.A	39.70	homo-tetramer	HHblits	X-ray	2.53Å	0.40	0.97	NAD(P)H dehydrogenase (quinone)
5f4b.1.A	41.62	homo-tetramer	BLAST	X-ray	2.50Å	0.40	0.96	NAD(P)H dehydrogenase (quinone)
5f51.1.A	41.62	homo-tetramer	BLAST	X-ray	2.53Å	0.40	0.96	NAD(P)H dehydrogenase (quinone)
4c76.1.A	11.11	monomer	HHblits	X-ray	1.96Å	0.25	0.83	FMN REDUCTASE (NADPH)
1czt.1.A	16.56	monomer	HHblits	X-ray	1.90Å	0.29	0.80	FLAVODOXIN
1yob.1.A	20.25	monomer	HHblits	X-ray	2.25Å	0.29	0.80	Flavodoxin 2
1yob.2.A	20.25	monomer	HHblits	X-ray	2.25Å	0.29	0.80	Flavodoxin 2
2gsu.1.A	11.18	homo-dimer	HHblits	X-ray	2.92Å	0.25	0.83	yhdA
2oys.1.A	14.46	homo-dimer	HHblits	X-ray	2.00Å	0.27	0.81	Hypothetical protein SP1951
1sq5.1.B	14.46	homo-dimer	HHblits	X-ray	1.50Å	0.27	0.81	conserved hypothetical protein
1sq5.1.A	14.46	homo-dimer	HHblits	X-ray	1.50Å	0.27	0.81	conserved hypothetical protein
1czk.1.A	16.67	monomer	HHblits	X-ray	1.90Å	0.28	0.79	FLAVODOXIN
1obo.1.A	16.05	monomer	HHblits	X-ray	1.20Å	0.28	0.79	FLAVODOXIN
1d04.1.A	15.34	monomer	HHblits	X-ray	1.85Å	0.28	0.80	FLAVODOXIN
1ftg.1.A	15.43	monomer	HHblits	X-ray	2.00Å	0.28	0.79	APOFLAVODOXIN
1d03.1.A	16.05	monomer	HHblits	X-ray	1.85Å	0.28	0.79	FLAVODOXIN
2v5u.1.A	15.43	monomer	HHblits	X-ray	1.99Å	0.28	0.79	FLAVODOXIN
3eyw.1.B	18.99	homo-dimer	HHblits	X-ray	2.40Å	0.30	0.77	C-terminal domain of Glutathione-regulated potassium-efflux system protein kefC fused to full length Glutathione-regulated potassium-efflux system ancillary protein keff
3eyw.1.A	18.99	homo-dimer	HHblits	X-ray	2.40Å	0.30	0.77	C-terminal domain of Glutathione-regulated potassium-efflux system protein kefC fused to full length Glutathione-regulated potassium-efflux system ancillary protein keff
319x.1.A	18.99	homo-dimer	HHblits	X-ray	2.10Å	0.30	0.77	Glutathione-regulated potassium-efflux system protein kefC, linker, ancillary protein keff
1dx9.1.A	14.81	monomer	HHblits	X-ray	2.05Å	0.28	0.79	FLAVODOXIN
1dx9.2.A	14.81	monomer	HHblits	X-ray	2.05Å	0.28	0.79	FLAVODOXIN
1dx9.3.A	14.81	monomer	HHblits	X-ray	2.05Å	0.28	0.79	FLAVODOXIN
1nni.1.A	10.71	homo-dimer	HHblits	X-ray	2.50Å	0.25	0.82	hypothetical protein yhda
3gfq.1.A	11.38	homo-dimer	HHblits	X-ray	3.00Å	0.25	0.81	FMN-dependent NADPH-azoreductase
3esx.1.A	15.53	monomer	HHblits	X-ray	2.31Å	0.28	0.79	Flavodoxin
3esx.2.A	15.53	monomer	HHblits	X-ray	2.31Å	0.28	0.79	Flavodoxin
3esy.1.A	14.20	monomer	HHblits	X-ray	2.39Å	0.28	0.79	Flavodoxin
3gfs.1.A	11.38	homo-tetramer	HHblits	X-ray	2.11Å	0.25	0.81	FMN-dependent NADPH-azoreductase
3esz.1.A	14.20	monomer	HHblits	X-ray	1.94Å	0.27	0.79	FLAVODOXIN
1rli.1.C	15.76	homo-tetramer	HHblits	X-ray	1.80Å	0.26	0.80	Trp Repressor Binding Protein

