BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 9(16)

In-silico characterization of β-(1, 3)-endoglucanase (ENGL1) from *Aspergillus fumigatus* by homology modeling and docking studies

Rizwan Ahmed[1], Swatantra Kumar Jain[2] & Praveen Kumar Shukla[1]*

[1]Medical Mycology lab, Division of Microbiology, CSIR-Central Drug Research Institute, Sitapur Road, Lucknow-226001, India; [2]Department of Biotechnology, Jamia Hamdard University, New Delhi, India; Praveen Kumar Shukla – Email: pk_shukla@cdri.res.in; Phone: 91-522-3290091; Fax: 91-535-2700857; *Corresponding author

Received August 18, 2013; Accepted August 20, 2013; Published September 23, 2013

Abstract:
During the past few years a significant rise in aspergillosis caused by filamentous fungus *Aspergillus fumigatus* has been recorded particularly in immunocompromised patients. At present, there are limited numbers of antifungal agents to combat these infections and the situation has become more complex due to emergence of antifungal resistance and side-effects of antifungal drugs. These situations have increased the demand for novel drug targets. Recent studies have revealed that the β-1,3-endoglucanase (ENGL1) plays an essential role in cell wall remodeling that is absolutely required during growth and morphogenesis of filamentous fungi and thus is a promising target for the development of antifungal agents. Unfortunately no structural information of fungal β-glucanases has yet been available in the Protein Databank (PDB). Therefore in the present study, 3D structure of β-(1,3)-endoglucanase (ENGL1) was modeled by using I-TASSER server and validated with PROCHECK and VERIFY 3D. The best model was selected, energy minimized and used to analyze structure function relationship with substrate β-(1,3)-glucan by C-DOCKER (Accelrys DS 2.0). The results indicated that amino acids (GLU 380, GLN 383, ASP 384, TYR 395, SER 712, and ARG 713) present in β-1,3-endoglucanase receptor are of core importance for binding activities and these residues are having strong hydrogen bond interactions with β-(1,3)-glucan. The predicted model and docking studies permits initial inferences about the unexplored 3D structure of the β-(1,3)-endoglucanase and may be promote in relational designing of molecules for structure-function studies.

Keywords: Homology modeling, β-(1,3)-endoglucanase, *Aspergillus fumigatus*, Docking, β-(1,3)-glucan.

Background:
A fumigatus is the most common opportunistic pathogen that causes life-threatening invasive aspergillosis (IA) in immunocompromised patients. The risk of developing IA results primarily from a dysfunction in host defense mechanism in combination with fungal attributes that permits the survival and growth of *A. fumigatus* during infection **[1]**. Due to the lack of effective and safe antifungal therapeutics, the fungal cell wall and the underlying associated biosynthetic enzyme of pathogenic moulds have been placed in the spotlight in recent research. It is now evident that the fungal cell wall is a conserved and highly dynamic structure and their components are essential for virulence and viability of fungal pathogens **[2]**. The identification and study of these multifunctional molecules that help and promote the growth of *A. fumigatus* in the host during infection, may divulge unique infection mechanisms that could lead to new control measures for aspergillosis. β-(1,3)-glucans is a predominant cell wall polysaccharide of most medically important fungi, including *A. fumigatus*, and is thought to be responsible for the shape and rigidity of the wall **[3]**. The organization of the β-(1,3)-glucans in the fungal cell wall is poorly known and it has been suggested that β-(1,3)-endoglucanase activities are essential for the continuous rearrangement of the wall β-(1,3)-glucans during fungal growth

