JOURNAL OF PURE AND APPLIED MICROBIOLOGY, May 2014. Vol. **8**(Spl. Edn. 1), p. 791-796

Molecular Docking of Cellobiose and Cellotetraose into An Endoglucanase from *Fusarium oxysporum*

Ibrahim Ali Noorbatcha*, Abdul Aziz Ahmad and Hamzah Mohd Salleh

BioProcess and Molecular Engineering Research Unit (BPMERU) Department of Biotechnology Engineering, Kulliyyah of Engineering, International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia.

(Received: 08 January 2014; accepted: 24 March 2014)

Endoglucanase is one of the three enzymes required to synergistically hydrolyze cellulose to sugar monomers which can be fermented into ethanol. Hence efforts to improve the efficiency of this enzyme, using experimental and computational approaches have been continuing vigorously. An endoglucanase, FOegI, which is same as the well-characterized endoglucanase I from *Fusarium oxysporum* (PDB ID: 3OVW) has been cloned into an appropriate expression vector, except for two mutations. As an attempt to improve the substrate binding characteristics of FOegI, computational molecular docking studies to understand the interactions in the active site regions and to suggest mutations to improve the binding characteristics was carried out in this study. The docking of ligands into active site was carried out using Lamarkian Genetic Algorithm with a fast, simplified potential of mean force (PMF) to evaluate the docking efficiency. The two substrate used in this study are cellobiose and cellotetraose. The active site is confirmed by comparing the docked structure with the available experimental structures. Cellotetraose was found to have stronger binding than cellobiose. Among the three in-silico mutants tested, T224E is found to have the strongest binding energy of -530.72 ±0.85with cellotetraose.

Key words: *Fusarium oxysporum*, Endoglucanase I, Molecular Docking.

One of the major fields of research in response to the quest towards preserving the environment is the research and development in biofuel as a clean energy substitutes for fossil fuel. In the United States alone, 180 million tons of lignocellulosic biomass is produced per year. This biomass is renewable, cheap and readily accessible. At global level, 10-50 billion tons are produced annually. Rice straw which accounts for half of the biomass produced worldwide is burned to waste harming environment and causing health issues[1,2]. Endoglucanase, exoglucanase and *â*-glucosidase are required to fully break lignocellulosic biomass to monosaccharides and these in turn can be the feedstock for ethanol production. Endoglucanase is also used in the textile industry in the bio-stoning process. Bio-stoning is an enzymatic process in which cellulase enzymes is widely used for denim garment treatment as an alternative to the pumice stone-washing process in order to achieve an aged look for denim fabrics[3].

Endoglucanase I (EGI) catalyzes hydrolysis of the β-1,4-glycosidic linkages of cellulose with a net retention of the anomeric configuration at C-1. This occursviaa double displacement reaction as described by Koshland[4].

* To whom all correspondence should be addressed.
Tel.: +603 61965453; Fax: +603 61964442;
E-mail: ibrahiman@iium.edu.my

Computational simulation is widely used to study the various aspect of structural behavior of biomolecules[5]. In general, molecular docking is done to achieve an optimized conformation for the protein and ligand such that the free energy of the overall system is minimized. Molecular docking can be a powerful computational tool, which will reduce the labor and cost if used prior to experimental screening[6]. The field of molecular docking is very popular especially in the past five years based on the number of publications[7].

In this study, binding strength of *Fusariumoxysporum*Endoglucanase I (FOegI) cloned in our laboratory, towards two ligands i.e. cellobiose and cellotetraose is estimated using computer modeling. In addition, the active site of the FOegI is predicted computationally. Besides that, binding strength of three *in-silico* mutants were studied here based on our group's work on thermostabilityof EGI.

MATERIALS AND METHODS

Structure

The crystal structure of FOegI is not available in the protein data bank. However the crystal structure of endoglucanase Ifrom *Fusariumoxysporum*acquired from the RCSB Protein Data Bank under PDB code 3OVW[8] have 99.5% sequence identity. It is a native structure without any ligand complexes with it. The actual structure has no mutation, however *in-silico* mutation introduced at position R41H and S216N as in Figure 1 to construct the same sequence with the available FOegI in our lab as this simulation results will be validated experimentally by comparing with the available experimental results. The structure has six 3_{10}-helixes, six alpha helixes, 35 beta sheets, 39 coils and 30 turns. The initial structure for docking was prepared according the procedure described by Noorbatcha*et al*. (2012).