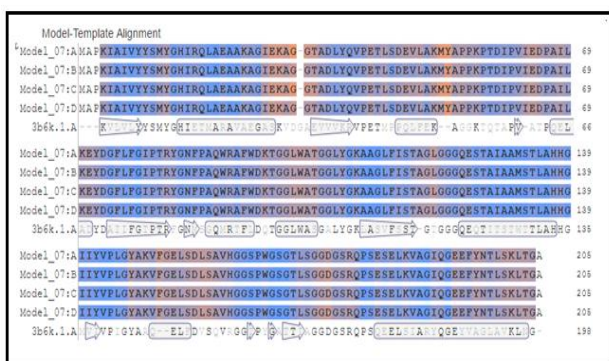


Fig. 5a.

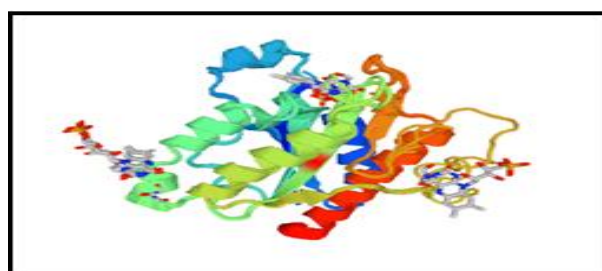


Fig. 6: Target-Template Alignment showed overall 754 templates out of them the selected templates of family Flavoprotein (3b6k.1.A) were used to build this model.

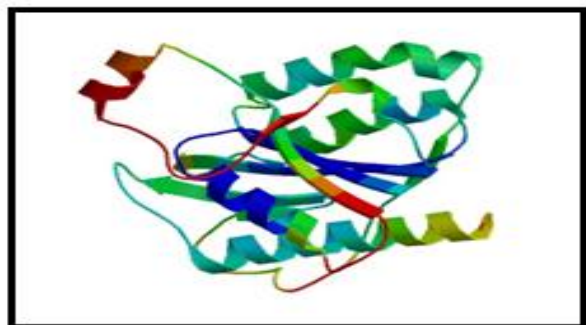


Fig. 7: Modeled protein image.

The stereo chemical quality of the 3D model (Fig.7) was validated by Ramachandran plot using RAMPAGE server. **Fig.8a** and **Table 2** shows that around 2.2% residues were present in the allowed regions and 97.8% residues in the favored region indicating that the quality of the model was good.

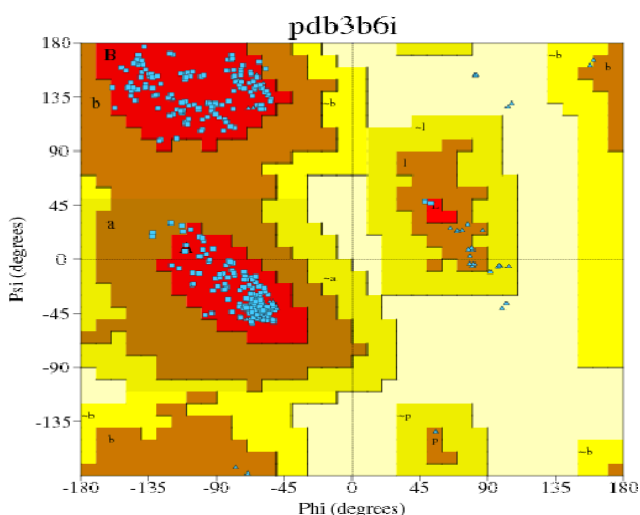


Fig. 8a: Ramachandran plot.

Table. 2: Ramachandran plot statistics.

