Asn124 of Cel5A from *Hypocrea jecorina* not only provides the N-glycosylation site but is also essential in maintaining enzymatic activity

Yuqi Qin* & Yinbo Qu

National Glycoengineering Research Center, and State Key Laboratory of Microbial Technology, Shandong University, Shandong 250100, China

INTRODUCTION

Cellulose is the most abundant polymer on earth. The enzymatic degradation of cellulose is essential for maintaining the global carbon cycle. The cellulase system of the soft rot fungus *Hypocrea jecorina* (syn. *Trichoderma reesei*) is the most studied and best understood of all cellulosytic systems (1). *H. jecorina* produces a suite of cellulosytic enzymes with distinctly different activities. These enzymes are classified into three major families, namely, endoglucanases (EGs; EC 3.2.1.4), cellobiohydrolases (CBHs; EC 3.2.1.91), and β-glucosidases (EC3.2.1.21) (2). According to the classical theory, EGs internally nick the cellulose, thus disrupting its crystallinity and generating new free ends in the polymer. CBHs act progressively from these free ends and release soluble cellobiose molecules. β-glucosidases subsequently hydrolyze cellobiose to glucose (3).

Most enzymes secreted by *H. jecorina* are glycoproteins with both O- and N-linked glycosylation sites (4, 5). N- and O-linked glycosylation could affect the catalytic efficiency, cellulose-binding affinity, and stability of cellulases (6). The extensive glycosylation in cellulytic enzymes in various species (bacteria, fungi, and insects) indicates that they may have a very subtle effect on enzyme catalysis (5, 7, 8). Increased level of N-glycosylation of the Cel7A (CBHI) of *H. jecorina* resulted in reduced activity and increased non-productive binding on cellulose (9). Deglycosylation of *Talaromyces emersonii* Cel7A expressed in *Saccharomyces cerevisiae* resulted in decreased hydrolytic efficiency on crystalline cellulose (10). Different glycoforms of cellobiohydrolase I (CBHI-B) in *Penicillium decumbens*, which have the same amino acid sequence but different N-glycosylation characteristics, display different biological functions (11). Glycosylation affects not only catalytic efficiency but also enzyme stability. N-glycosylation of EG of *Bacillus subtilis* protects itself from immobilized-papain attack and accounts for higher thermostability (8).

Cel5A (EGII) is one of the most abundant EGs from *H. jecorina* (12, 13). The lack of Cel5A production reduces EG activity in the culture supernatant by as high as 55% (14). The presence of both N- and O-linked glycans in Cel5A with only one N-glycosylation site (Asn124) was observed in its catalytic domain (4). The analysis of secretome produced by *H. jecorina* showed Cel5A heterogeneity. The molecular masses of Cel5A observed on 2D gels were higher than the expected masses calculated from the protein sequences, and the reason may be due to glycosylation (15).

Several mechanisms of glycosylation and deglycosylation effects on Cel7A (CBHI) are well characterized (9, 16-18). However, few data exist on the nature and effect of glycosylation on the effectiveness or stability of Cel5A. According to our previous report, no significant difference is observed in enzymatic activity between the native Cel5A from *H. jecorina* and the recombinant Cel5A expressed in *S. cerevisiae*, which has an original 57 kDa molecular mass. Endoglucanase H (Endo H)-treated rCel5A shows a 54 kDa molecular mass (19).
In the present study the function of N-linked glycosylation in Cel5A was examined using two mutants (rN124D and rN124H) that lack the potential N-linked glycosylation sites. We show that the Asn residue at 124 site not only provides the N-glycosylation site but is also essential in maintaining enzymatic activity. N-linked glycosylation of Cel5A affected the thermal stability of enzyme especially at lower pH value.

RESULTS

Asn residue at 124 site is essential for maintaining enzyme activity
To elucidate the functional role of N-glycosylation sites, two mutants (rN124D and rN124H) were constructed by site-directed mutagenesis. DNA sequencing of the entire gene confirmed the absence of secondary mutations in the mutants. The recombinant enzymes were purified, and SDS-PAGE analysis indicated that both recombinant mutants had lower molecular masses from approximately 57 kDa to 54 kDa (Fig. 1A and B). The reduction in size of the two mutants compared with the rCel5A was expected to be a result of deglycosylation. Zymograms showed three clear activity bands confirming the molecular mass change (Fig. 1B).

