



รายงานวิจัยฉบับสมบูรณ์
Final Report

โครงการ ความสัมพันธ์ระหว่างโครงสร้างและความจำเพาะต่อสับสเตรทของเอนไซม์
เบต้ากลูโคซิเดสในตระกูลไกลโคซิลไฮโดรเลส 1 และ 3 ที่ย่อยกลูโคโอลิโกแซคคาไรด์

**Structural Basis for Substrate-Specificity in Glucoooligosaccharide Hydrolyzing
 β -Glucosidases from Glycosyl Hydrolase Families 1 and 3**

โดย ศ. ดร. เจมส์ เกตุทัต-คาร์นส์ และคณะ
By Prof. Dr. James R. Ketudat-Cairns and Team

สิงหาคม 2556

รายงานวิจัยฉบับสมบูรณ์
Final Report

โครงการ ความสัมพันธ์ระหว่างโครงสร้างและความจำเพาะต่อสับสเตรทของเอนไซม์
เบต้ากลูโคซิเดสในตระกูลไกลโคซิลไฮโดรเลส 1 และ 3 ที่ย่อยกลูโคโอลิโกแซคคาไรด์

**Structural Basis for Substrate-Specificity in Glucoooligosaccharide Hydrolyzing
 β -Glucosidases from Glycosyl Hydrolase Families 1 and 3**

คณะผู้วิจัย สังกัด

Research Team and Affiliations

Prof. Dr. James R. Ketudat-Cairns, Head of Project,
Institute of Science, Suranaree University of Technology

Prof. Dr. M.R. Jisnuson Svasti, Advisor,
Laboratory of Biochemistry, Chulabhorn Research Institute &
Centre for Protein Structure and Function, Faculty of Science,
Mahidol University

Assist. Prof. Dr. Rodjana Opasiri, Coinvestigator in first year,
Institute of Science, Suranaree University of Technology

Dr. Sukanya Luang, Postdoctoral Researcher, 2011-2012
Institute of Science, Suranaree University of Technology

Dr. Salila Pengthaisong, Postdoctoral Researcher, 2012-2013
Institute of Science, Suranaree University of Technology

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัยและมหาวิทยาลัยเทคโนโลยีสุรนารี
(ความเห็นในรายงานนี้เป็นของผู้วิจัย
สกว. และมหาวิทยาลัยเทคโนโลยีสุรนารีไม่จำเป็นต้องเห็นด้วยเสมอไป)

Thailand Research Fund Contract Number BRG5380017
Structural Basis for Substrate-Specificity in Glucopoligosaccharide Hydrolyzing β -
Glucosidases from Glycosyl Hydrolase Families 1 and 3
Final Report

Grant Period: 1 November, 2010- 31 July, 2013.

Project Title: **Structural Basis for Substrate-Specificity in Glucopoligosaccharide**
Hydrolyzing β -Glucosidases from Glycosyl Hydrolase Families 1 and 3

(Thai): ความสัมพันธ์ระหว่างโครงสร้างและความจำเพาะต่อสับสเตรทของเอนไซม์เบต้ากลูโคซิเดสในตระกูลไกลโคซิลไฮโดรเลส 1 และ 3 ที่ย่อยกลูโคโอลิโกแซคคาไรด์

Head of Project: Dr. James R. Ketudat-Cairns

Institution: Institute of Science, Schools of Biochemistry and Chemistry,
Suranaree University of Technology
Nakhon Ratchasima 30000, THAILAND
Phone: 044 224304
Fax: 044 224185
Email: cairns@sut.ac.th, jrkcairns@yahoo.com

Advisor: Prof. Dr. M.R. Jisnuson Svasti Centre of Excellence in Protein Structure
and Function & Dept. of Biochemistry
Faculty of Science
Mahidol University
Phayathai, Bangkok 10400
Laboratory of Biochemistry
Chulabhorn Research Institute
Laksi, Bangkok

Abstract

Exoglucanases and β -glucosidases release glucose from the nonreducing end of oligosaccharides and glycosides and play important roles in biological glucoside turnover, in addition to have many potential applications. In order to understand the structural basis for binding oligosaccharides and glycoside substrates by plant enzymes of glycoside hydrolase (GH) families GH1 and GH3, we utilized a series of functional, mutational and X-ray crystallographic experiments on rice and barley enzymes. We solved the structure of Os4BGlu12 in apo form and in complexes with 2,4-dinitrophenyl- β -D-2-deoxy-2-fluoroglucoside (DNP2FG) and 2-fluoroglucoside (G2F) derived from this inhibitor. The Os4BGlu12 enzyme was also found to hydrolyze tuberonic acid glucoside (TAG) and other glycosides, and the structure of the Os4BGlu12 E179Q mutant with TAG was solved at low resolution. Similarly, Os3BGlu6 was found to hydrolyze gibberellin A4 β -D-glucose ester and a structure with a covalent glucose from this substrate was determined. Comparison of the structures of Os4BGlu12, Os3BGlu6, and Os3BGlu7 E176Q in complex with cellopentaose and molecular docking suggested that the Met251 of Os3BGlu6 may interfere with celooligosaccharide binding in Os3BGlu6, while the His253 in Os4BGlu12 is compatible with such binding. Mutation of the Os3BGlu6 Met251 to Asn improved oligosaccharide hydrolysis, whereas conversion of the corresponding Os3BGlu7 and Os4BGlu12 residues to Met decreased hydrolysis. The roles of apparent oligosaccharide-binding residues in the active site cleft of Os3BGlu7 in hydrolysis and in the transglycosylation of long oligosaccharides by the Os3BGlu7 E386G glycosynthase were also studied by mutagenesis, kinetic analysis and structure determination. The removal of Tyr341 had relatively small effects on hydrolysis, due to the binding of celooligosaccharides in different positions, depending on the inactivating mutation, while this residue was more critical for the transglycosylation of long glucooligosaccharides, which require proper alignment of the β -1,4-linked glucosyl acceptor chain. Although the structures of the Os3BGlu7 E386A, E386G and E386S in complexes with α -fluoroglucoside donor substrate showed little difference, QM/MM suggested the subtle differences in the interactions with a critical water molecule that helps to stabilize the leaving group can explain the higher relative activity for Os3BGlu7 compared to E386S and E386A, respectively. Although Os3BGlu7 β -glucosidase prefers 4-nitrophenyl β -D-glucoside (4NPGlc) over 4NP- β -D-mannoside (4NPMan), Os7BGlu26 β -mannosidase has the opposite preference. The structures of Os7BGlu26 and its complex with mannose were determined and showed the active site is very similar to Os3BGlu7 β -glucosidase, though slightly narrower. Docking of 4NPGlc and 4NPMan substrates into the active site and comparison to related β -glucosidase structures suggested differences in amino acids interacting with the catalytic acid-base between enzymes with significant and those with insignificant β -mannosidase activities. Mutation of Os7BGlu26 Tyr334, which is closely linked to the catalytic acid-base, resulted in significant changes in its activities for 4NPGlc and 4NPMan. Further studies with gluco- and manno-configured transition state analogs suggest that the Os7BGlu26 β -mannosidase may hydrolyze glucosides and mannosides via different transition state conformations. For GH3 enzymes, analysis of a large number of active site variants of barley HvExoI at the functional and structural levels has resulted in a clearer understanding of the broad specificity of the enzyme and the mechanism of product-substrate exchange, although finalization of these analyses is still in progress. We have supplemented this with characterization of rice GH3 isoenzymes OsExoI, expressed in *Pichia pastoris*, and OsExoII and OsExoIII, expressed in *Escherichia coli*. Aside from adding to our understanding of exoglucanase and β -glucosidase interactions with substrates for future enzyme engineering, these studies have contributed to 9 publications and scientific training for 5 students.

บทคัดย่อ

เอนไซม์เอกโซกลูคาเนสและบีต้ากลูโคซิเดสปล่อยกลูโคสจากปลาย nonreducing ของโอลิโกแซคคาไรด์และไกลโคไซด์และเอนไซม์นี้มีบทบาทสำคัญในการหมุนเวียนกลูโคไซด์ในทางชีวภาพรวมทั้งมีการประยุกต์ใช้งานที่มีศักยภาพมากมาย เพื่อให้เข้าใจถึงโครงสร้างพื้นฐานสำหรับการจับของสารตั้งต้นโอลิโกแซคคาไรด์และไกลโคไซด์ด้วยเอนไซม์ไกลโคไซด์ไฮโดรเลส (GH) จากพืชที่อยู่ในตระกูล GH1 และ GH3 เราใช้ชุดของการทดลองการทำงานโดยทำให้เกิดการกลายพันธุ์ของเอนไซม์จากข้าวและข้าวบาร์เลย์และหาโครงสร้างด้วยวิธี X-ray crystallography เราหาโครงสร้างของเอนไซม์Os4BGlu12 ในรูปแบบอิสระและโครงสร้างเชิงซ้อนกับสารยับยั้ง 2,4-dinitrophenyl-fluoroglucoside β -D-2-deoxy-2-(DNP2FG) และ 2-fluoroglucoside (G2F) ที่มาจากสารยับยั้งนี้ เอนไซม์ Os4BGlu12 สามารถย่อย tuberonic acid glucoside (TAG) และไกลโคไซด์อื่นๆ และสามารถหาโครงสร้างของเอนไซม์กลายพันธุ์ Os4BGlu12 E179Q ที่มี TAG ที่สามารถหักเหรังสีเอ็กซ์ได้ดีที่ความละเอียดต่ำ ในทำนองเดียวกันเอนไซม์ Os3BGlu6 สามารถย่อย gibberellin A4 β -D-glucose ester และเกิดพันธะโควาเลนต์เชิงซ้อนในโครงสร้างของเอนไซม์กับน้ำตาลกลูโคสจากสารตั้งต้นนี้ การเปรียบเทียบโครงสร้างของ Os4BGlu12 Os3BGlu6 และ Os3BGlu7 E176Q ที่เป็นโครงสร้างเชิงซ้อนกับ cellopentaose และ molecular docking ชี้ให้เห็นว่า Met251 ของ Os3BGlu6 อาจรบกวนการจับของ celooligosaccharide ใน Os3BGlu6 ในขณะที่ His253 ใน Os4BGlu12 สัมพันธ์กับการจับดังกล่าว การกลายพันธุ์ของ Os3BGlu6 Met251 ไปเป็น Asn ปรับปรุงการย่อยสลายโอลิโกแซคคาไรด์ ในขณะที่การกลายพันธุ์ที่สอดคล้องกันของ Os3BGlu7 และ Os4BGlu12 ไปเป็น Met พบว่าทำให้การย่อยสลายโอลิโกแซคคาไรด์ลดลง บทบาทของกรดอะมิโนที่จับกับโอลิโกแซคคาไรด์ที่ร่องของบริเวณเร่งของ Os3BGlu7 ในการย่อยสลายและในการเคลื่อนย้ายหมู่น้ำตาลของโอลิโกแซคคาไรด์สายยาวโดยเอนไซม์ไกลโคซิเนส E386G Os3BGlu7 จากการกลายพันธุ์ การวิเคราะห์การทำงานและการหาโครงสร้างพบว่าการจัด Tyr341 มีผลขนาดค่อนข้างเล็กต่อการย่อยสลาย เนื่องจากการจับของ celooligosaccharide ในตำแหน่งที่แตกต่างกันขึ้นอยู่กับกลายพันธุ์เพื่อยับยั้งการทำงานของเอนไซม์ ในขณะที่กรดอะมิโนนี้มีผลมากต่อการเคลื่อนย้ายหมู่น้ำตาลของกลูโคโอลิโกแซคคาไรด์สายยาว ซึ่งจำเป็นต้องมีการจัดตำแหน่งที่เหมาะสมของสาย β -1,4 ของตัวรับหมู่น้ำตาลกลูโคส แม้ว่าโครงสร้างของ Os3BGlu7 E386A E386G และ E386S ในโครงสร้างเชิงซ้อนกับสารตั้งต้นตัวให้หมู่น้ำตาลกลูโคส α -fluoroglucoside แสดงให้เห็นความแตกต่างเล็กๆ น้อยๆ แต่ QM/MM ชี้ให้เห็นว่าความแตกต่างในการมีปฏิสัมพันธ์กับโมเลกุลของน้ำที่สำคัญที่ช่วยในการรักษาเสถียรภาพของกลุ่มปล่อยสามารถอธิบายความสามารถในการทำงานที่มากกว่าของ E386G สำหรับ Os3BGlu7 เทียบกับ E386S และ E386A ตามลำดับ แม้ว่า Os3BGlu7 β -glucosidase ชอบ 4 nitrophenyl β -D-glucoside (4NPGlc) มากกว่า 4NP- β -D-mannoside (4NPMan) Os7BGlu26 β -mannosidase มีความชอบตรงข้ามกับ Os3BGlu7 โครงสร้างของ Os7BGlu26 และโครงสร้างเชิงซ้อนกับแมนโนสแสดงให้เห็นว่ามีบริเวณเร่งคล้ายกับ Os3BGlu7 β -glucosidase แม้ว่าจะแคบกว่าเล็กน้อย Docking ของสารตั้งต้น 4NPGlc และ 4NPMan ในบริเวณเร่งและการเปรียบเทียบความสัมพันธ์ของเอนไซม์บีต้ากลูโคซิเดสแสดงให้เห็นว่าความแตกต่างของกรดอะมิโนที่มีปฏิสัมพันธ์กับตัวเร่งปฏิกิริยากรด-เบสระหว่างเอนไซม์มีนัยสำคัญและไม่สำคัญต่อการทำงานเป็นเอนไซม์บีต้าแมนโนซิเดส การกลายพันธุ์ของ Os7BGlu26 Tyr334 ซึ่งอยู่ใกล้กับตัวเร่งปฏิกิริยากรด-เบส ส่งผลต่อการเปลี่ยนแปลงที่สำคัญในการทำงานสำหรับการย่อยของ 4NPGlc และ 4NPMan การศึกษาต่อไป gluco- และ manno- กำหนด transition state analogs ชี้ให้เห็นว่า Os7BGlu26 β -mannosidase อาจย่อยกลูโคไซด์และแมนโนไซด์ ผ่าน transition state conformations ที่แตกต่างกัน สำหรับเอนไซม์ในตระกูล GH3 การวิเคราะห์บริเวณเร่งของเอนไซม์กลายพันธุ์ของข้าวบาร์เลย์ HvExoI ในระดับการทำงานและโครงสร้าง ทำให้มีการเข้าใจที่ชัดเจนของความจำเพาะกว้างของเอนไซม์และกลไกของการแลกเปลี่ยนผลิตภัณฑ์-สารตั้งต้น แม้ว่าการสรุปจากการวิเคราะห์เหล่านี้ยังคงอยู่ในการรายงานความคืบหน้า เราได้เพิ่มเติมการศึกษาคุณสมบัติของข้าว GH3 isoenzymes OsExoI ที่แสดงออกใน *Pichia pastoris* และ OsExoII และ OsExoIII ที่