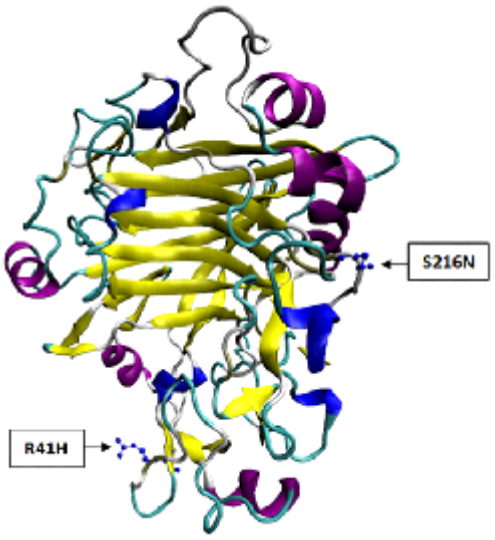


Fig. 1. Cartoon representation of FOegI coloured based on the residue names

Docking studies

Scigress version 3.0 (Fujitsu Limited) software used in this study. Genetic algorithm with maximum generation of 30,000 and local iteration of 300 was used. Ligand and active site was set to be both flexible. The active site sampling box size of 40Åx40Åx40Å with grid spacing of 0.375 was utilized and other parameters set as in the default setting. Such a box size selected to cover the major portion of the enzyme so that not only the active site but also the mutated residues are included in all the stages of energy minimization and the sampling for the docking.

As bond formation would cause changes in the atom types and possible substantial changes in the shape of the ligand and active sites, it is assumed that the binding is non-covalent. The potential of mean force (PMF) scoring, which is a knowledge-based approach that extracts pairwise atomic potentials from the structure information obtained from protein-ligand complexes contained in the Protein Data Bank, is used to calculate the strength of the binding of Cellobiose and cellotetraose (Figure 2) as ligands. The docking process is repeated at least three times and the average score is reported here.

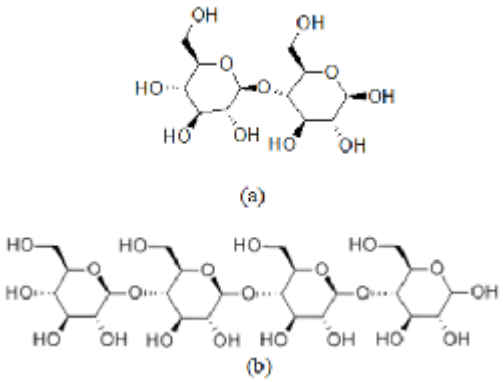


Fig. 2. 2D structure of cellobiose and cellotetraose

Mutation studies

The binding energy of three *in-silico* mutants was investigated. Based on molecular dynamics simulations study, three mutants were prepared via *in-silico* method. The mutants are T224E, G229A and T224E_G229A_S230F_S231E_N321R. These sites are selected asmolecular dynamics study suggests that the secondary structure conformation at the residues position 225 to 231 of EGFO changes significantly at higher temperature where as conformation of endoglucanase from humicolainsolens at these positions is maintained as the temperature is increased.

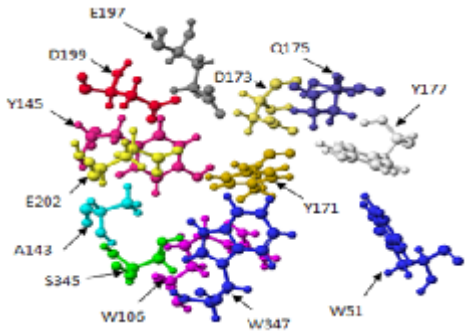


Fig. 3. Location of the active site within the range of 3Å from cellobiose after docking

RESULTS AND DISCUSSION

Active Site Determination

The active site is the collection of residues that involve dominantly in the catalytic reaction of the enzyme. In order to locate the active site of FOegIcomputationally, the residues which is in 3Å radius from the ligand in the crystal structure were selected and grouped as the active site as in Figure 3.

