



Received on 17 July, 2014; received in revised form, 26 September, 2014; accepted, 16 November, 2014; published 01 March, 2015

## CELLULOSE AS SOLE SUBSTRATE FOR BIO ETHANOL PRODUCTION: DEVELOPMENT OF STRUCTURAL MOLECULAR MODEL OF ENDOGLUCANASE III FROM *TRICHODERMA REESEI* USING BIOINFORMATIC TOOLS

B. Chandrasekhar and B. Bhima\*

Department of Microbiology, University College of Science, Osmania University, Hyderabad, India

### Keywords:

Bio Ethanol, Docking, Endoglucanase III, Homology modelling, *Trichoderma reesei*

### Correspondence to Author:

**Dr. B. Bhima**

Assistant Professor,  
Department of Microbiology,  
UCS, Osmania University,  
Hyderabad -500007, India.


**E-mail:** bhima.ou@gmail.com

**ABSTRACT:** One of our greatest challenges in improving nation's economy is to reduce the dependence on imported petroleum products. To accomplish this, we need a variety of alternative fuels, which includes ethanol produced from cellulosic materials. Cellulose is a rich natural biopolymer on earth and most dominating agricultural waste. Apart from the fossil fuels, bioethanol is used as chief renewable energy source found on earth. Hence, we developed the structural molecular model of the endoglucanase III from *Trichoderma reesei* by homology modelling using Modeller 9v12. Based on sequence similarity with protein of known structure (PDB ID: 1H1N) was chosen from PDB by using protein BLAST. The models were evaluated by PROCHECK. Resulted Ramachandran plot analysis showed that conformations for 99% of amino acid residues are within the most favored regions. A possible active site was identified in the model using Sybyl6.7. The predicted model permits initial inferences about the unexplored 3D structure of the endoglucanase III and may promote in relational designing of molecules for structure-function studies.

**INTRODUCTION:** Biofuels provide a potential route for avoiding the global political instability and environmental issues that arise from reliance on petroleum. At present, most of the biofuel is in the form of ethanol generated from sugar or starch, but this can meet only limited areas of global fuel requirements<sup>1</sup>. To reduce the energy cost, low-acid/temperature pretreatments<sup>2</sup> were attempted, but it is still necessary to add large amounts of cellulolytic enzymes to hydrolyze the cellulose material. In this current study Endoglucanases were engineered for broad specificity, pH, temperature and activity.

Endoglucanases cut at random at internal amorphous sites in the cellulose polysaccharide chain, that generates oligosaccharides of various lengths and consequently new chain ends. As endoglucanase (EG) is a key component of cellulase which it degrades the cellulose material, a polymer to be smaller monomers and have been used in various industries like textiles, fruits and vegetable processing, and recently in bioethanol production<sup>3</sup>.

Due to these essential applications of endoglucanase, and remaining cellulases various protein engineering approaches have been employed to enhance the functional properties of these enzymes to achieve better performance<sup>4</sup>. As ethanol production is cheaper than liquid petroleum, the production of ethanol depends on simultaneous saccharification of cellulose to glucose and fermentation of glucose to ethanol by a

<b>QUICK RESPONSE CODE</b> 	<b>DOI:</b> 10.13040/IJPSR.0975-8232.6(3).1132-36
	<b>Article can be accessed online on:</b> <a href="http://www.ijpsr.com">www.ijpsr.com</a>
<b>DOI link:</b> <a href="http://dx.doi.org/10.13040/IJPSR.0975-8232.6(3).1132-36">http://dx.doi.org/10.13040/IJPSR.0975-8232.6(3).1132-36</a>	

recombinant yeast strain as a microbial biocatalyst. Recombinant yeast strain expressing endoglucanase and  $\beta$ -glucosidase was able to produce ethanol from  $\beta$ -glucan, CMC and acid swollen cellulose<sup>5</sup>.

