ejpmr, 2017,4(10), 186-193

SJIF Impact Factor 4.161

EUROPEAN JOURNAL OF PHARMACEUTICAL AND MEDICAL RESEARCH

www.ejpmr.com

Research Article

ISSN 2394-3211

EJPMR

HOMOLOGY MODELING AND STRUCTURAL STUDIES OF CELL WALL BINDING PROTEIN B-1,3, GLUCANASE FROM *NEUROSPORA CRASSA*

K. Ashok Kumar, S. K. Gousia, M. Mary Moses and J. Naveena Lavanya Latha*

Department of Biotechnology, Krishna University, Machilipatnam-521001.

***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: The structure of protein determines its function and its interactions with other. The regions of proteins that interact with ligands, substrates or other proteins tend to be conserved both in sequence and structure. Experimentally identifying and characterizing metal ion binding sites is a time consuming and costly process. There are many computational methods developed to identify metal ion binding sites, but most use only sequence information. **Methods:** The work reported here contains, the stereo chemical quality of the protein model was checked by using *in silico* analysis with PROCHECK and QMEAN servers. **Results:** In this paper the protein with 4 metal binding sites shows highest metal binding probability for the metal namely calcium in sites 1 & 3 with the metal probability of **0.627722** & **0.757665;** magnesium in site 2 with the probability of **0.571190** and zinc in site 4 with the probability of **0.643216** was shown. The **85.9 %** residues in the core region of Ramachandran plot showing high accuracy of protein model and the QMEAN Z-score of **-3.03** indicates the overall model quality of protein. **Conclusions:** The result of the study may be a guiding point for further investigations on Lam G protein and metal binding sites.

KEYWORDS: Cell Wall Proteins, β-1,3, Glucanase, Metal Binding Sites, Homology Modeling, *Neurospora Crassa*.

INTRODUCTION

Cell wall that plays a key role in controlling shape and protecting the fungi and other organisms from the environment. It contains molecules that are involved in morphogenesis, reproduction, cell-cell and cell-matrix interactions[1]. The significant elements of the cell wall are glucan and chitin, which are connected with stringent tough nature, and mannoproteins. Cell wall proteins may vary in their form, production, or topological area inside the wall structure. Proteins might be altered by glycosylation (principally expansion of mannose buildups), phosphorylation, and ubiquitination. Cell wall proteins have been involved in grip to host tissues and ligands. Fibrinogen, supplement parts, and a few extracellular grid segments are among the host proteins bound by cell wall proteins[2]. In *Aspergillus fumigatus* stretching of 1-3 glucan brings about an expansion of acceptor destinations for chitin, galactomannan and a direct β 1-3/1-4 – glucan which substitutes the β- 1-6 glucan normally communicated in other growths. Glucans can likewise covalently tie to cell wall proteins (CWP). Infections connected with cell wall proteins - *Cryptococcus neoformans* and *Histoplasma capsulatum*. There are two assortments of *Histoplasma capsulatum* that are pathogenic to people, *H. capsulatum var. capsulatum* and *H. capsulatum var. duboisii*, and a third assortment that is an equine pathogen, *H. capsulatum var. farciminosum.*[1]

Calcium is an essential micronutrient required for the optimal growth and metabolism of all living organisms. Calcium is required for patterns of hyphal extension, branching, sexual and asexual reproduction in fungi.[3] A large number of proteins and enzymes require zinc as cofactor for functioning properly. Many transcription factors are the examples for zinc binding proteins.[4] In the present study, Homology modeling was applied to build a 3D structure of β-1, 3, glucanase purified and characterized from isolated cell wall preparations of *Neurospora crassa*. Metal binding sites were determined using CHED website and phylogenetic tree was also constructed.

