รายงานวิจัยฉบับสมบูรณ์

โครงการ การผลิตรีคอมบิแนนท์โปรตีน และการศึกษาหน้าที่ของกลูคาโนไฮโดรเลสจากข้าว

1. อ. ดร. รจนา โอภาสศิริ สำนักวิชาวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี
2. รศ. ดร. เจมส์ เกตุทัต-คาร์นส์ สำนักวิชาวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี
3. Dr. Takashi Akiyama National Agricultural Research Center for Hokkaido Region

สนับสนุนโดยสำนักงานคณะกรรมการการอุดมศึกษาและสำนักงานกองทุนสนับสนุนการวิจัย

สัญญาเลขที่ MRG4880066

รายงานวิจัยฉบับสมบูรณ์

โครงการ การผลิตรีคอมบิแนนท์โปรตีน และการศึกษาหน้าที่ของกลูคาโนไฮโดรเลสจากข้าว

1. อ. ดร. รจนา โอภาสศิริ	สำนักวิชาวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี
2. รศ. ดร. เจมส์ เกตุทัต-คาร์นส์	สำนักวิชาวิทยาศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี
3. Dr. Takashi Akiyama	National Agricultural Research Center for Hokkaido Region

สนับสนุนโดยสำนักงานคณะกรรมการการอุดมศึกษาและสำนักงานกองทุนสนับสนุนการวิจัย

บทคัดย่อ

รหัสโครงการ : MRG4880066

ชื่อโครงการ : การผลิตรีคอมบิแนนท์โปรตีนและการศึกษาหน้าที่ของกลูคาโนไฮโดรเลสจากข้าว

ชื่อนักวิจัย : อ. ดร. รจนา โอภาสศิริ

E-mail Address : opassiri@hotmail.com

ระยะเวลาโครงการ: 1 มิถุนายน 2548 – 31 พฤษภาคม 2550

ในการศึกษานี้ พบยีนของข้าวที่จัดอยู่ในกลุ่มเอนไซม์กลูคาโนไฮโดรเลสอย่างน้อย 139 ยีน ในฐานข้อมูลจีโนมข้าวและ Carbohydrate Active Enzyme database และยีนเหล่านี้ถูกจัดแยกไว้ใน 7 กลุ่มย่อยของ Glycosyl hydrolases family (GH) ได้นำเอนไซม์ในกลุ่มกลูคาโนไฮโดรเลส 2 ชนิด ที่อยู่ใน GH1 และ GH5 มาศึกษาการเร่งปฏิกิริยาและหน้าที่ ได้เพิ่มจำนวน cDNA ขนาดเต็มสายของยีน GH1 *Os4bglu12* β-glucosidase จากต้นอ่อนข้าวด้วยวิธี RT-PCR ลำดับกรดอะมิโนซึ่งแปลมาจากลำดับนิวคลีโอไทด์ของ *Os4bglu12* ประกอบด้วยกรดอะมิโน 510 ตัว ซึ่งกรดอะมิโนทางด้านปลายอะมิโนของ *Os4bglu12* มีลำดับคล้ายกับของ cell wall-bound β-glucosidase ที่เคยถูกแยกได้จากข้าว จำนวนถึง 40 ตัว จากทั้งหมด 44 ตัว โปรตีน Os4bglu12 ซึ่งต่ออยู่กับโปรตีนไทโอรีดอกซินที่ผลิตได้ใน *Escherichia coli* ในรูปรีคอมบิแนนท์โปรตีน สามารถย่อยกลูโคโอลิโกแซคคาไรด์ที่มีกลูโคสต่อกันอยู่ 3-6 หน่วย ด้วยพันธะไกลโคซิดิกแบบ β-(1,4) และไดแซคคาไรด์ที่มีกลูโคสต่อกันด้วยพันธะไกลโคซิดิกแบบ β-(1,3) ได้ดีมาก ได้เพิ่มจำนวน cDNA ของ GH5 glucan 1,3-β-glucosidases ที่ให้ชื่อว่า *GH5BG* จากต้นอ่อนข้าวด้วยวิธี RT-PCR คาดว่าโปรตีนนี้จะถูกส่งออกนอกเซลล์โดยดูจากลำดับกรดอะมิโนที่เป็น signal peptide GH5BG ซึ่งประกอบไปด้วยกรดอะมิโน 510 ตัว มีบริเวณสำคัญ 2 ส่วน ได้แก่ β-1,3-exoglucanase-like domain และ fascin-like domain ซึ่งไม่ค่อยพบในเอนไซม์พืช โปรตีน GH5BG ซึ่งต่ออยู่กับโปรตีนไทโอรีดอกซินที่ผลิตได้ในรูปรีคอมบิแนนท์โปรตีนสามารถย่อยกลูโคโอลิโกแซคคาไรด์ที่ต่อด้วยพันธะไกลโคซิดิกแบบ β-(1,4) และไดแซคคาไรด์ที่มีกลูโคสต่อกันด้วยพันธะไกลโคซิดิกแบบ β-(1,3) ได้ดีมาก ค่าคงที่ทางจลนศาสตร์ที่บ่งถึงประสิทธิภาพการเร่งปฏิกิริยา (k_{cat}/K_m) ของเอนไซม์นี้ในการย่อยกลูโคโอลิโกแซคคาไรด์แบบ β-(1,4) ที่มีกลูโคสต่อกันตั้งแต่ 3 ถึง 5 หน่วย มีค่าคงที่ ซึ่งแสดงให้เห็นว่าเอนไซม์นี้ควรจัดเป็น β-glucosidase

พบว่าทั้ง Os4bglu12 และ GH5BG สามารถย่อย *p*NP-β-glycosides ได้หลายชนิด ซึ่งแสดงให้เห็นว่าที่ subsite ตำแหน่ง -1 ของเอนไซมทั้งสองมีความจำเพาะต่อชนิดน้ำตาลที่เข้าจับต่ำ เอนไซม์ทั้งสองชนิดไม่สามารถย่อยกลูโคโอลิโกแซคคาไรด์และโพลิเมอร์ของกลูโคส ที่มีกลูโคสต่อกันด้วยพันธะไกลโคซิดิกแบบ β-(1,3) and 1,3,-1,4-β-glucans ได้ ผลการทดลองนี้แสดงให้เห็นว่าเอนไซม์ทั้งสองชนิดมีการทำงานแบบ exoglucanase หรือ β-glucosidase ซึ่งสอดคล้องกับบทบาทในเมทาโบลิซึมของผนังเซลล์

คำหลัก: กลูคาโนไฮโดรเลส, บีต้ากลูโคซิเดส, ข้าว, รีคอมบิแนนท์โปรตีน, ผนังเซลล์, โอลิโกแซคคาไรด์

Abstract

Project Code: MRG4880066
Project Title: Recombinant protein expression and functional characterization of glucanohydrolase from rice
Investigators: Dr. Rodjana Opassiri
E-mail Address: opassiri@hotmail.com
Project Period: 1 June 2005 – 31 May 2007

In this study, at least 139 rice genes homologous to glucanohydrolases which are classified as the members in 7 subfamilies of glycosyl hydrolase family (GH) has been identified in rice genome database and Carbohydrate Active Enzyme database. The catalytic activities and functions of two members of rice glucanohydrolases in GH1 and GH5 were analyzed. The full-length cDNA of the GH1 gene, *Os4bglu12* β-glucosidase, was cloned from rice seedlings by RT-PCR. The isolated *Os4bglu12* cDNA encoded a 510 amino acid long precursor protein which is identical at 40 of 44 amino acid residues with the N-terminal amino acid sequence of a cell wall-bound enzyme previously purified from germinating rice. A thioredoxin-Os4bglu12 fusion protein expressed in *Escherichia coli* efficiently hydrolyzed β-(1,4)-linked oligosaccharides of 3-6 glucose residues and β-(1,3)-linked disaccharide. The cDNA for GH5 glucan 1,3-β-glucosidases, designated *GH5BG*, was cloned from rice seedlings by RT-PCR. The protein was predicted to be extracellular as judged by the signal peptide sequence. A 510 amino acid mature protein of GH5BG contains two major domains, a β-1,3-exoglucanase-like domain and a fascin-like domain, which is not commonly found in plant enzymes. A recombinant thioredoxin-*GH5BG* fusion protein exhibited a marked preference for β-(1,4)-linked oligosaccharides and β-(1,3)-linked disaccharide. The catalytic efficiency (k_{cat}/K_m) values for hydrolysis of β-(1,4)-linked oligosaccharides by the enzyme remained constant as the degree of polymerization (DP) increased from 3 to 5, which indicated the enzyme can be classified as a β-glucosidase.

Os4bglu12 and GH5BG was found to hydrolyze many kinds of *p*NP β-glycosides which indicates the low stringency at the -1 subsite of both enzymes. Hydrolysis of β-(1,3)-linked oligosaccharides with DP more than 2, laminarin and 1,3, 1,4-β-glucans by both enzymes could not be detected. These results indicated that both enzymes has exoglucanase/β-glucosidase activity, consistent with a role in cell wall metabolism.

Keywords: Glucanohydrolase, β-glucosidase, rice, recombinant protein, cell wall, oligosaccharides

สรุปย่อ (Executive summary)

The purpose of this project is to study the catalytic activities and functions of glucanohydrolases in rice. This work was started from analysis of the coding DNA sequence that can be grouped into glucanohydrolases in rice genome database and Carbohydrate Active Enzyme database. Most of rice glucanohydrolases were found in 7 subfamilies of glycosyl hydrolase family (GH). In order to study the functions the new glucanohydrolases, 2 rice β-glucosidase cDNAs, *Os4bglu12* and *GH5BG*, were cloned and recombinant proteins expressed in *E. coli* as soluble active form.

To study the function of glucanonohydrolaes in GH1, the full-length cDNA of the gene that most closely matched the previously reported cell-wall-bound β-glucosidase protein sequence purified from germinating rice, *Os4bglu12* β-glucosidase, was cloned by RT-PCR with the primers derived from gene contig number AAAA02014151 in rice gemone database and the first-strand cDNA synthesized from the total RNA of rice seedlings as the template. The isolated *Os4bglu12* cDNA encoded a 510 amino acid long precursor protein which is identical at 40 of 44 amino acid residues with the N-terminal amino acid sequence of a cell wall-bound enzyme. To express the protein in *E. coli,* the cDNA encoding mature proteins of rice *Os4bglu12* was cloned by RT-PCR and inserted into pENTR-D/TOPO Gateway entry vector and transferred to the pET32a+/DEST Gateway expression vector for expression. The recombinant pET32a+/DEST-*Os4bglu12* plasmid was transformed into OrigamiB (DE3) *E. coli*. A thioredoxin-Os4BGlu12 fusion protein expressed in *E. coli* efficiently hydrolyzed β-(1,4)-linked oligosaccharides of 3-6 glucose residues and β-(1,3)-linked disaccharide. This substrate specficity was similar to those reported for rice cell-wall-bound, rice BGlu1, and barley β-glucosidases, but not identical, which suggests that small sequence differences likely determine specific functional properties. Hydrolysis of β-(1,3)-linked oligosaccharides with DP > 2, laminarin and 1,3, 1,4-β-glucans by this enzyme could not be detected. Hydrolysis of *p*NP-glycosides with different sugars was used to assess Os4BGlu12 glycone specificity, and it was found to hydrolyze many kinds of *p*NP β-glycosides.

To assess the role of glucanonohydrolaes in GH5, the cDNA for GH5 glucan 1,3-β-glucosidases, designated *GH5BG*, was cloned from rice seedlings by RT-PCR with cDNA synthesized from rice seedling RNA and primers derived from the GenBank accession number AC107314 and AK065000. *GH5BG* cDNA included a reading frame encoding a 510 amino acid precursor protein. The protein was predicted to be extracellular. The mature protein is a

member of a rice-specific subfamily of GH5 proteins that contain two major domains, a β-1,3-exoglucanase-like domain and a fascin-like domain, which is not commonly found in plant enzymes. To express the protein in *E. coli,* the cDNA encoding mature proteins of rice *Os4bglu12* β-glucosidase was cloned into the pET32a+/DEST Gateway expression vector and transformed to Origami B (DE3) *E. coli*. A recombinant thioredoxin-GH5BG fusion protein produced in *E. coli* showed high hydrolytic activity towards various kinds of *p*NP-glycosides, including *p*NP-β-D-glucoside, *p*NP-α-L-arabinoside, *p*NP-β-D-fucoside, *p*NP-β-D-galactoside, *p*NP-β-D-xyloside and *p*NP-β-D-cellobioside. These results indicate there is low stringency at the -1 subsite of GH5BG. The GH5BG exhibited a marked preference for β-(1,4)-linked oligosaccharides and β-(1,3)-linked disaccharide. The catalytic efficiency (k_{cat}/K_m) values for hydrolysis of β-(1,4)-linked oligosaccharides by the enzyme remained constant as the DP increased from 3 to 5, which indicated the enzyme can be classified as a β-glucosidase. This substrate specificity is significantly different from fungal GH5 exoglucanases, such as *Candida albicans* exo-β-(1,3)-glucanase for which a 3D structure has been solved, which indicates that the structural differences with these related enzymes may effect functional differences. Although GH5BG was designated a putative glucan 1,3-β-glucosidases based on sequence homology, its catalytic activity is somewhat like GH1 β-glucosidases, which show similar oligosaccharide preferences.

These results indicated that both enzymes has exoglucanase/β-glucosidase activity that might involve in the cleavage of oligosaccharides release from the hydrolysis of cell wall by endoglucanses, which is consistent with a role in cell wall metabolism.

1. บทนำ (Introduction)

ที่มาของงานวิจัยเรื่องนี้มาจากความสนใจในการศึกษาถึงความสัมพันธ์ระหว่างโครงสร้างและหน้าที่ของเอนไซม์ในกลุ่ม glucanohydrolases ซึ่งมีบทบาทในกระบวนการสร้างและสลายผนังเซลล์พืช ผนังเซลล์พืชทุกชนิดประกอบขึ้นมาจากโพลีแซคคาไรด์หลากหลายชนิด โดยทั่วไปจะประกอบด้วย cellulose ประมาณ 15-40% hemicellulose และ pectin ประมาณ 30-40% และ lignin ประมาณ 20% (Doy and Kosugi, 2004) cellulose ทำหน้าที่เป็นโมเลกุลโครงสร้าง ในขณะที่ hemicellulose และ pectin ทำหน้าที่เป็นโครงข่ายเชื่อมโยง celluloses เข้าด้วยกันด้วยพันธะไฮโดรเจน พืชแต่ละชนิดมีส่วนประกอบและการจัดเรียงตัวของโพลิเมอร์ที่ผนังเซลล์แตกต่างกัน และแม้แต่ในเนื้อเยื่อต่างๆ ของพืชชนิดเดียวกันก็ยังแตกต่างกัน

แหล่งของผนังเซลล์ซึ่งมีอยู่เป็นปริมาณมากที่เหลือทิ้งจากภาคเกษตรกรรม เช่น ฟางข้าว ฟางข้าวบาร์เลย์และซังข้าวโพด สามารถนำมาใช้เป็นวัตถุดิบในการแปรรูปเป็นสารผลิตภัณฑ์ของน้ำตาลได้ โครงสร้างพื้นฐานของผนังเซลล์พืชในกลุ่มธัญพืชและพืชตระกูลหญ้าแตกต่างจากพืชมีดอกอื่นๆ ตรงที่พืชเหล่านี้มี glycans ที่มีความซับซ้อนและหลากหลายมาก เรียก glycans เหล่านี้ว่า 1,3-1,4-β-D-glucans (เรียกโดยย่อว่า β-glucans) β-glucans เป็นโพลิเมอร์ที่ไม่แตกแขนงที่ประกอบไปด้วย cellotriose และ cellotetraose ในสัดส่วน 2.5 ต่อ 1 ซึ่งเชื่อมกันด้วยพันธะไกลโคซิดิกแบบ β-(1,3) (Carpita and McCann, 2000) β-glucans จะแทรกตัวและเชื่อมโยงมัด cellulose เข้าไว้ด้วยกัน ซึ่งช่วยให้ผนังเซลล์มีความแข็งแรง (Carpita, 1996)

เอนไซม์ที่เกี่ยวข้องกับการย่อยผนังเซลล์พืชเรียกรวมว่า glucanohydrolases ซึ่งเป็นเอนไซม์กลุ่มย่อยที่ถูกจัดไว้ใน glycosyl hydrolases family (GH) และพบเอนไซม์นี้ในสิ่งมีชีวิตกลุ่มอื่นทั้งสัตว์และจุลินทรีย์ด้วย glucanohydrolases แบ่งได้เป็น 3 ประเภท ได้แก่ endoglucanase, exoglucanases และ β-glucosidases เอนไซม์ endo-glucanses ย่อยพันธะที่อยู่ภายใน cellulose จนได้โอลิโกแซคคาไรด์สายสั้นลง การที่เอนไซม์นี้จะย่อยสับสเตรทได้ สับสเตรทจะต้องมีจำนวนกลูโคสอย่างน้อยสี่หน่วย (Hatfield and Nevins, 1987) เอนไซม์ exoglucanase ย่อยกลูโคสออกจากโอลิโกแซคคาไรด์ได้ทั้งจากปลายด้าน reducing หรือ non-reducing (Huber and Nevins, 1981) ส่วน β-glucosidase สามารถย่อยโอลิโกแซคคารด์จากปลาย reducing ends ได้เท่านั้น (Hrmova and Fincher, 2001a; Opassiri, 2004) มีการตั้งสมมุติฐานว่าเอนไซม์เหล่านี้ที่พบในพืชวงศ์หญ้าอาจถูกส่งออกไปที่บริเวณผนังเซลล์เพื่อย่อย cellulose และ β-glucans (Taiz, 1984; Buchanan et al., 2000) เอนไซม์ glucanase บางชนิดในพืชยังเข้าไปมีส่วนร่วมในระบบป้องกันเซลล์พืช โดยช่วยย่อยผนังเซลล์ของเชื้อราจนเป็นโอลิโกแซคคาไรด์สายสั้นๆ โอลิโกแซคคาไรด์เหล่านี้ทำหน้าที่เป็นตัวกระตุ้น (elicitor) การทำงานของยีนหลายชนิดที่ตอบสนองค่อการบุกรุกของเชื้อ ในระบบการส่งสัญญาณของเซลล์ (Yamaguchi et al., 2000).