แสดงออกในเชื้อ *Escherichia coli* นอกเหนือจากการเพิ่มความเข้าใจของเราเกี่ยวกับเอนไซม์
เอกโซกลูคาเนสและปฏิสัมพันธ์ของเอนไซม์บีต้ากลูโคซิเดสกับสารตั้งต้นสำหรับงานวิศวกรรม
เอนไซม์ในอนาคต การศึกษาเหล่านี้ได้ผลิตผลงานตีพิมพ์ในวารสารนานาชาติ 9 ฉบับและการ
ฝึกอบรมทางวิทยาศาสตร์สำหรับนักศึกษาระดับบัณฑิตศึกษา 5 คน

Executive Summary

Importance: β -Glucosidases and exoglucanases are important enzymes with many functions in plants and other organisms, and also play an important role in lignocellulosic biomass conversion. Understanding the mechanisms by which they bind to oligosaccharides and select their substrates is critical to identifying the functions of proteins identified from genomic sequencing and engineering enzymes for better application.

Objectives: The objectives of this work included:

- 1) Structural determination of inactive mutants of the GH1 enzymes Os3BGlu6 and Os4BGlu12 in complexes with oligosaccharide and glycoside substrates and comparison to the structures of Os3BGlu7 in complex with its substrates.
- 2) Determination of the roles of some active site residues in determining the differences in activities between Os3BGlu6, Os3BGlu7 and Os4BGlu12 and related rice GH1 isoenzymes by mutagenesis.
- 3) Elucidation of catalytic and structural differences between GH1 β -glucosidases and β -mannosidases
- 4) Investigation of the functional differences in GH3 exoglucanases from barley and rice and analysis of the structural basis of these difference, based on the barley Exo-I structure.
- 5) Determination of the functional roles of amino acids in the GH3 exoglucanase and β -glucosidase active sites by mutagenesis and evaluation of functional changes.

Methods and Results: The X-ray crystallographic structure of the Os4BGlu12 β -D-glucosidase was completed, including structures with 2,4-dinitrophenyl- β -D-2-deoxy-2-fluoroglucoside in Michaelis and covalent intermediate complexes, and a preliminary structure of the Os4BGlu12 E179Q mutant in complex with tuberonic acid glucoside. Comparison of the structures of Os3BGlu7 in complexes with cellooligosaccharides and those of Os3BGlu6 and Os4BGlu12 identified the Os3BGlu7 residue Asn245 as a residue hydrogen-bonding to oligosaccharides, while the corresponding Os3BGlu6 Met251 appeared to block such binding. Site directed mutagenesis of Os3BGlu6 Met251 to Asn created a 1,4-linked β -D-glucosyl binding site in Os3BGlu6, based on the kinetics of oligosaccharide hydrolysis, while converting the corresponding residues in Os3BGlu7 and Os4BGlu12 to Met resulted in decreased binding and hydrolysis. Mutagenesis of this residue and Tyr341 in Os3BGlu7 and its glycosynthase mutant resulted in differential effects on the hydrolysis and synthesis of oligosaccharides, and the structure of the Os3BGlu7 Y341A/E386G mutant showed cellotetraose in the active site in an alternate binding mode, suggesting a high level of plasticity in oligosaccharide binding by the Os3BGlu7 β -D-glucosidase. We also observed that the Os3BGlu7 E386G glycosynthase mutant accepts a diversity of acceptor and donor substrates, but forms extended oligosaccharides primarily with β -1,4-linked monosaccharides. Comparison of the X-ray crystal structures of Os3BGlu7 E386G with the less efficient glycosynthase mutants E386A and E386S showed the structures were very similar, but quantum mechanical/molecular mechanical (QM/MM) calculations were able to verify the E386G structure should give more efficient glycosyl transfer, due in part to the disposition of a critical water molecule that assists in fluoride departure for the α -D-fluoroglucoside donor substrate. Solving the X-ray crystal structures of the Os7BGlu26 β -D-mannosidase and its complex with mannose allowed identification of residues around the catalytic acid base, which site-directed mutagenesis confirmed affect the relative activities toward glucoside and mannoside substrates. For GH3 exoglucanases, we completed several mutants around the active site and the X-ray crystallographic structures of many complexes are in progress, while kinetic characterization of 3 rice GH3 exoglucanases was begun.

Output and Benefits: This work elucidated the structural basis of oligosaccharide hydrolysis, and contributed to 9 international publications and 5 graduate students' training.

BRG5380017

1. Project Name

Thai name: ความสัมพันธ์ระหว่างโครงสร้างและความจำเพาะต่อสับสเตรทของเอนไซม์เบต้ากลูโคซิเดสในตระกูลไกลโคซิลไฮโดรเลส 1 และ 3 ที่ย่อยกลูโคโอลิโกแซคคาไรด์
English name: Structural Basis for Substrate-Specificity in Glucopoligosaccharide Hydrolyzing β -Glucosidases from Glycosyl Hydrolase Families 1 and 3

2. Project Participants

2.1.1 Name: James R. Ketudat Cairns

2.1.2 Degree: Ph.D.

2.1.3 Position: Professor of Biochemistry

2.1.4 Work address: Institute of Science, Suranaree University of Technology,
111 University Avenue, Muang District, Nakhon Ratchasima 30000

2.1.5 Duties and responsibilities in project: Head of project, planning and oversight of project

2.2.1 Name: Rodjana Opassiri

2.2.2 Degree: Ph.D

2.2.3 Position: Instructor of Biochemistry

2.2.4 Work address: Institute of Science, Suranaree University of Technology,
111 University Avenue, Muang District, Nakhon Ratchasima 30000

2.2.5 Position and responsibilities: Coinvestigator, protein expression, first year of project.

2.3.1 Prof. Dr. M.R. Jisnusun Svasti

2.3.2 Degree: Ph.D.

2.3.3 Position: Laboratory Director (CRI); Emeritus Professor (MU)

2.3.4 Laboratory of Biochemistry, Chulabhorn Research Institute, Viphavadee-Rangsit Highway, Bangkok

Center for Protein Structure and Function, Dept. of Biochemistry, Faculty of Science, Mahidol University, Rama 6 Road, Phayathai, Bangkok 10400;

2.3.5 Duties and responsibilities: Advisor to project

2.4.1 Name: Ms. Salila Pengthaisong

2.4.2 Degree: Ph.D

2.4.3 Position: Student/Postdoctoral Fellow

2.4.4 Work address: Institute of Science, Suranaree University of Technology,
111 University Avenue, Muang District, Nakhon Ratchasima 30000

2.4.5 Position and responsibilities: Coinvestigator, kinetic and structural investigation of rice Os3BGlu7 glycosynthase and active site cleft mutants.

2.5.1 Name: Ms. Sukanya Luang

2.5.2 Degree: Ph.D

2.5.3 Postdoctoral Fellow, 2010-2012

2.2.4 Work address: Institute of Science, Suranaree University of Technology,
111 University Avenue, Muang District, Nakhon Ratchasima 30000

2.2.5 Position and responsibilities: Coinvestigator, protein expression, first year of project.

3. Area of research: Protein Structure and Function (plant enzymes)

4. Problem and Its Importance

The depletion of world oil resources has brought about renewed interest in conversion of plant carbohydrates into sugars for fuel production. Current production schemes that largely focus on starch are in conflict with production of food, the need for which is also increasing. However, cellulose and other β -glucans represent the largest source of carbohydrate material produced by photosynthesis, and are often found as waste products from food production. Although the processing of lignocellulosic materials is primarily done with enzymes from microorganisms, which utilize these materials as a food source at present, plants have also evolved a system for recycling the cell walls between their cells, which can be quite rapid during endosperm development and growth.

Degradation of cellulose involves three enzymes in microorganisms, cellulase, cellobiosidase and β -glucosidase (cellobiase), the activity of which is limiting in many fungal enzyme preparations (Ketudat Cairns and Esen, 2010). In plants, cellulases and other endoglucanases also break down cell walls with the help of exoglucanases and exo-acting β -glucosidases. Many of these cellooligosaccharide exo- β -glucosidases have activity toward other sugars, such as β -mannosides and β -galactosides, so they may play roles in breakdown of other cell wall oligosaccharides as well. The plant β -glucosidases and exoglucanases appear to fall into related families, with most falling in glycoside hydrolase (GH) families GH1 and GH3. Since these enzymes vary in specificity, recognizing different lengths of oligosaccharides, different linkages and different types of monosaccharides differentially, it is useful to understand the structural determinants for these differences in substrate specificity, so appropriate enzymes may be used and engineered for better usage.