METHODS

BLAST P, multiple sequence alignment and phylogenetic tree construction

The amino acid sequence obtained by MALDI-TOF/MS analysis of a protein intially refered as β-1, 3, glucanase from cell walls of *Neurospora crassa*. The protein sequences are scanned by using the BLAST P 2.2.24+ and can obtain the homologous protein sequences from the available protein sequences of various organisms. The target sequence was searched with BLAST[5] against

the primary amino acid sequence contained in the SMTL. A total of 6 templates were found. An initial HHblits profile has been built using the procedure outlined in a report by Remmert, et al[6], followed by 1 iteration of HHblits against NR20. The obtained profile has then been searched against all profiles of the SMTL. A total of 64 templates were found. Overall 70 templates were found (Table 1). Models are built based on the target-template alignment using Promod-II. 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 ProMod-II[7] does not give satisfactory results, an alternative model is built with MODELLER 9.15.[8] Phylogenetic tree was then constructed using phylogeny.fr (http://www.phylogeny.fr/) to determine the evolutionary relationships.[9,10,11]

Secondary structure prediction

Secondary structure of **Lam G** 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 **β-1,3, glucanase** was downloaded from the universal protein resource (Uniprot KB) (http://www.uniprot.org/)[12] (entry ID: Q1K7S6). The suitable template for homology modeling was identified through searching **β-1,3, glucanase** on PDB using the BLAST P algorithm.[13] The 3D structure of **β-1,3, glucanase** was downloaded from PDB (PDB ID: 3 wdw.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.[14] Verify 3D[15] and ERRAT[16] were used to assess the amino acid environment from the UCLA-DOE server (http://www.doe-mbi.ucla.edu/services).

Model Quality Estimation

The global and per-residue model quality has been assessed using the QMEAN scoring function.[17] For improved performance, weights of the individual QMEAN terms have been trained specifically for SWISS-MODEL.

Metal binding sites

In the majority of the metalloproteins, the residues involved in metal binding come close together in the tertiary structure to form the binding site, but are dispersed along the amino acid sequence.

RESULTS AND DISCUSSION

Scanning of protein sequence databases using BLAST P 2.2.24+ with the sequence obtained by MALDI-TOF/MS analysis of the purified protein (to be published elsewhere) revealed that the protein is a hypothetical protein from **β-1, 3, Glucanase** with an entry ID: Q1K7S6 was shown in (Fig.1). A phylogram constructed based on multiple sequence alignment using phylogeny.fr revealed that **LamG** was closely related to a conserved hypothetical protein from *Neurospora tetrasperma* was shown in (Fig.2).

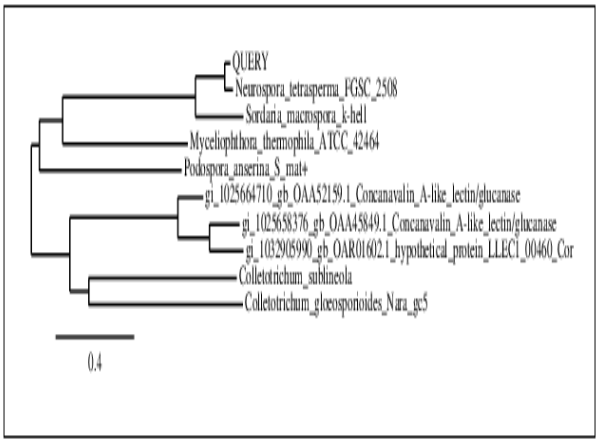

Fig. 1: Search for the β-1, 3, glucanase sequence in UNIPORT KB revealed that the sequence is a LamG protein from *Neurospora tetrasperma.*

Fig. 2: A phylogenetic tree of β-1, 3, glucanase, constructed using phylogeny.fr showing that β-1, 3, glucanase is closely related to *Neurospora tetrasperma.*

Secondary structure of the target protein was predicted by using SOPMA tool in Expasy (Fig. 3). The results indicate that LamG has 16.83%, α-helix thus making it stable for homology modeling.[18]

