

Binding Mode Study of β -glucosidase B from *P. Polymyxa* with Cellobiose and Laminaribiose

Nur Shima Fadhilah Mazlan and Nurul Bahiyah Ahmad Khairudin

Abstract—Beta-glucosidase (3.2.1.21) plays an essential role in the removal of non-reducing terminal glucosyl residues from sacharides and glycosides. Recently, beta-glucosidase has been of interest for biomass conversion that acts in synergy with two other enzymes, endo-glucanase and exo-glucanase. However, there is not much information regarding the molecular interactions of beta-glucosidase with cellobiose. Thus, this study reports on the binding modes between beta-glucosidase from glycoside hydrolase family 1 namely BglB with cellobiose and laminaribiose via molecular docking method. Further analysis on the hydrophobic interactions revealed the key residues involved in forming hydrogen bonds (h-bond) with the substrates. The important residues were reported to be Gln22, Glu167, Tyr298, Glu356, Glu402, and Trp410. These findings may provide valuable insights in designing beta-glucosidase with higher cellulose-hydrolyzing efficiency.

Index Terms— β -glucosidase, molecular docking, enzymatic hydrolysis

I. INTRODUCTION

Cellulose is a highly unbranched polymer consisting of glucose residues linked together by β -1, 4-glycosidic bonds with cellobiose as its structural unit. The glucose chains in cellulose are very tightly packed to form insoluble crystallite which is impenetrable to any molecules including water. The structure of lignocelluloses is very complex where cellulose forms a skeleton which surrounded by hemicelluloses and lignin like matrix compose.

Cellulose is one of the most abundant polysaccharide compound in nature and is thought to be a promisingly renewable biomass resource for alternative fuels [1]. The chemical formula of cellulose is $(C_6H_{10}O_5)_n$ and the structure of one chain of the polymer is shown in Fig. 1.

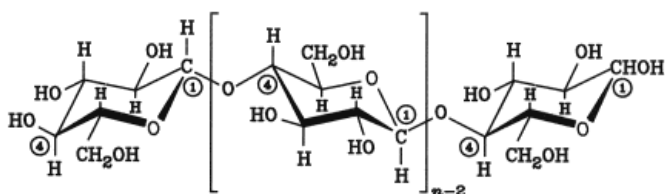


Fig. 1. Chemical structure of cellulose unit

The cellobiose unit in cellulose structure can be further degraded into glucose monomer unit through enzymatic hydrolysis which is performed by cellulases [2]. There are

three enzymes in cellulase system that work in synergy to hydrolyze cellulose into glucose monomer which are endocellulase (endo-1, 4- β -glucanohydrolase, EC 3.2.1.4), exocellulase (1, 4- β -D-glucan cellobiohydrolase, EC 3.2.1.91) and β -glucosidase (β -1, 4-glucosidase, EC 3.2.1.21).

The cellulose's biodegradation process starts with the randomly cut of β -1, 4-glucosidic linkages in the cellulose polysaccharides chain producing various lengths of oligosaccharides with the new chain ends. This action caused the breakdown of crystalline structure of cellulase. Afterward, exocellulase cleaves the reducing and non-reducing ends of this new oligosaccharides chain generating either glucose (glucanohydrolases) or cellobiose (cellobiohydrolase) as the major products. Finally, β -glucosidase hydrolyzes the remaining cellobiose or cellotetraose into glucose monomer unit [3].

Cellobiose which is an intermediate product is also a strong inhibitor for endoglucanase and exoglucanase and it becomes one of the key bottlenecks in enzymatic hydrolysis [4]. In order to prevent this inhibition process [5], cellobiose unit must be immediately removed. Thus, it is important to understand the catalytic activity of β -glucosidase in order to improve the efficiency of this enzyme. This will help in designing an enhanced β -glucosidases. However, little is known about the catalytic interactions between β -glucosidase and cellobiose.

The program utilized for molecular docking in this study is GOLD [6]. A large number of conformations of protein ligand complexes were generated by this docking program. Each conformation is positioned at active site in a variety of orientation known as pose whereas poses were ranked by the scoring function to determine the best overall poses. The scoring function incorporated in GOLD uses classical molecular mechanic pose force field. GOLD employed genetic algorithm on the whole molecule and hydrogen bonding and van der Waals interactions to determine the conformation and orientation of the ligand [7]. Since little information is known about the catalytic interactions between β -glucosidase and cellobiose, computational docking study was performed to explore the binding mode interaction of this protein and its ligands.

