ATLANTIS PRESS

Advances in Biological Sciences Research, volume 3

International Conference on Biological Engineering and Pharmacy (BEP 2016)

N-Glycosylation Sited at Residues of Catalytic Center Could Repress The Activity of Endoglucanase from *Rhizopus Stolonifer*

Bin TANG*
College of Biochemical Engineering
Anhui Polytechnic University
Wuhu, China
e-mail: tangbin@ahpu.edu.cn

Yingying ZHANG
School of Biotechnology
Jiangnan University
Wuxi, China
e-mail: zhangyingy.1106@163.com

Wei CHENG
College of Biochemical Engineering
Anhui Polytechnic University
Wuhu, China
e-mail: chengwei15@ahpu.edu.cn

Qingqing ZHANG
College of Biochemical Engineering
Anhui Polytechnic University
Wuhu, China
e-mail: zhangqq@ahpu.edu.cn

Song LI
College of Biochemical Engineering
Anhui Polytechnic University
Wuhu, China
e-mail: lisonga@ahpu.edu.cn

***Abstract*—As a common post-translational modification of proteins, glycosylation has been proven effectively influence the activity of enzyme. To improve the activity of endoglucanase (EGII) from *Rhizopus stolonifer* TP-02, we studied the effect of glycosylation on the structure and activity of EGII. Two potential glycosylation sites (N84 and N150) were predicted, of which N150 was located in the entrance of catalytic domain (CD). N84D and N150D was obtained by overlapping PCR and expressed in *Pichia pastoris* for purification. Compared with EGII, the specific activity of N84D was decreased by 28.7%, whereas the activity of N150D was increased by 56.9%. Dynamics simulations results indicated that the N-glycans located in the active center could repress the activity of EGII due to the steric hindrances that effectively hinder the cellulose chains to enter the CD domain.**

Keywords-N-glycosylation; endoglucanase; molecular dynamics simulation; Rhizopus stolonifer

I. INTRODUCTION

Glycosylation is a common type of the post-translational modifications in eukaryotic proteins [1, 2]. More than half of the proteins found in nature undergo glycosylation, of which three quarters are N-glycosylated [3, 4]. The N-linked glycosylation recognizes asparagine (N) residues with a conserved amino acid sequence N-X-S/T/C (where X can be any residue except proline) [5-7]. As a ubiquitous and complex protein post-translational modification, N-glycosylation affects the protein folding, secretion, intracellular transport and localization [8, 9]. It has been confirmed that N-glycosylation maintains the stability of proteins via the steric hindrance caused by N-linked glycan avoiding the hydrolysis of protease [10-12]. The importance of N-glycosylation in regulating the structure and function of cellulases has been reported, which indicated the unique and diversified role that N-glycosylation played [13-16].

As a major component of the cellulases, endoglucanase that catalyzes the hydrolysis of cellulose to yield cello-oligosaccharides by facilitating the random cleavage of the internal β-1,4-D-glycosidic linkages has been widely studied because of its application to a wide range of industrial fields [17, 18]. One of the bottlenecks to widespread application of endoglucanase is the enzymatic activity to hydrolyze the cellulosic material into sugars. Site-directed mutagenesis is used to modify the structure and improve the catalytic efficiency of endoglucanase [19, 20]. Furthermore, the structure-function relationship of endoglucanase is one of the hot topics, although it has yet been fully clarified.

In order to enhance the catalytic ability of endoglucanase, herein this paper, mutation of the glycosylation sites of the endoglucanase from *R. stolonifer* TP-02, together with the corresponding catalytic structure and effects on the enzymatic activity were studied.

II. MATERIALS AND METHODS

A. Strains and Culture Conditions

R. stolonifer TP-02 was isolated in our laboratory and stored in China General Microbiological Culture Collection Center (CGMCC No. 11119). Recombinant cloning vector pET28a-*eg2*, *Escherichia coli* DH5α, plasmid pPIC9K, and *Pichia pastoris* GS115 were stored in our laboratory.