Fig. 4 shows the active site residues determined experimentally. Eight out of 9 residues were accurately predicted as active site residues.

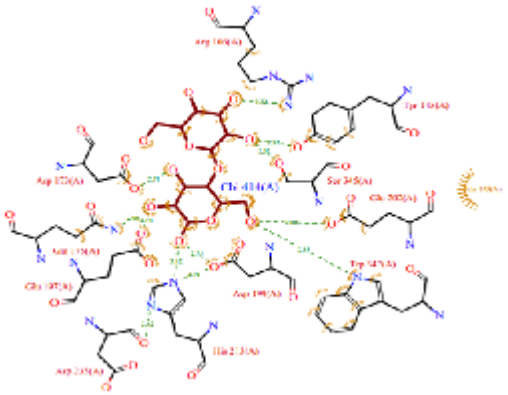
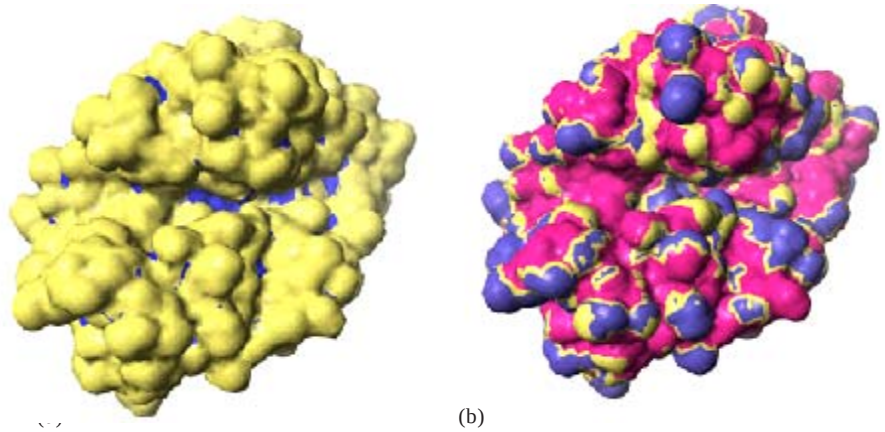


Fig. 4. Ligand interaction diagram of active site of EGfoIwith cellobiose

However, residue HIS213 could have been included if bigger than 3Å radius selected. It is interesting to note that HIS213 is not forming hydrogen bonds directly with the substrate. Sulzenbacher*et al.*[8] reported that the GLU197 is a catalytic nucleophile and GLU202 is the catalytic acid/base.

Fig.5a shows the crevice surface analysis from which the potential active site of the protein can be guessed. The surface that is away from the



(b)

Fig. 5. Surface analysis: (a) crevice surface - blue region indicating potential active site. (b) Accessible surface - purple-red and blue areas are hydrophilic while the cream area is hydrophobic regions

smooth outer surface is coloured in blue indicating potential active site. The hydrophobic residues are on the interior, hydrophilic groups are on the exterior and the hydrophilic regions dominate the outside of the protein as displayed in the figure 5b.

Docking Studies of FOegI

The docking scores of cellobiose and cellotetraose are given in Table 1. Cellotetraose showed stronger binding compared to cellobiose. The reasons for the stronger interactions and the role of different mutations on the binding characteristics are mainly due to hydrogen bonds. The three mutants with the mutated residues highlighted shown in figure 6(a)-(c). The mutant T224E is in the coil, N321R is in the alpha helix, G229A, S230F and S231E are in the turn of the structure.