Both recombinant mutants had reduced specific activity. The specific activity of rN124D was 194.9 U/mg, which was slightly lower (10%) than that of rCel5A (215.6 U/mg). However, the mutation rN124H displayed an apparent decrease in its specific activity (130.6 U/mg), approximately 60.6% compared with that of rCel5A (Fig. 1C). For both mutants, the zymograms of the crude fermentation showed the smear migrating above 85 kDa, which were identical to rCel5A. Given the deglycosylation, they might be heterozygous dimers or oligomers of recombinant protein but not hyper-mannosylated glycoproteins as we previously assumed (19) because CMCase activities were detectable regardless of the large molecular masses (Fig. 1B).

N-glycosylation affects the thermal stability of the enzymes especially at lower pH value
The effect of temperature on enzymatic hydrolysis of rN124D and rN124H at 70°C was studied at pH 5.0 and 4.0. At pH 5.0, the thermal stability of both mutants was identical to that of rCel5A. Both mutants lost approximately 50% of the initial activity after 1 h incubation (Fig. 2A). At pH 4.0, all three samples showed much less thermal stability than those at pH 5.0. Glycosylated rCel5A and deglycosylated rN124D and rN124H showed different thermal stabilities. Each sample respectively retained approximately 29%, 20%, and 41% of its original activity within 5 min (Fig. 2B).

Asn residue at site 124 of Cel5A is essential for maintaining enzyme conformation
The mutant rN124H indicated catalytic efficiency loss. Both mutants showed different thermal stabilities at lower pH value compared with that of rCel5A. To investigate whether the catalytic domains of the mutants were correctly folded, circular di-
N124 of Cel5A is essential for enzymatic activity
Yuqi Qin and Yinbo Qu

The far-UV CD spectrum of the purified rCel5A indicated a protein rich in β-sheet structure (59.7%) and random coil structure (39%). This result agrees with the high β-sheet content observed in the modeled three-dimensional structure of the enzyme (20). The CD spectra of rN124D indicated a secondary structure similar to that of the wild-type catalytic domain. The slight decrease in negative peak in the spectra range of approximately 202 nm to 205 nm corresponds to an increase in random coil conformation (21). However, the spectra region of rN124H differed greatly with that of rCel5A. Negative peaks of rN124H shifted to the right, with the positive peak in the spectra ranging between 195 and 200 nm. These peaks corresponded to a decrease in random coil but an increase in α- or β-helix conformation (22). The change in secondary structure by mutation was speculated to lead to reduction in hydrolysis activity of rN124H.

DISCUSSION

In Cel5A, there is O-glycosylation and the corresponding number of O-linked glycans ranged from 32 to 42 hexoses residues. However, the O-glycosylation site of Cel5A was not identified yet (4). Endo H is a highly specific endoglycosidase that cleaves asparagine-linked mannose-rich oligosaccharides, generating a truncated sugar molecule with one N-acetylgalactosamine residue retained on the asparagines (23). Endo H is assumed to work only on N-glycosylation of Cel5A, not on O-glycosylation, which always occurs on serine, threonine, and hydroxylysine residues. Therefore, rN124D and rN124H mutants were expected to have a lower molecular mass of approximately 54 kDa (Fig. 1A and 1B), consistent with rCel5A treated with endo H (19).

rN124D and rN124H showed approximately 10% and 40% reduced activity, respectively (Fig. 1C). Combined with our previous report that endo H-treated rCel5A also showed approximately 10% reduced activity (19), the reduced activity of rN124D can be ascribed to the loss of N-glycosylation. The activity reduction of rN124D is due to the protein structural change, which is the result of N-glycosylation loss by Asp substitution. The CD spectrum of rN124D, which indicates a secondary structure similar to that of rCel5A, suggested that deglycosylation could cause a slight conformational modification, which might be the reason for the reduced activity. The function of glycan chains in protein folding and structure is discussed in a number of reports (5, 24, 25). The long-range effects of local interactions between sugars and amino acid residues are observed in the polypeptide backbone. For example, fucose stabilizes hydrogen bonds on the far side of the β-sheet by making local interactions that determine the equilibrium state of the entire module for Locusta migratoria peptide C (26). On the contrary, the reduced activity of rN124H should not be completely ascribed to the loss of N-glycosylation completely but also to the substitution of His residue to Asn. According to the homology modeling structure of Cel5A obtained from the SWISS-MODEL (27), similar to other enzymes that belong to clan GH-A, Cel5A shares a (α/β)8 barrel structure, where the two catalytic residues (E218 and E329) are located inside the core of the β barrel (28, 29). The N124 site is located outside of the barrel (20, 30). Therefore, the glycosylation on this site is assumed to not intensively influence the overall structure of the recombinant enzyme.