Hrmova and Fincher (2001a) เปรียบเทียบโครงสร้างของ glycosyl hydrolases ของพืช และอธิบายความสัมพันธ์ระหว่างโครงสร้างกับความจำเพาะต่อสับสเตรทของเอนไซม์ endohydrolase, exohydrolase และ β-glucosidase การที่เอนไซม์ทั้งสามชนิดมีความจำเพาะต่อสับสเตรทแตกต่างกันเนื่องจากมีรูปร่างของบริเวณจับกับสับสเตรท (binding sites) ที่แตกต่างกัน โดยทั่วไป endohydrolases จะมีบริเวณจับกับสับสเตรทเป็นร่องยาว (grooves หรือ depressions) ตรงผิวนอกของโปรตีน ดังนั้น endohydrolase จึงสามารถจับกับบริเวณใดก็ได้ของสับสเตรทที่เป็นโพลิเมอร์ แล้วจึงย่อยพันธะที่อยู่ภายในโพลิเมอร์นั้นด้วยกรดอะมิโนที่เป็นหมู่เร่งที่ฝังตัวอยู่ตรงบริเวณนั้น บริเวณจับกับสับสเตรทของ exohydrolase มีรูปร่างเป็นแบบท่อหรือโพลงที่ลึกเข้าไปจากผิวโปรตีน (dead-end tunnel, slot, funnel) ความจำเพาะต่อสับสเตรทจึงขึ้นกับว่าสับสเตรทมีรูปร่างพอดีกับรูปร่างของ

บริเวณจับมากพอที่จะยื่นเข้าไปสัมผัสกับกรดอะมิโนหมู่เร่งที่อยู่ก้นของท่อนั้นได้มากเพียงใด (Hrmova and Fincher, 2001a, 2001b) บริเวณจับของ β-glucosidase มีรูปร่างเป็นช่องแคบและตื้นที่พอจะให้กลูโคสแค่ 2 หน่วยยื่นเข้าไปได้ โดยส่วนที่เหลือจะยื่นออกมาด้านนอก (Hrmova and Fincher, 2001a)

จากข้อมูลลำดับนิวคลีโอไทด์ในจีโนมของพืชหลายชนิด เช่น Arabidopsis, ข้าว, ข้าวบาร์เลย์ และข้าวโพด พบว่า glucanohydrolases ที่มีอยู่เป็นจำนวนมากมีความหลากหลายทางด้านโครงสร้างและลำดับกรดอะมิโนเป็นอย่างมาก และมีบทบาทหลากหลาย ได้แก่ การสังเคราะห์ ย่อยสลาย และเปลี่ยนแปลงโครงสร้างของโอลิโกแซคคาไรด์และโพลีแซคคาไรด์ (Hoj and Fincher, 1995; Davies and Henrissat, 2002) สามารถจำแนกยีนของ glucanohydrolases ประมาณ 400 ยีน ออกเป็นกลุ่มย่อยได้ 9 กลุ่ม endoglucanases ถูกจัดไว้ใน GH9, 10, 16, 17, 51และ 81 , exoglucanases ถูกจัดไว้ในกลุ่ม GH1, 3, และ 5 และ β-glucosidase ถูกจัดไว้ใน GH1 ปัจจุบันยังไม่สามารถศึกษากลไกการสร้างและสลายผนังเซลล์พืชที่เกิดขึ้นจริงๆ ในเซลล์พืชได้ เนื่องจากกระบวนนี้มีความซับซ้อนที่ต้องใช้เอนไซม์เข้าร่วมหลายชนิด ปัจจุบันมีเอนไซม์ glucanohydrolases เพียงบางชนิดในพืชได้ถูกนำมาศึกษาถึงบทบาทหน้าที่และความจำเพาะต่อสับสเตรต มียีนของ glucanohydrolases ของข้าวเพียงไม่กี่ยีนที่ได้มีการศึกษามา และเอนไซม์ที่ได้มีการศึกษาส่วนมากจะอยู่ในกลุ่ม GH 1, 3, 17 (Akiyama et al, 2004; Nishizawa et al, 2003; Opassiri et al, 2003; Zheng et al, 2003; Yamaguchi et al, 2002; Thomas et al, 2000; Rakwal et al, 1999) เป็นเรื่องยากที่จะคาดเดาความจำเพาะในการทำงานของเอนไซม์เหล่านี้จากลำดับกรดอะมิโนเพียงอย่างเดียว ในข้าวมียีนอีกอย่างน้อย 130 ชนิดที่ยังไม่ได้มีการศึกษาว่ามีกลไกการทำงานอย่างไร มีความจำเพาะต่อสับสเตรตแตกต่างกันหรือไม่ ยีนที่ควบคุมการสังเคราะห์เอนไซม์ดังกล่าวแสดงออกในช่วงเวลาหรือสภาวะใด เพื่อที่จะตอบคำถามดังกล่าวจึงมีความจำเป็นที่จะต้องแยกเอนไซม์ชนิดออกมาศึกษาคุณสมบัติต่างๆ ตลอดจนการศึกษาโครงสร้าง เพื่อเข้าใจถึงกลไกการทำงานของเอนไซม์ แล้วนำข้อมูลมาวิเคราะห์ถึงความแตกต่างและสัมพันธ์ในการทำงานระหว่างเอนไซม์แต่ละชนิด ความรู้ที่ได้จากการศึกษาดังกล่าวจะเป็นประโยชน์ในการนำเอนไซม์ชนิดนี้ไปใช้ร่วมกับเอนไซม์อื่นๆ ที่มีความจำเพาะกับสับสเตรตในรูปแบบที่แตกต่างกันไปใช้ในการย่อยสลายองค์ประกอบต่างๆ ในผนังเซลล์พืช เช่น cellulose และ glycans ซึ่งเป็น biomass ที่สามารถนำไปใช้ประโยชน์ในอุตสาหกรรมการผลิตสารต่างๆ เช่น แอลกอฮอร์, สารเชื้อเพลิง, กรดอะมิโน, กระดาษ, สารเพิ่มคุณค่าอาหารสัตว์ และยารักษาโรค เป็นต้น หรือทำวิศวกรรมเอนไซม์เพื่อให้มีหน้าที่เฉพาะตามที่ต้องการที่จะนำไปใช้ประโยชน์

2. วัตถุประสงค์ของโครงการ (Objectives)

1. เปรียบเทียบและหาความสัมพันธ์ของลำดับกรดอะมิโนของเอนไซม์ glucanohydrolases แต่ละชนิดใน subfamily ต่างๆ
2. เพิ่มจำนวนชิ้นส่วน cDNA ของยีนบางชนิดในกลุ่ม glucanohydrolase จากข้าวโดยใช้วิธี RT-PCR
3. ผลิตรีคอมบิแนนท์โปรตีนของ glucanohydrolase ของข้าว ในแบคทีเรียหรือยีสต์
4. ศึกษาสมบัติต่างๆ ทางชีวเคมี ของเอนไซม์และความจำเพาะในการย่อยสลาย glycans ชนิดต่างๆ

3. วิธีการทดลอง (Materials and Methods)

3.1 Database searching and sequence analysis

Identification of rice genes homologous to glucanohydrolases was done using the BLAST suite of programs (Altschul et al., 1997) in GenBank at NCBI (/www.ncbi.nlm.nih.gov), Beijng Genomic Institute (http://btn.genomics.org.cn/rice/) and Carbohydrate Active Enzyme Database (http://afmb.cnrs-mrs.fr/CAZY). Identification of homologous genes and cDNAs was done using tBLASTn and the known protein sequences derived from Entrez search at GenBang NCBI. Coding regions of genes were identified by BLASTx searches against the NCBI nr protein database. All gene/protein clone IDs were retrieved and collected in the catalog and grouped into each glycosyl hydrolase subfamily.

3.2 Plant materials and growth conditions

Rice (*Oryza sativa* L. spp. *indica* cv. KDML105 and spp. *japonica* cv. Yukihikari) seeds were germinated in the dark from day 0 to day 3 and in 12 h light-12 h dark from day 4 to day 7 at 28°C on germinating paper moistened with sterile distilled water. The whole seedlings were harvested and kept at -70°C for RNA isolation.

3.3 Cloning and recombinant protein expression of rice Os4bglu12 β-glucosidase cDNA

Total RNA was isolated from 100 mg 5-6-d-old rice seedlings using Trizol Reagent (Invitrogen , Carlsbad, CA). The total RNA (5 μg) was used as the template to synthesize the first-strand cDNA with SuperScript II reverse transcriptase according to the manufacturer's protocol (Invitrogen). Primers for amplifying the full-length coding sequence (CDS) cDNA (designated *Os4bglu12*) and a cDNA encoding the mature protein of rice *Os4bglu12* β-glucosidase were designed from the GenBank *indica* rice genome contig number AAAA02014151 and AK100820 and AK105375 cDNA sequences (Kikuchi *et al*., 2003). A 5' sense primer *Os4bglu12*_fullf (5'-TGTCCATGGCGGCAGCAG-3'), and the antisense primer *Os4bglu12*_3'UTRr (5'-AACTGGATTACTTCCATCTC-3') were used to amplify the full coding cDNA. The amplification was done with 30 cycles of 94°C, 30 s, 53°C 30 s and 72°C 4 min, using *Pfu* DNA polymerase (Promega, Madison, WI). A full-length product was cloned into the *Eco*R V site of pBlueScript II SK+ (Stratagene, La Jolla, CA), and sequenced.

To express the protein in *Escherichia coli,* the cDNA encoding mature proteins of rice *Os4bglu12* β-glucosidase was cloned by RT-PCR and inserted into pENTR-D/TOPO Gateway entry vector and transferred to the pET32a+/DEST Gateway expression vector for expression. The cDNA encoding mature proteins of the *Os4bglu12* was PCR amplified using cDNA cloned as the template with the *Os4bglu12*matNcoIf (5'-CACCATGGCCTACAATAGCGCCGGCG AG-3') and *Os4bglu12*stopr (5'-ATCATTTCAGGAGGAACTTCTTG-3') primers and *Pfu* DNA polymerase to introduce a direction cloning site at the 5' end. The amplification was done as above, but with 45°C annealing temperature. The PCR product was cloned into pENTR-D/TOPO Gateway entry vector, according to the supplier directions (Invitrogen) and sequenced.

The cDNA insert in the pENTR-D/TOPO vectors was subcloned into the pET32a+/DEST Gateway expression vector by LR Clonase recombination following the recommended protocols (Invitrogen) and thoroughly sequenced. The recombinant pET32a+/DEST-*Os4bglu12* plasmid was transformed into OrigamiB (DE3), and positive clones were selected with 15 μg/ml kanamycin, 12.5 μg/ml tetracycline and 50 μg/ml ampicilin LB-agar plate. For recombinant protein expression, the selected clones were grown in selecting LB medium at 37°C until the optical density at 600 nm reached 0.5-0.6, IPTG was added to a final concentration of 0.3 mM and the cultures were incubated at 20°C for 8 h. Induced cultures were harvested by centrifugation at 5000×g at 4°C for 10 min. The cell pellets were resuspended in freshly prepared extraction buffer (50 mM phosphate buffer pH 8.0, 200 μg/ml lysozyme, 1% Triton-X 100, 1 mM phenylmethylsulfonylfluoride (PMSF), 40 μg/ml DNase I), and incubate at room temperature for 30 min. The soluble protein was recovered by centrifugation at 12,000xg at 4°C for 10 min. All of the protein samples were subjected to SDS-PAGE by standard methods (Laemli, 1970).

3.4 Cloning and recombinant protein expression of GH5BG β-glucosidase cDNA

Total RNA was isolated from 100 mg 5-6-d-old rice cv. KDML105 seedlings with Trizol Reagent, and 5 μg of total RNA was used as the template to synthesize the first-strand cDNA with SuperScript II reverse transcriptase according to the manufacturer's protocol (Invitrogen, Carlsbad, CA). The GenBank rice genome contig accession number AC107314 (deduced protein sequence GenBank AC AAAM08614) and AK065000 cDNA sequences (Kikuchi *et al*., 2003) were used to design the primers to amplify a full-length coding sequence (CDS) cDNA and a cDNA encoding the mature protein of rice glycosyl hydrolase family 5 β-glucosidase (designated *GH5BG*). The 5' sense primer AK065000f (5'-GCTGAAAAATCTTC GTCTTCATC-3') and the antisense primer AAM08614EcoRIr (5'-CCATCCAACTGGAATTC

TCACTG-3') were used to amplify a 774 bp-5' PCR fragment. The 5' sense primer AAM08614EcoRIf (5'-CGCAGTGAGAATTCCAGTTG-3') and the antisense primer AK065000r (5'-CTTCACAAGAGAAAGTTACACTC-3') were used to amplify a 1016 bp-3' PCR fragment. The amplification for 5' and 3' PCR fragments was done with 30 cycles of 94°C, 30 s, 51°C 30 s and 72°C 4 min, using *Pfu* DNA polymerase (Promega, Madison, WI) with the first-strand cDNA as the template. Finally, the AK065000f and AK065000r primers were used to amplify a full CDS cDNA by using the 5' and 3' PCR cDNA fragments amplified above as template in overlapping PCR. The amplification was done as above, but with 45°C annealing temperature. A full-length product was cloned into the *Eco*R V site of pBlueScript II SK+ (Stratagene, La Jolla, CA), and sequenced. Protein analyses were done at the Expasy proteomics server (http://www.expasy. org/), predictions of signal sequences and cellular location were done were done with SignalP (Bendtsen, 2004) and PSORT (Nakai and Horton, 1999), respectively.

The cDNA encoding the predicted mature protein of the *GH5BG* was PCR amplified with the cloned full-length cDNA as the template, the AAM08614matNcoIf (5'-CACCATGGT CTCCGATGGGAGGACG-3') and AAM08614XhoIstopr (5'-CCCTCGAGCTAGCTTTTGA GAGAGATGATCC-3') primers and *Pfu* DNA polymerase to introduce an *Nco*I site at the 5' end and an *Xho*I site at the 3' end. The amplification was done as described above, but with 45°C annealing temperature. The cDNA product was digested with *Nco*I and *Xho*I, cloned into pENTR4 Gateway entry vector that had been digested with the same restriction enzymes, and subcloned into the pET32a+/DEST Gateway expression vector by LR Clonase recombination following the recommended protocols (Invitrogen) and thoroughly sequenced. The recombinant pET32a+/DEST-*GB5BG* plasmid was transformed into *E. coli* strain OrigamiB (DE3) by heat shock (Sambrook *et al.*, 1989), and positive clones were selected on 15 μg/ml kanamycin, 12.5 μg/ml tetracycline and 50 μg/ml ampicillin LB-agar.