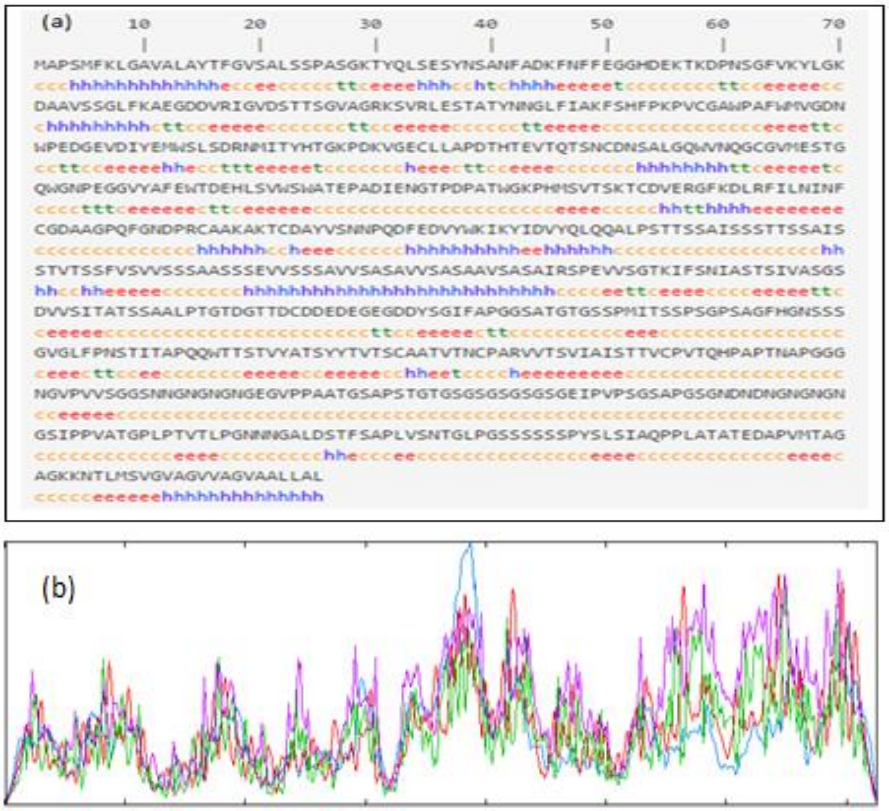

Fig.3: (a-b) secondary structure of β-1, 3, glucanase, (a) Sequence length: 725; Alpha helix(Hh): 122 is 16.83%, Extended strand (Ee) :169 is 23.31%, Beta turn (Tt) : 43 is 5.93%, Random coil (Cc): 391 is 53.93% and (b) Distribution of Secondary structure elements of β-1, 3, glucanase. Blue line-Alpha Helix, Red-Extended strand, Green-beta turn, Orange-random coil.

The first step in homology modeling involves identification of a suitable template. This was met by performing a BLAST P search against known protein structures deposited in PDB. The studies of Rost[19] and Yang and Honig[20] demonstrated that 3D structures will be similar if the sequence identity between target and template proteins is higher than 25%. Generally, a target which shares a sequence similarity of 30% or more to an experimentally solved protein structure (template) can only be employed for homology modeling. The crystal structure of Beta-1, 3-1, 4-glucanase (3wdw.1.A) with a sequence identity of 38.25% to the target sequence was selected based on BLAST P search against PDB database (Table.1). The sequence alignment between the template (3wdw.1.A) and the target was shown in (Fig.4).

Table. 1: BLAST results of target sequence of LamG against PDB for the identification of template for homology modeling.