II. MATERIAL AND METHODOLOGY

A. Preparation of Protein and Ligands

The x-ray crystal structures of β -glucosidase B (BglB) was obtained from the RCSB Protein Data Bank (<http://www.rcsb.org>) with accession code 209T [8] respectively as show in Fig. 2. Waters, cofactors and originally bounded ligand in BglB was manually removed from the original PDB file. A 2-dimensional (2D) structure of

Manuscript received July 10, 2013; revised September 15, 2013. This work was supported in MJIT Fellowship Scholarship.

The authors are with the Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia Kuala Lumpur, Jalan Semarak, 54100, Wilayah Persekutuan, Malaysia (e-mail: nsfadhilah2@live.utm.my, nurulbahiyah@ic.utm.my).

cellobiose and laminaribiose were obtained from PUBCHEM (<http://www.pubchem.org>) with chemical identification code 10712 and 439637 respectively as shown in Fig. 3.

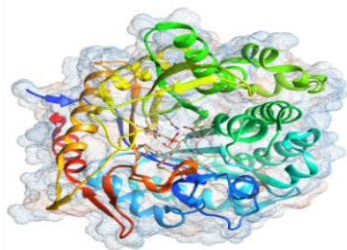


Fig. 2. Ribbon structures of β -glucosidase B. The pictures were generated from Chimera software are colored in preset mode of interactive 1 and publication 3 options; depth-cued and rounded ribbon option and covered by dot representation of surfaces. The existed substrates were shown in stick models with carbon atoms in grey and oxygen atoms in red.

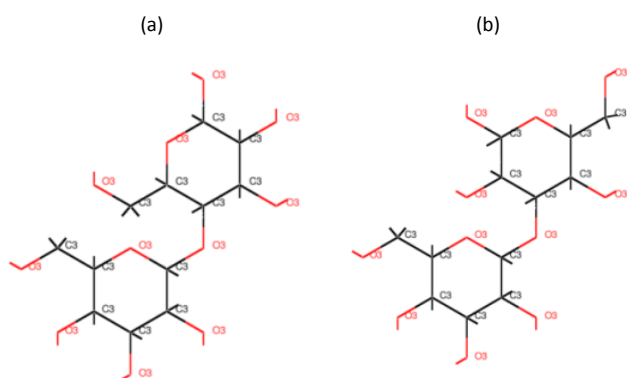


Fig. 3. Formula chemical structure of (a) cellobiose (CBI) and (b) laminaribiose (LBI) generated by Chimera, colored by hetero atom and labeled by atomic number

The three-dimensional cellobiose (CBI) and laminaribiose (LBI) structures were built using the Builder module in MOE version 2011.10 (Molecular Operating Environment). MOE software was installed in Macintosh machine running on Intel Core i5 processor with 2GB RAM and 500GB hard disk with Mac OSX Lion version 10.7.5 as operating system.

The multi conformation structural data of the ligands was generated using LowModeMD [9] in MOE 2011.10 with the default parameters; rejection limit = 100, RMSD gradient = 0.005, iteration limit = 10000, MM Iteration Limit, RMSD Limit = 0.25, energy windows = 7 and conformation limit = 10000. The top ten ligand conformations were saved in *.mol2* format and were used in docking simulation by GOLD.

B. Docking Procedure

The computational docking was performed in machine running on Intel Core i7 processor with 4GB RAM and 500GB hard disk with CentOS version 5 as the operating system. The active site was set at the centre of the protein. It was defined within 10 Å and ligand-binding interactions were analyzed using Gold Score (GS) scoring function. The default settings were applied for all other parameters: population size (100); selection-pressure (1:1); number of operation (10,000); number of island (1); niche size (2) and operator weight for migrate (0), mutate (100) and crossover (100).