[4]. These enzymes catalyze the hydrolysis of the substrate by two possible mechanisms: (a) β-(1,3)-exo-glucanases hydrolyze the substrate by sequentially cleaving glucose residues from the nonreducing end, (b) β-(1,3)-endo-glucanases cleave β-linkages at apparently random sites along the polysaccharide chain, releasing smaller oligosaccharides **[5]**. Because of their mode of action, β-(1,3)-endo-glucanases have a more prominent action on the cell wall than β-(1,3)-exo-glucanases, especially in the morphological events such as germination and branching which require plasticizing of the wall structure. The β-(1,3)-endo-glucanases of *A. fumigatus* is able to act directly on the β-(1,3)-glucans polymers of the entire cell wall allowing for hyphal branching as well as for germ tube emergence; and also seems to play role in formation numerous of free reducing and non-reducing ends necessary for the activity of β-(1,3)-glucanosy1 transferases identified in the periplasmic space of fungi **[6,7]**. Although these hydrolytic enzymes are believed to play cell wall remodeling roles during growth and morphogenesis in filamentous fungi, they may play multiple physiological roles, which have not been fully elucidated in *A. fumigatus* **[8]**. We have recently reported the immunodominant nature of β-(1,3)-endo-glucanases isolated from *A. fumigatus* by the application of immuno-proteomics approach and is considered to have diagnostic potential as well as can be targeted as vaccine candidate **[9]**. Information on fungal β-(1,3)-endo-glucanases is very limited and their exact role during cell wall ontogeny is not well known. Despite extensive studies of β-(1,3)-endo-glucanases, no structural information has been reported to date on their mode of interaction with oligosaccharide ligands. It is clear that computer-based homology modeling and docking studies can be useful in the identification of conserved residues essential for catalysis and structural domains responsible for the regulation of enzyme activity. In the present study the homology modeling and docking analysis of a β-(1,3)-endo-glucanases from *A. fumigatus* were performed. The results provided new insight into molecular interactions of active site residues with substrates for the enzymatic function.

Methodology:

Retrieval of target sequence

The 727 amino acid long sequence of ENGL1 from *Aspergillus fumigatus* was obtained from the Protein sequence database of NCBI and blasted **[10]** against Protein Data Bank (PDB) entries to find similar sequences.

Characterization of target sequence

Subcellular localization of protein using amino acid composition was achieved by MultiLoc **[11].** NetNGlyc 1.0 Server predicted N-Glycosylation sites using artificial neural networks. Physiochemical properties such as molecular weight, theoretical pI, total number of negatively (Asp+Glu) and positively (Arg+Lys) charged residues, extinction coefficients, instability index, aliphatic index and grand average of hydropathicity (GRAVY) **[12]** of the mature protein were computed using Expasy's ProtParam Proteomics server **[13]**.

Protein Topology prediction

Secondary structure of the protein ENGL1 from *Aspergillus fumigatus* was calculated with PSIPRED (http: //bioinf. cs.ucl.ac.uk/ psipred) and disulfide bonds were predicted by the Cys_REC tool (http://sunl.softberry.com/berry. phtml?topic).

Homology modeling

Homology modeling was used to generate a reliable 3D model of ENGL1 protein by MODELLER 9.10 **[14].** The model predicted by MODELLER in our study was of poor quality with inappropriate folded conformation as a good quality model could not be obtained even by using multiple templates. Hence, ENGL1 structure was predicted using the online protein structure and function prediction server I-TASSER **[15]** which uses the threading technique to predict the 3D models. By the use of this server, 5 best models based on multiple-threading alignments and iterative template fragment assembly simulations along with their confidence scores were generated**.** These 5 models were visualized by the Accelrys DS 2.0 software (Accelrys Inc., San Diego, CA, USA) and evaluated using different validation techniques.

Model refinement and evaluation

The stereo chemical quality and accuracy of the predicted model was evaluated with PROCHECK **[16]** by Ramachandran plot analysis **[17]**. The best model was selected on the basis of overall G-factor, number of residues in core, allowed, generously allowed and disallowed regions. The selected model was further put to analysis by VERIFY 3D **[18]** and finally the protein was analysed with Accelrys DS 2.0.

Docking Studies

The structure of substrate β-(1,3)-glucan was obtained from NCBI PubChem and the docking was performed using C-DOCKER. The C-DOCKER protocol is an implementation of the C-DOCKER algorithm **[19]** in the DS environment which allows running a refinement docking of any number of ligands with a single protein receptor. C-DOCKER is a grid-based molecular docking method that employs CHARMm. The ligand was added with H (Hydrogen) atoms using DS. The models were energy minimized with CHARMm force field before performing docking. The receptor is held rigid while the ligand is allowed to flex during the refinement and the docking was performed using default settings. At the end of docking, the best conformation of the ligand was analyzed for its binding interactions.