We have solved the structures of three rice (*Oryza sativa* L.) GH1 β -glucosidases that have different preferences for linkages and lengths of oligosaccharides and have worked on the characterization of the barley (*Hordeum vulgare* L.) GH1 oligosaccharide β -mannosidase and GH3 exoglucanase Exo I. The structures of these enzymes led to insights into how they bind oligosaccharide and glycoside substrates differently, and the hypotheses generated from these insights were assessed by mutagenesis and studies of the structural interactions of these enzymes with various substrates and inhibitors in this project. Rice is one of the most important crops in Thailand and a significant source of lignocellulosic waste. The study of these enzymes from rice and a related grass (barley) may help to develop better enzymes and strains of rice with more appropriate expression levels of these enzymes to allow efficient recycling of the rice straw. In addition, the enzymes appear to act on glycoside substrates, as well as oligosaccharides, assessing how they bind different glycoside substrates differently may lead to an improved understanding of their roles in rice development, resistance to disease, and growth that may allow these processes to be improved for higher yield. For instance, the Os4BGlu12 isoenzyme was found to hydrolyze the phytohormone β -D-glucosides tuberonic acid glucoside and sialic acid glucoside in collaboration with colleagues in Hokkaido University. Thus, this work had added to our knowledge of rice phytohormone metabolism, in addition to knowledge on binding, hydrolysis and synthesis of oligosaccharides.

5. Objectives

The purpose of this project was to understand the structural basis for substrate recognition and hydrolysis by plant enzymes that work on β -linked gluco-oligosaccharides, with the hopes this will lead to improved enzymes in the future. To meet this goal, we put forth the following objectives:

- 5.1. Structural determination of inactive mutants of the GH1 enzymes Os3BGlu6 and Os4BGlu12 in complexes with oligosaccharide and glycoside substrates and comparison to the structures of Os3BGlu7 in complex with its substrates.
- 5.2. Determination of the roles of some active site residues in determining the differences in activities between Os3BGlu6, Os3BGlu7 and Os4BGlu12 and related rice GH1 isoenzymes by mutagenesis.
- 5.3. Elucidation of catalytic and structural differences between GH1 β -glucosidases and β -mannosidases
- 5.4. Investigation of the functional differences in GH3 exoglucanases from barley and rice and analysis of the structural basis of these difference, based on the barley Exo-I structure.
- 5.5. Determination of the functional roles of amino acids in the GH3 exoglucanase and β -glucosidase active sites by mutagenesis and evaluation of functional changes.

6. Activities and Methods:

6.1. Crystal structure of Os4BGlu12 and characterization of the non-fusion-tagged Os4BGlu12 activity.

At the onset of this project, we completed the X-ray crystal structures of Os4BGlu12 in apo form and in complexes with a noncovalently bound 2,4-dinitrophenyl β -D-2-deoxy-2-fluoroglucoside and a covalently bound α -D-2-fluoroglucoside. Previously, we had produced Os4BGlu12 as an N-terminally thioredoxin and fusion-tagged protein in *Escherichia coli*, purified it by immobilized metal affinity chromatography (IMAC) followed by removal of the fusion tag by enterokinase digest and a second IMAC step to adsorb the tag, and crystallized it (Sansenya et al., 2010). Crystals of Os4BGlu12 and its complex with 2,4-dinitrophenyl β -D-2-fluoroglucoside (dNPG2F) were used to diffract X-rays, the diffraction patterns were processed to derive the structure factors, and the preliminary structures were solved by molecular replacement. The complex with dNPG2F was found when the protein was crystallized in the presence of dNPG2F and soaked in cryoprotectant with high DNPG2F concentration prior to freezing. In this project, we cocrystallized in the presence of DNPG2F, as before, but soaked the crystals in cryoprotectant solution without this slowly hydrolyzed substrate. The crystals were diffracted at the National Synchrotron Radiation Research Center, Hsinchu, Taiwan, and the structure solved as before, to find that the protein was covalently bound through its nucleophile with α -D-2-F-glucoside (G2F), as originally intended. We further refined the structures of the native enzyme, dNPG2F and G2F complexes with Refmac5 (Murshudov et al., 1999), checked their qualities with PROCHECK (Laskowsky et al., 1993) and Molprobity (Chen et al., 2010), submitted the structures and published them in Sansenya et al. (2011).

Because Os4BGlu12 had been identified as a tuberonic acid glucoside (TAG) β -glucosidase (Wakuta et al., 2011) and protein produced in *Pichia pastoris* had differences in activity compared to that reported for Os4BGlu12 fusion protein, we purified the Os4BGlu12 protein after removing the tag, as done for crystallography, and characterized the activity (Himeno et al., 2013). The protein activity was assessed by measuring the velocity for hydrolysis of following substrates (2 mM): 4-nitrophenyl (4NP) β -D-glucopyranoside

(4NPGlc), 2-nitrophenyl (2NP) β -D-glucopyranoside (2NPGlc), 4NP β -D-fucopyranoside, 4NP β -D-galactopyranoside, 4NP β -D-xylopyranoside, 4NP β -D-mannopyranoside (4NPMann), TAG, helicin, sialic acid glucoside (2-SAG), 3-hydroxysialic acid (3-SAG, also called 3-hydroxybenzoic acid), 4-hydroxy sialic acid (4-SAG or 4 hydroxybenzoic acid), cellobiose, cellotriose, cellotetraose, sophorose, laminaribiose, and gentiobiose. Reaction velocities toward 4-nitrophenyl β -glycosides and 2NPGlc were measured were determined based on A_{405} . In the reactions towards β -glucosides other than 4NPGlc and 2NPGlc, the enzyme reactions were stopped by adding 200 μ L of 2 M Tris-HCl buffer (pH 7.0), and D-glucose liberated was measured by the glucose oxidase-peroxidase method using a Sigma PGO kit (Opassiri et al., 2003). Kinetic parameters for various substrates were determined by fitting the initial velocities at various concentrations of substrates to the Michaelis-Menten equation with the Grafit version 7.0.2 computer program (Erithacus Software).

6.2. Mutagenesis of barley and rice β -glucosidases and β -glucosidase/ β -mannosidases.

6.2.1. Design of subsite 2 mutations. With the structures of the rice Os3BGlu7 (BGlu1), Os3BGlu6, and Os4BGlu12 β -glucosidases determined, along with previous analysis of their activity differences, the amino acids around the active sites of these enzymes were inspected to generate hypotheses for their roles in determining substrate specificity differences.

When the structures of the Os3BGlu7 E176Q mutant in complexes with cellotetraose and cellopentaose (Chuenchor et al., 2011) were superimposed on the Os3BGlu6 complex with octyl- β -D-thioglucoside and Os4BGlu12 in complex with 2,4-dinitrophenyl β -D-2-deoxy-2-fluoroglucoside, the Os3BGlu7 residue Asn245, which is important for binding the third glucosyl residue in cellooligosaccharides, was observed to be replaced by Met251, which extends into the active site and blocks binding of these oligosaccharides, in Os3BGlu6, while the corresponding His253 in Os4BGlu12 was positioned to hydrogen bond with the oligosaccharides. Therefore, computational docking with Autodoc was used to confirm that the oligosaccharides would form different binding complexes with these enzymes, and Os3BGlu6 Met251 was mutated to Asn (M251N), while the corresponding residues were mutated to Met in Os3BGlu7 (N245M) and Os4BGlu12 (H253M). Kinetic analysis of cellooligosaccharide hydrolysis was used to calculate the differences in glucosyl moiety binding energies caused by these mutants. Activity against *p*-nitrophenyl (*p*NP) β -D-glucoside (*p*NPGlc) was used to evaluate the overall effects of the mutations on enzyme activities.

6.2.2. Enzyme production and kinetic analysis. As previously described (Opassiri et al., 2003, 2006, Seshadri et al. 2009), the enzymes were expressed as thioredoxin fusion proteins in *Escherichia coli*, and mutations made by the QuikChange method (Stratagene). The activities toward 4NP glycosides will be determined by measuring the 405 nm absorbance of the released 4-nitrophenolate at high pH, while the glucose released from cellooligosaccharides can be measured by the glucose oxidase assay. In each case, appropriate standard curves were used to calculate the amounts of products released. Kinetic parameters were determined with triplicate activity measurements, after the initial velocity had been carefully checked. The substrate concentrations were chosen to bracket the apparent K_m at approximately 0.2 to 5 times the K_m . Inhibition constants for substrates that are not hydrolyzed by active mutants were determined by Dixon plots, after a series of Lineweaver-Burk plots at different inhibitor concentrations was inspected to ensure competitive inhibition. Subsite mapping calculations to check the effects on various glucosyl residue subsites were done as previously described (Hiromi et al., 1973; Opassiri et al., 2004), while $\Delta\Delta G$ for the effects of mutations on transition state binding were assessed as

described by Fersht et al (1987) from the equation $\Delta\Delta G_{S^*mut} = -RT[\ln(k_{cat}/K_m)_{mutant} - \ln(k_{cat}/K_m)_{wildtype}]$. To test the effects of the mutations on pK_a of important active site residues (catalytic acid/base and nucleophile), the optimum pH was determined for each mutated protein.

6.2.3. Os7BGlu26 β -mannosidase production and mutagenesis. To improve the expression of Os7BGlu26 β -D-mannosidase, the pET32a/Os7BGlu26 plasmid, which includes the *Os7BGlu26* cDNA in frame to produce an N-terminally thioredoxin and His-tagged Os7BGlu26 fusion protein (Kuntothom *et al.*, 2009), was transformed into *E. coli* strain Rosetta-gami(DE3) cells. The cells were cultured in low salt LB (Lennox) media containing 50 μ g/ml ampicillin, 15 μ g/ml kanamycin, 12.5 μ g/ml tetracyclin and 34 μ g/ml chloramphenicol. When the culture optical density at 600 nm reached 0.4-0.5, protein expression was induced with 0.3 mM IPTG for 24 h, at 20°C. Cell pellets were collected by centrifugation and suspended in the extraction buffer (50 mM Tris-HCl, pH 8.0, 150 mM sodium chloride, 200 μ g/ml lysozyme, 1% (v/v) Triton-X 100 mM, 1 mM PMSF, 4 μ g/ml DNase I) at approximately 25 °C, 30 min. After removing the insoluble debris by centrifugation, the protein was purified from the soluble extract by IMAC on cobalt-equilibrated IMAC resin (GE Healthcare). The resin was washed with the equilibration buffer (150 mM NaCl, 50 mM Tris-HCl, pH 8.0), followed by 20 mM imidazole in the equilibration buffer, and eluted with 250 mM imidazole in the equilibration buffer. The fractions of Os7BGlu26 containing β -D-glucosidase activity, as judged by 4NPGlc hydrolysis, were pooled and imidazole removed by dialysis in 150 mM NaCl, 20 mM Tris-HCl buffer, pH 8.0. The dialysed preparation was concentrated in a 30 kDa molecular mass cut off (MWCO) Centricon centrifugal filter (Millipore). The N-terminal fusion tag was removed from the Os7BGlu26 fusion protein by cleavage with 2 ng enterokinase (New England Biolabs) per 1 mg of a fusion protein at 23°C for 18 h, followed by a second round of IMAC. The flow-through fractions containing β -glucosidase activity were pooled and the protein purity was analysed by SDS-PAGE. The Os7BGlu26 was dialysed and concentrated with a 30 kDa MWCO Centricon filter to obtain approximately 6 mg/mL Os7BGlu26 protein in the 150 mM NaCl, 20 mM Tris-HCl buffer, pH 8.0.