Yeast cell-surface is a technique used in protein engineering that enables  $10^4$ – $10^5$  enzymes to be displayed on the surface of a yeast cell<sup>6</sup>. The surface displayed yeast can be used as a whole-cell biocatalyst without requiring enzyme-separation and purification processes, and it can immediately in take glucose degraded from cellulose and thus, it is hardly contaminated by other organisms. Furthermore, direct fermentation to ethanol from cellulose could be conducted by using cellulase surface displaying yeast<sup>7</sup>. This reduces the chance of contamination and high efficiency of product yield with low economic value of bioethanol production<sup>8</sup>.

## MATERIALS AND METHODS:

### Query Retrieval and Template recognition:

The protein sequence of endoglucanase of “*Trichoderma reesei*” was retrieved from NCBI database (National Center for Biotechnology Information) which have 418 amino acid residues, accession number (Genbank id) is AAA34213.1 for modeling. Homology modeling usually starts by searching the PDB (Protein Data Bank) of known protein structure using a query sequence. This search is generally performed by comparing the target sequence with the sequence of each of the structure in the database. The protein sequence of 1H1N (template) worked as a target sequence. For identification of appropriate template BLASTp was performed against Protein Data Bank, select templates for generating the putative 3D structures.

### Homology Modeling:

The method of comparative modeling requires identification of homologous sequences with existing structure. The first step of comparative/homology modeling, scanning and selection of template protein structure was carried out using the target sequence as query<sup>9</sup>. A sequence similarity search by BLAST program was carried out with query sequence against proteins whose solved structures had been deposited in the Protein Data Bank (PDB) to find related proteins as templates. BLASTp generally used for template

selection procedure [http://blast.ncbi.nlm.nih.gov/Blast]<sup>10</sup>. The BLAST analysis outputted potential template for modeling is 1H1N, the high degree of primary sequence identity (99%) between target and endoglucanase III indicates that this crystallographic structure was good model to be used as template for endoglucanase III enzyme.

The alignment of the endoglucanase III target and 1H1N is shown in (Figure 1). A total of 20 models were generated using program MODELLER 9.11. MODELLER implements comparative protein modeling by satisfaction of spatial restraints. Comparative initial alignment was done by using clustalX.<sup>11</sup>

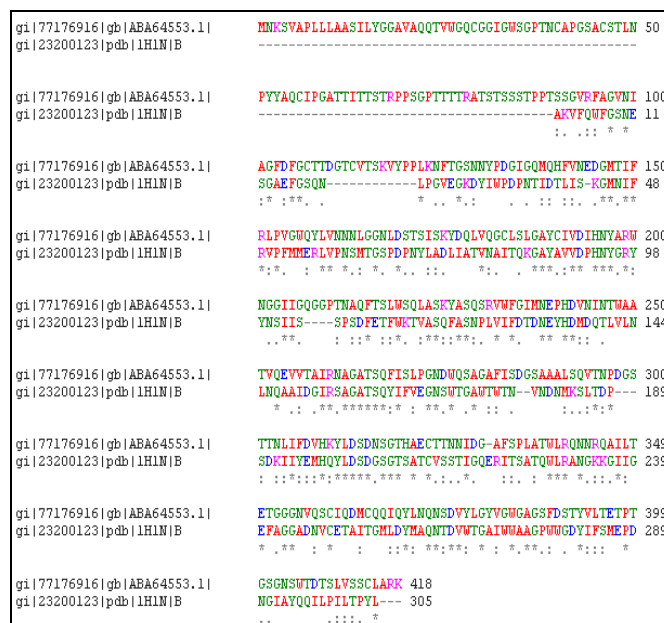


FIGURE 1: ALIGNMENT OF QUERY SEQUENCE (AAA34213.1) AGAINST THE TEMPLATE PDB (PDB ID: 1H1N)

### Model analysis:

Modeller generated the 20 models for the same target, out of which the best model can be selected for further analysis. We evaluated the model with the lowest value of the Modeller objective function and PROCHECK results. The stereochemical quality of the developed models for endoglucanase III enzyme using Ramachandran's plot calculation computed with PROCHECK<sup>12</sup>.