Results & Discussion:

All the information about a protein's biological function cannot be ascertained by mere knowledge of its primary sequence or the secondary structure**.** It is therefore, essential to know its tertiary structure**.** Additionally, the 3D structure of ENGL1 has not been reported in RCSB PDB Data bank. BLASTp similarity search was performed against PDB data base but no significant results with complete query coverage were obtained. 3D model of ENGL1 protein (Q9UVV0) was predicted by homology modeling using MODELLER 9.10 with multiple templates. The obtained 3D structure was of poor quality with inappropriate folded conformations. Therefore, the automated 3D structure of ENGL1 from *Aspergillus fumigatus* was predicted based on the sequence-to-structure-to-function paradigm using I-TASSER **(Figure 1).** The stereochemical quality of ENGL1 structure was checked by PROCHECK **(Table 1)** and the Ramachandran plot for ENGL1 has been illustrated in **Figure 1**. Altogether more than 90% of the residues were found to be in favored and allowed regions, which validate the quality of homology model. The overall G-factor for ENGL1 was found to be -0.22 which is greater than the acceptable value -0.50, and therefore suggested

that the modeled structure is acceptable. The modeled structure was also validated by another structure verification server Verify 3D which assigned a 3D-1D score of >0.2 for the modeled protein **Table 1 (see supplementary material)**. This implies that the model is compatible with its sequence and the 3D structure predicted by I-TASSER proved to be of good quality with proper folded conformation. Multiloc predicted the protein to be of extracellular in nature and furthermore four asparagines residues were found to be N-glycosylated based on analysis of results obtained by NetNGlyc server. It has been suggested that glycosylation of asparagines residues is required for correct folding of protein before being exported. The physiochemical properties of protein were computed using Expasy ProtParam tools and are presented in **Table 2 (see supplementary material)**. The computed isoelectric point (pI value) of protein determined to be 5.09, which is less than 7 (pI<7), revealing the acidic nature of protein**.** This computed isoelectric point (pI) may be useful for developing buffer system for purification by isoelectric focusing method. Total predicted negative residues of ENGL1 protein are more in comparison to positive residues. These results also supported the acidic nature of the target protein. A high extinction coefficient value of protein was obtained, indicating the presence of high concentration of Trp and Tyr which helps in the quantitative study of protein-protein and protein-ligand interactions in solution. Further, stability of the protein was studied by analyzing the values for instability index, aliphatic index and Grand average of hydropathicity (GRAVY) index. The predicted Instability index of the target protein was 36.65 which reveals that ENGL1 protein is thermostable because Instability index of a protein smaller than 40 makes it stable while more than 40 considered as unstable**.** The aliphatic index (AI) which is defined as the relative volume of a protein occupied by aliphatic side chain is regarded as the positive factor for the increase of thermal stability of globular proteins. The high aliphatic index of protein infers that this protein may be stable for a wide range of temperature. The very low GRAVY index of protein infers that this protein could result in a better interaction with water. The results of CYS_REC indicated that the percentage of Cysteine (C) was found to be very low in the target protein resulting in absence of disulphide bond linkages **Table 3 (see supplementary material)**. It is assumed that in absence disulphide bonds the stability of this protein may be due to extensive hydrogen bonding. The secondary structure of target protein ENGL1 were predicted and analyzed by PSIPRED view and were shown in (**Figure 2).** This protein has mixed secondary structures i.e. alpha-helices, beta-strands and coils. The predicted structure of protein ENGL1 from *Aspergillus fumigatus* was successfully deposited in Protein Model Database (PMDB) **[20]**. This 3D structure may be further used in characterizing the protein experimentally. The predicted model was flexibly docked with ligand using C-DOCKER where out of ten poses produced, the best ligand pose was selected based on C-DOCKER top score and the target structure was chosen for further analysis. The ligand poses were analyzed and a heat map was produced to count H bonds made by the poses to the receptor molecule and count close contacts (van der Waals clashes) between the poses and receptor molecule. Interestingly no Van der Waals clashes were observed between the ligand and the modeled structure. Docking results indicated that amino acid residues (GLU 380, GLN 383, ASP 384, TYR 395, SER 712, and ARG 713) **Figure 3** in the protein play an important role in maintaining a functional conformation and directly involved in ligand binding. These residues interacted with substrate molecule by forming hydrogen bonds and all the hydrogen bond distances between ENGL1 and β-glucan complex were observed within the range of 2.76 Å to 3.18 Å **Table 4 (see supplementary material)**. The negative and low value of binding energy as well as maximum hydrogen bonding illustrated the occurrence of strong and most favorable binding between protein and ligand molecule.