Fermentation of *P. pastoris* in the 10 L fermenter was referenced in *P. pastoris* expression kit (Invitrogen) and *Pichia* Fermentation Process Guidelines (Invitrogen). MM

TABLE I. SEQUENCES OF PRIMERS FOR AMPLIFICATION OF TARGET GENES

Names	Sequences(5'-3')
EGII	F:5'-CGGAATTCATGAAGTTTATTACTATTACG-3' R:5'ATAAGAATGCGGCCGCTATGGTGATGGTGAT GATGTTTTCTTGAACAACC3'
N84D	F:CATGCCCCTGAAAGCAACGGCGACAAAACTTCT GAAAGCGCTCAT R:ATGAGCGCTTTCAGAAGTTTTGTCGCCGTTGCTT TCAGGGGCATG
N150D	F:TGTAGCTGGCCCGGTAAGGCCGACGTCAGTTCT CCTGTCAAGTCC R:GGACTTGACAGGAGAACTGACGTCGGCCTTACC GGGCCAGCTACA

medium containing sodium carboxymethylcellulose (CMC-Na) was added to a final concentration of 1% for the CMC-Na. The concentration of Congo red was 10.

B. Predicted N-glycosylation Site and Site-directed Mutagenesis of Target Gene

The NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/) was used to predict the N-glycosylation and mutagenesis site of *eg2*. To obtain *eg* gene from the recombinant plasmid pET28a-*eg2*, we designed the primers shown in Table I (the restriction sites are underlined). The original gene was used as a template. Overlapping PCR technique was used to complete the mutagenesis of *eg2* glycosylation sites [21]. Two groups of four primers were designed (Table I, the mutated bases are underlined). The following PCR reaction conditions were adopted: initial denaturation at 94°C for 5 min, 30 cycles at 94°C for 50 s, 52°C for 50 s, and 72°C for 90 s, followed by 72°C for 10 min.

C. Bioinformatics Analysis and Molecular Dynamics Simulation of EGII and Its Mutants

Glycosylation sites analysis of EGII was predicted by using NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/). The PDB database was used to search templates for homology modeling. Bioinformatics software Discovery Studio 2.5 (DS2.5) was used to simulate the 3D structure of EGII. Molecular slicing was completed in Libdock modules of DS2.5. Molecular dynamics simulation was conducted to analyze and elaborately explain the catalytic processes of EGII and its mutant. The process of molecular dynamics simulation was as follows: firstly, the protein-ligand complexes were given force field by CHARMm in DS2.5 [22]; secondly, sodium and chloride were added into cations and anions, respectively; finally, the steps and time steps were 8000 and 0.002 ps, respectively. The set type was NPT.

D. Construction and Selection of The Recombinant Yeast

After purification and digestion with *Eco*RI and *Not*I, the PCR amplified fragment was linked by T4 DNA ligase at the corresponding cloning sites of pPIC9K vector. These products were transformed into *E. coli* DH5α competent cells, and the positive clones were screened to construct the expression vectors pPIC9K-N84D, pPIC9K-EGII, and pPIC9K-N150D. After linearized by *Sal*I, these three recombinant plasmids were transformed into *P. pastoris* competent cells by electroporation. Ice-cold 1 mol/L sorbitol (1 mL) was immediately added to the electroporation cuvette after pulsing and screened in MD agar plates at 30°C for 3 days. Sterilized toothpicks were used to pick monoclonal colonies, which were then inoculated into MD and MM agar plates. The MM agar plates contained 1% CMC-Na. After 3 days, the MM plates were stained with Congo red. The corresponding strains that had large clear zone of hydrolysis in MD agar plates were selected for subsequent tests.

E. Purification of Recombinant Proteins

The seed culture medium was inoculated into 10 L of fermenter containing fermentation basal salt medium (containing 4% glycerol). The medium was maintained under dissolved oxygen of more than 20%, temperature of 30°C, and pH of approximately 5.0. When the cell wet weight reached 200 g/L, glycerol feed was terminated and the supernatant entered into the methanol fed-batch phase. After being induced for 96 h, the supernatant was concentrated and precipitated. Buffer A (20 mM Tris-HCl + 0.5 M NaCl + 20mM C3H4N2, pH 7.0) was then used for dialysis for 24 h. ÄKTA prime plus (GE Healthcare) rapid purification system was employed to purify the target proteins. The sample was adsorbed and balanced on the Ni column [23]. Afterward, the sample was added with buffer B (20 mM Tris-HCl, 0.5 M NaCl, 200 mM-500 mM C3H4N2, pH 7.0) for linear gradient elution. The rate was 1 mL/min. Fractions were collected for activity assays and analysis by sodium dodecyl sulfate polyacrylamide gel electrophoresis.