Docking Result Analysis

From GOLD docking results, the top ten conformation pose were selected and saved in *.pdb* format for further use in protein ligand interaction analysis. The analysis of these complexes was performed using Protein Ligand Interaction Fingerprint (PLIF) and Ligand Interaction modules in MOE 2011.10.

III. RESULTS AND DISCUSSIONS

A. Docking Results

In this study, cellobiose and cellotetrose were docked into β -glucosidase active site using GOLD docking program. GOLD used genetic algorithm (GA) for docking flexible ligands into protein with flexible hydroxyl groups. In most cases the protein is considered to be rigid. This may lead a good choice since the binding pocket may contain amino acids that form hydrogen bonds with the ligands.

GOLD scoring function is a molecular-like function made up with four terms: (i) Protein ligand hydrogen bond energy (external H-bond); (ii) Protein-ligand vanderwaals energy (external vdw); (iii) Ligand internal vanderwaals energy (internal vdw); (iv) Ligand intramolecular hydrogen bond energy (internal H-Bond). By default, the external vdw score is multiplied by a factor of 1.375 during the computation of total fitness score [9]. Equation 1 shows the fitness function that has been optimized for the prediction of ligand binding position.

$$GF = S_{hb_ext} + S_{vdw_ext} + S_{hb_int} + S_{vdw_int} \quad (1)$$

where GF is Goldscore Fitness, S_{hb_ext} is the protein-ligand hydrogen bond and S_{vdw_ext} is the protein-ligand van der Waals score. S_{hb_int} is the contribution to the fitness due to the intermolecular hydrogen bonds. S_{vdw_int} is the contribution due to intermolecular strain in the ligand [7].

TABLE I: LIST OF HIGHEST GOLD SCORE VALUE

	BglB-CBI	BglB-LBI
Gold Score	81.74	80.59
S (hb_ext)	23.22	28.43
S (vdw_ext)	47.60	39.89
S(hb_int)	0.00	0.00
S(int)	6.93	-2.69

The list of highest gold score for both BglB-CBI and BglB-LBI complexes were tabulated in Table I. Both complexes gave slightly similar scores since both have similarity in term of chemical structure. For clearer view of binding mode, the poses of both protein and ligands complexes as shown in Fig. 4.

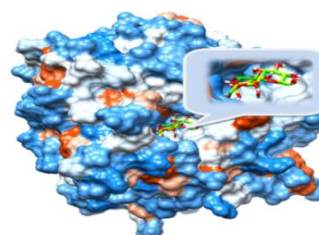


Fig. 4. Docked conformations of BglB with cellobiose (green) and laminaribiose (yellow) view in Chimera using hydrophobicity surface mode