To produce the protein, selected clones were grown in LB medium containing the same antibiotics as above at 37°C until the optical density at 600 nm reached 0.5-0.6. Then, IPTG was added to a final concentration of 0.5 mM and the cultures were incubated at 20°C for 12 h. Induced cultures were harvested by centrifugation at 3000 × g at 4°C for 10 min. The cell pellets were resuspended in freshly prepared extraction buffer (50 mM phosphate, pH 8.0, 200 μg/ml lysozyme, 1% Triton-X 100, 1 mM phenylmethylsulfonylfluoride, 40 μg/ml DNase I), and incubated at room temperature for 30 min. The soluble protein was recovered by centrifugation at 12,000 x g at 4°C for 10 min. The expressed thioredoxin-GH5BG fusion protein was purified by immobilized metal affinity chromatography (IMAC) with cobalt BD TALON column according to the manufacturer's instructions (Clontech, Palo Alto, CA). The fractions with *p*NPG hydrolysis activity were pooled and concentrated with 10 kDa-cut off centrifugal ultrafiltration membranes (YM-10, Amicon, Beverly, MA). All of the protein samples were analyzed by SDS-PAGE by standard methods (Laemmli, 1970).

3.5 Enzyme assays

The purified thioredoxin-Os4bglu12 and -GH5BG recombinant protein was tested against *p*NP (*p*NP) derivatives of monosaccharides and cellobiose to determine sugar specificity. In a 100 µL reaction assay volume, 1.47-2.94 pmol enzyme was incubated with substrate in 50 mM sodium acetate, pH 5.0, at 37°C, except for the assay with *p*NP-β-D-cellobioside, in which 29.4 pmol enzyme was used. At the end of the reaction time, 70 µL of 0.4 M sodium carbonate was added to stop the reaction, and the absorbance of the liberated *p*NP was measured at 405 nm. The enzymes were tested with oligosaccharides including cello-oligosaccharides with DP of 2-6, laminari-oligosaccharides DP 2-5 and gentiobiose. In a 50 µL reaction volume, 0.74 pmol enzyme was incubated with substrate in 50 mM sodium acetate, pH 5.0, for 5 min at 37°C, and the reaction was stopped by boiling. The release of the glucose was determined by the PGO (peroxidase/glucose oxidase) assay method (Babcock and Esen, 1994; Opassiri *et al*. 2003).

The enzyme was also tested for hydrolysis of polysaccharides. In the assay, 1-5 µg enzyme was incubated separately with 0.5% (w/v) laminarin and barley β-glucans in 50 mM sodium acetate (pH 5.0) at 37°C for 30-60 min. The reaction was stopped by the addition of *p*-hydroxybenzoic acid hydrozide reagent and the increase in reducing sugars was measured colorimetrically, as described by Lever (1972). One unit of β-glucosidase activity was defined as the amount of enzyme that produced 1 µmol of product per min. Protein assays were performed by the Bio-Rad protein assay kit (Bio-Rad, Richmond, CA) using bovine serum albumin as a standard. The pH optimum was determined by measuring the release of *p*NP from *p*NP-β-D-glucoside (*p*NPG) in different 50 mM buffers ranging in pH from 3.5 to 10 in 0.5 pH unit increments for 10 min (formate, pH 3.5-4.5; sodium acetate, pH 4.0-5.5; sodium phosphate, pH 5.5-8; Tris, pH 7.5-9.0; CAPS, pH 9.0-10). To find the temperature optimum, *p*NPG hydrolysis was measured in 50 mM sodium acetate, pH 5.0, at temperatures ranging from 5°C to 90°C in 5° increments for 10 min.

4. Result and discussion

4.1 Database searching and sequence analysis

The completion of the *Oryza sativa* L. spp. *japonica* Rice Genome Project and the complementary indica rice (*O. sativa* L. spp. *Indica*) genome project by the Beijing Genomic Institute and Carbohydrate Active Enzyme Database (http://afmb.cnrs-mrs.fr/CAZY) has allowed genome-wide analysis of gene families in this important crop. The sequence and mapping information provided to the public databases by these projects enabled us to identify glucanohydrolase genes. At least 139 genes homologous to glucanohydrolases which can be classified to be the members in 7 subfamilies of glycosyl hydrolase family in rice (Table 1).

Table 1 Summary of genes homologous to glucanohydrolase in glycosyl hydrolase family

Family	Gene ID	Protein ID	Putative protein
1	U28047	AAA84906	beta-glucosidase, beta-D-glucan exohydrolase
	AAAA02002143	BAD73293	beta-glucosidase
	AAAA02014151	CAE05483	beta-glucosidase, beta-D-glucan exohydrolase
	AAAA02010831	AAX95520	beta-glucosidase
	AAAA02022575	-	beta-glucosidase
	AAAA02034198	CNS08C89	beta-glucosidase
3	AL732339	CAJ86028	
	AL442109	CAJ86207	
	AP004150	BAD07748	beta-D-glucan exohydrolase
	AL606455	CAE02971	
	AC130602	AAU90111	beta-D-glucan exohydrolase
	AC135599	AAX96800	
	AC091811	AAL58966	exoglucanase
	AP003106	BAB56084	exo-1,3-beta-glucanase
	AP003295	BAB89846	exo-1,3-beta-glucanase
	AL662976	CAD41212	
	AC091811	AAL58963	
	AC091811	AAL58976	exohydrolase
5	AAL442104	CAH67677	Cellulase
	AAL442104	CAH67676	Cellulase
	AL606614	CAD41088	Cellulase
	AP004583	BAD61770	Cellulase, endo-beta-1,4-mannanase
	AP003284	BAB91747	Cellulase, (1-4)-beta-mannan endohydrolase
	AP004817	BAD17132	Cellulase, (1-4)-beta-mannan endohydrolase
	AP003076	BAB56016	Cellulase, (1-4)-beta-mannan endohydrolase
	AC096687	AAL79761	Cellulase, endohydrolase
	AC119290	AAV43969	glucan 1,3-beta-glucosidase
	AC136224	AAV44080	

Family	Gene ID	Protein ID	Putative protein
5	AAL60614	CAD41089	
	AP004023	BAD16839	cellulase
	AC107314	AAM08614	Glucan 1,3-beta-glucosidase
	APOO4213	BAD09304	1,3-beta-glucosidase
	AC096687	AAL79758	endohydrolase
	AL606614	CAD41090	
	AC107314	AAM08620	Glucan 1,3-beta-glucosidase
9	AL732351	CAJ86099	
	AL732344	CAH68191	
	AL606627	CAD41248	
	AP005092	BAD33331	Cellulase
	AP005319	BAD16147	endo-1,4-beta-glucanase
	AP003614	BAD53575	endo-1,4-beta-glucanase
	BX548155	CAE01493	
	AP004689	BAD05437	
	AP002094	BAC00553	endo-1,4-beta-glucanase
	AP005657	BAD10555	endo-1,4-beta-glucanase
	AP004027	BAD45673	endo-beta-1,4-glucanase
	AP004463	BAC55745	endo-beta-1,4-glucanase
	AL606457	CAE03241	
	AP006453	BAD26493	endo-beta-1,4-glucanase
	AP004846	BAD19513	cellulase
	AP003722	BAB92772	endo-beta-1,4-glucanase
	AP005112	BAD16040	endo-beta-1,4-glucanase
	AC118133	AAP03405	endo-beta-1,4-glucanase
	AC135914	AAT44235	endo-beta-1,4-glucanase
	AP002094	BAD81358	endo-beta-1,4-glucanase
	AP004885	BAD07956	endoglucanase 1
	AP005619	BAD46308	endo-beta-1,4-glucanase
16	AP004886	BAD07973	xyloglucan endo-1,4-beta-D-glucanase
	AP005859	BAD08162	xyloglucan endo-1,4-beta-D-glucanase
17	U72255	AAD10386	beta-1,3-glucanase
	U72253	AAD10384	beta-1,3-glucanase
	AB027430	BAA77785	beta-1,3-glucanase
	AF443600	AAL35900	endo- beta-1,3-glucanase
	U72251	AAD10382	beta-1,3-glucanase
	AF323610	AAK16694	glucanase
	AB027432	BAA77787	beta-glucanase
	U72254	AAD10385	beta-1,3-glucanase
	AY217745	AAO64485	beta-1,3-glucanase
	AL732352	CAJ86292	
	AP003884	BAD01673	
	AP005185	BAC83955	beta-1,3-glucanase
	AC146936	AAX95654	
	AP003804	BAC83070	elicitor inducible beta-1,3-glucanase

Family	Gene ID	Protein ID	Putative protein
17	U72248	AAD10379	beta-1,3-glucanase
	AP003259	BAC15778	endo- beta-1,3-glucanase
	AL662947	CAD40655	
	AP004073	BAD88030	endo- beta-1,3-glucanase
	AC092781	AAP68878	lichenase
	AC137611	AAT69611	
	AC134045	AAX95357	glucan endo-1,3-beta-glucosidase
	AC112160	AAU44050	beta-1,3-glucanase
	AP004073	BAB86248	endo- beta-1,3-glucanase
	AP005446	BAD36114	elicitor inducible beta-1,3-glucanase
	AP002972	BAD86947	elicitor inducible beta-1,3-glucanase
	AP005810	BAC84505	beta-1,3-glucanase
	AC0099324	AAO73270	beta-1,3-glucanase
	AP004073	BAB86250	beta-1,3-glucanase
	AP002542	BAD67869	beta-1,3-glucanase
	AC133930	AAP44659	beta-1,3-glucanase
	AC137064	AAX95269	glucan endo-1,3-beta-glucosidase
	AP005810	BAC84500	beta-1,3-glucanase
	AC108874	AAT01345	beta-1,3-glucanase
	AP003622	BAD32917	beta-1,3-glucanase
	AB027428	BAA77783	beta-1,3-glucanase
	AP004073	BAD88015	beta-1,3-glucanase
	AP004230	BAC20025	glucanase
	AP004031	BAD87992	beta-1,3-glucanase
	AP003799	BAC75843	beta-1,3-glucanase
	AP005788	BAD26208	elicitor inducible beta-1,3-glucanase
	AP004749	BAD54223	beta-1,3-glucanase
	AC104473	AAN60993	
	AP005810	BAC84503	beta-1,3-glucanase
	AP004869	BAD15845	beta-1,3-glucanase
	AP005786	BAC80125	beta-1,3-glucanase
	U72249	AAD10380	beta-1,3-glucanase
	AP003259	BAC15784	endo-1,3-beta-glucosidase
	AP003287	BAD82033	elicitor inducible beta-1,3-glucanase
	AC091787	AAP12947	beta-1,3-glucanase
	AP003221	BAB86422	beta-1,3-glucanase
	AP005294	BAD10251	beta-1,3-glucanase
	AP005110	BAD28425	beta-1,3-glucanase
	AC084765	AAP46217	glucanase
	AC105364	AAN05325	beta-1,3-glucanase
	AC098696	AAN04212	beta-1,3-glucanase
	AP003259	BAD87205	endo-1,3-beta-glucosidase
	AP004690	BAD05454	beta-1,3-glucanase
	AP004073	BAD88028	endo-1,3-beta-glucosidase
	AC090871	AAO37977	beta-1,3-glucanase

Family	Gene ID	Protein ID	Putative protein
17	AC133859	AAP50997	beta-1,3-glucanase
	AC137064	AAX95270	endo-1,3-beta-glucosidase
	AP003259	BAD87200	endo- beta-1,3-glucanase
	AP003888	BAC75423	beta-1,3-glucanase
	AP005810	BAC84487	beta-1,3-glucanase
	AC097112	AAT39206	beta-1,3-glucanase
	U72252	AAD10383	beta-1,3-glucanase
	AB027429	BAA77784	beta-1,3-glucanase
	AP004073	BAB86249	beta-1,3-glucanase
	AP005840	BAD31779	beta-1,3-glucanase
	AP004058	BAD16854	beta-1,3-glucanase
	AP005092	BAD33320	endo-1,3-beta-glucosidase
	U72250	AAD10381	endo-1,3-beta-glucosidase
	AP004031	BAB85436	glucanase
	AP004340	BAC83528	endo-1,3-beta-glucosidase
	AP004031	BAB85424	endo- beta-1,3-glucanase
	AC092778	AAP12921	beta-1,3-glucanase
	AB070742	BAC02926	beta-1,3-glucanase
	AF030167	AAB86556	glucanase
	AF030166	AAB86541	glucanase
81	AP005702	BAD26084	beta-glucan binding protein