After inspecting the Os7BGlu26 structure (see below), its superposition with the covalent 2-fluoroglucoside complex structures of Os3BGlu6, Os3BGlu7 and Os4BGlu12 and docking of substrates into the active site, mutations were designed and created by Quikchange mutagenesis, as described above. Due to the need to consider different glucosyl and mannosyl moiety structures for the docking, we collaborated with Prof. Dr. Carme Rovira to dock four different conformations (1S_3 , 1S_5 , 2S_0 , 3S_1) of glucosyl and mannosyl groups in 4NPGlc and 4NPMAN into the active site of the Os7BGlu26 structure with the *Autodock* 4.2 program (Morris *et al.*, 2009). Based on the docked structures of the glycosides with the 1S_3 and 1S_5 sugars that had lowest energies, the surrounding amino acids were evaluated. Based on these observations the following variants were generated: Os7BGlu26 E179Q and E179A (catalytic acid-base mutants), Os7BGlu26 Y134W, Os7BGlu26 Y134F, Os7BGlu26 C182A, and Os7BGlu26 C182T. These mutants were assessed for the effects of the mutations on the kinetics of 4NPGlc and 4NPMAN hydrolysis, as described above.

6.3. Crystal structures of mutants and substrate complexes.

The structures of the Os3BGlu6, Os3BGlu7, and Os4BGlu12 mutants that had the catalytic acid-base or nucleophile removed to allow substrate binding without hydrolysis or had unusual activities that suggested rearrangements of residues in the active site or different substrate binding modes were solved by X-ray crystallography. The proteins were crystallized by screening in hanging drop screens around the same conditions as previously

described for the wildtype proteins (Chuenchor et al., 2006, 2008, Seshadri et al., 2009; Sansenya et al., 2011).

6.3.1. Os3BGlu7 E386G glycosynthase and its active site cleft mutants.

In the past, we made glycosynthase mutations of Os3BGlu7 and showed that this enzyme can synthesize long oligosaccharides starting from 4-nitrophenyl-cellobioside (4NPC2) and α -D-glucosyl fluoride (α -GlcF), which appeared to be related to its long celooligosaccharide binding cleft (Hommalai et al., 2007). To analyze the relationship of the long cleft to glycosynthase products, we solved the structure of the Os3BGlu7 E386G glycosynthase with cellotetraose and cellopentaose, as well as those of two binding cleft mutants of this glycosynthase, E386G/S334A and E386G/Y341A. The crystals cocrystallized with 2 mM cellotetraose were soaked with saturated cellotetraose in cryoprotectant prior to cooling (Pengthaisong et al., 2012a). To analyze the differences in relative transglycosylation activities of different glycosynthase mutants, the E386G, E386A and E386S glycosynthase mutant proteins were also crystallized in apo forms and with 2-fluoroglucoside. Attempts were also made to cocrystallize these mutants with both 2-fluoroglucoside and 4NP-cellobioside or cellotriose or to soak these compounds into the crystals as specified for cellotetraose above. We then evaluated the differences of activity of the Os3BGlu7 E386G, E386A, and E386S glycosynthases by sending their crystal structures with α -fluoroglucoside to Prof. Liu and his student to conduct Quantum Mechanics/Molecular Mechanics (QM/MM) simulations on the glycosynthase reactions for these three enzymes to evaluate the activation energy and apparent transition state structure in each case (Wang et al., 2013).

In addition, we evaluated the contamination of wild type Os3BGlu7 in the Os3BGlu7 E386G glycosynthase by cyclophelitol inhibition. We also evaluated the activity of a less mutable glycosynthase, Os3BGlu7 E386G2, in which the codon was changed from GAA to GGC instead of the original GGA, so that it required 2 mutations or misread bases to revert back to encode the Glu catalytic nucleophile. Evaluation of the products formed with the new mutant and with inclusion of cyclophelitol allowed identification of the true glycosynthase products (Pengthaisong et al., 2012b).

In order to assess the functional effects of the mutations, the same mutations were made in the Os3BGlu7 wildtype enzyme and an acid-base mutant that has transglycosylation activity with 4NPGlc donor and acetate or azide acceptor, Os3BGlu7 E176Q. The kinetics of the hydrolysis of Os3BGlu7 mutants hydrolysis of cellotriose, cellotetraose and cellopentaose and the inhibition of Os3BGlu E176Q use of 4NPGlc by these oligosaccharides was evaluated to determine the effects on that active site cleft binding of oligosaccharides. Further mutations of the active site cleft and secondary oligosaccharide binding residues Gln187, Arg178, and Trp337 were also made alone and in combination with the other mutations and evaluated for effects on oligosaccharide binding and hydrolysis and synthesis of 4NP-oligosaccharides by the glycosynthase combination mutants.

6.3.2. Binding of substrates by Os3BGlu6 and Os4BGlu12. In order to determine the interactions between the proteins and their substrates, the catalytic acid-base of Os3BGlu6 was mutated from glutamate to glutamine (E178Q) and aspartate (E178D), and its catalytic nucleophile similarly mutated to glutamine (E389Q) and aspartate (E389D), as previously done for Os3BGlu7 (Chuenchor et al., 2011). The Os4BGlu12 acid-base was similarly mutated to glutamine (E179Q). The crystals were grown as described for the wildtype enzymes (Seshadri et al., 2009; Sansenya et al., 2011) in the presence and absence of natural substrates. The Os3BGlu6 E178Q mutant was complexed with gibberellin GA₄ glucose ester (GA₄GE) and 4NPGlc, while the Os3BGlu6 E389D mutant was also produced in the presence of GA₄GE. The Os4BGlu12 E179Q mutant was crystallized in the presence of and the crystals soaked in trycin 5-O-glucoside and 7-O-glucoside, which we and Dr. Prasat Kittakoop had purified from rice extracts, as well as tuberillic acid glucoside (TAG) provided

by Prof. H. Matsui of Hokkaido University. Before rapid cooling of the crystals in liquid nitrogen, the crystals were soaked in mother liquor and/or cryoprotectant solutions that contain the substrates at high concentrations.

6.3.3. Crystal structure solution. The crystals were soaked in cryoprotectant that is optimized for the crystals, starting from around 18% glycerol in 110% precipitant. They were then flash cooled in liquid nitrogen and stored in liquid nitrogen or a dry shipper until data collection. Preliminary data was collected on the rotating anode source in-house system at the Synchrotron Light Research Institute. Higher resolution data collection was collected at the National Synchrotron Radiation Research Center (NSRRC) in Hsinchu, Taiwan, on the PX 13B beamline. The data was processed with HKL2000 (Otwinowski and Minor, 1997), then the scaled data was converted to .mtz files for phase solution in the CCP4 computer suite (CCP4, 1994). The structures from crystals that are isomorphous with wildtype protein crystals were solved by rigid body refinement in Refmac5 (Murdushov et al., 1997), while those from nonisomorphous crystals were solved by molecular replacement with MolRep (Vagin and Teplyakov, 1997) in the CCP4 suite. The structures were refined by iterative model building in coot (Emsley and Cowtan, 2004) and refinement with Refmac5. Final structures were checked for reasonable structures with PROCHECK (Laskowski et al., 1993), Molprobability (Chen et al., 2010) and other available quality assessment programs and submitted to the Protein Data Bank at the time of paper submission.

6.4. Attempts to determine the structure of related GH1 enzymes and Os7BGlu26 structure.

Since our laboratory is trying to characterize members of other GH1 clusters and GH1 β -mannosidases, several proteins were screened for crystallization by the microbatch method, including Os1BGlu4 cytosolic β -D-glucosidase, Os7BGlu26 β -D-mannosidase, HvBII β -D-mannosidase, and Os9BGlu31 transglucosidase, although only Os7BGlu26 β -D-mannosidase was successfully and reproducibly crystallized. Crystallization was screened by microbatch under oil with commercial screening kits from Hampton Research, Emerald Biosciences, and others and the PEG/salt screen of Dr. Jirandorn Yuvaniyama (Mahidol University). Conditions that yield crystals were optimized in hanging drop vapor diffusion experiments by varying precipitant concentrations, protein concentration, pH, and other parameters, and microseeding were used to improve crystal quality for Os7BGlu26.

After optimization of the crystallization conditions, the crystals were grown in 0.8 M K₂Na tartrate, 0.1 M Na HEPES, pH 7.5. The Os7BGlu26 β -mannosidase, the proteins may also be cocrystallized with 2-deoxy-2-fluoromannoside (from Stephen G. Withers, University of British Columbia, Canada), but it was rapidly hydrolyzed. The enzyme was also crystallized and soaked with this inhibitor or 4NPGlc, 4NPMann, D-mannose or D-glucose for crystallization trials and/or optimization. For the complex with ligand, the crystals were soaked in 100 or 400 mM D-mannose in the precipitant solution. Prior to data collection, the crystals were soaked in cryoprotectant containing the precipitant, 100 or 400 mM D-mannose and 20% (v/v) glycerol. The crystals were flash vitrified and stored in liquid nitrogen. In addition, the Os7BGlu26 β -D-mannosidase was mutated at the catalytic acid-base to change glutamate to glutamine (Os7BGlu26 E179Q), and attempts were made to crystallize this protein with various mannoside and glucoside substrates, however, attempts to determine these structures are still underway. The structures of Os7BGlu26 alone and with mannose and various inhibitors were solved, refined and verified as described in Section 6.3.

6.5. Analysis of putative transition state analog inhibitors binding and structures.

Because the structure of the Os7BGlu26 active site was very similar to that of the Os3BGlu7 β -D-glucosidase and analysis of substrate and product binding did not generate a complete rationale for their differences, we needed to consider the structures bound to the transition state. In other glycoside hydrolase families, β -mannosidases have been shown to hydrolyze β -D-mannoside substrates via a conformational trajectory from a 1S_5 skew boat, in the Michaelis complex, to a $B_{2,5}$ boat, near the transition state, to a 0S_2 skew boat, at the covalent intermediate, while β -D-glucosides follow a 1S_3 skew boat to 4H_3 half chair or 4E envelope transition state to a 4C_1 chair covalent intermediate trajectory (Davies et al., 2003). Putative transition state analog inhibitors were obtained from Prof. Spencer Williams, including glucoimidazole, mannosimidazole, glucotetrazole, and isofagamine, Prof. Yves Blerót, who provided the locked boat inhibitors shown in Figure 1, and Prof. Maria Hrmova, who provided phenethylglucoimidazole, which she had from a previous study. In addition, we acquired the commercially available putative inhibitors nojirimycin sulphate, 1-deoxy-nojirimycin and 1-deoxy-mannojojirimycin, shown in Figure 1. After an initial test of inhibition by these compounds at 0.1 mM concentrations, competitive inhibition of Os7BGlu26, Os3BGlu7 and HvBII was verified by inspection of Lineweaver-Burk plots for those compounds that inhibited, and the competitive K_i values were determined with Dixon plots. Furthermore, binding to Os3BGlu7 was assessed for glucoimidazole, mannoimidazole and isofagamine by microcalorimetry. Attempts were made to do a similar microcalorimetric analysis for Os7BGlu26, but the protein precipitated in the assay buffer at the concentrations required for the analysis. Furthermore, the mannoimidazole, glucoimidazole and isofagamine were soaked into Os7BGlu26 and Os3BGlu7 crystals, and X-ray diffraction data collection. Preliminary structures were calculated, although structure refinement and finalization is still underway.