F. Enzymatic Assays

Endoglucanase activity was measured by the release of the reducing sugar during hydrolysis of CMC-Na or phosphoric acid swollen cellulose (PASC) using the DNS method [24, 25]. Enzyme solution (1 mL) was incubated with 1 mL of 1% (w/v) CMC-Na or 2% (w/v) PASC in sodium acetate buffer (0.1 M, pH 4.8) at 50°C for 30 min. Inactivated enzyme solution was used as a control group. One unit of endoglucanase activity releases 1μmol of reducing glucose per min.

CMC-Na and PASC were used as substrate to measure kinetic constants. They were added into the enzyme solutions with a concentration between 1 and 10 mg·mL-1. Lineweaver-Burk plots were used to estimate Km and Vmax.

III. RESULTS AND DISCUSSION

A. Glycosylation Sites Prediction and Structural Simulation of EGII

As previously reported, EGII contains a fungal cellulose binding domain CBM1 (29-57) and a catalytic domain that belongs to glycoside hydrolase family 45 (129-305) [26]. The results of glycosylation sites analysis showed that there

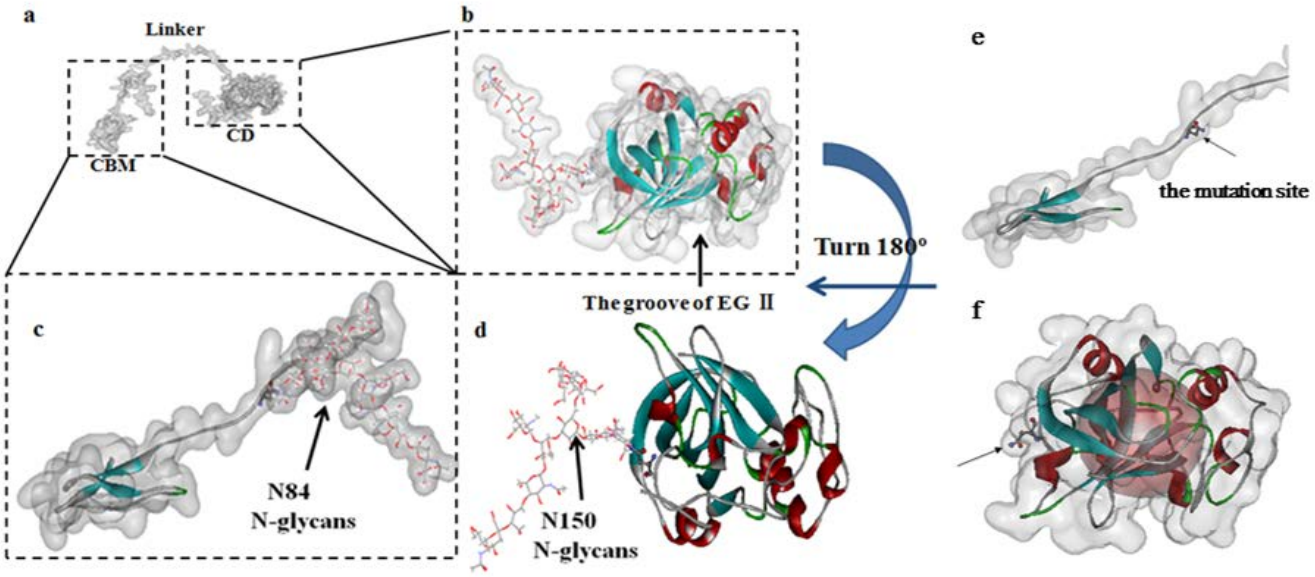

Figure1. The structures of EGII (a-d), N84D (e) and N150D (f). Red spherical region was the catalytic activity center of CD.

are four potential glycosylation sites (N84, N111, N150 and the characteristic sequence motifs N-X-T (very often), N-X-S (often) or N-X-C (very rare), where X could be any residue except proline [6]. Hence, N84 and N150 were most likely glycosylated.