Template	Seq Identity	Oligo-state	Found by	Method	Resolution	Seq Similarity	Coverage	Description
3wdu.1.A	40.85	monomer	BLAST	X-ray	2.25Å	0.41	0.39	Beta-1,3-1,4-glucanase
3wdt.1.A	40.85	monomer	BLAST	X-ray	1.98Å	0.41	0.39	Beta-1,3-1,4-glucanase
3wdy.1.A	40.49	monomer	BLAST	X-ray	1.94Å	0.41	0.39	Beta-1,3-1,4-glucanase
3wdw.1.A	40.49	monomer	BLAST	X-ray	1.80Å	0.41	0.39	Beta-1,3-1,4-glucanase
3wdt.1.A	38.60	monomer	HHblits	X-ray	1.98Å	0.40	0.39	Beta-1,3-1,4-glucanase
3wdu.1.A	38.60	monomer	HHblits	X-ray	2.25Å	0.40	0.39	Beta-1,3-1,4-glucanase
3wdy.1.A	38.60	monomer	HHblits	X-ray	1.94Å	0.40	0.39	Beta-1,3-1,4-glucanase
3wdw.1.A	38.25	monomer	HHblits	X-ray	1.80Å	0.39	0.39	Beta-1,3-1,4-glucanase
2w39.1.A	37.50	monomer	BLAST	X-ray	1.10Å	0.39	0.39	PUTATIVE LAMINARINASE
2wlq.1.A	37.14	monomer	BLAST	X-ray	1.40Å	0.39	0.39	PUTATIVE LAMINARINASE
2wlq.1.A	35.97	monomer	HHblits	X-ray	1.40Å	0.38	0.38	PUTATIVE LAMINARINASE
2w39.1.A	36.46	monomer	HHblits	X-ray	1.10Å	0.38	0.38	PUTATIVE LAMINARINASE
4atf.1.A	13.44	monomer	HHblits	X-ray	1.90Å	0.27	0.35	BETA-AGARASE B
3dgt.1.A	22.36	monomer	HHblits	X-ray	1.50Å	0.31	0.33	Endo-1,3-beta-glucanase
3iln.1.A	24.14	monomer	HHblits	X-ray	1.95Å	0.32	0.32	Laminarinase