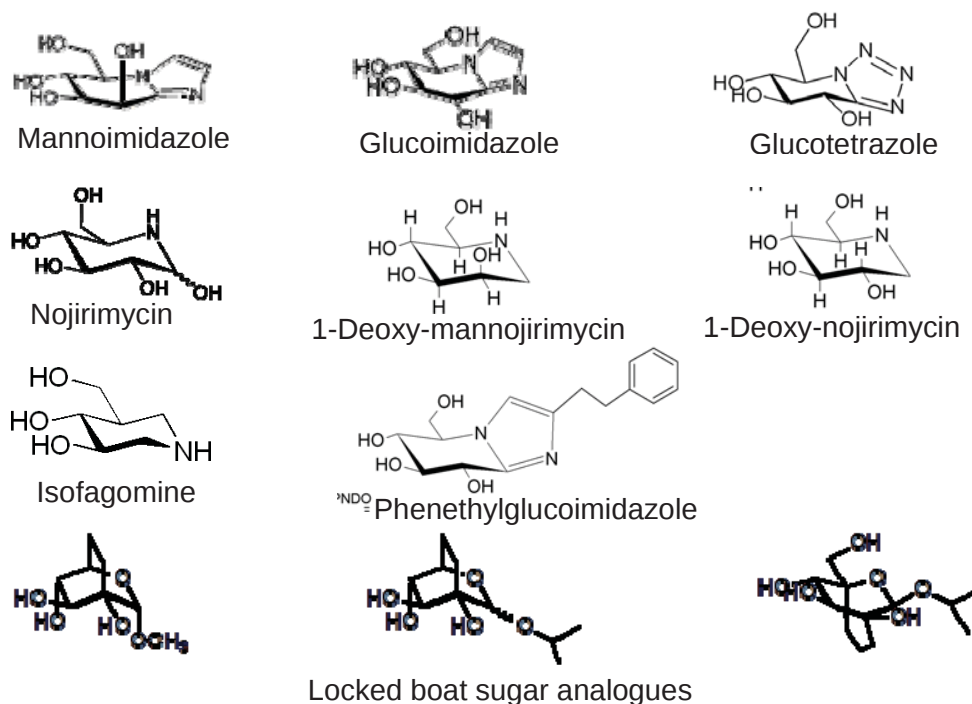


Figure 1 Structures of transition-state analogue candidates studied.

6.6. Expression, functional and structural studies of GH3 exoglucanases.

In previous work, we had made several mutations of HvExoI, which we previously expressed from an optimized cDNA as a secreted protein in *Pichia pastoris* (Luang et al., 2010a). It can be purified by an initial precipitation or chromatography step (ie hydrophobic interaction), followed by immobilized metal affinity chromatography (IMAC) and desalting in a centrifugal concentrator. The protein was deglycosylated with Endoglycosidase H for crystallization or left in glycosylated form for most enzymatic work. In the previous project, we made the mutations: D95N, D95A, E220Q, E220A, E161A, E161Q, R158A, K206V, and R158A/E161A to test residues in the -1 subsite for substrate binding and W286A and W434A in the +1 subsite. In this project, we made additional mutants by site-directed mutagenesis, including D285G, D285S, K206A, E220D, W286F, W434F, W286F/W434F, W286F/W434A and W286A/W434A, studied their activities and crystallized some of these alone or in complex with thio-linked oligosaccharide inhibitors.

Previously, we crystallized the recombinant ExoI mutant and those of a few mutants and collected preliminary X-ray diffraction data (Luang et al., 2010b). In this project, we have collected more data sets and refined the structures to near completion, including for structures for the HvExoI mutants R158A, E161A, E161Q, R158A/E161A, W286A, W434A and W286A/W434A alone and with thio-linked glucooligosaccharides with β -1,3, 1,4- and 1,6- linkages. In most cases, crystal seeds generated serially from the original native crystal were used, and all crystals were isomorphous with these crystals (Hrmova et al., 1998; Varghese et al., 1999).

In the previous project, we had cloned a cDNA for the rice GH3 gene *OsExoII* (gene locus Os03g0749300) and made a synthetic cDNA for *OsExoI* (Os03g0749500), so we continued to express the proteins from these genes. The *OsExoII* protein was expressed as a thioredoxin fusion protein with a six histidine tag from a pET32a-derived vector in *E. coli* strain Origami(DE3), while the optimized *OsExoI* was also expressed as a secreted protein in *Pichia pastoris* from pPICZ α BNH8 (Toonkool et al., 2006). The activity of the purified proteins were tested against 4NP-glycosides by release of 4NP and against the polysaccharides laminarin, barley (1,3)(1,4)- β -glucan and lichenan and gluco-oligosaccharides by quantifying released glucose with the glucose oxidase assay.

An additional cDNA for the third GH3 exoglucanase, designated *OsExoIII* (gene locus Os01g0771900, cDNA Genbank accession AK066300), was cloned into a pET32a-derived *E. coli* expression vector. The protein was expressed and a preliminary purification conducted with tests for activity. Attempts were also made to clone the Os03g0749100 (putative exoglucanase) and Os02g0752200 (putative β -D-xylosidase) locus cDNA, but without success.

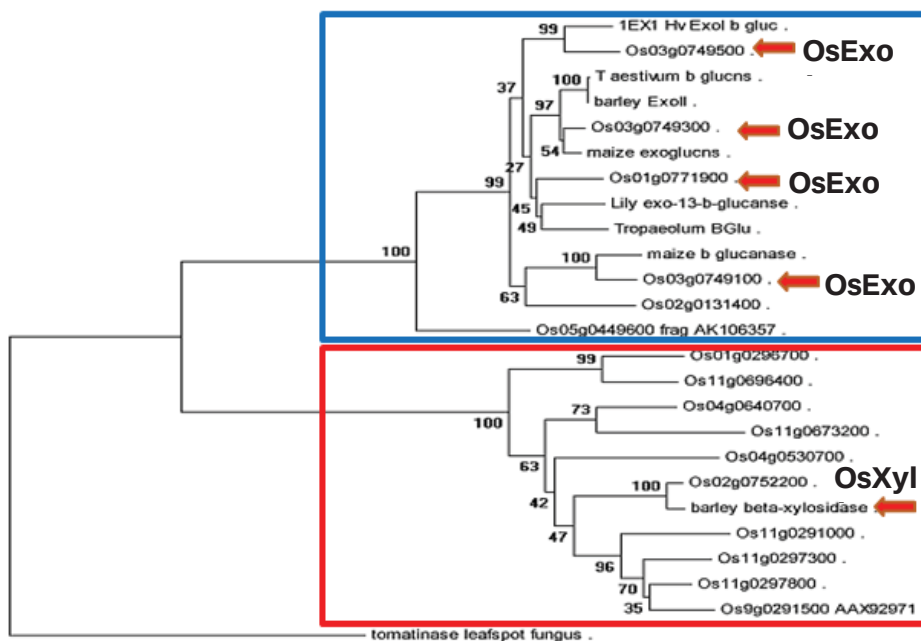


Figure 2. Phylogenetic tree of GH3 enzymes in rice and related enzymes. The tree is based on a multiple protein sequence alignment generated with the MUSCLE algorithm implemented in the MEGA5 program suite, and drawn by the distance-based neighbor joining method. The rice enzymes are denoted by their gene's locus number. The proteins/genes targeted in this research are labeled and they or the corresponding barley enzyme from which the putative activity was derived are highlighted with arrows. The upper box represents putative exoglucanases, while the lower box indicates β-D-xylosidase/α-L-arabinosidase-like enzymes. The fungal tomatinase enzyme was used as an outgroup to root the tree.

7. Results:

7.1. Crystal structure of Os4BGlu12 and its complexes with 2,4DNPG and G2F

During this project, we completed the structures of Os4BGlu12 β -glucosidase and its Michaelis complex with 2,4-dinitrophenyl 2-deoxy-2-fluoro- β -D-glucoside, and generated a structure with a covalent intermediate with α -linked 2-deoxy-2-fluoroglucoside (G2F), as reported in Sansenya et al. (2011). The final data and structure statistics for these three structures are given in Table 1. The structure was found to have two protein molecules in the asymmetric unit, and these two molecules bound a Zn molecule asymmetrically between them, as shown in Figure 3. As can be seen in the figure, the overall structure of each protein is a typical $(\beta/\alpha)_8$ -barrel (or triose phosphate isomerase, TIM, barrel), as seen in other GH1 enzyme structures. The four variable loops that determine the different shapes of GH1 active sites are highlighted. One difference between Os4BGlu12 and other plant GH1 enzymes is that it has two disulfide bonds stabilizing the structure of the variable loop C, while most GH1 enzymes that have been described have a single disulfide bond in this loop (see the two noted bonds in Figure 3). At first, this was considered to correlate to the higher stability of Os4BGlu12 compared to Os3BGlu7, but heating the two enzymes in the absence and presence of reducing agents showed that the Os3BGlu7 stability was more dependent on the presence of the disulfide bond than Os4BGlu12 (Sansenya et al., 2011).

The noncovalent complex was found indicated the residues surrounding the sugars in the active site (Figure 4). All of the residues surrounding the glucosyl residue were conserved with other GH1 enzymes with known structures and made essentially the same hydrogen bonds. In the aglycone-binding site, the residues were less conserved, although the conserved tryptophan (Trp365), which serves as a platform to bind sugars and nonpolar aglycons is conserved. His252 made a polar contact through a water to the nitrosyl group of 2,4-dinitrophenyl, suggesting it may mediate polar contacts to substrate aglycones, similar to the corresponding Asn245 of Os3BGlu7 (Figure 5 C), which hydrogen bonds to oligosaccharide substrates. The covalent complex of the Os4BGlu12 enzyme with α -D-2-fluoroglucoside, the nucleophile was displaced less than the nucleophile in similar complexes of Os3BGlu6 and Os3BGlu7. Since this small displacement was similar to that seen in the covalent complex with myrosinase, a thioglucosidase, the hydrolysis of the thio-linked substrates 4NP- β -D-thioglucoside and octyl- β -D-thioglucoside was tested and it was found that Os4BGlu12 had significant hydrolysis activity for these substrates, while Os3BGlu6 and Os3BGlu7 did not (Sansenya et al., 2011). Thus, the intermediate structure seemed to correlate with the catalytic ability of the enzyme.

Table 1 Refinement statistics for the structures of Os4BGlu12 apoenzyme and its complexes with G2F and DNP2FG.