To study the conformational change of EGII without glycosylation, the 3D structural model of EGII and its mutants were constructed by homology modeling (Fig. 1). Two glycosylation sites, namely N84 and N150, were marked. The N84 was located at the linker close to CBM, whereas the N150 was located at the entrance of the catalytic activity center of EGII. Considering to the stability of structure, asparagine (N) was anticipated to be replaced by aspartic acid (D). Fig. 1(e) and Fig. 1(f) displayed the deglycosylation conformation.

B. Mutagenesis and Expression of Recombinant EGII

EGII, N84D, and N150D were successfully expressed in *P. pastoris* screening by the MM agar plates that contained N295). However, N-glycosylation was predicted to occur at 1% CMC-Na. The CMC-Na could be hydrolyzed by endoglucanase to produce oligosaccharides that react with Congo red forming different sizes of transparent circles around the clones [27]. The sizes of these transparent circles were positively correlated with endoglucanase activity. Accordingly, the activity of N150D might be higher than EGII, of which the transparent circle was large (Fig. 2). In contrast, the deglycosylation of N84 might reduce the normal activity of EGII.

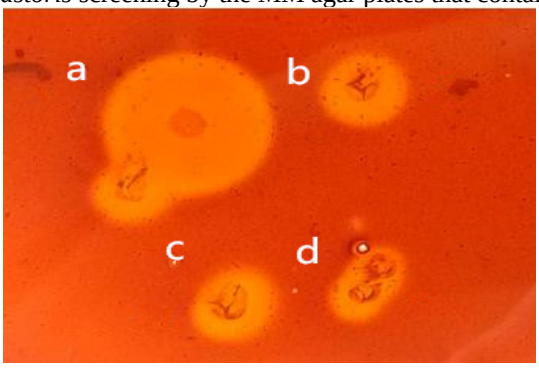

Figure2. Congo red staining of pPIC9K-N150D (a), pPIC9K-EGII (b, c), and pPIC9K-N84D (d).

C. Purification and Enzymatic Characteristics of Recombinant Proteins

Given that the target gene contains 6-His tag in its C-terminal, protein could specifically bind to the metal Ni. The nickel column purification system (ÄKTA prime plus) was used to purify the target proteins (Fig. 3). Further enzymatic characteristics of three purified proteins (EGII, N84D, and N150D) were studied. The results showed that specific

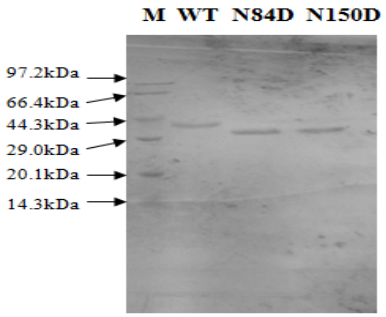

Figure 3. SDS-PAGE analysis of purified EGII, N84D and N150D.

TABLE II. PURIFICATION OF THREE RECOMBINANT PROTEINS

Purification	Total protein (mg)			Total activity (IU)			Specific activity(IU/mg)			Activity recovery (%)		
	N84D	EGII	N150D	N84D	EGII	N150D	N84D	EGII	N150D	N84D	EGII	N150D
zymotic fluid	ND	ND	ND	1105.9	1432.6	2014.2	ND	ND	ND	100	100	100
Concentrate	277.1	424.2	630.6	886.9	1187.6	1639.6	3.2	2.8	2.6	80.2	82.9	81.4
HisTrap	16.8	15.2	15.4	449	573.6	914.4	26.7	37.7	59.4	40.6	48.3	45.4

The substrate used for enzymatic assays was CMC-Na; ND: not determined.

TABLE III. KINETIC CONSTANTS OF THREE PROTEINS

Substrate	Enzyme	K_m (mg·mL-1)	V_{max} (μmol/min)
CMC-Na	N84D	2.15±0.06	2.39±0.05
	EGII	2.02±0.02	2.74±0.08
	N150D	1.51±0.03	3.79±0.07
PASC	N84D	1.63±0.04	4.08±0.05
	EGII	1.42±0.02	4.59±0.06
	N150D	1.06±0.02	5.76±0.07

activity of N84D was decreased by 28.7% compared with EGII, whereas the activity of N150D increased by 56.9% (Table II).