3azx.1.A	21.70	monomer	HHblits	X-ray	1.80Å	0.30	0.32	Laminarinase
3b01.1.A	21.70	monomer	HHblits	X-ray	1.87Å	0.30	0.32	Laminarinase
3rq0.1.A	15.97	monomer	HHblits	X-ray	2.02Å	0.29	0.33	Glycosyl hydrolases family protein 16
3wz1.1.A	13.93	monomer	HHblits	X-ray	1.60Å	0.26	0.34	Agarase
4dfs.1.A	21.74	monomer	HHblits	X-ray	3.75Å	0.31	0.32	Glycoside hydrolase, family 16
2vy0.1.A	19.91	monomer	HHblits	X-ray	2.16Å	0.30	0.32	ENDO-BETA-1,3-GLUCANASE
3atg.1.A	20.96	monomer	HHblits	X-ray	1.66Å	0.30	0.32	GLUCANASE
4bq1.1.A	21.88	monomer	HHblits	X-ray	1.50Å	0.31	0.31	ENDO-1,3-BETA-GLUCANASE, FAMILY GH16
4bow.1.A	20.98	monomer	HHblits	X-ray	1.35Å	0.31	0.31	ENDO-1,3-BETA-GLUCANASE, FAMILY GH16
2hyk.1.A	22.42	monomer	HHblits	X-ray	1.30Å	0.31	0.31	Beta-1,3-glucanase
1o4y.1.A	14.22	monomer	HHblits	X-ray	1.48Å	0.26	0.32	beta-agarase A
4ate.1.A	17.57	monomer	HHblits	X-ray	1.10Å	0.29	0.31	BETA-PORPHYRANASE A
4awd.1.A	11.79	monomer	HHblits	X-ray	2.40Å	0.26	0.32	BETA-PORPHYRANASE
4crq.1.A	19.35	monomer	HHblits	X-ray	1.50Å	0.30	0.30	ENDO-1,3-BETA-GLUCANASE, FAMILY GH16
4wzf.1.A	13.96	monomer	HHblits	X-ray	1.70Å	0.28	0.31	1,3-beta-glucanase
1urx.1.A	13.60	monomer	HHblits	X-ray	1.70Å	0.25	0.31	BETA-AGARASE A
5dxd.1.A	17.97	monomer	HHblits	X-ray	1.70Å	0.29	0.30	Putative beta-glucanase
1dyp.1.A	18.52	monomer	HHblits	X-ray	1.54Å	0.29	0.30	KAPPA-CARRAGEENASE
3juu.1.A	11.16	monomer	HHblits	X-ray	1.80Å	0.26	0.31	porphyranase B
4xxp.1.A	15.35	monomer	HHblits	X-ray	1.60Å	0.29	0.30	Putative uncharacterized protein (Rv0315 ortholog)
4pq9.1.A	14.88	monomer	HHblits	X-ray	1.20Å	0.29	0.30	Beta-1,3-glucanase
3ilf.1.A	16.06	monomer	HHblits	X-ray	1.80Å	0.28	0.30	porphyranase A
4xdq.1.A	14.22	monomer	HHblits	X-ray	1.35Å	0.28	0.30	Glycoside hydrolase family protein
4w65.1.A	15.02	monomer	HHblits	X-ray	1.38Å	0.29	0.29	Glycosyl hydrolase family protein
2ayh.1.A	14.71	monomer	HHblits	X-ray	1.60Å	0.28	0.28	1,3-1,4-BETA-D-GLUCAN 4-GLUCANOHYDROLASE
3wvj.1.A	12.14	monomer	HHblits	X-ray	1.95Å	0.27	0.28	Beta-glucanase
1u0a.1.A	14.71	monomer	HHblits	X-ray	1.64Å	0.28	0.28	Beta-glucanase
3i4i.1.A	11.65	monomer	HHblits	X-ray	1.89Å	0.27	0.28	1,3-1,4-beta-glucanase
1mac.1.A	14.71	monomer	HHblits	X-ray	2.30Å	0.28	0.28	1,3-1,4-BETA-D-GLUCAN 4-GLUCANOHYDROLASE
3o5s.1.A	14.85	monomer	HHblits	X-ray	2.20Å	0.28	0.28	Beta-glucanase
1gbg.1.A	16.08	monomer	HHblits	X-ray	1.80Å	0.29	0.27	(1,3-1,4)-BETA-D-GLUCAN 4 GLUCANOHYDROLASE
2uwc.1.A	14.84	monomer	HHblits	X-ray	2.30Å	0.27	0.25	CELLULASE
2uwa.1.A	14.84	monomer	HHblits	X-ray	1.80Å	0.27	0.25	CELLULASE
2uwb.1.A	15.82	monomer	HHblits	X-ray	2.00Å	0.28	0.24	CELLULASE
2vh9.1.A	15.34	monomer	HHblits	X-ray	2.10Å	0.28	0.24	CELLULASE
1umz.1.A	14.94	monomer	HHblits	X-ray	1.80Å	0.27	0.24	XYLOGLUCAN ENDOTRANSGLYCOSYLASE
1un1.2.A	14.94	monomer	HHblits	X-ray	2.10Å	0.27	0.24	XYLOGLUCAN ENDOTRANSGLYCOSYLASE
1zm1.1.A	17.18	homo-dimer	HHblits	X-ray	2.30Å	0.28	0.22	Beta-glucanase
3d6e.1.A	16.67	monomer	HHblits	X-ray	2.40Å	0.28	0.22	Beta-glucanase
3hr9.1.A	17.18	monomer	HHblits	X-ray	1.70Å	0.28	0.22	Beta-glucanase
3h0o.1.A	16.56	monomer	HHblits	X-ray	1.40Å	0.27	0.22	Beta-glucanase
3axd.1.A	17.90	monomer	HHblits	X-ray	1.53Å	0.28	0.22	Beta-glucanase
1mve.1.A	16.56	monomer	HHblits	X-ray	1.70Å	0.27	0.22	Truncated 1,3-1,4-beta-D-glucanase
2r49.1.A	16.15	monomer	HHblits	X-ray	2.20Å	0.27	0.22	Beta-glucanase
1ups.1.A	22.52	monomer	HHblits	X-ray	1.82Å	0.30	0.21	GLCNAC-ALPHA-1,4-GAL-RELEASING ENDO-BETA-GALACTOSIDASE