Dataset	Os4BGlu12	Os4BGlu12_G2F	Os4BGlu12_DNP2FG
PDB code	3PTK	3PTM	3PTQ
Beamline	BL13B1	BL13B1	BL13B1
Wavelength (Å)	1.00	1.00	1.00
Space group	P4 ₃ 2 ₁ 2	P4 ₃ 2 ₁ 2	P4 ₃ 2 ₁ 2
Unit-cell parameters (Å)	a = 112.7 b = 112.7 c = 182.8	a = 114.0 b = 114.0 c = 184.1	a = 114.1 b = 114.1 c = 184.5
Resolution range (Å)	30-2.50	30-2.40	30-2.45
Resolution outer shell(Å)	2.50-2.49	2.49-2.40	2.51-2.45
No. Unique reflections	39533	48334	43131
No. Observed reflections	408534	329557	381316
Completeness (%)	100.0 (99.9)	99.6 (99.9)	100.0 (99.9)
Average redundancy per shell	9.7 (10.0)	6.8 (6.2)	8.4 (8.6)
<i>I</i> / σ (<i>I</i>)	19.6 (6.4)	17.5 (3.4)	21.7 (4.4)
R _(merge) (%)	10.6 (41.7)	10.4 (53.7)	9.4 (49.9)
R _{factor} (%)	20.25	19.58	20.90
R _{free} (%)	24.76	23.43	25.07
No. of residues in protein	486 (2 molecules)	486 (2 molecules)	486 (2 molecules)
No. protein atoms	7720 (2 molecules)	7720 (2 molecules)	7720 (2 molecules)
No. Ligand atoms	16 (Tris)	22 (G2F)	48 (DNP2FG)
No. Other hetero atoms	1 (Zn)	37 (GOL, Zn)	31 (GOL, Zn)
No. waters	228	511	280
<i>Mean B-factor</i>			
Protein	32.47	28.31	33.18
Ligand	52.70 (Tris)	21.47 (G2F)	62.69 (DNP2FG)
Other hetero atoms	24.23(Zn)	51.20 (GOL), 22.40 (Zn)	45.04 (GOL), 22.75 (Zn)
Waters	32.18	30.40	31.60
r.m.s. bond deviations (length)	0.009	0.006	0.010
r.m.s angle deviations (degrees)	1.192	1.015	1.264
Ramachandran plot			
Residues in most favorable regions (%)	88.4	89.5	89.4
Residues in additional allowed regions (%)	10.5	9.6	9.4
Residues in generously allowed regions (%)	1.1	0.9	1.2
Residues in disallowed regions (%)	0	0	0

Numbers in parentheses are statistics for the outer resolution shell. Ramachandran values were determined from PROCHECK (Laskowski et al., 1993). Table is from Sansenya et al. (2011).

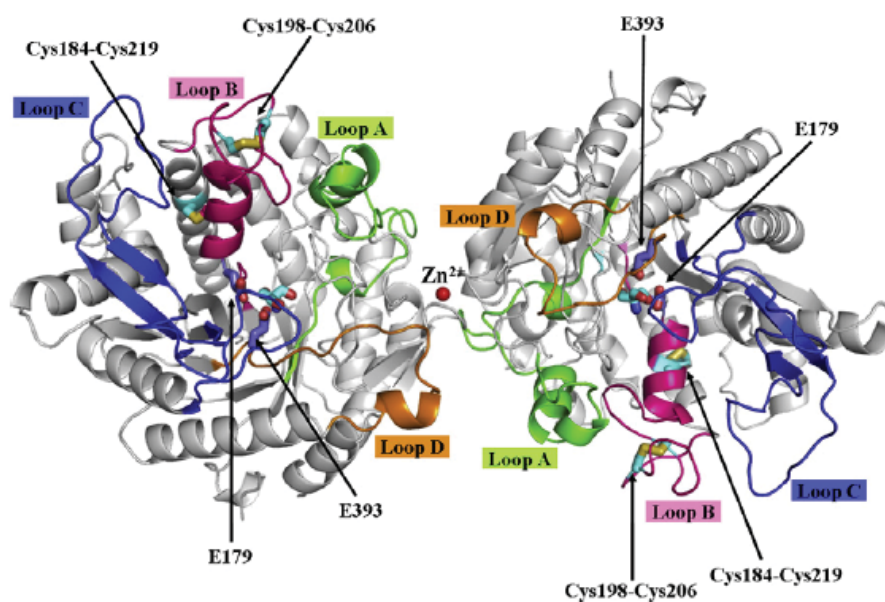


Figure 3 Structure of the asymmetric unit of Os4BGlu12. Two molecules of Os4BGlu12 β -glucosidase were seen in the asymmetric unit, asymmetrically binding a Zn ion between them. The four variable loops are labeled, along with the two disulfide bonds and catalytic acid/base (E179) and nucleophile (E393). The Zn ion binding is not physiological, since the protein is monomeric in solution. (Sansenya et al., 2011).

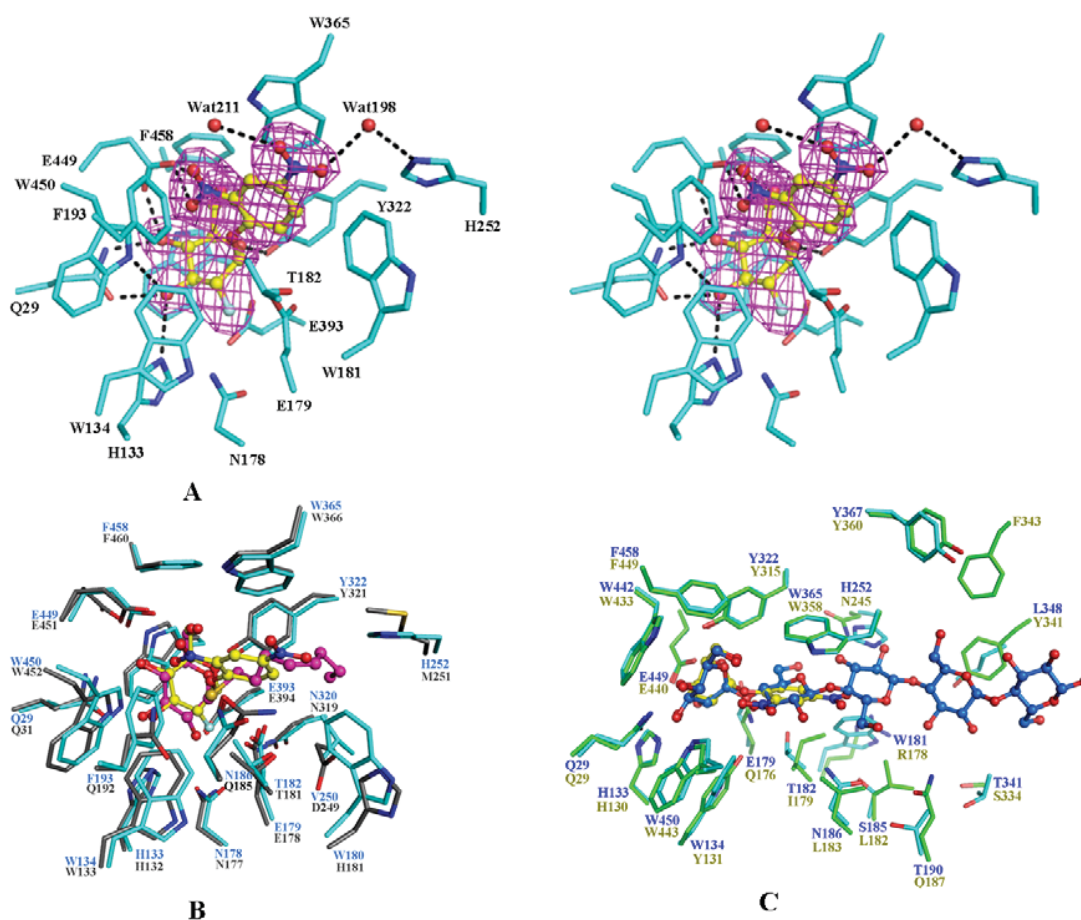


Figure 4. Protein-ligand interactions in the active site of the Os4BGlu12 DNP2FG complex. (A) Electron density (2F_{obs}-F_{calc} map contoured at the 1 σ level) of DNP2FG, calculated with a model from which DNP2FG was omitted. (B) Superimposition of Os4BGlu12_DNP2FG (cyan carbons in the electronic version) with Os3BGlu6 with *n*-octyl-β-D-thioglucopyranoside (gray carbons). The ligands are shown in ball and stick representations, with pink carbons for DNP2FG and yellow carbons for *n*-octyl-β-D-thioglucopyranoside in the electronic version. (D) Superimposition of the protein ligand binding of Os4BGlu12 with DNP2FG (yellow carbons) and Os3BGlu7 (3F5K) with cellopentaose (purple carbons) and the ligand carbons drawn in green and cyan for DNP2FG and cellopentaose, respectively (in the electronic version). (Sansenya et al., 2011).

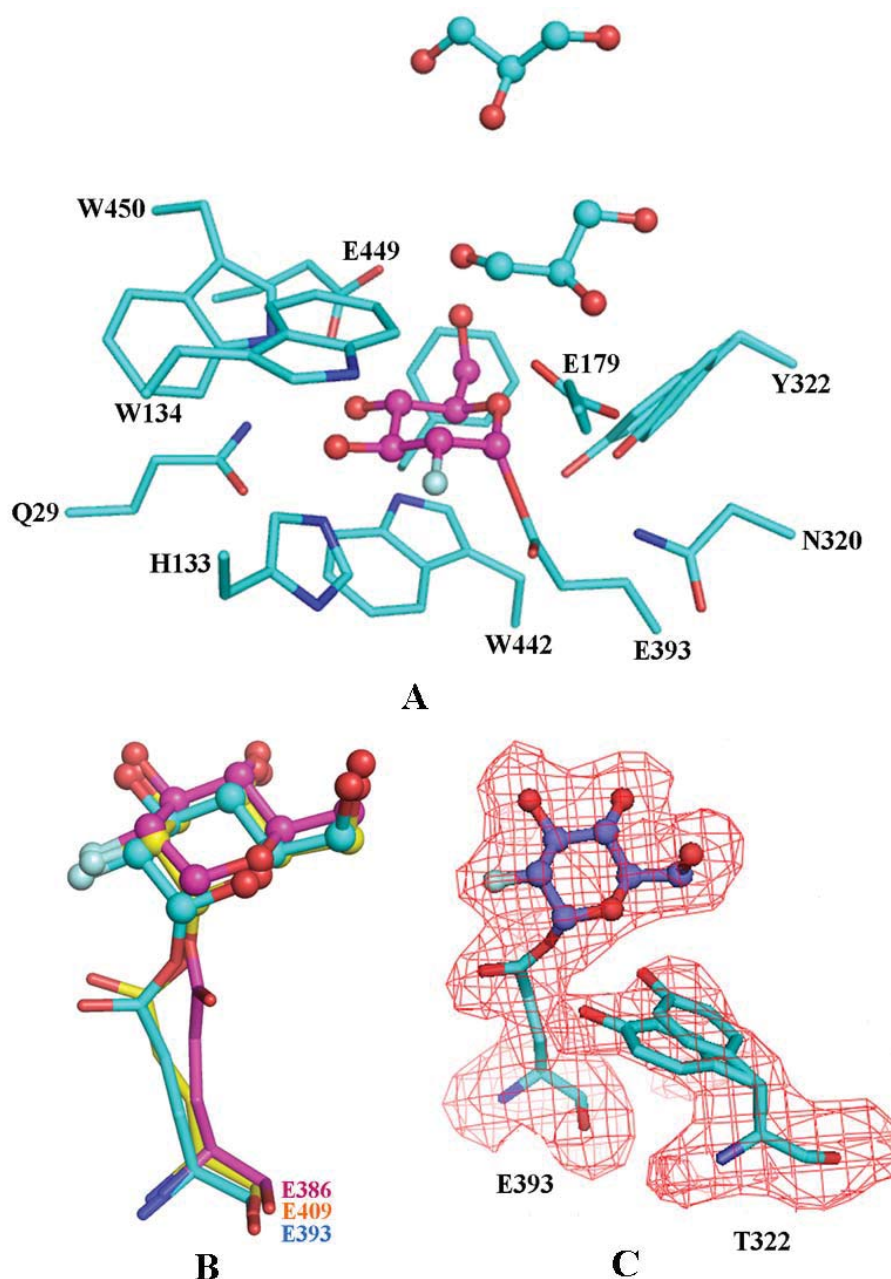


Figure 5. The active site region of the covalent complex of Os4BGlu12 with G2F. (A) Protein ligand binding of the monomer A (Os4BGlu12/G2F complex) amino acids surrounding the ligand are represented as sticks with carbons colored cyan. The ligands are represented in ball and stick and include G2F and glycerol of the complex. (B) Superimposition of nucleophile and sugar residues of glycosyl-enzyme intermediates of the Os4BGlu12/G2F complex with those of *S. alba* myrosinase and rice Os3BGlu7. In the electronic version, carbons are shown in cyan for Os4BGlu12/G2F, in pink for rice Os3BGlu7 (2RGM) and in yellow for *S. alba* myrosinase (1E73). (C) $F_{\text{obs}} - F_{\text{calc}}$ electron density map of glycosyl-enzyme intermediate showing the two conformations of Tyr322 in the structure of Os4BGlu12 with G2F. The map is contoured at the 3 σ level. (Sansenya et al., 2011)