Amino acid residues and regions (%)	Percentage
Residues in most favored regions [A,B,L]	97.8%
Residues in the allowed [a,b,l,p]	2.2%
Residues in the outlier regions	0%

Ramachandran Plot statistics for FMN_reductase homology model using RAMPAGE server

The quality of estimated model is based on the QMEAN scoring function were normalized with respect to the number of interactions.^[20] The QMEAN score of the model was 0.59 and the Z-score was -2.69, which was very close to the value of 0 and this shows the fine quality of the model, because the estimated reliability of the model was expected to be in between 0 and 1 and this could be inferred from the density plot for QMEAN scores of the reference set^[21-22] (**Fig.9A**). A comparison

between normalized QMEAN score (0.40) and protein size in non-redundant set of PDB structures in the plot revealed different set of Z-values for different parameters such as C-beta interactions (-1.43), interactions between all atoms (-2.33), solvation (-2.94), torsion (0.71) (**Fig.9B**).

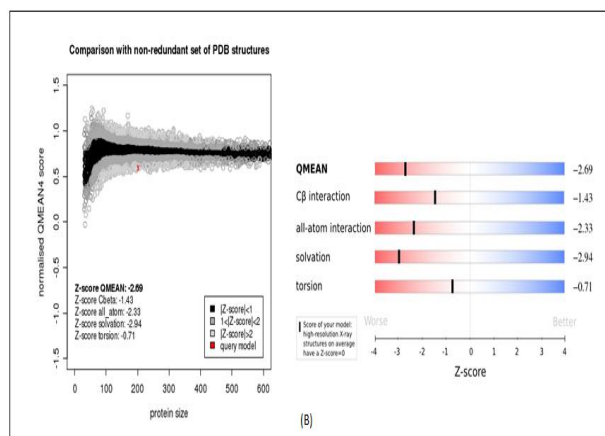
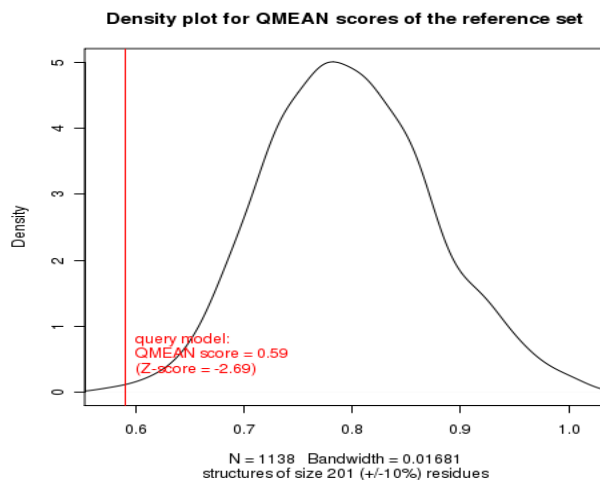


Fig.9: (A) The density plot for QMEAN showing the value of Z-score and QMEAN score (B) plot showing the QMEAN value as well as Z-score.

Metalloproteins are proteins capable of binding one or more metal ions, which may be required for their biological function, for regulation of their activities or for structural purposes. Metal-binding properties remain difficult to predict as well as to investigate experimentally at the whole-proteome level. Consequently, the current knowledge about metalloproteins is only partial. In this paper the protein with 4 metal binding sites shows highest metal binding probability for the metal namely calcium in sites 1,2 & 3 with the metal probability of **0.398702**, **0.627722** and **0.627722** ; magnesium in site 4 with the probability of **0.571190** were showed in (Table 3).

Table 3: Metal Binding Sites.

Protein	Metal Binding Sites	Metal binding probability								Metal Binding Pockets	
		CA	CO	CU	FE	MG	MN	NI	ZN	Amino acids	Position
Cell wall protein FMN-Reductase from <i>N.crassa</i>	Site-1	0.39870 2	0.01387 7	0.01665 2	0.34527 7	0.058975	0.03989 5	0.00797 9	0.11864 4	Y*	83
										G*	84
										N	85
										G	122
										G	123
	Site-2	0.62772 2	0.01675 1	0.02010 1	0.06323 3	0.071190	0.04815 7	0.00963 1	0.14321 6	S	172
										P*	180
										E*	184
										L	185
										A*	67
	Site-3	0.62772 2	0.01675 1	0.02010 1	0.06323 3	0.071190	0.04815 7	0.00963 1	0.14321 6	K*	70
										E*	71
										D*	65
	Site-4	0.12772 2	0.01675 1	0.02010 1	0.06323 3	0.571190	0.04815 7	0.00963 1	0.14321 6	A	67
										I	68
										E*	71

CONCLUSION

Communicable cell wall comprises of complex fibrillar material inserted in polysaccharide and different complexes, and utilitarian multi-functional proteins and glycoproteins. The cell wall might be highly protective, or generally exposed to the environment, as a result of different constituents in the signaling system. The cell membrane further permits the cell to communicate with the environment, empowering production, identification and assembling the cell constituents in a proper manner.