4.2 Cloning and expression of Os4bglu12 β-glucosidase

Os4bglu12 β-glucosidase cDNAs cloning and sequence analysis

The protein product for *Os4bglu12* gene has highest sequence similarity to the previously described cell wallbound β-glucosidase purified from rice seedlings (Akiyama et al., 1998). Therefore, it was chosen for expression to test if the protein would have the expected activity. The sequence of the *Os4bglu12* β-glucosidase mRNA from rice was confirmed by RT-PCR cloning and sequencing, using rice cultivar KDML105 cDNA as the template. A specific PCR product of 1635 bp was produced, and its sequence overlapped that of the *indica* rice contig AAAA02014151. The reconstructed cDNA sequence of *Os4bglu12* included a 1530-nucleotide long open reading frame encoding a 510 amino acid long precursor protein (Figure 1). The Signal P program predicted the protein to contain a 24 amino acid signal sequence and a 486 amino acid mature protein. The deduced Os4bglu12 N-terminal amino acid sequence was identical to the N-terminal amino acid sequence of the previously purified cell-wall-bound rice β-glucosidase at 40 of 44 residues (Akiyama et al., 1998).

```
ATGGCGGCAGCAGGGGCAATGCCCGGTGGCCTTCTCCTCACGTTCCTCCTC
M  A  A  A  G  A  M  P  G  G  L  L  L  T  F  L  L
CTTGCTGTCGTCGCTTCCGGCGCCTACAATAGCGCCGGCGAGCCGCCGGTC
L  A  V  V  A  S  G  A  Y  N  S  A  G  E  P  P  V
AGCCGGAGAAGCTTCCCCAAGGGGTTCATCTTCGGGACAGCCTCGTCGTCG
S  R  R  S  F  P  K  G  F  I  F  G  T  A  S  S  S
TATCAGTACGAGGGTGGCGCAGCGGAGGGCGGCAGAGGACCAAGCATCTGG
Y  Q  Y  E  G  G  A  A  E  G  G  R  G  P  S  I  W
GACACCTTCACACATCAGCACCCAGAGAAAATCGCCGACAGAAGCAACGGG
D  T  F  T  H  Q  H  P  E  K  I  A  D  R  S  N  G
GATGTGGCTTCGGATTCCTACCATCTCTACAAGGAAGATGTGCGCCTCATG
D  V  A  S  D  S  Y  H  L  Y  K  E  D  V  R  L  M
AAGGATATGGGAATGGATGCATACAGGTTCTCCATCTCATGGACAAGAATC
K  D  M  G  M  D  A  Y  R  F  S  I  S  W  T  R  I
CTTCCAAATGGAAGTCTGAGGGGTGGAGTCAACAAAGAAGGCATAAAGTAC
L  P  N  G  S  L  R  G  G  V  N  K  E  G  I  K  Y
TACAATAATTTGATCAATGAGCTATTATCGAAAGGGGTGCAACCGTTTATT
Y  N  N  L  I  N  E  L  L  S  K  G  V  Q  P  F  I
ACCCTTTTCCACTGGGACTCACCTCAGGCGTTGGAAGATAAATATAACGGA
T  L  F  H  W  D  S  P  Q  A  L  E  D  K  Y  N  G
TTTCTTAGCCCTAATATCATAAATGACTTTAAGGACTATGCTGAAATCTGC
F  L  S  P  N  I  I  N  D  F  K  D  Y  A  E  I  C
TTCAAGGAGTTTGGTGACAGAGTGAAAAATTGGATCACCTTCAATGAGCCT
F  K  E  F  G  D  R  V  K  N  W  I  T  F  N  E  P
TGGACTTTCTGCTCTAATGGCTATGCAACTGGCCTGTTTGCACCAGGCCGT
W  T  F  C  S  N  G  Y  A  T  G  L  F  A  P  G  R
TGTTCGCCCTGGGAGAAGGGAAATTGCAGTGTTGGAGATTCAGGAAGGGAG
C  S  P  W  E  K  G  N  C  S  V  G  D  S  G  R  E
CCTTACACTGCATGCCATCATCAACTACTTGCCCACGCGGAAACTGTTCGG
P  Y  T  A  C  H  H  Q  L  L  A  H  A  E  T  V  R
TTGTACAAAGCGAAATATCAGGCCTTACAAAAAGGGAAGATTGGAATAACT
L  Y  K  A  K  Y  Q  A  L  Q  K  G  K  I  G  I  T
CTGGTCTCGCACTGGTTTGTTCCCTTCTCCCGCTCCAAATCCAACAATGAT
L  V  S  H  W  F  V  P  F  S  R  S  K  S  N  N  D
GCTGCAAAGCGTGCTATAGACTTCATGTTTGGATGGTTTATGGATCCCCTC
A  A  K  R  A  I  D  F  M  F  G  W  F  M  D  P  L
ATTAGAGGCGACTACCCCCTAAGCATGAGAGGATTGGTTGGGAATCGCTTG
I  R  G  D  Y  P  L  S  M  R  G  L  V  G  N  R  L
CCACAGTTCACTAAAGAACAGTCTAAGTTGGTCAAGGGTGCATTTGACTTT
P  Q  F  T  K  E  Q  S  K  L  V  K  G  A  F  D  F
ATTGGACTTAACTACTACACTGCGAACTATGCTGATAACCTTCCTCCATCA
I  G  L  N  Y  Y  T  A  N  Y  A  D  N  L  P  P  S
AATGGCCTTAACAACAGCTATACCACCGATTCTCGAGCTAATCTTACCGGT
N  G  L  N  N  S  Y  T  T  D  S  R  A  N  L  T  G
GTACGAAATGGCATCCCCATAGGACCGCAGGCTGCTTCACCTTGGCTTTAC
V  R  N  G  I  P  I  G  P  Q  A  A  S  P  W  L  Y
GTCTACCCTCAAGGCTTCCGTGACTTGCTACTTTATGTTAAGGAGAACTAT
V  Y  P  Q  G  F  R  D  L  L  L  Y  V  K  E  N  Y
GGCAATCCTACCGTCTACATCACTGAAAATGGCGTTGATGAATTCAACAAT
G  N  P  T  V  Y  I  T  E  N  G  V  D  E  F  N  N
AAGACCTTACCACTCCAAGAAGCCTTGAAGGATGACGCTAGGATAGAATAC
K  T  L  P  L  Q  E  A  L  K  D  D  A  R  I  E  Y
TACCACAAGCACCTCCTTTCCCTGCTAAGTGCTATAAGGGATGGAGCAAAC
Y  H  K  H  L  L  S  L  L  S  A  I  R  D  G  A  N
GTGAAGGGATACTTTGCATGGTCGCTGCTTGATAACTTCGAGTGGTCGAAC
V  K  G  Y  F  A  W  S  L  L  D  N  F  E  W  S  N
GGCTACACAGTTCGATTTGGGATAAACTTCGTAGATTACAATGACGGAAGG
G  Y  T  V  R  F  G  I  N  F  V  D  Y  N  D  G  R
AAGAGATACCCCAAGAACTCTGCCCATTGGTTCAAGAAGTTCCTCCTGAAA
K  R  Y  P  K  N  S  A  H  W  F  K  K  F  L  L  K
```

Figure 1 The full-length cDNA sequence and deduced amino acid sequence of rice *Os4bglu12*

A few reports described the expression patterns of β-glucosidases in rice plants. Based on enzyme activity, gibberellic acid glucoside and pyridoxine glucoside β-glucosidases are found in rice bran (Schlieman, 1984; Iwami and Yasumoto, 1986), and the cell wall-bound enzyme is found in seedling (Akiyama et al., 1998). Northern blot analysis showed that *Os3bglu7* and *Os9bglu30* (rice *bglu 2* in Opassiri et al., 2003) β-glucosidase genes are highly expressed in seedling shoots, but only *Os3bglu7* is expressed in flowers (Opassiri et al., 2003). Microarray analysis indicated that the transcripts of the ESTs BE607353 and BG101702, whose sequences are homologous to *Os3bglu7* and *Os4bglu12* β-glucosidases genes, respectively, are upregulated in response to high salinity stress in salt-tolerant rice (var Pokkali), but not in the salt-sensitive cultivar IR29(Kawasaki et al., 2001). Subtractive hybridization cDNA library screening indicated that the transcript level of the EST contig BPHiw028, homologous to *Os4bglu12*, is upregulated in response to brown planthopper (Wang et al., 2005). The presence of tricin-O-glucoside, a probing stimulant for planthopper (Adjei-Afriyie et al., 2000), suggests that the role of this enzyme is to release an active flavonol for defense. However, these studies did not show the specific roles of these enzymes in rice cells in response to such stresses. Therefore, identification of natural substrates for the enzymes is needed to understand the functions of these enzymes.

Functional expression of recombinant Os4bglu12

The *Os4bglu12* cDNA CDS including the stop codon was inserted into pET32a(+)/ DEST. The construct was used to transform *OrigamiB* (DE3) *E. coli*. Comparison of the protein profile of induced cultures with the *Os4bglu12* insert with those of empty plasmid controls by SDS-PAGE showed the thioredoxin-Os4bglu12 fusion protein as an intense band at 69 kDa on SDS-PAGE. The fusion protein was purified by IMAC, and a band corresponding to 69 kDa was observed in SDS-PAGE (Figure 2). The enzyme was found to hydrolyze *p*NPG with optimal activity at pH 5.0 and 37°C. The enzyme activity with *p*NPG at 70°C and 80°C drops about 17% and 39%, respectively, from the optimal activity at 37°C in a 10 min assay. It was stable at 4°C for several months.

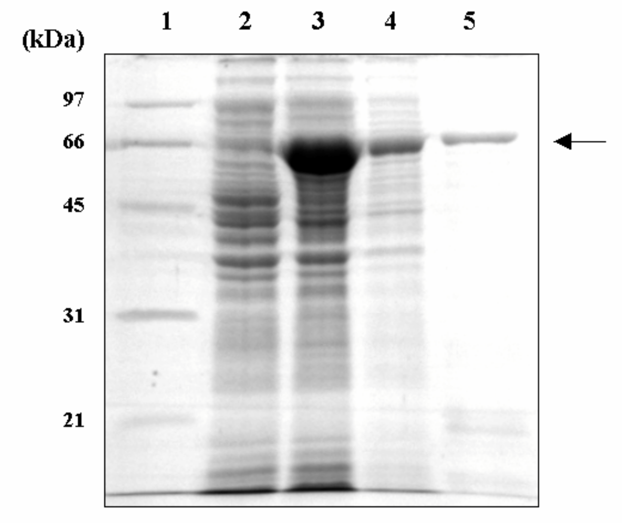


Figure 2 SDS-PAGE profiles of Os4bglu12 recombinant protein expressed in *OrigamiB* (DE3) *E. coli* after incubation in the presence of 0.3 mM IPTG, at 20°C for 8 h. Lanes: 1, standard marker (Bio-RAD); 2, total protein in *E. coli* cells containing pET32a(+) without an insert; 3, total protein of *E. coli* cells containing pET32a(+)/DEST-*Os4bglu12*; 4, soluble fraction of *E. coli* cells containing pET32a(+)/DEST-*Os4bglu12*; 5, purified Os4bglu12 recombinant protein. The arrow points to the position of thioredoxin fusion protein monomer.

Os4bglu12 substrate specificity

The activity of the purified rice Os4BGlu12 β-glucosidase towards natural and artificial glycosides is summarized in Table 2. Since the reliable kinetic parameters could not be retreived dedrived from the cellobiose product inhibition, so the relative activities were used to determine the substrate specificity of this enzyme. The Os4BGlu12 hydrolyzed the β-1,3-linked glucose disaccharide laminaribiose, but not cellobiose (β-1,4) or gentiobiose (β-1,6). It showed high hydrolytic efficiency at different rates with β-(1,4)-linked oligosaccharides with DP of 3-6. Hydrolysis of β-(1,3)-linked oligosaccharides with DP > 2, laminarin and barley 1,3, 1,4-β-glucans by this enzyme could not be detected. The rate of hydrolysis of oligomeric substrates by Os4BGlu12 tended to remain approximately constant with increasing DP, which is a characteristic often observed with β-glucosidases (Reese *et al.*, 1968). On the TLC, Os4BGlu12 showed hydrolytic activity towards 5 mM laminaribiose and cello-oligosaccharides but the transglycosylation activity could not be detected (Figure 3).

Hydrolysis of *p*NP-glycosides with different sugars was used to assess Os4BGlu12 glycone specificity, and it was found to hydrolyze many kinds of *p*NP β-glycosides. It hydrolyzed *p*NPG and *p*NP-β-D-fucopyranoside with 2-3 fold lower relative activity than oligosaccharides. It also hydrolyzed *p*NP-β-D-galactopyranoside, *p*NP-β-D-xyloside, and *p*NP-α-L-arabionoside, at 45%, 45% and 26% the rate of *p*NPG, respectively. Hydrolysis of *p*NP-β-D-mannoside, *p*NP-β-D-cellobioside, *p*NP-α- D-glucopyranoside, and *p*NP-β-L-fucopyranoside were not detectable. High hydrolysis of β-xyloside is similar to white clover β-glucosidase, but otherwise rare in GH1 enzymes, as noted by Marana (2006*),* who placed the clover enzyme in a separate functional grouping based on this activity.

Rice Os4BGlu12, rice BGlu1 (Opassiri *et al.* 2003), cell-wall-bound rice β-glucosidase (Akiyama *et al.*, 1998) and barley βII β-glucosidase (Hrmova *et al.*, 1998) are enzymes that hydrolyze β-linked glucose oligosaccharides, but not polysaccharides. However, the specificity for glycones and substrate chain lengths of these enzymes are different. In contrast to barley and rice cell-wall-bound enzyme, Os4BGlu12 did not hydrolyze β-(1,3)-linked oligosaccharides longer than laminaribiose, and hydrolyzed various *p*NP-derivatives of monosaccharides. This substrate preference was not expected, since it was initially expected that *Os4bglu12* was the gene for the cell-wall-bound β-glucosidase, and the sequence differences might be due to cultivar differences or sequencing errors. The substrate preference of Os4BGlu12 is somewhat similar to rice BGlu1, however, there were many differences between these enzymes. In contrast to rice BGlu1, Os4BGlu12 hydrolyzed β-(1,4)-linked oligosaccharides and laminaribiose at higher rates than *p*NPG, and did not hydrolyze cellobiose, gentiobiose, *p*NP β-D-mannoside, and *p*NP-β-D-cellobioside. Their sequence differences are likely to reflect the differences in substrate binding to the active site between these enzymes. The amino acids identified by Czjzek *et al.* (2000) as critical for glucose binding (Q38, H142, E191, E406, E464 and W465 in maize Bglu1) are conserved in rice Os4BGlu12, rice BGlu1, and barley β-glucosidase. Interestingly, the rice BGlu1 protein sequence was closest to barley BGQ60 at some of substrate binding residues that line the active site cleft and interact with the substrate aglycone of maize βglu1 (W378, F198, F205, and F466) (Czjzek *et al.*, 2000), suggesting it may have a similar substrate-specificity. However, these above mentioned amino acid residues were different from those in the Os4BGlu12 enzyme and reflect the different substrate specificities for some oligosaccharides and glycones. For instance, BGlu1 and barley BGQ60 cluster with tomato and arabidopsis β-mannosidase and can hydrolyze β-mannoside, while Os4BGlu12 does not, and they also hydrolyze longer chain 1,3-linked oligosaccharides (Hrmova *et al.*, 1996; Opassiri *et al.*, 2004). All three enzymes prefer shorter 1,3-linked oligosaccharides, with Os4bglu12 being the most extreme, only hydrolyzing the dimer with this linkage. This likely reflects the bent shape of oligosaccharides with the 1,3-linkage, which is somehow incompatible with the active site for longer chains. Elucidation of the tertiary structures of these enzymes would help to clarify the enzyme-substrate binding mechanism leading to these preferences. Elucidation of the tertiary structures of these enzymes would help to clarify the enzyme-substrate binding mechanism.

Table 2 Substrate specificity of the purified rice Os4bglu12

Substrate	**Relative activity[a] (%)**
Laminaribiose[b]	238
Laminaritriose	0
Laminaritetraose	0
Laminaripentaose	0
Cellobiose	0
Cellotriose[b]	231
Cellotetraose[b]	301
Cellopentaose[b]	279
Cellohexaose[b]	295
Gentiobiose	0
Laminarin	0
Barley 1,3, 1,4-β-glucans	0
*p*NP-β-D-glucoside	100
*p*NP-β-D-fucoside	118
*p*NP-β-D-galactoside	45
*p*NP-β-D-xyloside	45
*p*NP-α-L-arabionoside	26
*p*NP-β-D-mannoside	0
*p*NP-β-D-cellobioside	0
*p*NP-α-D-glucoside	0
*p*NP-β-L-fucoside	0

[a] Percent activity relative to glucose or *p*NP released from *p*NP-β-D-glucoside. The assay contained 1 mM substrate in 50 mM sodium acetate (pH 5.0) buffer at 37°C.
[b] Note that the values for oligosaccharides are in terms of total glucose released.