### Structure Visualization and structural analysis

Visualization and Structural analysis of final protein model were carried out with Swiss PDB

Viewer (spdbv)<sup>13</sup> software. Structural alignment and Root Mean Square Deviation (RMSD) calculation with several models and template were carried out with spdbv.

#### Active site analysis:

The SITE ID module of Sybyl 6.7<sup>13</sup> is a suite of programs for identifying and characterizing binding sites and functional residues from protein structure. In this study, the active site search program was used to search the protein binding sites by locating binding pockets studied structure. Through comparing the conserved residues in family of the studied protein and combining the search results and other experiment results, the binding sites of Endoglucanase III were predicted. Those results were used to guide the following docking experiment.

#### Molecular Docking Studies:

FlexX automated docking software used for docking simulation<sup>14</sup> installed on Silicon Graphics Inc octane2 workstation using the package SYBYL 6.7<sup>15</sup> to investigate the binding mode between the inhibitors and Endoglucanase III receptor. The flexible docking program uses incremental construction algorithm, which employs the preparation of receptor by adding hydrogen's and assigning Gasteiger Huckel charges, and saved it in PDB file. The ligand cellulose was minimized by assigning Gasteiger charges and non polar hydrogen's. The incremental construction algorithm is known to be the most efficient and reliable method of FlexX.

### RESULTS AND DISCUSSION:

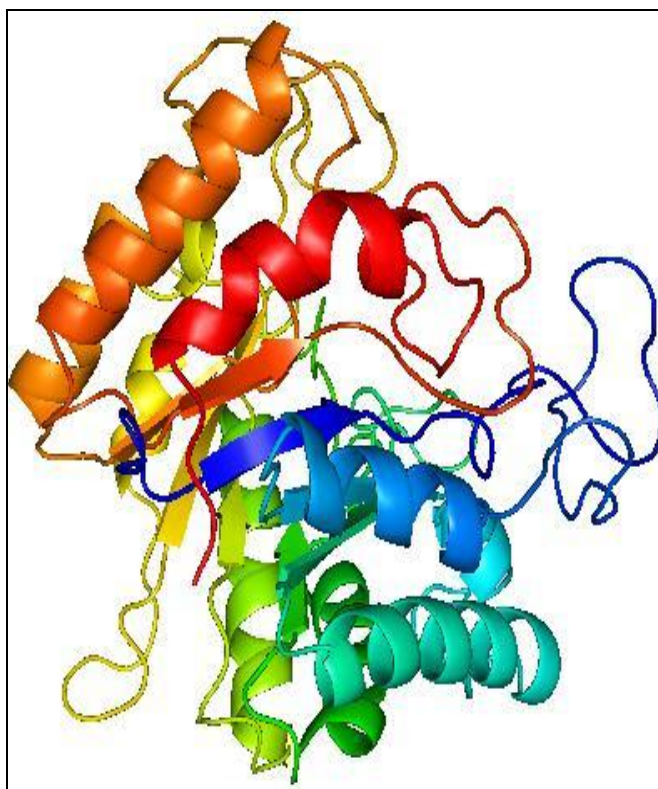
#### Model building, refinement and evaluation:

In the present study, the primary sequence identity between endoglucanase (target) and *Trichoderma reesei* indicates that ~90% sequence identity. Medium accuracy model, obtained with a template-target sequence identity of 50-60%. The sequence analysis result based on BLASTp, a number of models had been generated for each binary complex and the final model selected based on MODELLER objective function.

In the current research work, experimental structure of endoglucanase of *Trichoderma reesai* is not available in Protein Data Bank (PDB), homology

modeling approach was used in order to derive their structure. BLAST search results had shown higher similarity with crystal structure of IHIN, while the template was selected on the basis of higher sequence identity. It shows 99.0% sequence identity, with maximum conserved residues. Comparative modeling predicts the 3-D structure of endoglucanase III model of given protein sequence (target), based primarily on this alignment to the template. The developed model was also checked for phi and psi torsion angles using Procheck for Ramachandran plots.