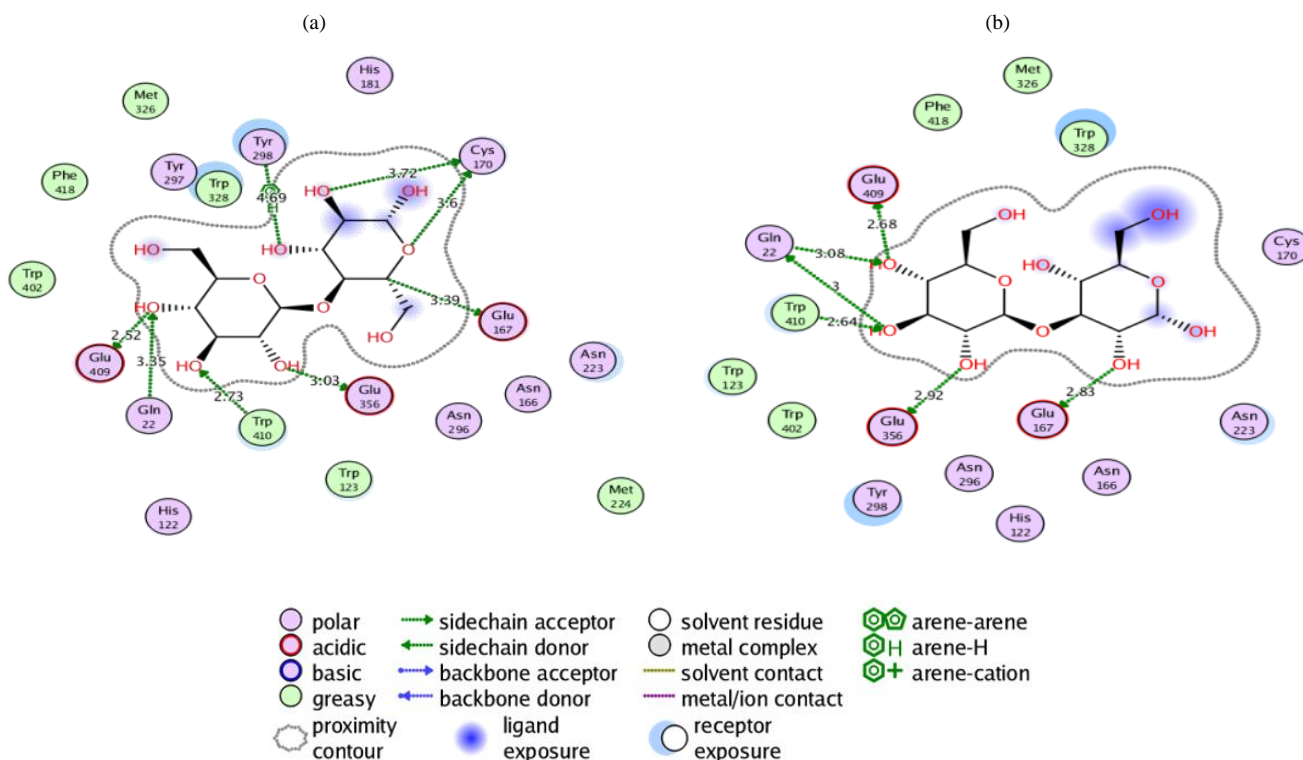


Fig. 5. The 2D depiction rendered from MOE 2011.10 shown the ligand interaction of docking complexes conformation (a) BglA-cellobiose and (b) BglB-laminaribiose

B. Protein Ligand Interaction

Molecular interaction especially hydrogen bonds plays an important role in forming enzyme-substrates complexes. Both cellobiose and laminaribiose were stabilized by BglB residues via hydrogen bond and pi-pi interactions. By utilizing ligand interaction mode in MOE 2011.10, the protein ligand interaction had been mapping out in Fig. 5. MOE was able to visualize solvent exposed ligand atoms as well as side chain and backbone acceptor and donor interactions.

Seven hydrogen bonds and one pi-pi interaction were observed in BglB-cellobiose complex as shown in Table II which involve residues Cys170, Glu356, Glu409, Glu167, Trp410, Gln22 and Tyr298.

TABLE II: LIST OF INTERACTION FOR BGLB-CELLOBIOSE

Ligand Atom	Receptor Atom	Interaction	Distance (Å)	Energy (kcal/mol)
O3	SG CYS170	H-Donor	3.6	-1.1
O5	OE2 GLU356	H-Donor	3.03	-2.9
O7	OE2 GLU409	H-Donor	2.52	-2.6
O8	SG CYS170	H-Donor	3.72	-1.0
C15	OE2 GLU167	H-Donor	3.39	-0.7
O6	NE1 TRP410	H-Acceptor	2.73	-1.2
O7	NE2 GLN22	H-Acceptor	3.35	-0.3
O4	6-Ring TYR298	H-pi	4.69	-0.7

Whilst another six hydrogen bonds founds in BglB-laminaribiose complex as shown in Table III interacting with residues Glu356, Glu167, Gln22, Glu409

and Trp410. From protein ligand interaction report, cellobiose has more number of interactions with BglB compare to laminaribiose.

TABLE III: LIST OF INTERACTION FOR BGLB-LAMINARIBIOSE

Ligand Atom	Receptor Atom	Interaction	Distance (Å)	Energy (kcal/mol)
O5	OE2 GLU356	H-Donor	2.92	-3.4
O6	OE2 GLU167	H-Donor	2.83	-3.7
O7	OE1 GLN22	H-Donor	3.00	-1.2
O8	OE2 GLU409	H-Donor	2.68	-4.4
O7	NE1 TRP410	H-Acceptor	2.64	0.3
O8	NE2 GLN22	H-Acceptor	3.08	-1.5

Glutamate is important for reducing the energy barrier of the glycosylation step [10]. The glutamate side-chain can adapt its position to an axial O4 and has the ability to recognize galacto-configured substrates due to its conformational freedom characteristic [11]. In this study, Glu409 was involved to form hydrogen bond for both cellobiose and laminaribiose ligands. This glutamate residue is highly conserved among family glycoside hydrolase 1 [12]. Thus it is ensure that this residue plays important residue in catalytic mechanism of cellobiose hydrolysis.