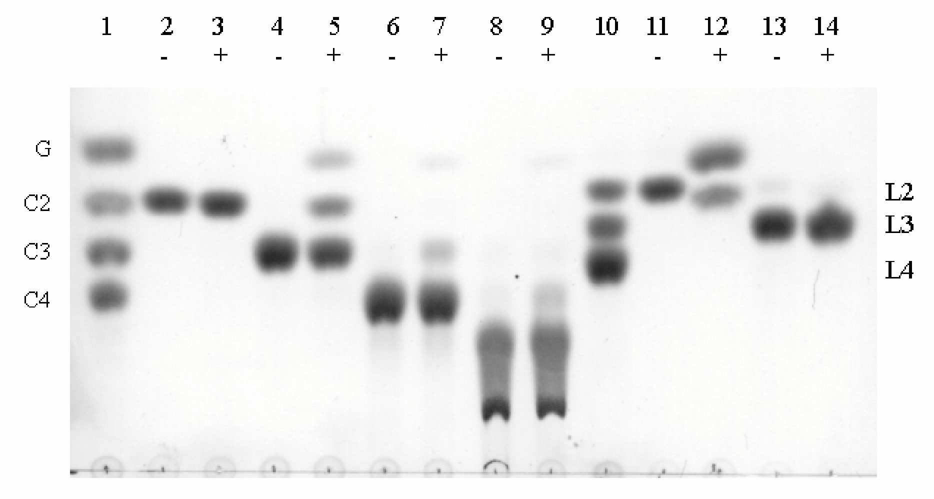


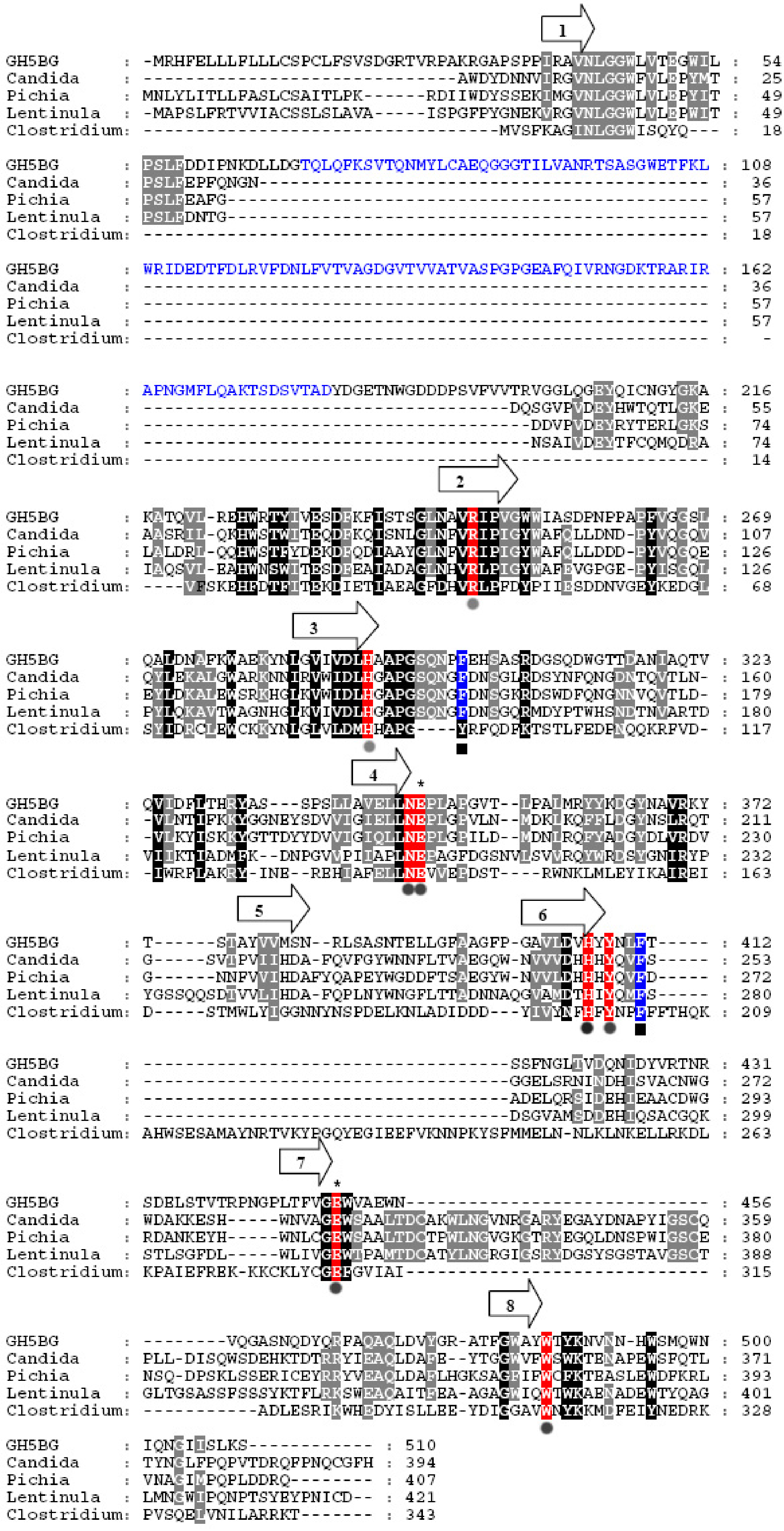
Figure 3 Hydrolysis of oligosaccharide substrates by Os4bglu12 detected by TLC. The Os4bglu12 was incubated with 5 mM substrates for 30 min and the products were detected after TLC by the carbohydrate staining. Samples were incubated with (+) or without (-) enzyme. Lanes: 1, glucose (G) and cello-oligosaccharides of DP 2–4 (C2-C4) marker; 2 and 3, cellobiose; 4 and 5, cellotriose; 6 and 7, cellotetraose, 8 and 9, cellopentaose, 10 laminari-oligosaccharides of DP 2–4 (L2-L4) marker; 11 and 12, laminaribiose; 13–14 laminaritriose.

4.3 Cloning and expression of glycosyl hydrolase family 5 β-glucosidase

GH5BG cDNA cloning and sequence analysis

With the completion of high quality drafts of the rice genome, analysis of GH5 members in rice has been reported in the CAZY homepage (http://www.cazy.org/CAZY/). Twenty GH5 genes putatively encoding 7 cellulases, 9 endo-β-mannanases, 3 glucan 1,3-β-glucosidases, and one 1,3-β-glucanase have been identified in rice databases (see CAZY). The putative glucan 1,3-β-glucosidases encoded by these genes include Genbank accessions AAM08614, AAM08620, and AAV43969. The simplest putative glucan 1,3-β-glucanase genes and AAM08614 was chosen for investigation.

A GH5 glucan-1,3-β-glucosidase cDNA, designated GH5BG, was cloned from rice seedlings by RT-PCR with KDML105 rice seedling RNA and primers derived from the GenBank accession number AC107314 (rice genomic contig from which AAM08614 is derived) and AK065000 (full-length cDNA, Kikuchi et al., 2003) sequences. A specific PCR product of 1680 bp was produced, and its sequence overlapped those of AC107314 and



```
GH5BG      : -MRHFELLLFLLLCSPCLFSVSDGRTVRPAKRGAPSPPIRAVNLGGWLVTEGWIL :  54
Candida    : -------------------------------AWDYDNNVIRGVNLGGWFVLEPYMT :  25
Pichia     : MNLYLITLLFASLCSAITLPK------RDIIWDYSSEKIMGVNLGGWLVLEPYIT :  49
Lentinula  : -MAPSLFRTVVIACSSLSLAVA-----ISPGFPYGNEKVRGVNLGGWLVLEPWIT :  49
Clostridium: ----------------------------------MVSFKAGINLGGWISQYQ--- :  18

GH5BG      : PSLFDDIPNKDLLDGTQLQFKSVTQNMYLCAEQGGGTILVANRTSASGWETFKL : 108
Candida    : PSLFEPFQNGN------------------------------------------- :  36
Pichia     : PSLFEAFG---------------------------------------------- :  57
Lentinula  : PSLFDNTG---------------------------------------------- :  57
Clostridium: ------------------------------------------------------ :  18

GH5BG      : WRIDEDTFDLRVFDNLFVTVAGDGVTVVATVASPGPGEAFQIVRNGDKTRARIR : 162
Candida    : ------------------------------------------------------ :  36
Pichia     : ------------------------------------------------------ :  57
Lentinula  : ------------------------------------------------------ :  57
Clostridium: ------------------------------------------------------ :   -

GH5BG      : APNGMFLQAKTSDSVTADYDGETNWGDDDPSVFVVTRVGGLQGEYQICNGYGKA : 216
Candida    : -----------------------------------DQSGVPVDEYHWTQTLGKE :  55
Pichia     : -------------------------------------DDVPVDEYRYTERLGKS :  74
Lentinula  : -------------------------------------NSAIVDEYTFCQMQDRA :  74
Clostridium: ------------------------------------------------------ :  14

GH5BG      : KATQVL-REHWRTYIVESDFKFISTSGLNAVRIPVGWWIASDPNPPAPFVGGSL : 269
Candida    : AASRIL-QKHWSTWITEQDFKQISNLGLNFVRIPIGYWAFQLLDND-PYVQGQV : 107
Pichia     : LALDRL-QQHWSTFYDEKDFQDIAAYGLNFVRIPIGYWAFQLLDDD-PYVQGQE : 126
Lentinula  : IAQSVL-EAHWNSWITESDFEAIADAGLNHVRLPIGYWAFEVGPGE-PYISGQL : 126
Clostridium: ----VFSKEHEDTFITEKDIETIAEAGFDHVRLPFDYPIIESDDNVGEYKEDGL :  68

GH5BG      : QALDNAFKWAEKYNLGVIVDLHAAPGSQNPFEHSASRDGSQDWGTTDANIAQTV : 323
Candida    : QYLEKALGWARKNNIRVWIDLHGAPGSQNGFDNSGLRDSYNFQNGDNTQVTLN- : 160
Pichia     : EYLDKALEWSRKHGLKVWIDLHGAPGSQNGFDNSGKRDSWDFQNGNNVQVTLD- : 179
Lentinula  : PYLQKAVTWAGNHGLKVIVDLHGAPGSQNGFDNSGQRMDYPTWHSNDTNVARTD : 180
Clostridium: SYIDRCLEWCKKYNLGLVLDMHHAPG----YRFQDFKTSTLFEDPNQQKRFVD- : 117

GH5BG      : QVIDFLTHRYAS---SPSLLAVELLNEPLAPGVT--IPALMRYYKDGYNAVRKY : 372
Candida    : -VLNTIFKKYGGNEYSDVVIGIELLNEPLGPVLN--MDKLKQFFLDGYNSLRQT : 211
Pichia     : -VLKYISKKYGTTDYYDVVIGIQLLNEPLGPILD--MDNLRQFYADGYDLVRDV : 230
Lentinula  : VIIKTIADMEK--DNPGVVPIIAPLNEPAGFDGSNVLSVVRQYWRDSYGNIRYP : 232
Clostridium: -IWRFLAKRY-INE--REHIAFELLNEVVEPDST----RWNKLMLEYIKAIREI : 163

GH5BG      : T------STAYVVMSN--RLSASNTELLGFAAGFP-GAVLDVHYYNLFT----- : 412
Candida    : G-----SVIPVIIHDA-FQVFGYWNNFLTVAEGQW-NVVVDHHHYQVFS----- : 253
Pichia     : G------NNFVVIHDAFYQAPEYWGDDFTSAEGYW-NVVLDHHHYQVFD----- : 272
Lentinula  : YGSSQQSDIVVLIHDA-FQPLNYWNGFLTTADNNAQGVAMDTHIYQMFS----- : 280
Clostridium: D-----STMWLYIGGNNYNSPDELKNLADIDDD---YIVYNFHFYNPFFFTHQK : 209

GH5BG      : -----------------------------------SSFNGLTVDQNIDYVRTNR : 431
Candida    : -----------------------------------GGELSRNINDHISVACNWG : 272
Pichia     : -----------------------------------ADELQRSIDEHIEAACDWG : 293
Lentinula  : -----------------------------------DSGVAMSDDEHIQSACGQK : 299
Clostridium: AHWSESAMAYNRTVKYPGQYEGIEEFVKNNPKYSFMMELN-NLKLNKELLRKDL : 263

GH5BG      : SDELSTVTRPNGPLTFVGEWVAEWN----------------------------- : 456
Candida    : WDAKKESH-----WNVAGEWSAALTDCAKWLNGVNRGARYEGAYDNAPYIGSCQ : 359
Pichia     : RDANKEYH-----WNLCGEWSAALTDCTPWLNGVGKGTRYEGQLDNSPWIGSCE : 380
Lentinula  : STLSGFDL-----WLIVGEWIPAMTDCATYLNGRGIGSRYDGSYSGSTAVGSCT : 388
Clostridium: KPAIEFREK-KKCKLYCGEFGVIAI----------------------------- : 315

GH5BG      : --------VQGASNQDYQRFAQAQLDVYGR-ATFGWAYWTYKNVNN-HWSMQWN : 500
Candida    : PLL-DISQWSDEHKTDTRRYIEAQLDAFE--YTGGWVFWSWKTENAPEWSFQTL : 371
Pichia     : NSQ-DPSKLSSERICEYRRYVEAQLDAFLHGKSAGFIPWCEKTEASLEWDFKRL : 393
Lentinula  : GLTGSASSFSSSYKTFLRKSWEAQAITFEA-AGAGWIQWTWKAENADEWTYQAG : 401
Clostridium: -----------ADLESRIKWHEDYISLLEE-YDIGGAVWNYKKMDFEIYNEDRK : 328

GH5BG      : IQNGIISLKS------------ : 510
Candida    : TYNGLFPQPVTDRQFPNQCGFH : 394
Pichia     : VNAGIMPQPLDDRQ-------- : 407
Lentinula  : LMNGWIPQNPTSYEYPNICD-- : 421
Clostridium: PVSQEIVNILARRKT------- : 343
```

Figure 4 Alignment of the protein sequences of rice GH5BG with exo-β-1,3-glucanases and endo-β-1,4-glucanase. GH5BG is rice GH5BG, Candida is exo-β-1,3-glucanase from *Candida albicans* (AC CAA39908), Lentinula is exo-β-1,3-glucanase from *Lentinula edodes* (AC AB192344), Pichia is exo-β-1,3-glucanase from *Pichia pastoris* (AC AY954499), and Clostridium is endo-β-1,4-glucanase from *Clostridium thermocellum* (AC AAA23220). The alignment was generated with the ClustalX implementation of ClustalW (Jeanmougin et al., 1998; Thompson et al., 1994) and analyzed and manually adjusted with Gendoc (Nicholas and Nicholas, 1997). Alignment of the *C. thermocellum* sequence relied on the structural alignment of the 1CEC structural model with the *C. albicans* Exg 1CZ1 structure. The positions of the β-strands of the central $(\beta/\alpha)_8$ barrel are indicated by arrows above the alignment. Red bars (•) mark invariant GH family 5 residues and the black and grey shading highlight other identities between sequences, the asterisks identify the two catalytic glutamate residues, the blue bars (■) indicate the two phenylalanine found at the +1 subsites of *C. albicans* Exg. The region of rice GH5BG homologous to fascin is indicated by blue text.

AK065000. The full-length cDNA sequence contains a 1530-nucleotide open reading frame (ORF) encoding a 510 amino acid precursor protein. The protein sequence was predicted to contain a 19 amino acid long prepeptide and a 491 amino acid long mature protein, and to be secreted out of the cell. Its predicted pI is 5.28. The mature protein includes two domains, a fascin-like domain (amino acids 70-180) and a glucan-1,3-β-glucosidase domain (amino acids 37-60 and 208-496) (Figure 4). In human, fascin acts as a protein that tightly bundles F-actin and it is important for the formation of actin-based protrusions and for maintaining cytoplasmic F-actin bundles (Adams, 2004).

In addition to the putative rice GH5 glucan-1,3-β-glucosidases AAM08620 and AAV43969, the GH5BG amino acid sequence is highly similar to the putative 1,3-β-glucanase BAD10703 with 49% identity, but has only 28-33% identity with glucan-1,3-β-glucosidases of fungi. Comparison of the deduced amino acid sequence of rice GH5BG with those of fungal GH5 exoglucanases revealed that Glu-347, which lies in the conserved NEP motif, is likely to be the catalytic acid/base and Glu-450, which lies in the conserved GEW motif, is likely the catalytic nucleophile (Cutfield et al., 1999; Sakatomo et al., 2005; Xu et al., 2006). Similar to other GH5 members, rice GH5BG contains eight invariant residues, these being Arg-247, His-291, Asn-346, Glu-347, His-406, Tyr-408, Glu-450, and Trp-486. Phe-300 and Phe-411 of

GH5BG are also conserved with Phe-144 and Phe-258 in *C. albicans* Exg, which were found to be located at the entrance to the active site pocket by crystallographic structural analysis (Cutfield et al., 1999).

Functional expression of recombinant GH5BG and substrate specificity

The *GH5BG* cDNA was expressed in redox-deficient, Origami (DE3), as a catalytically active thioredoxin fusion protein. Induced cultures of *E. coli* Origami (DE3) expressing GH5BG thioredoxin fusion proteins had intense bands at 68 in total cell lysates in SDS-PAGE (Figure 5). The optimal pH of the enzyme to hydrolyze *p*NPG is pH 5.0. The pH optima of this enzyme is in a broad range from 30-50ºC.

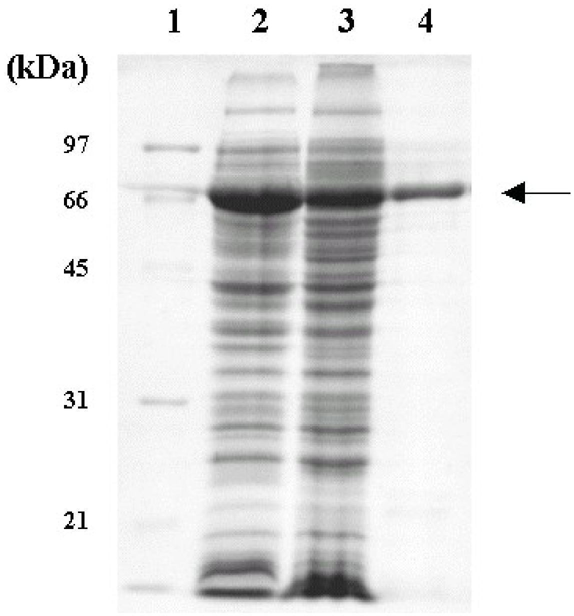


Figure 5 SDS-PAGE of GH5BG-thioredoxin fusion protein expressed in *E. coli* strain OrigamiB (DE3) after incubation in the presence of 0.5 mM IPTG, at 20ºC for 12 h. Lanes: 1, standard marker (Bio-RAD); 2, total protein of *E. coli* cells containing pET32a(+)/DEST-*GH5BG*; 3, soluble fraction of *E. coli* cells containing pET32a(+)/DEST-*GH5BG*; 4, purified thiredoxin-GH5BG. The arrow points to the thioredoxin-GH5BG.