Although rCel5A, rN124D, and rN124H shared an identical thermal stability at the optimum pH of 5.0, the two deglycosylated samples showed different thermal stabilities compared with the glycosylated rCel5A at pH 4.0 when they were incubated at 70°C. Glycosylated rCel5A showed higher thermal stability than deglycosylated rN124D. The result is expected because many previous studies showed that the increased levels of glycosylation could improve the thermotolerance of glycoproteins (31-33). Although N124 was located outside the (α/β)8 barrel, the covalent binding of glycans to the protein surface may inherently enhance the thermal and kinetic stabilities of proteins. N-linked glycosylation was hypothesized to cause a decrease in dynamic fluctuations throughout the entire molecule, which led to an increase in thermal and structural stabilities (34).

However, the function of glycosylation in improving the thermal stability of glycoprotein is ambiguous. Some reports show that high levels of glycosylation even reduce the thermostability capacity of certain proteins (35, 36). However, the decrease of thermal stability for rCel5A compared with rN124H at pH 4.0 was not assumed to be due to its glycosylation. The substitution of His residue to Asn had an essential function in improving the thermal stability for N124H at lower pH value. Although many reports presented that His residues may be responsible for maintaining the pKa of the
yet and should be investigated in the future. The CD spectrum difference between rN124H and rCel5A suggests that the substitution of His residue to Asn resulted in a secondary structure change in the protein, which may be due to the increased thermal stability and reduced enzyme activity. However, the molecular basis is not understood yet and should be investigated in the future.

MATERIALS AND METHODS

Strains, plasmids, and culture conditions
S. cerevisiae H158 [MATa leu2-3 leu2-112, ura3-52 trp1-289 his4-519 pbl1 cir+] was used as a host for the expression vector pAJ401, which carries a PGK1 gene as promoter and URA3 gene as selective marker (VTT Biotechnology, Finland). The cloning vector pUC18, which is carrying the wild-type Cel5A cDNA (19), was used as a template for PCR.

Site-directed mutagenesis
Overlap extension PCR was used to generate site-directed mutagenesis mutant (40). The N-linked glycosylation site, Asn124, was mutated by creating codons for His (CAC) and Asp (GTC).

SDS-PAGE and zymograms
SDS-PAGE analysis was performed to detect the purified protein. Zymograms were prepared according to the general procedure described by Medve (43). The 12% separating gel contained 0.15% sodium carboxymethyl cellulose (CMC-Na). The gel was washed in a solution containing 0.5 M NaAc and 25% isopropanol to remove the sodium dodecyl sulfate. The proteins were renatured in a 50 mM acetate buffer (pH 5.0) containing 5 mM mercaptoethanol by rocking the gel overnight at 4°C. The gel was transferred to a 50 mM NaAc-HAc buffer (pH 5.0) and incubated at 50°C for 2 h. The gel was stained in 0.2% Congo Red for 1 h and destained in 1 M NaCl. Clear bands against a red background indicated CMC-Na hydrolysis.

Activity measurement and effect of temperature on enzyme activity
The activities of the purified recombinant enzymes were determined by measuring the increase in reducing sugar liberated from 0.5% of CMC-Na in 50 mM acetate (pH 5.0) using the dinitrosalicylic acid reagent method (44). The hydrolysis time of CMC-Na was 30 min at 50°C. Effect of temperature on enzyme activity was tested by incubating the enzymes at 70°C in 50 mM acetate buffer at pH 5.0 and pH 4.0, followed by measuring the residual activity at different incubation time points.

CD spectroscopy
Far-UV CD spectra of the purified recombinant enzymes were measured in 10 mM acetate buffer (pH 5.0) at 25°C with a Jasco J-810 spectropolarimeter (Hachioji City, Japan) with a 3 mm-path-length cuvette and a protein concentration of 0.1 mg/ml. The protein far-UV spectra were recorded over a wave-length range of 190 nm to 240 nm. For each measurement, a total of six spectra were collected, averaged, and corrected by
N124 of Cel5A is essential for enzymatic activity
Yuqi Qin and Yinbo Qu

subtraction of a buffer blank.

ACKNOWLEDGEMENTS

We thank the grants from National Basic Research Program of China (Grant no. 2011CB707403) and National Natural Science Foundation, China (Grant No. 31030001 and 31370086).

REFERENCES

1. Kubicek, C. P. (2013) Systems biological approaches towards understanding cellulase production by Trichoderma reesei. J. Biotechnol. 163, 133-142.