C. Active Residue Analysis

The Protein Ligand Interaction Fingerprints (PLIF) application is a method for summarizing the interacting residues between protein and ligands. The top ten ranked conformation complexes from each ligand as shown in Fig. 6 were used as input in PLIF program to determine common active residue between these complexes.

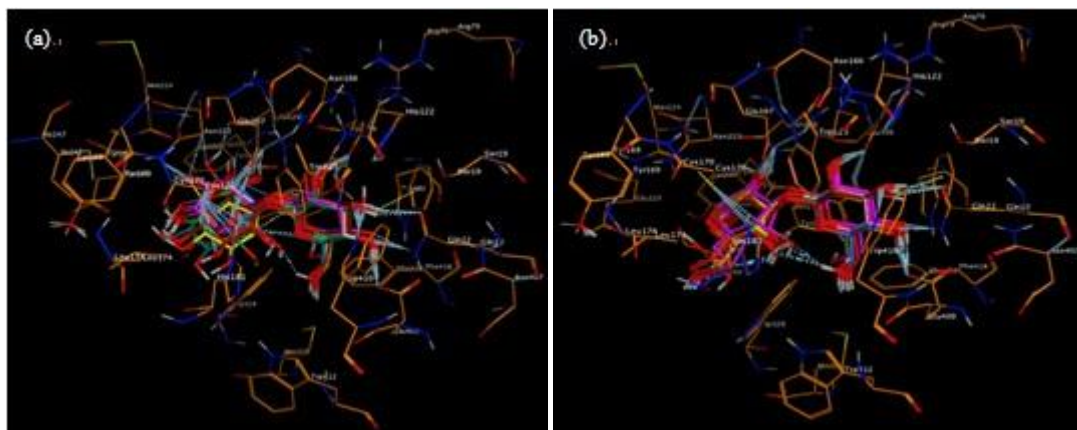


Fig. 6. Visualization of top ten ligand structure conformation,(a) BglB-cellobiose and (b) BglB-laminaribiose in binding pocket are shown in ball and stick representation whereas light blue dash line represent hydrogen bond interactions. The picture is visualized by MOE 2011.10.