The activity of the purified GH5BG towards artificial glycosides and oligosaccharides is summarized in Table 3. The GH5BG hydrolyzed a broad range of *p*NP-glycosides of the monosacharides, β-D-glucoside, β-D-fucoside, β-D-galactoside, β-D-xyloside, and α-L-arabionoside, but it could not hydrolyze *p*NP-β-D-mannoside. Among *p*NP-glycosides, *p*NP-β-D-fucoside was hydrolyzed twice as efficiently as *p*NPG, whereas *p*NP-β-D-galactoside and *p*NP-α-L-arabinoside were hydrolyzed with 27% and *p*NP-β-D-xyloside with 15% the efficiency of *p*NPG. These results indicate there is low stringency at the -1 subsite of GH5BG, in which the non-reducing glycosyl moiety is bound, which is somewhat similar to many GH1 and GH3 β-glucosidases such as rice BGlu1 (Opassiri et al., 2003, 2004). The enzyme could hydrolyze glycosides of β-D-glucose, β-D-fucose, β-D-galactose, β-D-xylose, and α-L-arabionoside, which are epimers with equatorial OH-2, but not β-D-mannoside, which has an axial projection at OH-2. Therefore, the epimerization of OH-2 is critical for binding of the sugar residue to the -1 subsite. However, the equatorial or axial projection at OH-4 and the elimination of CH_2OH-6 of D-glucose to H- in D-xylose or CH_3- in D-fucose are not critical to the capacity of the substrates to bind to the active site.

For disaccharides, GH5BG hydrolyzed laminaribiose (β-1,3), but had very low activity with cellobiose (β-1,4) or gentiobiose (β-1,6). It also hydrolyzed cello-oligosaccharides with degrees of polymerization (DP) of 3-6 at a relatively high rate. On the other hand, GH5BG could not hydrolyze laminari-oligosaccharides with DP 3-5, laminarin, or barley 1,3-1,4-β-glucans. The catalytic efficiency (k_{cat}/K_m) values for hydrolysis of β-(1,4)-linked oligosaccharides by GH5BG remained approximately constant as the DP increased from 3 to 5 and decreased for cellohexaose. Having rates of hydrolysis of oligomeric substrates which remain approximately constant or decrease with increasing DP length is a characteristic often observed with β-glucosidases, unlike polysaccharide exohydrolases, in which the hydrolytic rate increases with oligosaccharide length (Reese et al, 1968).

Although GH5BG was designated a putative glucan 1,3-β-glucosidases based on sequence homology, its catalytic activity is somewhat like GH1 β-glucosidases, which show similar oligosaccharide preferences. Rice BGlu1 GH1 exoglucanase/β-glucosidase (Opassiri et al., 2004) and rice Os4bglu12 GH1 β-glucosidase (Opassiri et al., 2006) are the enzymes that prefer to hydrolyze β-1,4-linked oligosaccharides and a broad range of *p*NP-glycosides, but with differences in catalytic efficiency compared to GH5BG.

Table 3 Kinetic parameters of rice GH5BG in the hydrolysis of *p*NP-glycosides, disaccharides and oligosaccharides

Substrate	**k_{cat} (s^{-1})**	**K_m (mM)**	**k_{cat}/K_m (s^{-1} mM^{-1})**
Cellobiose	4.3 ± 0.8	16.4 ± 1.9	0.27 ± 0.02
Cellotriose	41 ± 5	4.53 ± 0.01	9.1 ± 1.2
Cellotetraose	38 ± 2	4.09 ± 0.17	9.3 ± 0.9
Cellopentaose	35.5 ± 0.4	3.4 ± 0.4	10.4 ± 0.4
Cellohexaose	9.7 ± 0.8	2.2 ± 0.5	4.5 ± 0.5
Laminaribiose	36 ± 5	7.0 ± 1.1	5.05 ± 0.07
Laminaritriose	n.d.[a]	n.d.[a]	n.d.[a]
Laminaritetraose	n.d.[a]	n.d.[a]	n.d.[a]
Laminaripentaose	n.d.[a]	n.d.[a]	n.d.[a]
Laminarin	n.d.[a]	n.d.[a]	n.d.[a]
Barley 1,3, 1,4-β-glucans	n.d.[a]	n.d.[a]	n.d.[a]
*p*NP-β-D-glucoside	36.1 ± 0.7	0.47 ± 0.03	77 ± 4
*p*NP-β-D-fucoside	24.5 ± 0.5	0.17 ± 0.07	144 ± 3
*p*NP-β-D-galactoside	27 ± 3	1.30 ± 0.10	20.7 ± 0.5
*p*NP-β-D-xyloside	3.2 ± 0.3	0.27 ± 0.05	11.9 ± 1.5
*p*NP-α-L-arabinoside	2.88 ± 0.08	0.14 ± 0.02	21 ± 3
*p*NP-β-D-cellobioside	2.07 ± 0.09	6.23 ± 0.17	0.34 ± 0.01
*p*NP-β-D- mannoside	n.d.[a]	n.d.[a]	n.d.[a]

[a] means not detectable

Similar to rice BGlu1, GH5BG had cleaved the β-glucosidic bond between the 2 glucose residues in *p*NP-β-D-cellobioside, thereby releasing glucose and *p*NPG, which was then rapidly hydrolyzed (data not shown), but the catalytic efficiency of GH5BG for hydrolysis of *p*NP-β-D-cellobioside is about 10 times lower than that of rice BGlu1 (Opassiri et al, 2003, 2004). Although GH5BG could hydrolyze both cellotriose and *p*NP-β-D-cellobioside, it hydrolyzed *p*NP-β-D-cellobioside about 26 times less efficiently than cellotriose. This suggests that the *p*NP-β-D-cellobioside, unlike cellotriose, cannot bind well to the third subsite in the active site cleft of GH5BG.

Opassiri et al. (2004) reported subsite affinity differences between barley and rice GH1 β-glucosidases, which each have 6 subsites for binding β-1,4-linked glucosyl residues. Barley βII β-glucosidase has a relatively low K_m (2.67 mM) and high k_{cat} (11.58 s^{-1}) for cellobiose, but cellotriose has a nearly 6-fold lower k_{cat} (Hrmova et al., 1998). In contrast, the k_{cat} of rice BGlu1 increases 10 fold as the DP increases from 2 to 3. The difference corresponded to a high affinity of subsite +2 for glucose residue 3 in BGlu1, while a negative interaction was seen at this site in the barley enzyme. Conversely, the binding affinity at site +1 is negligible for BGlu1, consistent with its high K_m and low k_{cat} for cellobiose. Because of this, the affinity of cellobiose binding nonproductively in subsites +2 and +3 is expected to be higher than binding in the productive position at the -1 and +1 subsites, so cellobiose is a better inhibitor than substrate. It is likely that GH5BG also binds to cellobiose in a nonproductive position better than in the productive position. Once the third glucosyl residue's interaction is added in cellotriose, the binding to the productive position becomes more favorable.

5. Conclusions

In this study, At least 139 genes homologous to glucanohydrolases which can be classified to be the members in 7 subfamilies of glycosyl hydrolase family in rice has been grouped. In addtion, the 2 rice β-glucosidase cDNAs, *Os4bglu12* and *GH5BG*, were cloned and recombinant proteins expressed in E. coli as soluble active form.

Os4bglu12 contig containing the deduced amino acid sequence which appeared to be very similar to the previously purified and characterized cell wall-bound β-glucosidase was cloned by RT-PCR. Recombinant Os4BGlu12 protein hydrolyzed β-linked oligosaccharides and *p*NP-glycosides. The specificity of Os4BGlu12 for oligosaccharides and *p*NP-glycosides was different from the previously characterized GH1 β-glucosidases/exoglucanases cell-wall-bound rice β-glucosidase, rice BGlu1, and barley βII β-glucosidase.

A putative GH5 glucan-1,3-β-glucosidase contains a fascin-like domain at the N-terminus was cloned from germinating rice. A recombinant thioredoxin-GH5BG produced in *E. coli* showed high hydrolytic activity toward various kinds of *p*NP-glycosides and exhibited a marked preference for β-(1,4)-linked oligosaccharides and laminaribiose (β-(1,3)-linked disaccharide). The substrate specificity of GH5BG is different from fungal GH5 exo-β-(1,3)-glucanases, which is likely due to differences in the structures of the loops around the active site, in addition to the structures and placement of the amino acid resides around the active site.

These results indicated that both enzymes has exoglucanase/β-glucosidase activity that might involve in the cleavage of oligosaccharides release from the hydrolysis of cell wall by endoglucanses, which is consistent with a role in cell wall metabolism.

Abbreviations

BGI, Beijing Genomic Institute; CDS, coding sequence; *C. albicans* Exg, Exo-β-(1,3)-glucanases of *C. albicans*; DP, degree of polymerization; EST, expressed sequence tag; GH, Glycosyl hydrolases; GH1, glycosyl hydrolase family 1; GH5, glycosyl hydrolase family 5; IMAC, immobilized metal affinity chromatography; pI, isoelectric points; MW, molecular weights; ORFs, open reading frames; *p*NP, *p*-nitrophenol; *p*NPG, *p*-nitrophenyl-β-D-glucoside; PGO, peroxidase/glucose oxidase.

6. เอกสารอ้างอิง (References)

Adams, J. C. (2004). Roles of fascin in cell adhesion and mobility. *Curr. Opin. Cell Biol.* 16: 590-596.

Adjei-Afriyie, F., Kim, C., Takemura, M., Ishikawa, M., and Horiike, M. (2000). Isolation and identification of the probing stimulants in rice plat for the white-black planthopper, Sogatella furcifera (Homoptera: Delphacidae). *Biosci. Biotechnol. Biochem.* 64: 443-446.

Akiyama T, Pillai MA, and Sentoku N. (2004). Cloning, characterization and expression of OsGLN2, a rice endo-1,3-β-glucanase gene regulated developmentally in flowers and hormonally in germinating seeds. *Planta.* 220: 80-86.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 25: 3389-3402.

Babcock, G. W. and Esen, A. (1994). Substrate specificity of maize β-glucosidase. *Plant Sci.* 101: 31-39.

Bendtsen, J. D., Nielsen, H., von Heijne, G. and Brunak, S. (2004). Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* 340: 783-795.

Buchanan, B.B., Gruissem, W. and Jones R.L. (2000). Biochemistry and Molecular Biology of Plants. John Wiley and Sons, Inc., New Jersey.

Carpita N.C. (1996). Structure and biogenesis of the cell walls of grasses. *Annu. Rev. Plant. Mol. Biol*. 47: 445-476.

Carpita, N. and McCann, M. (2000). The cell wall. In *Biochemistry and Molecular Biology of Plants*, BB. Buchanan, W. Gruissem and RLJones (eds.). American Society of Plant Physiologists, Maryland. pp. 52-108.

Cutfield, S. M., Davies, G. T., Murshudov, G., Anderson, B. F., Moody, P. C., Sullivan, P. A. and Cutfield, J. F. (1999). The structure of the exo-β-1,3-glucanase from *Candida albicans* in native and bound forms: relationship between a pocket and groove in family 5 glycosyl hydrolases. *J. Mol. Biol.* 294, 771-783.

Czjzek, M., Cicek, M., Zamboni, V., Burmeister, W.P., Bevan, D.R., Henrissat, B. and Esen, A. 2000. The mechanism of substrate (aglycone) specificity in β-glucosidases is revealed by crystal structures of mutant maize β-glucosidase-DIMBOA, -DIMBOAGlc, and -dhurrin complexes. *Proc. Natl. Acad. Sci*. USA. 97: 13555-13560.

Davies, G. J. and Henrissat, B. (2002). Plant glyco-related genomics. *Biochem Soc Trans*. 30: 291-297.

Doy, R. H. and Kosugi, A. (2004). Cellulosomes: Plant-cell-wall-degrading enzyme complexes. *Nature Reviews Microb*. 2: 541-551.

Hatfield, R. D. and Nevins, D. J. (1987). Hydrolytic activity and substrate specificity of an endoglucanase from *Zea mays* seedling cell wall. *Plant Physiol*. 83: 203-207.

Hoj, K. R. and Fincher G. B. (1995). Molecular evolution of plant β-glucan-endohydrolases. *Plant J*. 7: 367-379.

Hrmova, M., Harvey, A. J., Wang, J., Shirley, N. J., Jones, G. P., Stone, B. A., Hoj, P. B. and Fincher, G. B. (1996). Barley β-D-glucan exohydrolases with β-D-glucosidase activity. *J. Biol. Chem*. 271: 5277-5286.

Hrmova, M., MacGregor, E. A., Biely, P., Stewart, R. J. and Fincher, G. B. (1998) Substrate binding and catalytic mechanism of a barley β-D-glucosidase/(1,4)-β-D-glucan exohydrolase. J. Biol. Chem. **273**, 11134-11143

Hrmova, M. and Fincher, G. B. (2001a). Plant enzyme structure: Explaining substrate specificity and the evolution of function. *Plant Phyiol*. 125: 54-57.

Hrmova, M. and Fincher, G. B. (2001b). Structure-function relationship of β-glucan exohydrolases from higer plants. *Plant Mol. Biol.* 47: 73-91.

Huber, D. J. and Nevins, D. J. (1981). Partial purification of endo- and exo-β-D-glucanase enzymes from *Zea mays* L. seedlings and their involvement in cell wall autohydrolysis. *Planta*. 151: 206-214.

Iwami, K. and Yasumoto, K. (1986). Synthesis of pyridoxineβ-glucosidase and its in situ absorbtion in rat small intestine. Nutr. Res. 6: 407-414.

Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G. and Gibson, T. J. (1998) Multiple sequence alignment with Clustal X. *Trends Biochem. Sci*. 23: 403-405.

Kawasaki, S., Borchert, C. Deyholos, M. Wang, H., Brazille, S., Kawai, K. Galbraith, D., and Bonert, H. J. (2001). Gene expression profiles during the initial phase of salt stress in rice. *Plant Cell*. 13: 889-905.

Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Otomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y. (2003). Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. *Science*. 301: 376-379.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 227: 680-685.

Lever, M. (1972). A new reaction for colorimetric determination of carbohydrates. *Anal. Chem*. 47: 273-279.

Marana, S. R.(2006). Molecular basis of substrate specificity in family 1 glycoside hydrolases. *IUBMB Life*. 58: 63-73.

Nakai, K. and Horton, P. (1999). PSORT: a program for detecting the sorting signals of proteins and predicting their subcellular localization. *Trends Biochem. Sci.* 24: 34-35.

Nicholas, K. B. and Nicholas, H. B. (1997). Genedoc: a tool for editing and annotating multiple sequence alignments [Online]. Available: http://www.psc.edu/biomed/ gendoc.

Nishizawa, Y., Saruta, M., Nakazono, K., Nishio, Z., Soma, M., Yoshida, T., Nakajima, E., and Hibi, T. (2003). Characterization of transgenic rice plants over-expressing the stress-inducible β-glucanase gene Gns1. *Plant Mol. Biol.* 51: 143-152.

Opassiri, R., Ketudat Cairns, J. R., Akiyama, T., Wara-Aswapati, O., Svasti, J. and Esen, A. (2003). Characterization of a rice β-glucosidase highly expressed in flower and germinating shoot. *Plant Sci.* 165: 627-638.

Opassiri, R., Hua, Y., Wara-Aswapati, O., Akiyama, T., Esen, A. and Ketudat Cairns, J. R. (2004). β-Glucosidase, exo-β-glucanase and pyridoxine transglucosylase activities of rice BGlu1. *Biochem. J.* 379: 125-131.

Opassiri, R., Pomthong, B., Onkoksoong, T., Akiyama, T., Esen, A. and Ketudat Cairns J. R. (2006). Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12 β-glucosidase. *BMC Plant Biol.* 6: 33.

Rakwal R, Agrawal GK, and Yonekura M. (1999). Separation of proteins from stressed rice (*Oryza sativa* L.) leaf tissues by two-dimensional polyacrylamide gel electrophoresis: induction of pathogenesis-related and cellular protectant proteins by jasmonic acid, UV irradiation and copper chloride. *Electrophoresis.* 20: 3472-3478.