1ups.2.A	22.52	monomer	HHblits	X-ray	1.82Å	0.30	0.21	GLCNAC-ALPHA-1,4-GAL-RELEASING ENDO-BETA-GALACTOSIDASE
1cpm.1.A	15.79	monomer	HHblits	X-ray	2.00Å	0.28	0.21	CIRCULARLY PERMUTED
1cpn.1.A	17.33	monomer	HHblits	X-ray	1.80Å	0.29	0.21	CIRCULARLY PERMUTED
1axk.1.A	18.24	homo-dimer	HHblits	X-ray	2.10Å	0.29	0.20	GLUXYN-1
1axk.1.B	18.24	homo-dimer	HHblits	X-ray	2.10Å	0.29	0.20	GLUXYN-1
1ajk.1.A	15.75	monomer	HHblits	X-ray	1.80Å	0.28	0.18	CIRCULARLY PERMUTED (1-3,1-4)-BETA-D-GLUCAN 4-GLUCANOHYDROLASE
4asm.1.A	17.74	monomer	HHblits	X-ray	1.50Å	0.29	0.17	BETA-AGARASE D
1o4z.1.A	10.28	homo-dimer	HHblits	X-ray	2.30Å	0.26	0.15	beta-agarase B
1ajo.1.A	15.63	monomer	HHblits	X-ray	2.07Å	0.28	0.13	CIRCULARLY PERMUTED (1-3,1-4)-BETA-D-GLUCAN 4-GLUCANOHYDROLASE CPA16M-127
1ajo.1.A	16.44	monomer	HHblits	X-ray	2.07Å	0.29	0.10	CIRCULARLY PERMUTED (1-3,1-4)-BETA-D-GLUCAN 4-GLUCANOHYDROLASE CPA16M-127

Fig. 4: Alignment between target (LamG) and template β-1, 3-glucanase (3wdw.1.A).

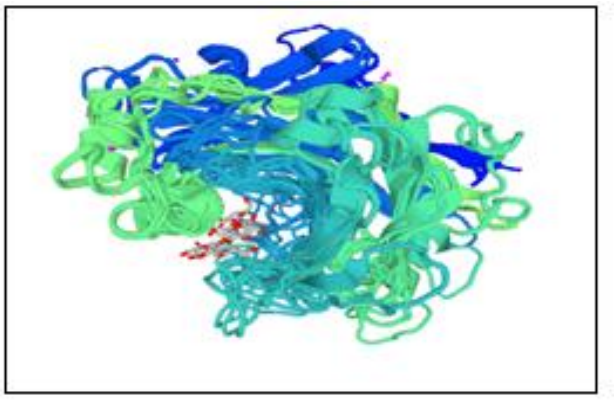

Fig. 5: Target-Template Alignment showed overall 70 templates out of them the selected template Beta-1, 3-1, 4-glucanase (3wdw.1.A) were used to build this model.

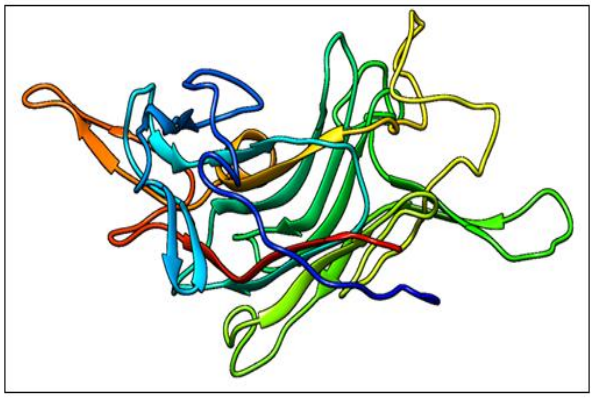

Fig. 6: Modeled protein image created by aligning the Target-template models.

The stereo chemical quality of the 3D model was validated by Ramachandran plot using RAMPAGE server. Fig.7 and Table 2 shows that around 13.0% residues were present in the allowed regions, 85.9 % residues in the favored region and only 1.1% residues were present in the outlier region indicating that the quality of the model was good.