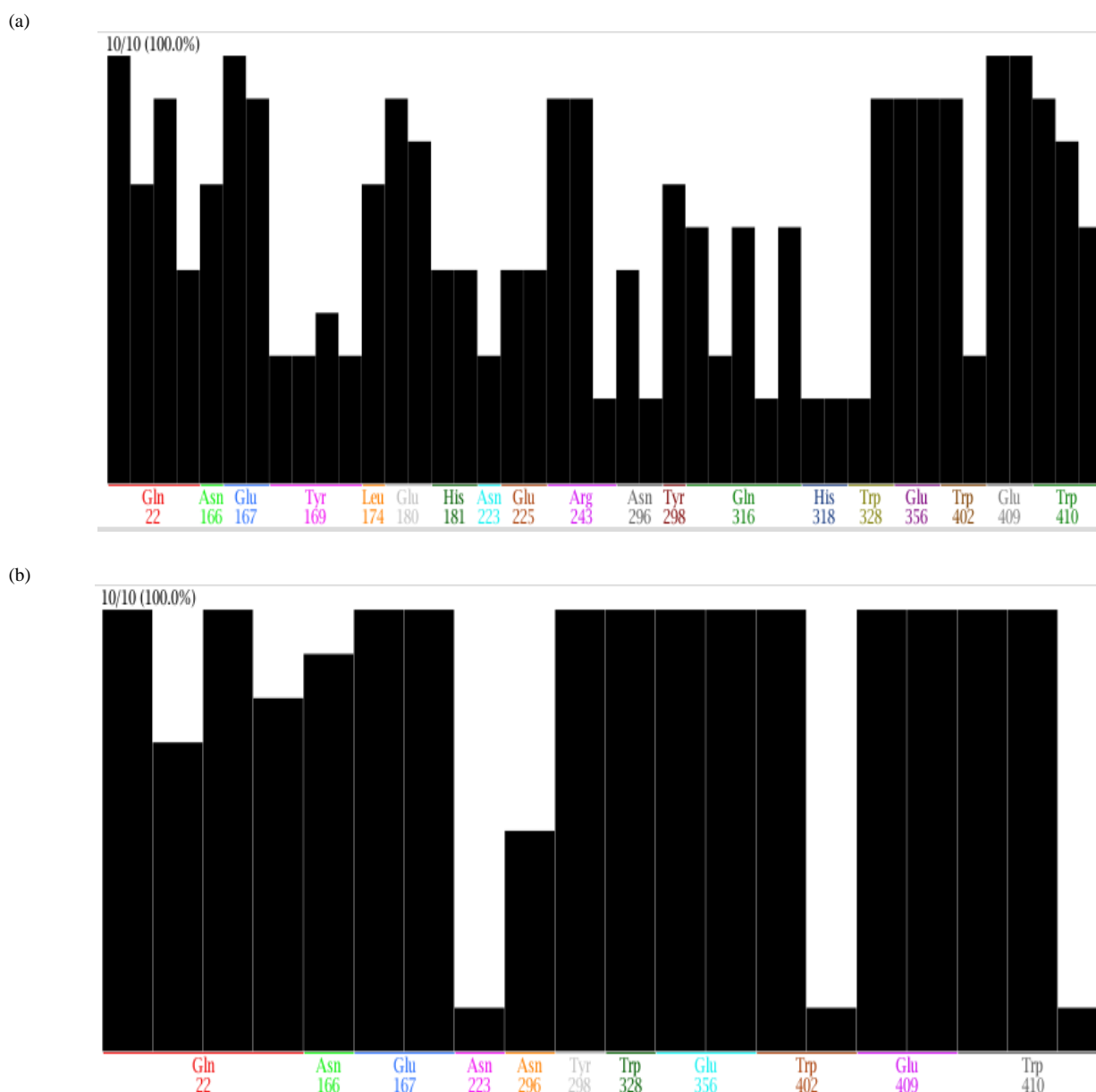


Fig. 7. PLIF diagram show common interacting residue between top ten ranked conformation protein-ligand complexes where,(a) BglB-cellobiose and (b) BglB-laminaribiose

From Fig. 7(a), the common active residue in BglB-cellobiose complexes with score 90% and above are observed to be Gln22, Glu167, Tyr298, Glu356, Trp402 and

Trp410. Whilst, from Fig. 7(b), the common active site in BglB-laminaribiose are Gln22, Glu167, Tyr298, Trp328, Glu356, Trp402, Glu409 and Trp410.

BglB-cellobiose and BglB-laminaribiose complexes share common active residue namely Gln22, Glu167, Tyr298, Glu356, Trp402 and Trp410 which suggesting these residues play important role in hydrolyzing substrates.

It has been reported that the residues Glu167 acts as a protonated agent of interglycosidic oxygen atom in hydrolysis mechanism, Glu356 acts as nucleophilic in stabilizing the transition state and Glu409 as catalytic residue [8]. Whilst, Tyr298 which is hydrogen bonded to Glu356 involved in recognition of the inhibitor. In agreement with P. Isorna *et al* (2007), it is observed that Glu167, Glu356, Glu409 and Tyr298 play an important role in catalytic function for both BglB-cellobiose and BglB-laminaribiose complexes.

IV. CONCLUSION

The molecular docking studies were conducted in order to further understand the binding modes of β -glucosidases. The docking analysis resulted in identification of important residue in protein-ligand interaction with respect to the binding site for β -glucosidase B (BglB). The important residues are reported to be Gln22, Glu167, Tyr298, Glu356, Glu402, and Trp410. The result from this study will eventually contribute new ideas to experimentalist to design better enzyme for more efficient enzymatic hydrolysis process with higher yields and lower production cost.