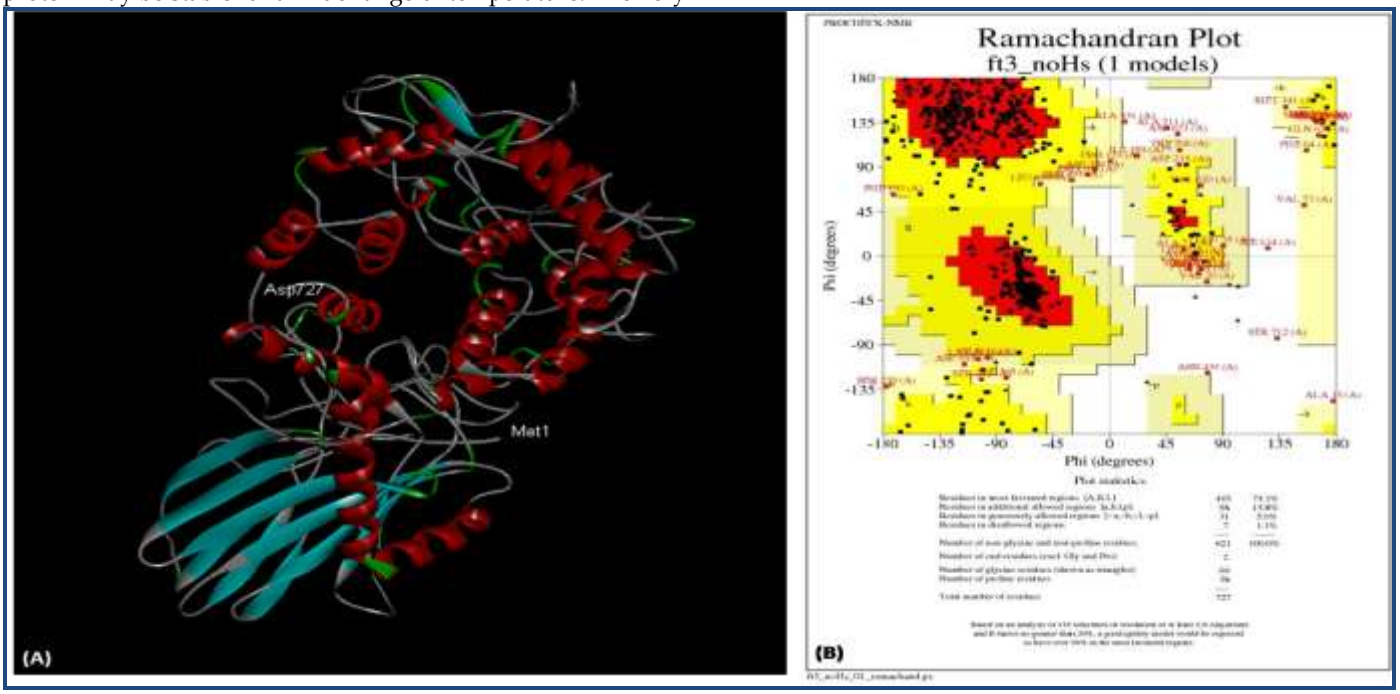

Figure 1: A) Homology model of ENGL1: Helix, Beta sheets and turns are in red, cyan and gray colour respectively; **B)** The Ramachandran plot of modeled structure validated by PROCHECK program.