Table 1. Docking Scores of FOegI

Ligand	Docking Score (kcal/mole)
Cellobiose	-438.17±3.8
Cellotetraose	-422.95±0.49

Table 2. Docking Scores of mutant FOegI

Mutant	Docking Score (kcal/mole)	
	Cellobiose	Cellotetraose
T224E	-424.49 + 0.68	-530.72 + 0.85
G229A	-338.35 + 1.30	-392.99 + 1.10
T224E_G229A_ S230F_ S231E_ N321R	-340.55 + 0.95	-430.12 + 0.17

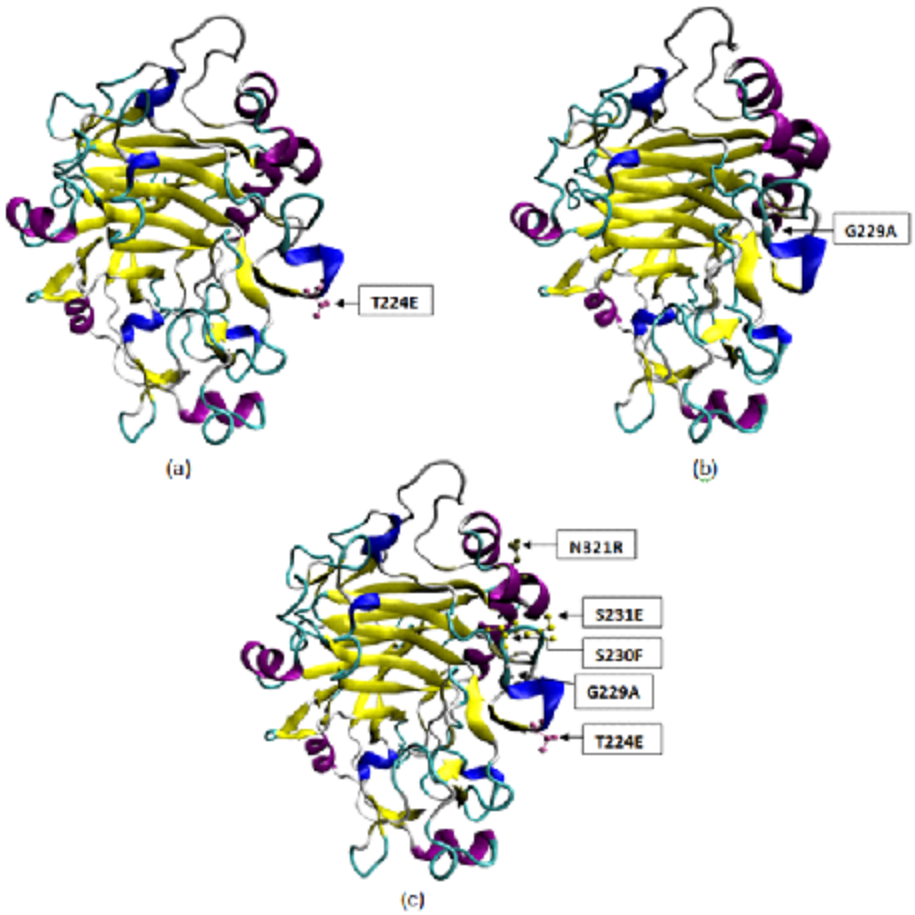


Fig. 6. The location of the three mutants (a) T224E (b) G229A (c) T224E_G229A_S230F_S231E_N321R

In general, docking of cellotetraose to all the three mutants gives lower value compared to cellobiose as ligand indicating stronger binding between mutant FOegl and cellotetraose as shown in Table 2. Mutant T224E having the highest binding energy among all the mutant with cellotetraose. Hydrogen bonds plays vital role in increasing the binding energy.As far as H-bond is concerned, some reports say that the H-bond distance can be greater than 3.2 Å[9,10]. Nevertheless, in this study H-bonds distances less than 2.5Åis considered to be pertinent as the relatively longer H-bonds in distance are too weak to argue for the purpose of strong binding. The protein forms five hydrogen bonds (less than 2.5Å) with cellotetraose and 3 hydrogen bonds with cellobiose as shown in figure 7 and 8 respectively. The observed hydrogen bond is contributing to the difference in the binding energy between cellobiose and cellotetraose.