ACKNOWLEDGMENT

N.S.F. Mazlan would like to thank Prof. Shunsuke Aoki from Department of Bioscience and Bioinformatics, Kyushu Institute of Technology, Iizuka-shi, Fukuoka, Japan for guidance and providing facilities in his laboratory.

REFERENCES

- [1] J. Maclellan, *Strategies to Enhance Enzymatic Hydrolysis of Cellulose in Lignocellulosic Biomass*, pp. 31-35, 2010.
- [2] M. Wang, K. Liu, L. Dai, J. Zhang, and X. Fang, "The structural and biochemical basis for cellulose biodegradation," *Journal of Chemical Technology and Biotechnology*, vol. 88, no. 4, pp. 491-500, April. 2013.
- [3] M. Saharay, H. Guo, and J. C. Smith, "Catalytic mechanism of cellulose degradation by a cellobiohydrolase, CelS," *PloS one*, vol. 5, no. 10, pp. e12947, Jan. 2010.

- [4] A. N. U. Nutt, *Hydrolytic and Oxidative Mechanisms Involved in Cellulose Degradation*, 2006.
- [5] G. Zhao, X.-Y. Dong, and Y. Sun, "Ligands for mixed-mode protein chromatography: Principles, characteristics and design," *Journal of biotechnology*, vol. 144, no. 1, pp. 3-11, Oct. 2009.
- [6] W. Luo, J. Pei, and Y. Zhu, "A fast protein-ligand docking algorithm based on hydrogen bond matching and surface shape complementarity," *Journal of molecular modeling*, vol. 16, no. 5, pp. 903-13, May 2010.
- [7] M. L. Verdonk, J. C. Cole, M. J. Hartshorn, C. W. Murray, and R. D. Taylor, "Improved protein-ligand docking using GOLD," *Proteins*, vol. 52, no. 4, pp. 609-23, Sep. 2003.
- [8] P. Isorna, J. Polaina, L. L. Garca, F. J. Canada, B. Gonzalez, and J. S. Aparicio, "Crystal structures of *Paenibacillus polymyxa* beta-glucosidase B complexes reveal the molecular basis of substrate specificity and give new insights into the catalytic machinery of family I glycosidases," *Journal of molecular biology*, vol. 371, no. 5, pp. 1204-18, Aug. 2007.
- [9] V. Mohan, A. C. Gibbs, M. D. Cummings, E. P. Jaeger, and R. L. DesJarlais, "Docking: successes and challenges," *Current Pharmaceutical Design*, vol. 11, no. 3, pp. 323-33, Jan. 2005.
- [10] M. Tiwari, K.-M. Lee, D. Kalyani, R. Singn, H. Kim, J.-K. Lee, and P. Ramachandran, "Role of Glu445 in the Substrate Binding of β -Glucosidase," *Process Biochemistry*, no. 2010, Sep. 2012.
- [11] J. S. Aparicio, J. A. Hermoso, M. M. Ripoll, J. L. Lequerica, and J. Polaina, "Crystal structure of beta-glucosidase A from *Bacillus polymyxa*: insights into the catalytic activity in family I glycosyl hydrolases," *Journal of Molecular Biology*, vol. 275, no. 3, pp. 491-502, Jan. 1998.
- [12] R. Opasiri, B. Pomthong, T. Onkoksoong, T. Akiyama, A. Esen, and J. R. K. Cairns, "Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12 beta-glucosidase," *BMC Plant Biology*, vol. 6, pp. 33, Jan. 2006.



Nur Shima Fadhilah Mazlan graduated from Universiti Teknologi Malaysia, Johor Bahru with Bachelor of Chemical (Bioprocess) Engineering in September 2011. She is currently pursuing master degree in Malaysia Japan International Institute of Technology, Kuala Lumpur. During her master study, she spent four months attachment at Kyushu Institute of Technology, Iizuka-shi, Fukuoka, Japan. Her interest study is molecular docking and molecular dynamics simulation of enzyme system.



Nurul Bahiyah Ahmad Khairudin is a senior lecturer of Bioprocess engineering at Malaysia Japan International Institute of Technology, Kuala Lumpur. She received her PhD from Universiti Sains Malaysia. Her research focuses on work in molecular dynamics simulations of protein and other biomolecules.