Km can determine the approximate affinity of enzymes and substrate. A smaller value of Km means a greater affinity. The Km value of N150D was lower than that of EGII, suggesting that the N150D have a greater affinity (Table III). However, a contrary result was observed in N84D. The Vmax of N150D was higher than that of EGII, whereas the Vmax of N84D was lower than that of EGII. All these results indicated that the N-glycans of EGII can affect its substrate affinity and catalytic efficiency.

D. Dynamics Simulations of Recombinant Proteinss

In order to analyze the effect of N-glycosylation on the structure and function of EGII, dynamics simulations were adopted. The results indicated that removing the N-glycans

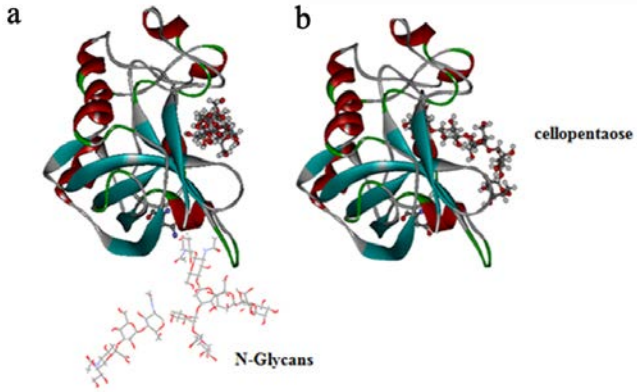

Figure 4. The docking renderings of EGII/N84D (a) and N150D (b) with substrates.

of N150 could result large change of the overall conformation of CD, particularly the changed contact angles might influenced the interaction capabilities of enzyme and substrates. The glycosylation site N150, which was situated at the loop of CD, may inhibit the activity of the loop that located at the entrance of the catalytic activity center. Thus, this phenomenon may change the path of the cellulose entering into the catalytic activity center, affecting its catalytic efficiency. As Fig. 5 showed, the style of cellulose chain entered into the catalytic activity center of different CD conformations was different. Similar to most endoglucanases, all the cellulose chains entered into the catalytic activity center of EGII via its unique “groove” structure (Fig. 4a). However, this particular entry mode was changed in N150D (Fig. 4b). Meanwhile, the degrees of interaction between the enzyme and substrates were changed.

Notably, kinds of interaction forces are existed between cellulose chains and the cellulases, including van der Waals forces, hydrogen bonding, and electrostatic forces, wherein the hydrogen bond plays a major role. The effect of hydrogen bonds on the cellulose chain was more obvious in N150D than EGII and N84D, which means the cellulose chains were more likely to cleaved (Fig. 5). Moreover, the heat map of hydrogen bonds primarily reflects the key hydrogen bonds in the catalytic process, of which deeper red color indicates stronger hydrogen bond force. Fig. 6 showed the strengthening effect of the key amino acids after mutation and the number of major amino acids involved in increasing catalysis, resulting in the superiority of the entire system of N150D to EGII. The key sites in N150D included Y135, C139, Q140, D173, R229, D231, and D242, of which three aspartic acids are the catalytic sites. Furthermore, The total energy was enhanced from -5.8×105 kJ·mol^{-1} to -4.4×105 kJ·mol^{-1} (Fig. 7).