All together 98% of the residues were found to be in core and allowed regions, which validate the quality of predicted homology models. Visualization and analysis of the model using Swiss PDB Viewer (spdbv) reveals that there are no steric hindrances between the residue and thus modeled structure are stable. Target structure-template structure superimposition was done in order to calculate Root Mean Square Deviation (RMSD) between the target and template sequence. This implies good quality of the modeled structure (**Figure 2**) and ramachandran plot (**Figure 3**) of Endoglucanase III.



**FIGURE 2: MODELLED PDB STRUCTURE OF THE ENDOGLUCANASE III**



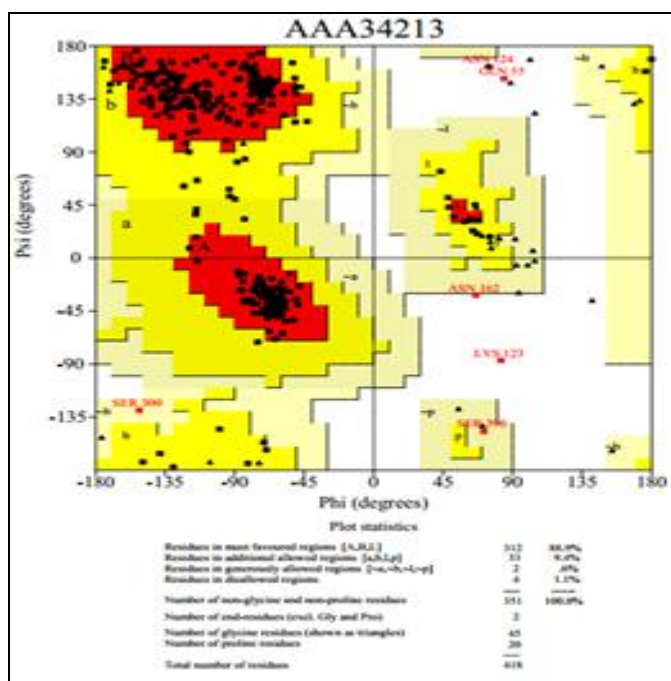


FIGURE 3: RAMACHANDRAN PLOT FOR MODELLED ENDOGLUCANASE III.

#### Molecular Docking Studies:

Modeled endoglucanase III enzyme were docked against cellulose (Figure 4). The ligand interacts with seven amino acids and showing around 20 interactions. The low binding energy values of endoglucanase enzyme indicate that the ligands are in most favorable region of the protein and that the enzyme has more affinity towards the ligand cellulose. Gly92, Asn99, Gly235, Gly379, Arg417 and Lys418. Among all these seven interactions Arg417 shows highest number of interactions (nine interactions), Gly92 shows 5 interactions and rest of the amino acids shows 2-3 interactions. All the interactions are shown in Figure 4.

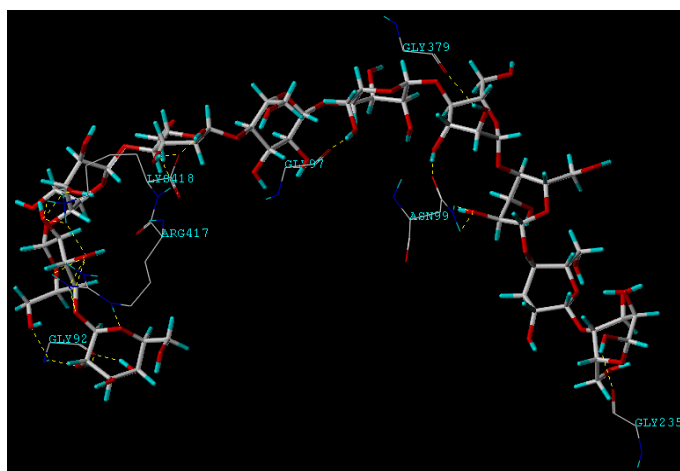


FIGURE 4: CELLULOSE INTERACTIONS WITH ENDOGLUCANASE III.