Reese, E. T., Maguire, A. H. and Parrish, F. W. (1968). Glucosidases and exo-glucanases. Can. J. *Biochem.* 46: 25-34.

Sakamoto, Y., Irie, T. and Sato, T. (2005). Isolation and characterization of a fruit body-specific exo-β-1,3-glucanase-encoding gene, exg1, from *Lentinula edodes*. *Curr. Genet*. 47: 244-252.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, New York.

Schliemann, W. (1984). Hydrolysis of conjugated gibberellins by β-glucosidases from dwarf rice (Oryza sativa L. cv. Tan-ginbozu). *J. Plant Physiol*. 116: 123-132.

Taiz L. (1984). Plant cell expansion: regulation of cell wall mechanical properties. *Annu. Rev. Plant physiol.* 35: 585-657.

Thomas BR, Inouhe M, Simmons CR, and Nevins DJ. (2000). Endo-1,3;1,4-β-glucanase from coleoptiles of rice and maize: role in the regulation of plant growth. *Int J Biol Macromol.* 27: 145-149.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 22: 4673-4680.

Wang, X., He, R., He, G. (2005). Construction of suppression subtractive hybridization libraries and identification of brown planthopper-induced genes. *J. Plant Physiol*. 162: 1254-1262.

Xu, Z., Shih, M. and Poulton. J. E. (2006). An extracellular exo-β-(1,3)-glucanase from *Pichia pastoris*: purification, characterization, molecular cloning and functional expression. *Protein Exp. Puri.* 47: 118-127.

Yamaguchi T, Nakayama K, Hayashi T, Tanaka Y, and Koike S. (2002). Molecular cloning and characterization of a novel β-1,3-glucanase gene from rice. *Biosci. Biotechnol. Biochem.* 66: 1403-1406.

Zheng BS, Yang L, Zhang WP, Mao CZ, Wu YR, Yi KK, Liu FY and Wu P. (2003). Mapping QTLs and candidate genes for rice root traits under different water-supply conditions and comparative analysis across three populations. *Theor. Appl. Genet*. 107: 1505-1515.

7. Output จากโครงการวิจัยที่ได้รับทุนจาก สกว.

7.1 ผลงานวิจัยที่ตีพิมพ์ในวารสารวิชาการระดับนานาชาติ

1) **Opassiri R**, Pomthong B, Onkoksoong T, Akiyama T, Esen A, and Ketudat Cairns JR. 2006. Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12 β-glucosidase. BMC Plant Biology. 6: 33. (unofficial Impact factor2007 3.44, official Impact factor due June 2008)

2) Opassiri R, Pomthong B, Akiyama T, Nakphaichit M, Onkoksoong T, Ketudat-Cairns M, and Ketudat Cairns JR. 2007. A stress-induced rice β-glucosidase represents a new subfamily of glycosyl hydrolase family 5 containing a fascin-like domain. Biochemical Journal. (in review) (Impact factor2005 4.224)

7.2 กิจกรรมอื่น ๆ ที่เกี่ยวข้อง

ได้แก่

(1) การเสนอผลงานในที่ประชุมนานาชาติ

Opassiri, R., Pomthong, B., Onkoksoong, T., Akiyama, T., and Ketudat Cairns, J.R. 2006. Functional genomic analysis of rice glycosyl hydrolase family 1. The 20th IUBMB International Congress of Biochemistry and Molecular Biology and 11th FAOBMB Congress, Kyoto, Japan, 18-23 June 2006.

(2) การเชื่อมโยงทางวิชาการกับนักวิชาการอื่นๆ ทั้งในและต่างประเทศ

- รศ. ดร. สุมนทิพย์ บุญนาค คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น เกี่ยวกับคำแนะนำและเทคนิควิธีในการศึกษาบทบาททางสรีรวิทยาของยีนในข้าว
- Dr. Takashi Akiyama Department of Low-Temperature Science, National Agricultural Research Center for Hokkaido Region ประเทศญี่ปุ่นเกี่ยวกับเทคนิคในการวิเคราะห์การทำงานของเอนไซม์ glucan hydrolase ในการย่อยสลายสับสเตรทกลุ่ม glycans และ oligosaccharides ชนิดต่าง ๆ

7.3 การประยุกต์ใช้ความรู้ที่ได้จากงานวิจัย

1) ข้อมูลที่ได้จากงานวิจัยนี้เป็นพื้นฐานในการทำงานวิจัยที่เกี่ยวข้องกับการศึกษาเอนไซม์ในกลุ่มกลูคาร โนไฮโดรเลสชนิดอื่นๆ ต่อไป

2) เอนไซม์ที่ผลิตได้ในรูปแบบรีคอมบิแนนท์โปรตีนในแบคทีเรียสามารถผลิตได้ในปริมาณมากและมีต้นทุนไม่สูงมาก จึงน่าเหมาะที่จะนำเอนไซม์นี้ไปใช้ในการศึกษาการย่อย biomass เพื่อนำไปใช้เป็นวัตถุดิบในการผลิตสารผลิตภัณฑ์จากน้ำตาลได้

ภาคผนวก

บทความสำหรับเผยแพร่

ผนังเซลล์พืชทุกชนิดประกอบขึ้นมาจากโพลีแซคคาไรด์หลากหลายชนิด พืชแต่ละชนิดมีส่วนประกอบและการจัดเรียงตัวของโพลิเมอร์ที่ผนังเซลล์แตกต่างกัน แหล่งของผนังเซลล์ซึ่งมีอยู่เป็นปริมาณมากที่เหลือทิ้งจากภาคเกษตรกรรม เช่น ฟางข้าวและข้าวบาร์เลย์ ซังข้าวโพด สามารถนำมาใช้เป็นวัตถุดิบในการแปรรูปเป็นสารผลิตภัณฑ์ของน้ำตาลได้ โครงสร้างพื้นฐานของผนังเซลล์พืชในกลุ่มธัญพืชและพืชตระกูลหญ้าแตกต่างจากพืชมีดอกอื่นๆ ตรงที่พืชเหล่านี้มีโพลิเมอร์ของน้ำตาลที่มีความซับซ้อนและหลากหลายมาก เอนไซม์ในกลุ่มกลูคาโนไฮโดรเลส เป็นเอนไซม์ที่มีบทบาทในกระบวนการสร้างและสลายผนังเซลล์พืชโดยเกี่ยวข้องกับการย่อยสลายโพลิเมอร์ของน้ำตาลกลูโคส เอนไซม์เหล่านี้ถูกจัดไว้เป็นสมาชิกกลุ่มย่อยของกลุ่มเอนไซม์ไกลโคซิลไฮโดรเลส กลูคาโนไฮโดรเลสแบ่งได้เป็น 3 ประเภท ได้แก่ เอนโดกลูคาเนส เอกโซกลูคาเนส และบีต้ากลูโคซิเดส เอนไซม์ เอนโดกลูคาเนสทำหน้าที่ย่อยพันธะที่อยู่ภายในโมเลกุลเซลลูโลสซึ่งเป็นส่วนประกอบของผนังเซลล์พืช จนได้โอลิโกแซคคาไรด์สายสั้นลง เอนไซม์เอกณโซกลูคาเนสจะย่อยกลูโคสออกจากด้านหัวหรือท้ายของโอลิโกแซคคาไรด์ก็ได้ ส่วนบีต้ากลูดคซิเดสย่อยกลูโคสจากปลายท้ายของโอลิโกแซคคาไรด์ได้เท่านั้น

ในการศึกษานี้พบยีนของข้าวที่จัดอยู่ในกลุ่มเอนไซม์กลูคาโนไฮโดรเลสอย่างน้อย 139 ยีน เพื่อที่จะทำให้เกิดความเข้าใจตลอดจนการนำเอนไซม์นี้มาใช้ประโยชน์ จึงได้นำเอนไซม์ในกลุ่มกลูคาโนไฮโดรเลส 2 ชนิด ที่อยู่ในไกลโคซิลไฮโดรเลสกลุ่มที่ 1 และ 5 ชนิดที่เป็นเอนไซม์บีต้ากลูโคซิเดส มาศึกษาหน้าที่และการเร่งปฏิกิริยา ได้เพิ่มจำนวนชิ้นส่วน cDNA ของบีต้ากลูโคซิเดสของข้าวทั้งสองชนิด ซึ่งได้ให้ชื่อเอนไซม์นี้ว่า Os4bglu12 และ GH5BG ด้วยวิธี RT-PCR และได้ใช้เทคนิคการตัดต่อ cDNA ดังกล่าวเข้ากับดีเอ็นเอพาหะแล้วนำเข้าสู่เซลล์แบคทีเรียเพื่อให้เกิดการผลิตเอนไซม์ในรูปแบบโปรตีนสายผสมให้ได้ปริมาณมาก (เรียกรีคอมบิแนนท์โปรตีน) เอนไซม์ที่ผลิตถูกนำมาศึกษาคุณสมบัติทางชีวเคมีและความจำเพาะต่อสารตั้งต้น พบว่าเอนไซม์ทั้งสองชนิดนี้มีความสามารถในการย่อยพันธะที่เชื่อมต่อระหว่างโมเลกุลกลูโคสของกลูโคโอลิโกแซคคาไรด์ที่มีกลูโคสต่อกันอยู่ 3-6 หน่วย ด้วยพันธะไกลโคซิดิกแบบ β-(1,4) และไดแซคคาไรด์ที่มีกลูโคสต่อกันด้วยพันธะไกลโคซิดิกแบบ β-(1,3) ได้ดีมาก และเอนไซม์ทั้งสองชนิดยังย่อย *p*NP-β-glycosides ได้หลายชนิด ซึ่งแสดงให้เห็นว่าเอนไซม์น่าจะสามารถย่อยโมเลกุลน้ำตาลชนิดอื่นนอกเหนือจากกลูโคสได้อีก จากการทดลองนี้ทำให้คาดว่าเอนไซด์นี้น่าจะมีบทบาทสำคัญในการย่อยสลายผนังเซลล์ โดยนำโอลิโกแซคคาไรด์หรือไดแซคคาไรด์ที่เกิดจากการย่อยผนังเซลล์ด้วยเอนโดกลูคาเนสหรือเอกโซกลูคาเนส มาย่อยต่อจนกระทั่งได้น้ำตาลโมเลกุลเดี่ยว

การศึกษาเอนไซม์โดยสกัดออกมาจากข้าวโดยตรงจำเป็นต้องใช้หลายขั้นตอน เอนไซม์ที่ได้มีปริมาณน้อยและมักจะถูกปนเปื้อนกับเอนไซม์อื่นๆที่มีสมบัติและขนาดใกล้เคียงกัน แต่เอนไซม์ที่ได้จากการนำ cDNA ของยีนที่สนใจมาใช้ผลิตเอนไซม์ในรูปรีคอมบิแนนท์โปรตีนจะสามารถถูกผลิตขึ้นได้ในปริมาณมากและมีความบริสุทธิ์สูง ดังนั้นจึงสามารถนำเอนไซม์ไปใช้ในการศึกษาสมบัติต่างๆ ตลอดจนการศึกษาโครงสร้างผลึกซึ่งอาจนำไปสู่การค้นพบกลไกการทำงานของเอนไซม์และการทำวิศวกรรมเอนไซม์ในขั้นต่อไป เอนไซม์ที่ได้จากการศึกษานี้อาจจะสามารถนำไปประยุกต์ใช้ในอุตสาหกรรมการย่อยสลายโพลิเมอร์ของน้ำตาลที่จะนำไปใช้ในการผลิตสารต่างๆ ได้ เช่น แอลกอฮอล์, กรดอะมิโน เป็นต้น

สัญญาเลขที่ MRG4880066

Reprint

BMC Plant Biology

Research article

Open Access

Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12 β-glucosidase

Rodjana Opassiri[1], Busarakum Pomthong[1], Tassanee Onkoksoong[1], Takashi Akiyama[2], Asim Esen[3] and James R Ketudat Cairns*[1]

Address: [1]Institute of Science, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand, [2]Department of Low-Temperature Science, National Agricultural Research Center for Hokkaido Region, Sapporo 062-8555, Japan and [3]Department of Biology, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0406, USA

Email: Rodjana Opassiri - opassiri@hotmail.com; Busarakum Pomthong - busarakum_p@yahoo.com; Tassanee Onkoksoong - onkoksoong@yahoo.com; Takashi Akiyama - takiyama@affrc.go.jp; Asim Esen - aevatan@vt.edu; James R Ketudat Cairns* - cairns@sut.ac.th

* Corresponding author

Published: 29 December 2006

Received: 19 September 2006

Accepted: 29 December 2006

BMC Plant Biology 2006, **6**:33 doi:10.1186/1471-2229-6-33

This article is available from: http://www.biomedcentral.com/1471-2229/6/33

Abstract

Background: Glycosyl hydrolase family 1 (GH1) β-glucosidases have been implicated in physiologically important processes in plants, such as response to biotic and abiotic stresses, defense against herbivores, activation of phytohormones, lignification, and cell wall remodeling. Plant GH1 β-glucosidases are encoded by a multigene family, so we predicted the structures of the genes and the properties of their protein products, and characterized their phylogenetic relationship to other plant GH1 members, their expression and the activity of one of them, to begin to decipher their roles in rice.

Results: Forty GH1 genes could be identified in rice databases, including 2 possible endophyte genes, 2 likely pseudogenes, 2 gene fragments, and 34 apparently competent rice glycosidase genes. Phylogenetic analysis revealed that GH1 members with closely related sequences have similar gene structures and are often clustered together on the same chromosome. Most of the genes appear to have been derived from duplications that occurred after the divergence of rice and *Arabidopsis thaliana* lineages from their common ancestor, and the two plants share only 8 common gene lineages. At least 31 GH1 genes are expressed in a range of organs and stages of rice, based on the cDNA and EST sequences in public databases. The cDNA of the *Os4bglu12* gene, which encodes a protein identical at 40 of 44 amino acid residues with the N-terminal sequence of a cell wall-bound enzyme previously purified from germinating rice, was isolated by RT-PCR from rice seedlings. A thioredoxin-Os4bglu12 fusion protein expressed in *Escherichia coli* efficiently hydrolyzed β-(1,4)-linked oligosaccharides of 3–6 glucose residues and laminaribiose.

Conclusion: Careful analysis of the database sequences produced more reliable rice GH1 gene structure and protein product predictions. Since most of these genes diverged after the divergence of the ancestors of rice and *Arabidopsis thaliana*, only a few of their functions could be implied from those of GH1 enzymes from *Arabidopsis* and other dicots. This implies that analysis of GH1 enzymes in monocots is necessary to understand their function in the major grain crops. To begin this analysis, Os4bglu12 β-glucosidase was characterized and found to have high exoglucanase activity, consistent with a role in cell wall metabolism.

Background

β-glucosidases (3.2.1.21) are glycosyl hydrolases that hydrolyze the β-O-glycosidic bond at the anomeric carbon of a glucose moiety at the nonreducing end of a carbohydrate or glycoside molecule. These enzymes are found essentially in all living organisms and have been implicated in a diversity of roles, such as biomass conversion in microorganisms [1] and activation of defense compounds [2,3], phytohormones [4,5], lignin precursors [6], aromatic volatiles [7], and metabolic intermediates by releasing glucose blocking groups from the inactive glucosides in plants [8]. To achieve specificity for these various functions, β-glucosidases must bind to a wide variety of aglycones, in addition to the glucose of the substrate.

The β-glucosidases that have been characterized to date fall predominantly in glycosyl hydrolase families 1 and 3 [9], with family 1 enzymes being more numerous in plants. Glycosyl hydrolase family 1 (GH1) contains a wide range of β-glycosidases, including β-galactosidases, β-mannosidases, phospho-β-galactosidases, phospho-β-glucosidases, and thioglucosidases, in addition to β-glucosidases. The plant enzymes in this family generally fall in a closely related subfamily, but, despite their high sequence similarity, display a wide range of activities. Besides β-glucosidases with diverse specificities, these plant enzymes include thio-β-glucosidases or myrosinases, β-mannosidases, disaccharidases, such as primeverosidase and furcatin hydrolase, and hydroxyisourate hydrolase, which hydrolyzes an internal bond in a purine ring, rather than a glycosidic linkage [7,9-11]. In addition, many enzymes in this group are capable of releasing multiple kinds of sugars from aglycones, such as isoflavonoid β-glucosidases, which can release the disaccharide acuminose and malonyl glucose, in addition to glucose itself, from isoflavonoids [12,13]. Other β-glucosidases in this subfamily may have high specificity for glucosides or glucosides and fucosides, or may hydrolyze other glycosides, such as β-galactosides, β-mannosides, and β-xylosides, as well. Primeverosidase has high specificity for primeverosides, with no hydrolysis of glucosides [7], while furcatin hydrolase can hydrolyze glucosides as well as disaccharide glycosides [10]. Clearly, plant family 1 glycosyl hydrolases show a range of sugar specificities.