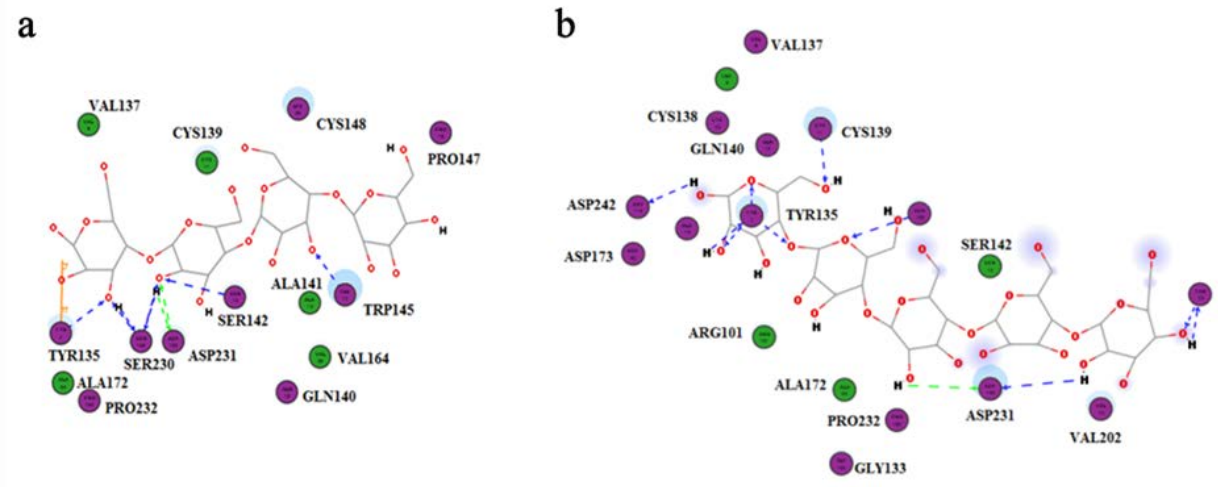

Figure 5. The 2D depiction of protein-ligand complexes. (a): The docking rendering of EGII/N84D with cellotetraose. (b): The docking rendering of N150D with cellopentaose. The amino acids that participated in the interactions of hydrogen bonding, electrostatic or polar were marked with purple. The amino acids that involved in van der Waals interactions were marked with green.