2. Lynd, L. R., Weimer, P. J., van Zyl, W. H. and Pretorius, I. S. (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol. Mol. Biol. Rev. 66, 506-577.

3. Béguin, P. and Aubert, J. P. (1994) The biological degradation of cellulose. FEMS Microbiol. Rev. 13, 25-58.

4. Hui, J. P. M., White, T. C. and Thibault, P. (2002) Identification of glycan structure and glycosylation sites in cellobiohydrolase II and endoglucanases I and II from Trichoderma reesei. Glycobiology. 12, 837-849.

5. Beckham, G. T., Dai, Z., Matthews, J. F., Mornany, M., Payne, C. M., Adney, W. S., Baker, S. E. and Himmel, M. E. (2012) Harnessing glycosylation to improve cellulase activity. Curr. Opin. Biotechnol. 23, 338-345.

6. Zhou, F., Olman, V. and Xu, Y. (2009) Large-scale analyses of glycosylation in cellulases. Genomics Proteomics Bioinformatics. 7, 194-199.

7. Xia, D., Wei, Y., Zhang, G., Zhao, Q., Zhang, Y., Xiang, Z. and Lu, C. (2013) cDNA cloning, expression, and enzymatic activity of a novel endogenous cellulase from the beetle Bactocera horsfieldi. Gene 514, 62-68.

8. Zhao, X. H., Wang, W., Wang, F. Q. and Wei, D. Z. (2012) A comparative study of β-1, 4-endoglucanase (possessing β-1, 4-exoglucanase activity) from Bacillus subtilis LH expressed in Pichia pastoris GS115 and Escherichia coli Rosetta (DE3) Bioreour. Technol. 110, 539-545

9. Jeoh, T., Michener, W., Himmel, M. E., Decker, S. R. and Adney, W. S. (2008) Implications of cellobiohydrolase glycosylation for use in biomass conversion. Biotechnol Biofuels 1, 10.

10. Voutilainen, S. P., Murray, P. G., Tuohy, M. G. and Koivula, A. (2010) Expression of Talaromyces emersonii cellobiohydrolase Cel7A in Saccharomyces cerevisiae and rational mutagenesis to improve its thermostability and activity. Protein Eng. Des. Sel. 23, 69-79.

11. Gao, L., Gao, F., Wang, L., Geng, C., Chi, L., Zhao, J. and Qu, Y. (2012) N-glycoform diversity of cellobiohydrolase I from Penicillium decumbens and synergism of non-hydrolytic glycolyme in cellulose degradation. J. Biol. Chem. 287, 15906-15915.

12. Jun, H., Kieselbach, T. and Jönsson, L. J. (2011) Enzyme production by filamentous fungi: analysis of the secretome of Trichoderma reesei grown on unconventional carbon source. Microb. Cell. Fact. 10, 68.

13. Medve, J., Karlsson, J., Lee, D. and Tjemeld, F. (1998) Hydrolysis of microcrystalline cellulose by cellobiohydrodralase I and endoglucanase II from Trichoderma reesei: adsorption, sugar production pattern, and synergism of the enzymes. Biotechnol. Bioeng. 59, 621-634.

14. Suominen, P. L., Mäntylä, A. L., Karhunen, T., Hakola, S. and Nevalainen, H. (1993) High frequency one-step gene replacement in Trichoderma reesei. II. Effects of deletions of individual cellulase genes. Mol. Gen. Genet. 241, 523-30.

15. Herpoël-Gimbert, I., Margeot, A., Dolla, A., Jan, G., Mollé, D., Lignon, S., Mathis, H., Signollet, J. C., Monot, F. and Asther, M. (2008) Comparative secretome analyses of two Trichoderma reesei RUT-C30 and CL647 hyper-secretory strains. Biotechnol Biofuels 1, 18.

16. W. S. (2008) Implications of cellobiohydrolase II and endoglucanases I and II from Trichoderma reesei. Glycobiology. 12, 837-849.

17. Zhou, F., Olman, V., and Xu, Y. (2009) Large-scale analyses of glycosylation in cellulases. Genomics Proteomics Bioinformatics. 7, 194-199.

18. Stals, I., Sandra, K., Geysens, S., Contreras, R., Van Beeumen, J. and Claeyssens, M. (2004) Factors influencing glycosylation of Trichoderma reesei cellulases. I. Postsecretorial changes of the O- and N-glycosylation pattern ofCel7A. Glycobiology. 14, 713-724.