7.2. Mutational analysis of subsite +2 binding residue effect on oligosaccharide binding.

When the structures of Os3BGlu6 and Os4BGlu12 were compared to the structure of Os3BGlu7 E176Q bound to oligosaccharides, it was clear that Os3BGlu7 Asn245 made direct hydrogen bonds to β -1,3- and β -1,4-linked oligosaccharides in Os3BGlu7, and His253 could play a similar role in Os4BGlu13, but Met251 in Os3BGlu6 could not and might block cellotriose binding (Figure 6 A). We conducted computational docking to identify the probable cellotriose binding position in Os3BGlu6 and Os4BGlu12. As seen in Figure 6 B-D, the position of the cellotriose docked into Os4BGlu12 was similar to that of cellotetraose and cellopentaose in their complexes with Os3BGlu7, while the nonreducing end of the cellotriose docked into Os3BGlu6 was displaced and flipped over compared to the oligosaccharide in Os3BGlu7, with poorer binding energy (Sansenya et al., 2012). We investigated the kinetic effect of mutation of Os3BGlu6 Met251 to Asn (Figure 2A), which showed improved hydrolysis of laminaribiose and cellobiose, cellotriose, cellotetraose and cellopentaose (Table 2). We furthermore showed that mutation of the corresponding residues in Os3BGlu7 (N245M) and Os4BGlu12 (H252M) reduced oligosaccharide hydrolysis by these enzymes (Table 2). In the case of Os3BGlu7 N245M, the effect was larger than that previously seen in Os3BGlu7 N245V (Figure 7), corroborating some role for the steric hindrance of the Met suggested by the Os3BGlu6 structure (Ketudat Cairns et al., 2012).

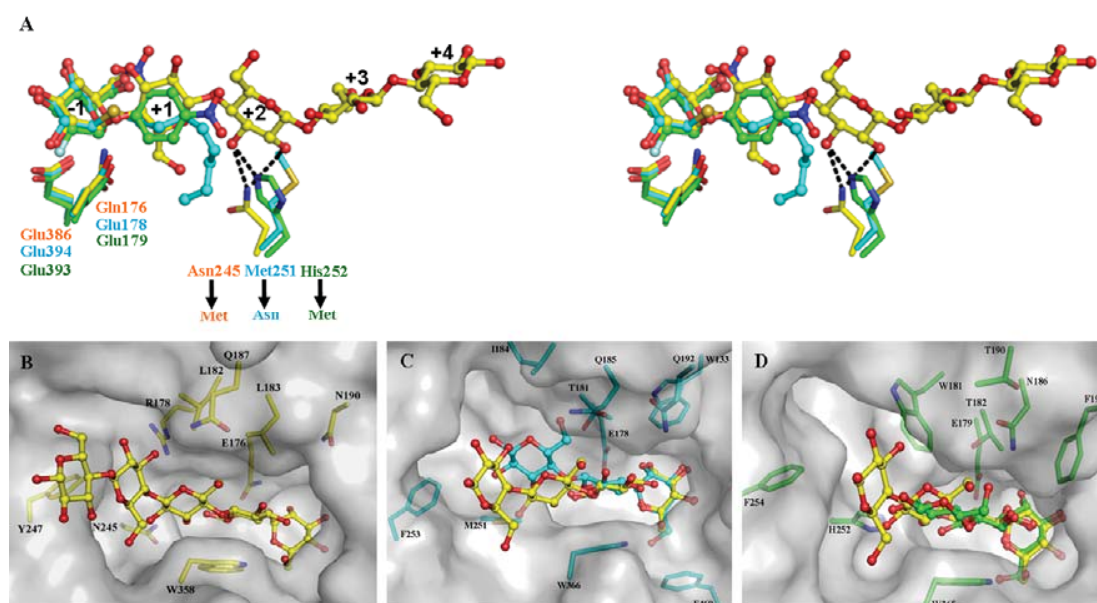


Figure 6. Comparison of celloligosaccharide binding in the active sites of Os3BGlu7, Os3BGlu6 and Os4BGlu12. A. shows a superposition of the active sites of the crystal structures of Os3BGlu7 complexed with cellopentaose, Os3BGlu6 complexed with octyl thioglucoside, and Os4BGlu12 complexed with 2,4-dinitrophenyl β -D-2-deoxy-2-fluoroglucoside in stereo view. The mutations made in the respective enzymes are written under the residues. B. The active site cleft of the structure of Os3BGlu7 in complex with cellopentaose. C. The active site cleft of Os3BGlu6 with cellotriose docked into the active site with Autodoc superimposed with the cellotetraose ligand in the X-ray crystal structure of its complex with Os3BGlu7. D. The active site cleft of Os4BGlu12 with cellotriose docked into the active site superimposed with the cellotetraose from its structure with Os3BGlu7, as in C. The figure was made with Pymol and is from Sansenya et al. (2012).

Table 2 Kinetic constants of Os3BGlu6, Os3BGlu6M251N, Os3BGlu7 Os3BGlu7N245M and Os4BGlu12, Os4BGlu12H252M for hydrolysis of oligosaccharides.

Substrate	Kinetic parameters ^a	Os3BGlu6	Os3BGlu6 M251N	Os3BGlu7 ^b	Os3BGlu7 N245M	Os4BGlu12	Os4BGlu12 H252M
Laminaribiose (β -1,3-linked Glc ₂)	K_m (mM)	8.7 $\pm$ 0.07	3.50 $\pm$ 0.20	2.05 $\pm$ 0.10	4.47 $\pm$ 0.28	5.1 $\pm$ 0.2	7.24 $\pm$ 0.44
	k_{cat} (s ⁻¹)	1.64 $\pm$ 0.06	9.90 $\pm$ 0.50	31.9 $\pm$ 3.1	9.58 $\pm$ 0.34	23 $\pm$ 1	10.58 $\pm$ 0.33
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	0.18	2.83	15.7	2.41	4.5	1.46
Cellobiose	$\Delta\Delta G_{S^*mut}$ (kJ mol ⁻¹) ^c		-6.9		+4.7		+1.3
	K_m (mM)	17.52 $\pm$ 0.27	14.3 $\pm$ 1.2	31.5 $\pm$ 1.6	46.17 $\pm$ 2.5	27.02 $\pm$ 2.13	22.68 $\pm$ 1.87
	k_{cat} (s ⁻¹)	0.27 $\pm$ 0.004	1.97 $\pm$ 0.09	1.52 $\pm$ 0.13	1.18 $\pm$ 0.09	5.92 $\pm$ 0.37	2.42 $\pm$ 0.11
Celotriose	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	0.015	0.14	0.05	0.03	0.22	0.11
	$\Delta\Delta G_{S^*mut}$ (kJ mol ⁻¹)		-5.6		+1.3		+1.7
	K_m (mM)	25.14 $\pm$ 0.77	3.20 $\pm$ 0.20	0.72 $\pm$ 0.002	1.51 $\pm$ 0.10	5.01 $\pm$ 0.49	11.17 $\pm$ 0.40
Cellobiotetraose	k_{cat} (s ⁻¹)	0.41 $\pm$ 0.014	1.32 $\pm$ 0.12	18.1 $\pm$ 0.4	2.16 $\pm$ 0.13	44.28 $\pm$ 0.29	41.92 $\pm$ 0.13
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	0.017	0.41	25.4	1.43	8.82	3.58
	$\Delta\Delta G_{S^*mut}$ (kJ mol ⁻¹)		-8.0		+7.2		+2.3
Cellobiotetraose	K_m (mM)	20.25 $\pm$ 0.28	3.86 $\pm$ 0.23	0.28 $\pm$ 0.01	1.78 $\pm$ 0.16	5.94 $\pm$ 0.53	23.35 $\pm$ 0.16
	k_{cat} (s ⁻¹)	0.54 $\pm$ 0.004	1.62 $\pm$ 0.10	17.3 $\pm$ 0.6	5.48 $\pm$ 0.20	89.0 $\pm$ 2.5	87.82 $\pm$ 0.20
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	0.027	0.42	61.1	3.08	14.98	5.76
Cellobiotetraose	$\Delta\Delta G_{S^*mut}$ (kJ mol ⁻¹)		-6.9		+7.5		+2.4
	K_m (mM)	27.03 $\pm$ 0.49	4.57 $\pm$ 0.40	0.24 $\pm$ 0.01	1.32 $\pm$ 0.11	5.53 $\pm$ 0.35	20.25 $\pm$ 0.11
	k_{cat} (s ⁻¹)	0.25 $\pm$ 0.003	0.62 $\pm$ 0.04	16.9 $\pm$ 0.1	6.66 $\pm$ 0.50	105 $\pm$ 9	62.13 $\pm$ 0.50
Cellobiotetraose	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	0.009	0.14	72	5.05	19.0	3.07
	$\Delta\Delta G_{S^*mut}$ (kJ mol ⁻¹)		-6.9		+6.7		+4.6
	K_m (mM)	ND ^a	ND	0.11 $\pm$ 0.01	1.52 $\pm$ 0.08	ND	ND
Cellobiotetraose	k_{cat} (s ⁻¹)	ND	ND	16.9 $\pm$ 0.3	7.73 $\pm$ 0.40	ND	ND
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	ND	ND	153	5.10	ND	ND
	$\Delta\Delta G_{S^*mut}$ (kJ mol ⁻¹)				+8.6		

^a ND means not determine

^b Kinetics constants were taken from Opassiri *et al.*, 2004.

^c The change in Gibbs free energy for transition state binding by the mutants compared to the wildtype enzyme was calculated as $\Delta\Delta G_{S^*mut} = -RT[\ln(k_{cat}/K_m)_{mutant} - \ln(k_{cat}/K_m)_{wildtype}]$ (Fersht *et al.*, 1987), so a negative value indicates stronger binding by the mutant, while a positive value indicates poorer transition state binding.