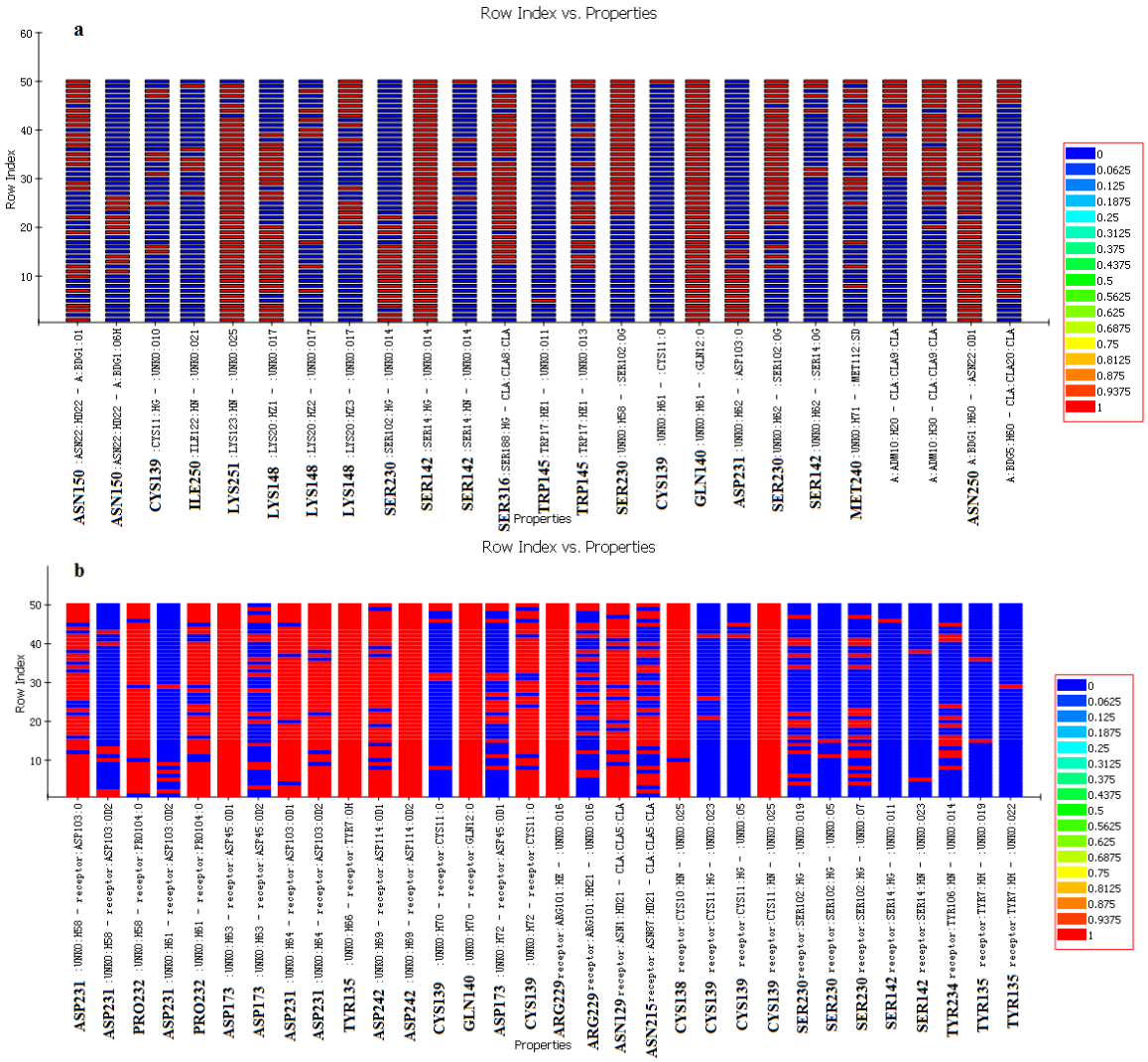

Figure 6. Heat map of hydrogen bonds that interacted between EGII/N68D (a) or N134D (b) and substrate.

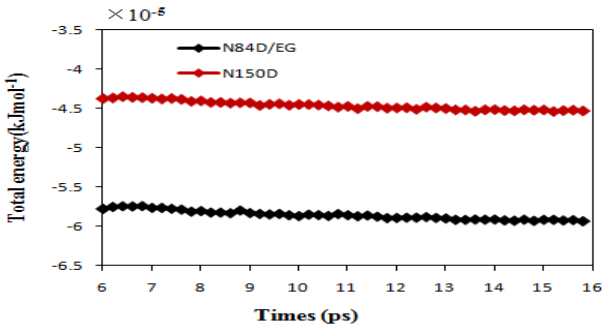

Figure7. Total energy as a function of simulation time. N84D/EGII (black line); N150D (red line)

IV. CONCLUSION

In general, N-glycans have an important role in stabilizing the conformation. The N-glycans of N84 might support the CBM to firmly grasp the cellulose chains. The N-glycans of N150 might contribute to stabilize the CD conformation. However, this stability formed steric hindrances, which hinder the natural binding of CD and cellulose chains repressing the activity of EGII.

ACKNOWLEDGMENT

This work was financially supported by the National Natural Science Foundation of China (No.31270135).

REFERENCES

[1] G. Reuter, and H. J. Gabius, "Eukaryotic glycosylation: whim of nature or multipurpose tool?," CMLS-Cell Mol. Life S., 55, 368-422, 1999.

[2] R. G. Spiro, "Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds," Glycobiology, 12, 43R-56R, 2002.

[3] R. Apweiler, H. Hermjakob, and N. Sharon, "On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database," BBA-Gen Subjects, 1473, 4-8, 1999.

[4] S. Zou, S. Huang, I Kaleem, and C. Li, "N-glycosylation enhances functional and structural stability of recombinant β-glucuronidase expressed in *Pichia pastoris*," J. Biotechnol., 75-81, 2013.

[5] P. Burda, and M. Aebi, "The dolichol pathway of N-linked glycosylation," BBA-Gen Subjects, 1426, 239-257, 1999.

[6] C. Caragea, J. Sinapov, A. Silvescu, D. Dobbs, and V. Honavar, "Glycosylation site prediction using ensembles of Support Vector Machine classifiers," BMC Bioinformatics, 8, 438, 2007.

[7] Y. Gavel, and G. V. Heijne, "Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering," Protein Eng., 3, 433-442, 1990.

[8] K. T. Pilobello, and L. K. Mahal, "Deciphering the glycocode: the complexity and analytical challenge of glycomics," Curr. Opin. Chem. Biol., 11, 300-305, 2007.

[9] N. H. Thuan, and J. K. Sohng, "Recent biotechnological progress in enzymatic synthesis of glycosides," J. Ind. Microbiol. Biot., 40, 1329-1356, 2013.

[10] M. M. Chen, A. I. Bartlett, P.S. Nerenberg, C. T. Friel, C. P. R. Hackenberger, C. M. Stultz, S. E. Radford, and B. Imperiali "Perturbing the folding energy landscape of the bacterial immunity protein Im7 by site-specific N-linked glycosylation," P. Natl. Acad. Sci. USA., 107, 22528-22533, 2010.

[11] E. K. Culyba, J. L. Price, S. R. Hanson, A. Dhar, C. Wong, M. Gruebele, E. T. Powers, and J. W. Kelly, "Protein native-state stabilization by placing aromatic side chains in N-glycosylated reverse turns," Science, 331, 571-575, 2011.