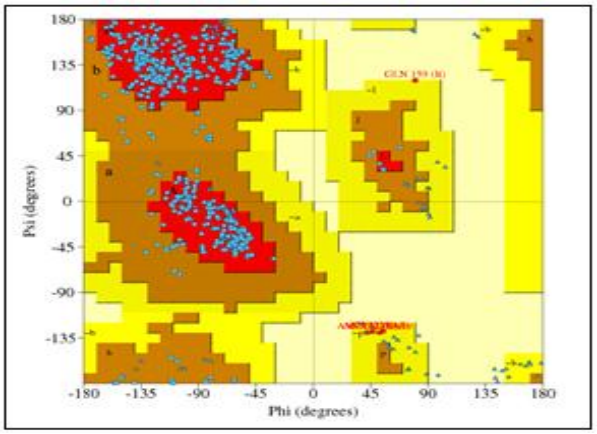

Fig. 7: Ramachandran plot of the PDB model from Fig.6 using RAMPAGE.

Table 2: Ramachandran Plot statistics for Lam G homology model using RAMPAGE server.

Amino acid residues and regions (%)	Percentage
Residues in most favored regions [A,B,L]	85.9 %
Residues in the allowed [a,b,l,p]	13.0 %
Residues in the outlier regions	1.1 %

The quality of estimated model is based on the QMEAN scoring function were normalized with respect to the number of interactions.[21] The QMEAN score of the model was 0.58 and the Z-score was -3.03, which was very close to the value of 0 and this shows the fine quality of the model[22,23] 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 (Fig.8A). 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 (0.20), interactions between all atoms (-0.90), solvation (-0.88), torsion (-2.62) showed in (Fig. 8B).

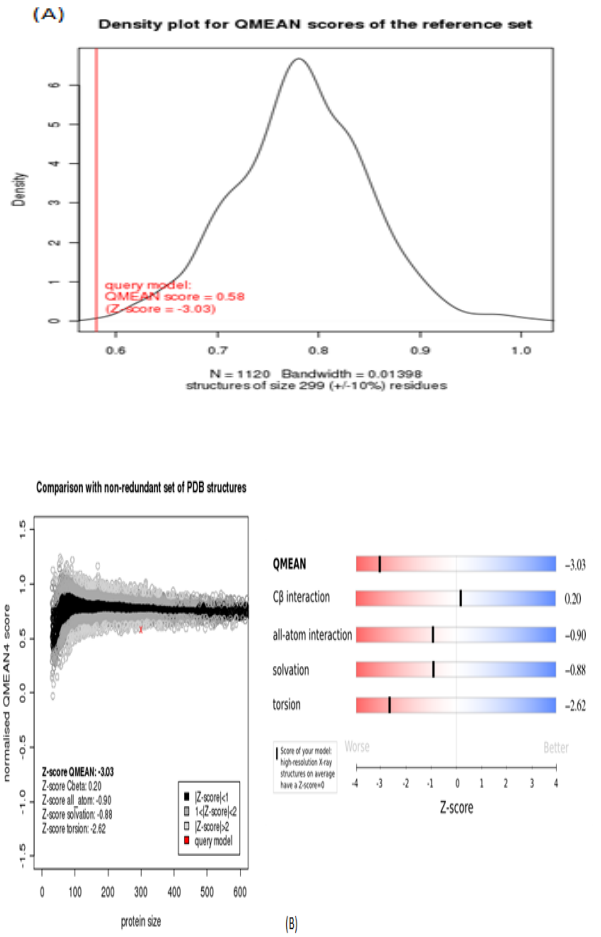

Fig. 8: (A) The density plot for QEMEAN showing the value of Z-score and QMEAN score (B) plot showing the QMEAN value as well as Z-score.

There are about approximately 1,100 genes, or more than 10% of the total predicted in the genome[24]. Of which, 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 & 3 with the metal probability of **0.627722** & **0.757665 ;** magnesium in site 2 with the probability of **0.571190** and zinc in site 4 with the probability of **0.643216** showed in (Fig.9).

Table 3: Metal binding site prediction using CHED server.