Plant family 1 glycosyl hydrolases tend to show high specificity for their aglycones, though many hydrolyze synthetic, nonphysiological substrates, like *p*-nitrophenol (*p*NP)-β-glycosides [14]. The aglycones span a wide range of structures, including sugars [15-17], hydroxaminic acids [18], isoflavonoids [12,13], rotenoids [19], alkaloids [20,21] hydroxyquinones [3], cyanogenic nitriles [2], etc. It is the specificity for these aglycones which is thought to specify the function of most of these enzymes [14]. Since many β-glucosidases function in plants, it is important that these enzymes specifically hydrolyze their own substrates and not other substrates with which they may come into contact. It seems evident that the substrate specificity, localization of the enzymes with respect to potential substrates, and the activities of the substrates and hydrolysis products will determine the roles of these enzymes.

Xu *et al.* [22] described 47 GH1 genes in the *Arabidopsis* genome, including 7 apparent thioglucosidases, and one enzyme that had high β-mannosidase activity, in agreement with the prediction from its similarity to tomato β-mannosidase. With the completion of high quality drafts of the rice genome, a thorough analysis of GH1 can be conducted in rice. To date, only a few rice β-glucosidase isozymes have been functionally characterized, with the activities described being hydrolysis of gibberellin glucosides, pyridoxine glucosides and oligosaccharides [16,17,23,24].

To assess the functions of GH1 in rice, genes homologous to GH1 β-glucosidase genes have been identified from the rice genome, and their structures, predicted protein products and evidence of expression evaluated. In addition, we have cloned a β-glucosidase from germinating rice based on genomic data, and assessed its biochemical properties after expression in *E. coli*.

Results and discussion

Glycosyl hydrolase family 1 β-glucosidase family

The completion of the *Oryza sativa* L. spp. *japonica* Rice Genome Project and the complementary *indica* rice (*O. sativa* L. spp. *indica*) genome project by the Beijing Genomic Institute (BGI) has allowed genome-wide analysis of gene families in this important crop [25,26]. The sequence and mapping information provided to the public databases by these projects enabled us to identify the genes for glycosyl hydrolase family 1 members (putative β-glucosidases) in rice, determine their gene structures and genomic organization, and model their protein products and phylogenetic relationships. In this study, we used the DNA sequences of *japonica* rice in the Monsanto Rice Genome Sequencing Project, the Torrey Mesa Research Institute and GenBank at NCBI and the *indica* rice sequences of the BGI as the starting point to examine the sequences homologous to GH1 members by manual annotation. By examination of the gene structures and prediction based on the knowledge of other plant GH1 genes, we rectified any errors in gene structures from the automatic annotation by the Rice Genome Sequencing Project contigs. Thereafter, the GH1 members of *indica* rice were compared with those of *japonica* rice to identify which genes are orthologues (see Table 1). Finally, all contig sequences were searched against the completed sequences of the 12 rice chromosomes in GenBank to map each contig position on the chromosomes and identify the new GH1 members that were not present in the other databases. A new systematic code for the genes based on their chromosome location was devised with the chromosome number followed by a bglu number counting from the top of chromosome 1 through the bottom of chromosome 12 (Table 1). To avoid confusion, previously published synonyms for all family members are provided in Table 1. The retrieved gene sequences were searched against the dbEST and *japonica* rice full-length cDNA databases to determine the mRNA expression patterns of each gene in rice.

Forty β-glucosidase genes, including 34 full-length genes, 2 pseudogenes, 2 gene fragments, and 2 intronless genes, were identified, as listed in Table 1. Thirty-six out of 40 genes are found in both *japonica* and *indica* rice with 98–100% sequence identity. The *Os11bglu35* gene was present only in *japonica* rice sequences, while *Os11bglu37*, *Osbglu39* and *Osbglu40* were only found in *indica* rice. The thirty-eight mapped GH1 genes are distributed over all chromosomes, except chromosome 2 (Table 1). The *Osbglu39* and *Osbglu40* sequences have not been mapped to any chromosome, and it is possible they represent contamination of endophytic genes remaining in the *indica* genome draft. Twenty-two out of 40 gene sequences are derived from the automated annotation in the public databases and 18 genes are derived from manual annotation. We corrected 4 of 22 automated annotation contigs that had misassigned one or more intron-exon boundaries. *Os11bglu35* and *Os11bglu37* appear to be pseudogenes, since they have premature stop codons and cannot produce full-length proteins.

The size of rice GH1 is not unexpected, since a search of the *Arabidopsis thaliana* genome identified 47 glycosyl hydrolase family 1 homologues, including 8 probable pseudogenes and 3 intronless genes, which are distributed throughout all five chromosomes [22]. The slightly larger size of the family in *Arabidopsis* may be due to the presence of myrosinases, which are not found in rice, and a larger number of pseudogenes. The large size of both rice and *Arabidopsis* GH1 may reflect different substrate specificity and expression patterns in rice tissues and/or in response to environmental conditions among the GH1 members.

The presence of many GH1 genes in rice suggests they may hydrolyze an array of possible substrates, depending on their substrate specificity and localization with respect to the substrates. Although a number of glycosides that could serve as potential substrates for rice GH 1 β-glucosidases have been purified from rice tissues, there have been few reports about the hydrolysis of these substrates by the enzymes. The major glycosides found in various tissues of rice include glycosylsterols, flavonoid glucosides, hormone glucosides, a vitamin glucoside, and pantonic acid glucoside. Glycosylsterols found in rice are glycosyl-sitosterol, -campesterol and -stigmasterol in rice bran [27] and β-sitosterol-3-O-β-D-glucoside in rice hulls [28]. The major flavonoid glucosides present in rice include 1) anthocyanins, such as cyanidin-O-β-D-glucoside and peonidin-O-β-D-glucoside, in black rice [29,30]; 2) tricin-O-glucoside in rice hulls, bran, leaf and stem [28,31]; and 3) hydroxycinnamate sucrose esters, such as 6'-O-feruloylsucrose and 6'-O-sinapoylsucrose in germinated brown rice [32]. Hormone glucosides found in rice include gibberellin glucosides in ungerminated seeds and anther [23,33], salicylic glucoside [34] and indole-3-acetic acid (IAA)-glucoside [35]. Pyridoxine-β-D-glucoside was found in rice bran, callus and seedling [36-38]. Another glycoside, namely R(-) pantoyllactone-β-D-glucoside, was found in the shoots but not the roots of rice seedlings [39].

Many compounds (including glycosides) have been found in rice tissues in response to environmental stresses and in transgenic rice plants. Recently, it was found that there is a high accumulation of IAA-glucoside in tryptophan-overproducing transgenic rice [35] and of salicylic glucoside in rice overproducing NH1, a key regulator of salicylic acid mediated systematic acquired resistance, in transgenic rice [34]. The level of pyridoxine glucoside was reported to be increased by the application of pyridoxine to rice callus and germinating seeds [37,38]. Markham *et al.* [40] reported that exposing UV-tolerant rice to high UV-B levels increased the levels of flavone glucosides. These results may indicate that the presence of high amounts of some metabolic compounds is corrected by converting them to the glucoside-conjugated forms. It still needs to be shown whether or not these compounds are later reactivated by β-glucosidases.

Protein sequence alignment and phylogenetic analysis

The open reading frames (ORFs) of thirty-seven gene-derived cDNAs (excluding *Os11bglu36*, *Osbglu39* and *Osbglu40*, which are more closely related to bacterial GH1 genes) showed a high level of shared deduced amino acid sequence identity to each other and other known plant β-glucosidase sequences. All deduced β-glucosidase protein sequences contain the putative catalytic acid/base and nucleophilic glutamate residues, except Os4bglu14 and Os9bglu33, in which the acid/base glutamate is replaced with glutamine, as seen in thioglucosidases. The catalytic acid/base and nucleophile consensus sequences are: W-X-T/I-F/L/I/V/S/M-N/A/L/I/D/G-E/Q-P/I/Q and V/I/L-X-E-N-G, respectively, with relative frequencies of amino acids at each position shown in Figure 1. These sequences are similar to the consensus sequences previously derived from known GH1 β-glucosidase sequences [41,42]. The

Table 1: Summary of identified genes homologous to glycosyl hydrolase family 1 glucosidase

Gene name	BGI ID (AAAA...)[a]	RGP GenBank ID[c]	Gene locus ID/position[e]/Chr[f]	Gene pattern	Corresponding cDNAs[g]	Number ESTs[h]	Tissue libraries[i]	Comment
Os1bglu1	02002143 (F) 02002142 (aa 110–189[b])	AP003217 (F) (BAD73293[d]) AP008207 (F)	AP008207/17752382 bp-17760802 bp/chr 1	2	AK069177 (F) AK060988 (n)	13	sh, pn, wh-TL, 2 wk lf-ABF3	
Os1bglu2	02004130 (aa 1–105) 02004129 (aa 106–561)	AP003570 (F) AP004331 (F) AP008207 (F)	AP008207/34595732 bp-34582220 bp/chr 1	1	-	4	pn-FW, wh-TL, 35 d lf-Dr, 3 wk lf-Bl	
Os1bglu3	02004130 (F) 02004127 (aa 134–288)	AP003570 (F) AP004331 (F) AP008207 (F)	AP008207/34604232 bp-34599017 bp/chr 1	1	AK067934 (F?) AK063065 (n)	4	sh, 2 wk lf-ABF3, 3 wk lf-Bl	
Os1bglu4	02004468 (aa 1–414) 02004470 (aa 426–479)	AP003349 (F) (BAD82183) AP003418 (F) (BAD82346) AP008207 (F)	AP008207/38998348 bp-39003033 bp/chr 1	1	-	9	sh, pn-FW, pn-FW-Dr, 3 wk lf-Bl	
Os1bglu5	02004619 (F)	AP003272 (F) (BAD87322) AP004330 (F) (BAD88178) AP008207 (F)	AP008207/40834604 bp-40840341 bp/chr 1	1	AK070499 (F) AK119221 (F)	23	sh, st-IM, pn, pn-FW, wh-TL, wh-BT, wh-TF, 2 wk lf-AtJMT, lf-Dr, 3 wk lf-Ls	
Os3bglu6	02008013 (F)	AC146619 (F) AP008209 (F)	AP008209/5850657 bp-5844110 bp/chr 3	1	AY129294 (F) AK119546 (F)	14	sh, pn-FW, cl-Co, 3 wk lf-Bl	
Os3bglu7	02010831 (aa 1–99) 02006516 (aa 100–504)	AC091670 (F) (AAX95519) AC133334 (F) (AAS07254) AP008209 (F)	AP008209/28041529 bp-28037050 bp/chr 3	2	OSU28047 (F)AK100165 (F) AK103027 (F) AK105026 (F) AK059920 (n)	326	cl, sh, rt-SD, st-IM, pn, pn-FW, wh-TL, cl-ABA, cl-NAA, cl-BAP, cl-Cd, cl-heat, cl-Co; sh-UV, sh-Co, 35 d lf-Dr, 3–4 wk rt-Sa, 2 wk lf-ABF3, 2 wk cl-HDAC1, 3 wk lf-Bl, lf-M-Bl	*bglu1*
Os3bglu8	02010831 (F)	AC091670 (F) (AAX95520) AC133334 (F) (AAS07251) AP008209 (F)	AP008209/28050325 bp-28045526 bp/chr 3	2	AK120790 (F) AK105850 (n)AK059517 (n)	77	sh, pn, pn-FW, wh-TL, wh-TF, cl-BAP, sh-Co, 2 wk lf ABF3	
Os4bglu9	02014146 (F)	AL731582 (F) AP008210 (F)	AP008210/23697091 bp-23691010 bp/chr 4	1	AK066908 (F?)	11	sh, lf-IM, 3–4 wk rt-Sa	
Os4bglu10	02014151 (F)	AL731582 (F) (CAE05481) AP008210 (F)	AP008210/23708851 bp-23703582 bp/chr 4	1	AK065793 (F) AK062029 (F) AK073031 (n) AK068304 (n)	17	sh, lf-M, wh-TL, 2 wk lf ABF3, 2 wk lf-AtJMT, 3 wk lf-Bl	
Os4bglu11	02014151 (F)	AL731582 (F) (CAE05482) AP008210 (F)	AP008210/23717688 bp-23710742 bp/chr 4	1	-	4	sh-Co	
Os4bglu12	02014151 (F)	AL731582 (F) (CAE05483) AP008210 (F)	AP008210/23728066 bp-23723058 bp/chr 4	1	AK062776 (n) AK100820 (n) AK105375 (n)	30	cl, sh, 2 wk lf and rt, sp, wh-TL, wh-TF, 1 wk rt-Sa, sd-Co, pn-FW-Dr, 2 wk cl-HDAC1, 2 wk sd-Ph, 3 wk lf-Bl, lf-BT-Xa	
Os4bglu13	02014151 (F) 02014154 (aa 465–520)	AL73182 (F) (CAE05485) AP008210 (F)	AP008210/23742711 bp-23738108 bp/chr 4	1	AK070962 (F)	22	sh, pn, wh-TL, wh-TF, 3 wk lf-Wd, 3 wk lf-Bl, lf-M-Bl	
Os4bglu14	02014354 (F)	AL606622 (F) (CAE03397) AP008210 (F)	AP008210/25617791 bp-25613930 bp/chr 4	3	AK067841 (F)	1	sh	
Os4bglu15	02014355	AL606622 (n) (CAE003399) AP008210 (n)	AP008210/25626016 bp-25623565 bp/chr 4	-	-	0		Gene fragment, lacks exon 1–8
Os4bglu16	02014360 (aa 1–69) 02014359 (aa 70–516)	AL606622 (F) (CAE54544) AL606659 (F) (CAE01908) AP008210 (F)	AP008210/25631832 bp-25640157 bp/chr 4	3	AK066850 (F?) AK068772 (F?)	14	rt-SD, sh, pn, pn-FW, wh-TL, cl-Co, 3 wk rt-Sa, 3 wk lf-Bl, lf-M-Bl	
Os4bglu17	02014358	AL606622 (n) AL606659 (n) AP008210 (n)	AP008210/25646002 bp-25648366 bp/chr 4	-	-	0		Gene fragment lacks exon 9–13
Os4bglu18	02014362 (aa 1–46) 02014361 (aa 47–505)	AL606622 (F) (CAE01910) AL606659 (F) (CAE54546) AP008210 (F)	AP008210/25667349 bp-25654991 bp/chr 4	3	AK058333 (n)	10	sh, pn-FW, 3 wk lf-Bl	
Os5bglu19	02017035 (F) 02016858 (aa 1–272)	AC121366 (F) (AAS79738) AC135927 (F) AC137618 (F) AP008211 (F)	AP008211/17386160 bp-17389960 bp/chr 5	1	AK105546 (F?)	5	pn-FW, pn-FW-Dr, 2 wk lf- AtJMT, 3 wk lf-Wd	
Os5bglu20	02016859 (F) 02017035 (F)	AC121366 (F) AC137618 (F) AP008211 (F)	AP008211/17403620 bp-17407871 bp/chr 5	1	AK120998 (F?)	0		
Os5bglu21	02016862 (F)	AC121366 (F) AC137618 (F) AP008211 (F)	AP008211/17421799 bp-17427364 bp/chr 5	1	-	0		
Os5bglu22	02016869 (F) 02016867 (aa 1–61)	AC121366 (F) AC137618 (F) (AAV31358) AP008211 (F)	AP008211/17450999 bp-17456012 bp/chr 5	1	AK071469 (F)	39	sh, lf-M, pn-FW, cl-BAP, cl-NAA, 3 wk lf-Ls, 3 wk lf-Bl, lf-M-Bl	