19. Qin, Y., Wei, X., Liu, X., Wang, T. and Qu, Y. (2008) Purification and characterization of recombinant endoglucanase of Trichoderma reesei expressed in Saccharomyces cerevisiae with higher glycosylation and stability. Protein Expr. Purif. 58, 162-167.

20. Lo-Leggio, L. and Larsen, S. (2002) The 1.62 A structure of Thermosascus aurantius endoglucanase: completing the structural picture of subfamilies in glycoside hydrolase family 5. FEBS Lett. 523, 103-108.

21. Greenfield, N. and Farman, G. D. (1969) Computed circular dichroism spectra for the evaluation of protein conformation. Biochemistry. 8, 4108-4116.

22. Holzwarth, G. and Doty, P. (1965) The ultraviolet circular dichroism of polysaccharides. J. Am. Chem. Soc. 87, 218-228

23. Tarentino, A. L., Plummer, T. H. Jr. and Maley, F. (1974) The release of intact oligosaccharides from specific glycoproteins by endo-beta-N-acetylglucosaminidase H. J. Biol. Chem. 249, 818-824.

24. Hellenius, A., Trombetta, E. S., Herbert, D. N. and Simons, J. F. (1997) Calnexin, calreticulin and the folding of glycoproteins. Trends. Cell. Biol. 7, 193-200.

25. Mansell, T. J., Guarino, C. and Delisa, M. P. (2013) Engineered genetic selection links in vivo protein folding and stability with asparagine-linked glycosylation. Biotechnol. J. doi: 10.1002/biot.201300026.

26. Kaushik, S., Mohanty, D. and Surolia, A. (2012) Molecular dynamics simulations on pars intercerebralis major pep-
tide-C (PMP-C) reveal the role of glycosylation and disulfide bonds in its enhanced structural stability and function. 

27. Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003) SWISS-MODEL: an automated protein homology-modelling server. Nucleic Acids Res. 31, 3381-3385.

28. Macarrón, R., Acébal, C., Castillón, M. P., Dominguez, J. M., de la Mata, I., Pettersson, G., Tomme, P. and Claeyssens, M. (1993) Mode of action of endoglucanase III from Trichoderma reesei. Biochem. J. 289, 867-873.

29. Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003) SWISS-MODEL: an automated protein homology-modelling server. Nucleic Acids Res. 31, 3381-3385.

27. Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003) SWISS-MODEL: an automated protein homology-modelling server. Nucleic Acids Res. 31, 3381-3385.

28. Macarrón, R., Acébal, C., Castillón, M. P., Dominguez, J. M., de la Mata, I., Pettersson, G., Tomme, P. and Claeyssens, M. (1993) Mode of action of endoglucanase III from Trichoderma reesei. Biochem. J. 289, 867-873.

29. Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003) SWISS-MODEL: an automated protein homology-modelling server. Nucleic Acids Res. 31, 3381-3385.

28. Macarrón, R., Acébal, C., Castillón, M. P., Dominguez, J. M., de la Mata, I., Pettersson, G., Tomme, P. and Claeyssens, M. (1993) Mode of action of endoglucanase III from Trichoderma reesei. Biochem. J. 289, 867-873.

29. Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003) SWISS-MODEL: an automated protein homology-modelling server. Nucleic Acids Res. 31, 3381-3385.

28. Macarrón, R., Acébal, C., Castillón, M. P., Dominguez, J. M., de la Mata, I., Pettersson, G., Tomme, P. and Claeyssens, M. (1993) Mode of action of endoglucanase III from Trichoderma reesei. Biochem. J. 289, 867-873.

29. Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003) SWISS-MODEL: an automated protein homology-modelling server. Nucleic Acids Res. 31, 3381-3385.

28. Macarrón, R., Acébal, C., Castillón, M. P., Dominguez, J. M., de la Mata, I., Pettersson, G., Tomme, P. and Claeyssens, M. (1993) Mode of action of endoglucanase III from Trichoderma reesei. Biochem. J. 289, 867-873.

29. Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003) SWISS-MODEL: an automated protein homology-modelling server. Nucleic Acids Res. 31, 3381-3385.

28. Macarrón, R., Acébal, C., Castillón, M. P., Dominguez, J. M., de la Mata, I., Pettersson, G., Tomme, P. and Claeyssens, M. (1993) Mode of action of endoglucanase III from Trichoderma reesei. Biochem. J. 289, 867-873.

29. Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003) SWISS-MODEL: an automated protein homology-modelling server. Nucleic Acids Res. 31, 3381-3385.