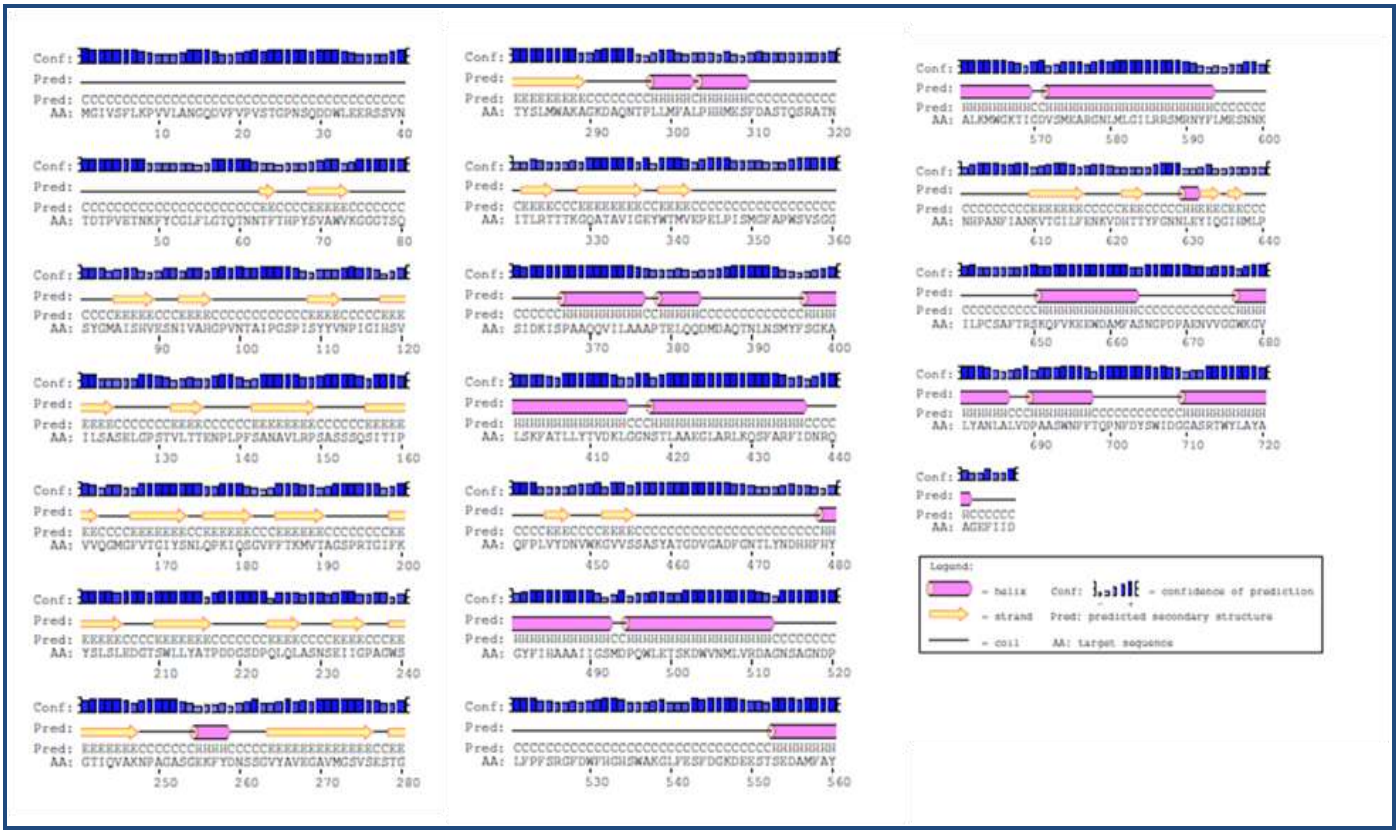

Figure 2: Secondary structure elements as predicted by PSIPRED

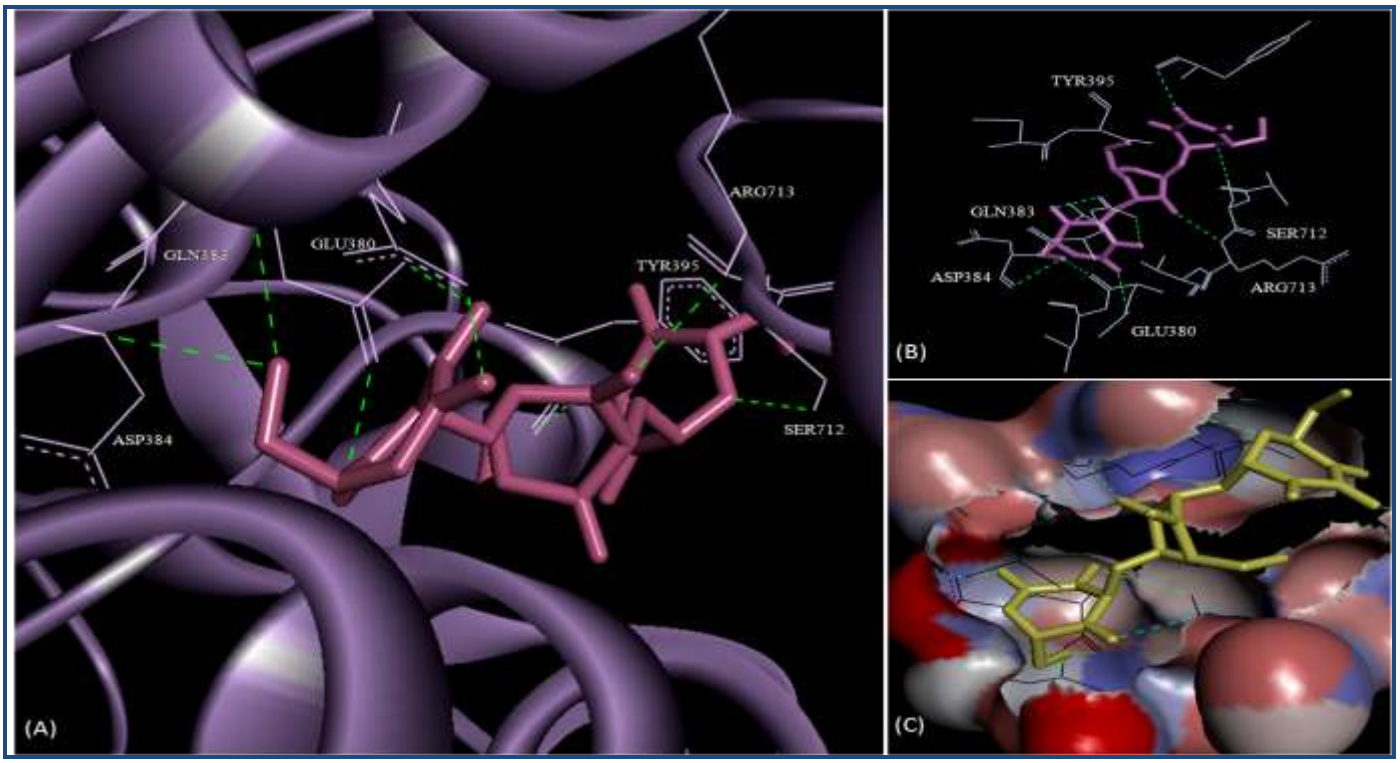

Figure 3: A) Docking of ligand in protein cleft: The ligand (pink) and interacting residues(purple) are shown in stick format and green dotted lines represent H-bond; **B)** Active site residues (purple) and their interaction with ligand(pink); **C)** Molecular surfaces: showing channel for substrate binding

Conclusion:

Precise evaluation and modeling of proteins is a major goals and key aspect of computational Biology. In the present study we have generated for the first time a good quality and reliable homology 3D-model of protein ENGL1 from *A. fumigatus*. Different parameters such as isoelectric point, molecular weight,

and total number of positive and negative residues, extinction coefficient, instability index, aliphatic index and grand average hydropathy (GRAVY) were computed for this protein in order to determine its physiochemical characteristics. This protein was found to be deficient in amino acid cystein that results in lack of disulphide linkages. In the absence of disulphide bond, extensive hydrogen bonding is believed to be responsible for stability of these proteins. Further Docking analysis with substrate identified important amino acids for catalytic function. The structure generated in the present study may serve as cornerstone for functional analysis of experimentally derived crystal structures and can be of great importance especially in unlocking the full potential of β-1,3- endoglucanase enzymes.