Protein	Metal Binding Sites	Metal binding probability								Metal Binding Pockets	
		CA	CO	CU	FE	MG	MN	NI	ZN	Amino acids	Position
CWP 1,3, β-glucase from N.crassa	Site-1	**0.627722**	0.016751	0.020101	0.063233	0.071190	0.048157	0.009631	0.143216	W	135
										E*	146
										D*	148
										E*	151
	Site-2	0.127722	0.016751	0.020101	0.063233	**0.571190**	0.048157	0.009631	0.143216	S*	208
										T*	209
										W	212
										P	215
	Site-3	**0.757665**	0.007761	0.009313	0.029297	0.102833	0.022313	0.004463	0.066356	A	40
										N	41
										F	42
										G	83
										D*	84
										K*	321
										Y*	322
	Site-4	0.127722	0.016751	0.020101	0.063233	0.071190	0.048157	0.009631	**0.643216**	C*	203
										A	235
										T	236
										E*	237

SIGNIFICANT STATEMENT/CONCLUSION

In this paper we have showed the different tools like SOPMA for secondary structure prediction, Phylogenetic tree was then constructed to determine the evolutionary relationships. Model validation was done by using Ramchandran plot and metal binding sites. More than 70,000 protein structures are currently found in the Protein Data Bank, and approximately one-third contain metal ions essential for function. The result of the study may be a guiding point for further investigations on Lam G protein and metal binding sites.

ACKNOWLEDGMENT

The authors are grateful to Krishna University authorities for providing the facilities required for the present work. KAK grately acknowledges the grant received by UGC-RGNF (F1-17.1/2016-17/RGNF-2015-17-SC AND-20792/(SA-III)

REFERENCES

1. Pinto, M.R.; Barreto-Bergter, E.; Taborda C.P. “glycoconjugates and polysaccharides of fungal cell wall and activation of immune system”. Brazilian Journal of Microbiology, 2008; 39: 195-208.
2. Lajean chaffin, W jose luis lo pez-ribot,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 1998; 130–180.
3. Thommandru Naga Sowjanya and Pamarthi Maruthi Mohan. “A calcium binding protein from cell wall of Neurospora crassa”. Journal of Basic Microbiology, 2009; 49: 371-376.
4. Sarah Mac Pherson, Marc Larochelle and Bernard Turcotte. “A Fungal Family of Transcriptional Regulators: the Zinc Cluster Proteins”. Microbiology and Molecular biology reviews, sept. 2006; 583–604.
5. 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.
6. 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.
7. Guex, N. and Peitsch, M.C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis, 1997; 18: 2714-2723.
8. Sali, A. and Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol, 1993; 234: 779-815.
9. Dereeper A., Audic S., Claverie J.M., Blanc G. BLAST-EXPLORER helps you building datasets for phylogenetic analysis. BMC Evol Biol., Jan 12 2010; 10: 8.
10. Dereeper A.*, Guignon V.*, Blanc G., Audic S., Buffet S., Chevenet F., Dufayard J.F., Guindon S., Lefort V., Lescot M., Claverie J.M., Gascuel O. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 2008 Jul 1;36(Web Server issue):W465-9. Epub, Apr 19 2008; (PubMed) *: joint first authors.
11. Chevenet F., Brun C., Banuls AL., Jacq B., Chisten R. TreeDyn: towards dynamic graphics and annotations for analyses of trees. BMC Bioinformatics, 10 2006; 7: 439.
12. The UniProt Consortium, Reorganizing the protein space at the Universal Protein Resource (UniProt). Nucleic Acids Res., 2012; 40: D71-D75.

13. Altschul, S.F., W. Gish, W. Miller, E.W. Myers and D.J. Lip man, Basic local alignment search tool. J. Mol. Biol., 1990; 215: 403-410.
14. Lovell, S.C., I.W. Davis, W.B. Arendall III, P.I.W. de Bakker and J.M. Word et al., Structure validation by Cα geometry: n,ψ and Cβ deviation. Proteins: Struct., Funct., Genet., 2002; 50: 437-450.
15. 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. Toward 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. Benkert, P., Biasini, M. and Schwede, T. "Toward the estimation of the absolute quality of individual protein structure models". Bioinformatics, 2011; 27: 343-350.
22. 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.
23. 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.