[12] J. L. Price, D. L. Powers, E. T. Powers, and J. W. Kelly, "Glycosylation of the enhanced aromatic sequon is similarly stabilizing in three distinct reverse turn contexts," P. Natl. Acad. Sci. USA., 108, 14127-14132, 2011.

[13] G.T. Beckham, Z. Dai, J.F. Matthews, M. Momany, C.M. Payne, W.S. Adney, S.E. Baker, and M.E. Himmel, "Harnessing glycosylation to improve cellulase activity," Curr. Opin. Biotech., 23, 338-345, 2012.

[14] T. Jeoh, W. Michener, M. E. Himmel, S. Decker, and W. S. Adney, "Implications of cellobiohydrolase glycosylation for use in biomass conversion," Biotechnol. Biofuels, 1, 10, 2008.

[15] W. Wei, L. Chen, G. Zou, Q. Wang, X. Yan, J. Zhang, C. Wang, and Z. Zhou, "N-glycosylation affects the proper folding, enzymatic characteristics and production of a fungal β-glucosidase," Biotechnol. Bioeng., 110, 3075-3084, 2013.

[16] Y. D. Wei, K. S. Lee, Z. Z. Gui, H. J. Yoon, I. Kim, Y. H. Je, S. M. Lee, G. Z. Zhang, X. Guo, H. D. Sohn, and B. R. Jin, "N-linked glycosylation of a beetle (*Apriona germari*) cellulase Ag-EGase II is necessary for enzymatic activity," Insect Biochem. Molec., 36, 435-441, 2006.

[17] M. Huron, D. Hudebine, N. L. Ferreira, and D. Lachenal "Mechanistic modeling of enzymatic hydrolysis of cellulose integrating substrate morphology and cocktail composition," Biotechnol. Bioeng., 113, 1011-1023, 2016.

[18] C. M. Payne, B. C. Knott, H. B. Mayes, H. Hansson, M. E. Himmel, M. Sandgren, J. Ståhlberg, G. T. Beckham "Fungal cellulases," Chem. Rev., 115, 1308-1448, 2015.

[19] S. J. Kim, J. E. Joo, S. D. Jeon, J. EHyeon, S. W. Kim, Y. S. Um, and S. O. Han "Enhanced thermostability of mesophilic endoglucanase Z with a high catalytic activity at active temperatures," Int. J. Biol. Macromol., 86, 269-276, 2016.

[20] J. Zhang, H. Shi, L. Xu, X. Zhu, and X. Li, "Site-directed mutagenesis of a hyperthermophilic endoglucanase Cel12B from *Thermotoga maritima* based on rational design," PLOS ONE, 10, e0133824, 2015.

[21] S. N. Ho, H. D. Hunt, R. M. Horton, J. K. Pullen, and L. R. Pease "Site-directed mutagenesis by overlap extension using the polymerase chain resction," Gene 77, 51-59, 1989.

[22] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus, "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations," J. Comput. Chem., 4, 187-217, 1983.

[23] K. L. Franken, H. S. Hiemstra, K. E. van Meijgaarden, Y. Subronto, J. den Hartigh, T. H. M. Ottenhoff, and J. W. Drijfhout, "Purification of his-tagged protein by immobilized chelate affinity chromatography: the benefits from the use of organic solvent," Protein Expres. Purif., 18, 95-99, 2000.

[24] G. L. Miller, R. Blum, W. E. Glennon, and A. L. Burton "Measurement of carboxymethylcellulase activity," Anal. Biochem., 1, 127-132, 1960.

[25] E. Vlasenko, M. Schülein, J. Cherry, and F. Xu, "Substrate specificity of family 5, 6, 7, 9, 12, and 45 endoglucanases," Bioresource Technol., 101, 2405-2411, 2010.

[26] B. Tang, Y. Zhang, Y. Yang, Z. Song, and X. Li, "Expression and functional analysis of a glycoside hydrolase family 45 endoglucanase from *Rhizopus stolonifer*," World J. Microb. Biot., 30, 2943-2952, 2014.

[27] R. M. Teather, and P. J. Wood, "Use of congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen," Appl. Environ. Microb., 43, 4777-4780, 1982.