**CONCLUSION:** In the present study, As Endoglucanase is the key enzyme that degrades polymer cellulose to trimer and tetramers in which combination of two enzymes (cellobiohydrolase &  $\beta$ -glucosidase) convert it to bioethanol. The 3D model of Endoglucanase III has been constructed from *Trichoderma reesei*, using the Homology modelling approach. Molecular docking studies of cellulose with Endoglucanase III exhibited binding interactions and warrants further studies needed for the development of potent endoglucanase III inhibitor. This data provides molecular insights for the consideration of cellulose as potential candidate against Endoglucanase III target in *Trichoderma reesei*.

#### REFERENCES:

1. Sticklen MB: Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol. Nature Reviews Genetics. 2008; 9: 433-443.
2. Pasha C, Kuhad RC, Rao LV: Strain improvement of thermotolerant *Saccharomyces cerevisiae* VS3 strain for better utilization of lignocellulosic substrates. Journal of Applied Microbiology. 2007; 103: 1480-1489.
3. Aehle W: Enzymes in industry. Leiden: Wiley-VCH Verlag GmbH and Co. 2004; pp: 155-244.
4. Noorbacha IA, Waesoho S, and Salleh HM: Structural and dynamics behavior of native endoglucanase from fusarium oxysporum. Aust. J. Basic & Appl. Sci. 2012; 6: 89-92.
5. Jeon E, Hyeon JE, Suh DJ, Suh YW, Kim SW, Song KH, Han SO: Production of Cellulosic Ethanol in *Saccharomyces cerevisiae* Heterologous Expressing *Clostridium thermocellum* Endoglucanase and *Saccharomycopsis fibuligera*  $\beta$ -glucosidase Genes. Mol. Cells. 2009; 28: 369-373.
6. Kuroda K, Ueda M: Cell surface engineering of yeast for applications in white biotechnology. Biotechnology Letters. 2011; 33:1-9.
7. Nakatani Y, Yamada R, Ogino C, Kondo A: Synergetic effect of yeast cell-surface expression of cellulase and expansin-like protein on direct ethanol production from cellulose. Microbial Cell Factories. 2013; 12: 66.
8. Nakanishi A, Bae J, and Kuroda K, Ueda M: Construction of a novel selection system for endoglucanases exhibiting carbohydrate-binding modules optimized for biomass using yeast cell-surface engineering. AMB Express. 2012; 2: 56
9. <http://blast.ncbi.nlm.nih.gov/Blast>
10. Mulukala NSK, Gunda SK, Shaik M: Comparative modeling of Rab6 proteins: identification of key residues and their interactions with guanine nucleotides. J Mol Model. 2013; 19: 1891-900.
11. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG: Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23: 2947-2948.
12. Laskowski RA: PROCHECK: A program to check the stereochemical quality of protein structures. J.Appl. Cryst. 1993; 26: 283-291.
13. <http://www.expasy.org/spdbv/>

14. Gunda SK, Narasimha SK, Shaik M: P56 (Ick) kinase inhibitor studies: a 3D QSAR approach towards designing new drugs from flavonoid derivatives. Int J Comput Biol Drug Des. 2014; 7: 278-294.

15. Tripos, Sybyl 6.7, Tripos Inc., 1699 S. Hanley Rd., St. Louis, MO 63144

**How to cite this article:**

Chandrasekhar B and Bhima B: Cellulose as Sole Substrate for Bio Ethanol Production: Development of Structural Molecular Model of Endoglucanase III from *Trichoderma Reesei* Using Bioinformatic Tools. Int J Pharm Sci Res 2015; 6(3): 1132-36.doi: 10.13040/IJPSR.0975-8232.6(3).1132-36.

All © 2013 are reserved by International Journal of Pharmaceutical Sciences and Research. This Journal licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License.

This article can be downloaded to **ANDROID OS** based mobile. Scan QR Code using Code/Bar Scanner from your mobile. (Scanners are available on Google Playstore)