Acknowledgement:
Authors are thankful to Director CSIR-CDRI for providing facilities and Rizwan Ahmed is thankful to Indian Council of Medical Research for Senior Research Fellowship. CDRI Communication No. 8536

References:

[1] Schaffner A *et al. J Clin Invest*. 1982 **69**: 617 [PMID: 7037853]
[2] Gastebois A *et al. Future Microbiol*. 2009 **4**: 583 [PMID: 9492968]
[3] Beauvais A *et al. J Bacteriol*. 2001 **183**: 2273 [PMID: 11244067]
[4] Cabib E *et al. Annu Rev Biochem*. 1982 **51**: 763 [PMID: 7051965]
[5] Pitson SM *et al. Enzyme Microb Technol*. 1993 **15**:178[PMID: 7763458]
[6] Mouyna I *et al. Front Microbiol*. 2013 **17**: 81 [PMID: 23616783]
[7] Hartland RP *et al. J Biol Chem*. 1996 **271**: 26843 [PMID: 8900166]
[8] Martin K *et al. FEMS Microbiol Rev*. 2007 **31**:168 [PMID: 17313520]
[9] Kumar A *et al. J Proteomics*. 2011 **74**: 1104 [PMID: 21536156]
[10] Altschul SF *et al. J Mol Biol*. 1990 **215**: 403 [PMID: 2231712]
[11] Hoeglund A *et al. Bioinformatics* 2006 **22:** 1154 [PMID: 16428265]
[12] Kyte J & Doolittle RF, *J Mol Biol*. 1982 **157**: 105 [PMID: 7108955]
[13] Wilkins MR *et al*. *Methods Mol Biol*. 1999 **112**: 531 [PMID: 10027275]
[14] Sali A & Blundell TL, *J Mol Biol*. 1993 **234**: 779 [PMID: 8254673]
[15] Zhang Y *BMC Bioinform*. 2008 **9**: 40 [PMID: 18215316]
[16] Laskowski RA *et al. J Biomol NMR*. 1996 **8:** 477 [PMID: 9008363]
[17] Ramachandran GN *et al. J Mol Biol*. 1963 **7:** 95 [PMID: 13990617]
[18] Eisenberg DR *et al. Methods Enzymol*. 1997 **277**: 396 [PMID: 9379925]
[19] G Wu *et al. J Comp* Chem. 2003 **24**: 1549 [PMID: 12925999]
[20] Castrignanò T *et al. Nucleic Acids Res*. 2006 **34**: 306 [PMID: 16381873]

Edited by P Kangueane
Citation: **Ahmed *et al.*** Bioinformation 9(16): 802-807 (2013)

BIOINFORMATION

open access

Supplementary material:

Table 1: Evaluation results of the I-TASSER models of the tertiary structure by PROCHECK and VERIFY 3D.

	Model 1	Model 2	Model 3	Model 4	Model 5
Core Region	81.6%	75.2%	78.1%	78.9%	71.8%
Allowed Region	12.4%	17.2%	15.8%	15.5%	18.2%
Generous Region	4.5%	5.5%	5.0%	3.4%	6.6%
Disallowed Region	1.4%	2.1%	1.1%	2.3%	3.4%
VERIFY 3D	54.40%	70.88%	72.39%	56.04%	61.40%

Table 2: properties of ENGL1 predicted by ProtParam program

S.No.	Parameters	β-(1-3)-endoglucanase
1.	Source name	Aspergillus fumigatus
2.	Accession No.	Q9UVV0
3.	Sequence length	727 aa
4.	Molecular weight	78927.5
5.	Theoretical pI	5.09
6.	R*	66
7.	+R*	44
8.	Instability index	36.65
9.	Aliphatic index	75.03
10.	GRAVY	- 0.128

* –R: total number of negative residues. +R: total number of positive residues

Table 3: Presence of disulphide (ss) bond as predicted by Cys_Rec.

Accession number	CYS_REC	Score
Q9UVV0	Cys_52 (probably not SS-bounded)	-0.2
	Cys_644 (not SS-bounded)	-30.2

Table 4: Interface interaction (hydrogen bonding) results for docking of Model of ENGL1 and β-(1, 3)-glucan**.**

	ENGL1	Distance (Å)	β-(1,3)-glucan
(1)	GLU 380 (O)	2.76	O20
(2)	GLU 380 (OE1)	2.93	O21
(3)	GLN 383 (OE1)	3.17	O19
(4)	GLN 383 (NE2	3.15	O17
(5)	ASP 384 (O)	3.18	O20
(6)	TYR 395 (O)	2.94	O24
(7)	SER 712 (OG)	3.06	O26
(8)	ARG 713 (N)	3.02	O7