REFERENCES

- McGinnis MR, Tying SK. Introduction to Mycology. In: Baron S, editor. Medical Microbiology. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston; 1996.
- Fang J and Beattie Ds, " Novel FMN-containing rotenone-insensitive NADH dehydrogenase from *Trypanosoma brucei* mitochondria: isolation and characterization", *Biochemistry*, Mar. 5, 2002; 41(9): 3065-72.
- Kirill N Degtyarenko," Structural domains of P450-containing monooxygenase systems", *Protein Engineering*, 1995; 8(8): 737-747.
- C.O. Chardonnet, C.E. Sams a, W.S. Conway." Calcium effect on the mycelial cell walls of *Botrytis cinerea*", *pergamon, Phytochemistry*, 1999; 52: 967-973.
- W. lajean chaffin, jose´ luis lopez-ribo, manuel casanova, daniel gozalbo and josep. martinez. "Cell wall and secreted proteins of *candida albicans*: identification, function, and expression". *Microbiology and molecular biology reviews*, 1092-2172/98/\$04.0010 mar. 1998; 130-180.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, 1997; 25: 3389-3402.
- Remmert, M., Biegert, A., Hauser, A. and Soding, J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat Methods*, 2011; 9: 173-175.
- Guex, N. and Peitsch, M.C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, 1997; 18: 2714-2723.
- Sali, A. and Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*, 1993; 234: 779-815.
- Dereeper A., Audic S., Claverie J.M., Blanc G. BLAST-EXPLORER helps you building datasets for phylogenetic analysis. *BMC Evol Biol*, Jan. 2010; 12, 10: 8.
- Chevenet F., Brun C., Banuls AL., Jacq B., and Chisten R. TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinformatics*. 2006, Oct 10; 7:439.
- The UniProt Consortium, Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res.*, 2012; 40: D71-D75.
- Altschul, S.F., W. Gish, W. Miller, E.W. Myers and D.J. Lipman, Basic local alignment search tool. *J. Mol. Biol.*, 1990; 215: 403-410.
- Lovell, S.C., I.W. Davis, W.B. Arendall III, P.I.W. de Bakker and J.M. Word et al., Structure validation by α geometry: n, ψ and $C\beta$ deviation. *Proteins: Struct. Funct. Genet*. 2002; 50: 437-450.
- Bowie, J.U., R. Luthy and D. Eisenberg, A method to identify protein sequences that fold into a known three-dimensional structure. *Science*, 1991; 253: 164-170.

16. Colovos, C. and T.O. Yeates, Verification of protein structures: Patterns of nonbonded atomic interactions. *Prot. Sci.*, 1993; 2: 1511-1519.
17. Benkert, P., Biasini, M. and Schwede, T. Towards the estimation of the absolute quality of individual protein structure models. *Bioinformatics*, 2011; 27: 343-350.
18. Combet C., Blanchet C., Geourjon C. and Deleage G. NPS@: Network Protein Sequence Analysis. *TIBS*, March, 2000; 25(3): 291: 147-150.
19. Rost, B., Twilight zone of protein sequence alignments. *Protein Eng.*, 1999; 12: 85-94.
20. Yang, A.S. and B. Honig, An integrated approach to the analysis and modeling of protein sequences and structures. III. A comparative study of sequence conservation in protein structural families using multiple structural alignments. *J. Mol. Biol.*, 2000; 301: 691-711.
21. Wiederstein, M. and M.J. Sippl., ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, 2007; 35: 407-410.
22. Prajapat, R., A. Marwal and R.K. Gaur, 2014. Recognition of Errors in the Refinement and validation of three-dimensional structures of AC1 proteins of begomovirus strains by using ProSA-Web, *Journal of Viruses.*, 2014: doi.org/10.1155/2014/752656.