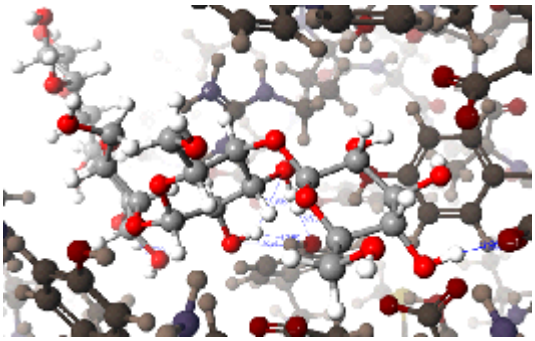


Fig. 7. EGfoIMutant T224E showing 5 hydrogen bonds between cellotetraose and the protein

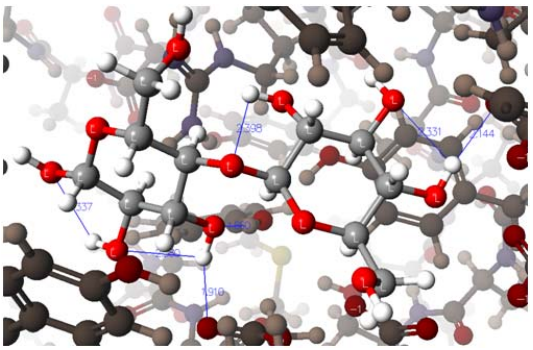


Fig. 8. EGfoIMutant T224E showing 3 hydrogen bonds between cellobiose and the protein

CONCLUSIONS

Cellobiose and cellotetraose were successfully docked into FOegI. Binding strength of cellobiose and cellotetraose in the active site of FOegI was evaluated. Cellotetraose was found to bind more strongly compared to cellobiose.Among all the mutant studies, Mutant T224E is found to have stronger binding with cellotetraose. The difference in binding strength is ascribed to the increase in the number of hydrogen bonds.

ACKNOWLEDGEMENTS

This work is partially supported by Ministry of Higher Education (MOHE) Malaysia and International Islamic University Malaysia (IIUM) under Academic Training Scheme.

REFERENCES

1. Kayaba, H. Activation of eosinophils by rice-husk dust exposure: a possible mechanism for the aggravation of asthma during rice harvest. *Tohoku J. Exp. Med.*2004; **204**: 27–36.
2. Kim, S. & Dale, B. E. Global potential bioethanol production from wasted crops and crop residues. *Biomass Bioenergy*, 2004*;* **26**: 361–375.
3. Arja, M.-O. Cellulases in the Textile Industryin*Industrial Enzymes - Structure, Function and Applications*, Eds. Polaina, J. &MacCabe, A. P., 51–63 (Springer Netherlands, 2007).
4. Koshland, D. E. Stereochemistry and the Mechanism of Enzymatic Reactions. *Biol. Rev.*1953; **28**: 416–436.
5. Noorbatcha, I. A., Waesoho, S. & Salleh, H. M. Structural and dynamics behavior of native endoglucanase from *Fusariumoxysporum*. *Aust. J. Basic Appl. Sci.* 2012; **6**: 89–92.
6. Ali, H. I., Fujita, T., Akaho, E. &Nagamatsu, T. A comparative study of AutoDock and PMF scoring performances, and SAR of 2-substituted pyrazolotriazolopyrimidines and 4-substituted pyrazolopyrimidines as potent xanthine oxidase inhibitors.*J. Comput. Aided Mol. Des.*2009; **24**: 57–75.
7. Yuriev, E. &Ramsland, P. A. Latest developments in molecular docking: 2010-2011 in review. *J. Mol. Recognit.*2013; **26**: 215–239.
8. Sulzenbacher, G., Schülein, M. & Davies, G. J. Structure of the endoglucanase I from Fusariumoxysporum: native, cellobiose, and 3,4-

epoxybutyl beta-D-cellobioside-inhibited forms, at 2.3 A resolution. *Biochemistry (Mosc.)*1997; **36**: 5902–5911.

9. Morris, G. M., Goodsell D.S., Halliday R.S., Huey,R., Hart, W.E., Belew, R.K, & Olson, A.J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.*1998; **19**: 1639–1662.
10. Okamoto, K.,Matsumoto, K., Hille R, Eger, B.T., Pai, E.F., Nishino, T. The crystal structure of xanthine oxidoreductase during catalysis: Implications for reaction mechanism and enzyme inhibition. *Proc. Natl. Acad. Sci. U. S. A.*2004; **101**: 7931-7936.