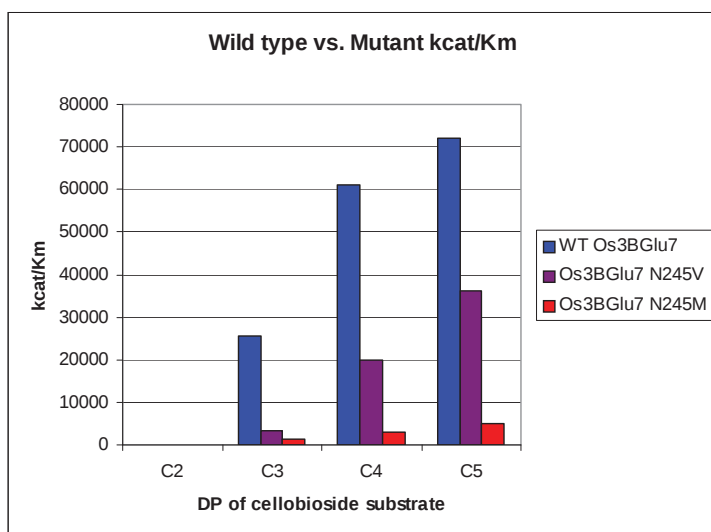


Figure 7. Comparison of the k_{cat}/K_m for cellooligosaccharides of different degrees of polymerization by Os3BGlu7 and its N245V and N245M mutants. C2 represents cellobiose, for which the activities were too low to see on this scale, while C3 is cellotriose, C4 is cellotetraose and C5 is cellopentaose. (Ketudat-Cairns et al., 2012).

7.3. Characterization of Os4BGlu12 activity after removal of the N-terminal fusion tag.

During characterization of the Os4BGlu12 β -glucosidase by our colleagues in Prof. H. Matsui's group in Hokkaido University, it was noted that the reported activity from the protein expressed in *E. coli* was much lower and had some differences. The Matsui group had previously identified Os4BGlu12 and its close homologue Os4BGlu13 as tuberonic acid glucoside (TAG) β -glucosidases (TAGG) (Wakuta et al., 2010, 2011). Since all the previous work was using the protein that still had the N-terminal thioredoxin, His₆ and S-tags on it, we prepared protein without the tag, as is used for crystallography, and tested the activity toward various substrates for a more appropriate comparison. As can be seen, the Os4BGlu12 enzyme without the N-terminal tag had 4.9-fold higher activity than that with the N-terminal tag removed for 4NPGlc and similarly higher activity toward almost all substrates. However, the activity of the tag-free *E. coli* protein was still lower than the protein produced in *P. pastoris* and purified from rice for most substrates.

The overall glycon (nonreducing end sugar) specificity seemed to be similar for the enzyme produced in all three systems, but there were some striking differences in aglycon specificities in the glucosides, such as 2NPGlc being a better substrate for the enzymes from rice and *P. pastoris* and some natural substrates also showing differences between the different enzymes. In this respect, there was variation in terms of which recombinant enzyme was more similar to native. For instance, the native and the *E. coli* derived enzymes had relatively more activity toward cellooligosaccharides than the enzyme from *P. pastoris*. In contrast, the native and *P. pastoris* enzymes had high activity toward salicin compared to what could be detected with the *E. coli* derived enzyme. It is not clear whether, for some substrates like salicin, the difference is actually derived from differences in the glucose oxidase assays used to measure the glucose released in the two labs involved in the study, since compounds with antioxidant activities may have different effects in different assays. Because of this, salicin was not included in the final paper (Himeno et al., 2013). Overall, the recombinant enzymes give a reasonable overall view of the native enzyme activity with a few significant differences.

Table 3. Hydrolytic Velocities of Os4BGlu12 toward Various Substrates

	rBGlu12P		rBGlu12E ^a		rBGlu12E new		Native BGlu12	
Substrate	Relative v	(%)	Relative v	(%)	Relative v	(%)	Relative v	(%)
4NP β-D-Glucoside	182 ± 6	100	12.4	100	60.7 ± 1.1	100	171 ± 7	100
2NP β-D-Glucoside	182 ± 4	100	N. D.	N. D.	23.0 ± 1.6	34.1	300 ± 14	175
4NP β-D-Fucoside	300 ± 6	165	13.1	106	96.7 ± 9.9	159	138 ± 6	80.7
4NP β-D-Galactoside	45.9 ± 0.8	25.2	6.14	49.5	43.0 ± 4.5	70.8	184 ± 1	108
4NP β-D-Xyloside	6.77 ± 0.38	3.72	4.66	37.6	13.4 ± 0.32	22.1	15.3 ± 0.2	8.95
4NP β-D-Mannoside	1.55 ± 0.03	0.852	N. H.	N. H.	0.37 ± 0.03	1.2	1.18 ± 0.18	0.69
TAG	44.5 ± 1.5	24.5	N. D.	N. D.	50.1 ± 1.0	82.5	25.3 ± 0.3	14.8
Arbutin	2.08 ± 0.42	1.14	N. D.	N. D.	N.H.	-	7.00 ± 1.53	4.09
Helicin	179 ± 4	98.4	N. D.	N. D.	186.1 ± 3.0	306	274 ± 5	160
Salicin	88.9 ± 3.0	48.8	N. H.	N. H.	0.16 ± 0.009	0.26	194 ± 3	113
SAG	201 ± 8	110	N. D.	N. D.	60.3 ± 1.4	99	118 ± 4	69.0
3-SAG	180 ± 10	98.9	N. D.	N. D.	61.0 ± 3.3	100	130 ± 10	76.0
4 -SAG	164 ± 8	90.1	N. D.	N. D.	104.7 ± 4.9	173	157 ± 4	91.8
Cellobiose	0.0925 ± 0.138	0.0508	0.116	0.935	0.52 ± 0.003	0.86	N. H.	N. H.
Cellotriose	7.33 ± 0.07	4.02	3.17	25.6	15.5 ± 0.49	25.5	36.8 ± 3.9	21.5
Cellotetraose	13.8 ± 0.4	7.58	3.84	31.0	26.1 ± 1.2	43	45.6 ± 2.9	26.7
Sophorose	3.32 ± 0.37	1.82	1.42	11.5	4.28 ± 0.33	7.05	19.9 ± 0.5	11.6
Laminaribiose	11.6 ± 0.6	6.37	5.63	45.4	28.8 ± 0.24	47.4	61.7 ± 6.2	36.1
Gentiobiose	0.117 ± 0.008	0.0643	N. H.	N. H.	0.45 ± 0.006	0.74	0.86 ± 0.15	0.502

Relative activities are shown for Os4BGlu12 enzymes recombinantly expressed in *P. pastoris* (rBGlu12P), recombinantly expressed in *E. coli* with the N-terminal tag (rBGlu12E, Opassiri et al., 2010) and with the tag removed (rBGlu12E new), as well as native enzyme purified from rice panicle (Native BGlu12). The activities were measured with 2 mM substrates and are given as μmol product/min/mg protein. Data are mean ± SD of three independent experiments. Percents are relative to 4NPGlc. ^acalculated from the kinetic parameters shown in Opassiri et al. (2010). N. D., not determined. N. H., not hydrolyzed. The table is modified from Himeno et al. (2013).

7.4. Structures of Os3BGlu6 and Os4BGlu12 Mutants with Substrates.

Although Os3BGlu7 and Os4BGlu12 hydrolyze oligosaccharides well, they are also known to hydrolyze glycoside substrates, as is Os3BGlu6 (Opassiri et al., 2003; 2010; Seshadri et al.; 2009, Wakuta et al., 2011; Himeno et al., 2013). During the course of another project identifying gibberellin glucosyl ester β-glucosidases, we identified Os3BGlu6 as the enzyme expressed in our lab that has the highest activity toward gibberellin GA₄ glucosyl ester (GA₄GE). When we attempted to solve the crystal structure of the Os3BGlu6 E178Q (catalytic acid/base) mutant with GA₄GE and 4NPGlc, we found that both substrates gave

glucose covalently bound to the catalytic nucleophile in the covalent intermediate of the two step retaining hydrolysis mechanism (Figure 8, Hua et al., 2013). These structures gave a clear view of how the 2-hydroxyl of the glucosyl moiety interacts with the surrounding amino acid residues and a critical water molecule, since previous structures of plant GH1 covalent complexes had 2-fluoroglucose, so these interactions were somewhat unclear.

Since we did not get the desired Michaelis complex of GA₄GE with the Os3BGlu6 E178Q mutant, we mutated the acid/base to Ala (Os3BGlu6 E178A) and the nucleophile to Asp (E392D) and Gln (E392Q). A crystal of Os3BGlu6 E394D with GA₄GE diffracted X-rays to 2.25 Å resolution, and showed density for the ligand in the active site. However, the density for the GA₄GE was somewhat disordered, so its exact position could not be clearly determined. Although the E392Q mutant crystals did not give adequate diffraction in the first attempt, it is hoped they can be used in the future, since this mutant showed the lowest rate of GA₄GE cleavage.

Since Wakuta et al. (2011) identified Os4BGlu12 as a tuberonic acid glucoside (TAG) β-glucosidase, TAGG2, we tried to solve the structure of Os4BGlu12 E179Q with tuberonic acid and tuberonic acid glucoside. Unfortunately, Os4BGlu12 crystals give quite variable and not very high resolution diffraction, but we managed to collect one dataset with TAG at 3.2 Å resolution. This structure was solved by rigid body refinement with the free Os4BGlu12 structure, and it showed strong electron density for the TAG in the active site. The model was built and refined and showed the general position of TAG binding, and then the structure was shared with Prof. Matsui's group in Hokkaido.

Aside from these datasets, we also collected several datasets with crystals of Os3BGlu6 E178Q, Os3BGlu7 E176Q and Os4BGlu12 E179Q crystals soaked with saturated trycin-5-O-glucoside and trycin-7-O-glucoside. However, due to the low solubility of these compounds, their electron densities were not observed in the data.

7.5. Structural and Functional Analysis of Os3BGlu7 Glycosynthases.

In this, we completed our analysis of what types of donor and acceptor substrates the Os3BGlu7 E386G glycosynthase can make, by using cyclophellitol to inhibit native activity. The TLC analysis of Os3BGlu7 E386G glycosynthase indicated that it was producing hydrolysis products and transglycosylation products of the acceptor substrate, such as 4NP-Glc-Glc, when 4NPGlc was used as the acceptor and α-fluoro-β-D-fucoside was used as the donor substrate. Since the glycosynthase should not transfer sugars from this donor, it suggested contamination with wildtype Os3BGlu7. Preincubation of the enzyme with the mechanism-based covalent inhibitor cyclophellitol for 2 h was shown to prevent formation of the hydrolysis and wild type transglucosylation products verifying that the Os3BGlu7 E386G was contaminated with wild type enzyme, either from mutational reversion or misreading of the codon on the ribosome. A new mutant, E386G2, with the codon GGC instead of GGA, which requires two bases to be changed or mismatched to bind to Glu tRNA decreased the wild type activity to the point that it was insignificant in the glycosynthase assay. The Os3BGlu7 E386G2 glycosynthase could use a range of donor and acceptor substrates, verifying the broad specificity of this enzyme, however, it required an equatorial 4-hydroxyl on the acceptor substrate in order to efficiently catalyze multiple transfers (Pengthaisong et al., 2012b).

To determine the structural reason for Os3BGlu7 E386G is a more efficient glycosynthase than Os3BGlu7 E386S and E386A, we determined the structures of these Os3BGlu7 glycosynthase mutants alone and with α-fluoroglucoside. Since these structures were very similar and showed a similarly positioned water molecule to help with fluoride departure (Figure 9), we began collaboration with Prof. Yongjun Liu and Jinhu Wang on

determining the structural reasons for different catalytic rates of our Os3BGlu7 E386G, E386S and E386A glycosynthases by QM/MM simulations starting from our crystal structures for the three enzymes and their complexes with donor substrate. The studies were able to calculate activation energies at the transition state that replicated the order of reaction rates in the three enzymes. They also showed that distances between the fluorine atom that acts as a leaving group and a conserved ordered water molecule followed the energy order, suggesting this is a key interaction in the glycosyl transfer reaction (Figure 10)

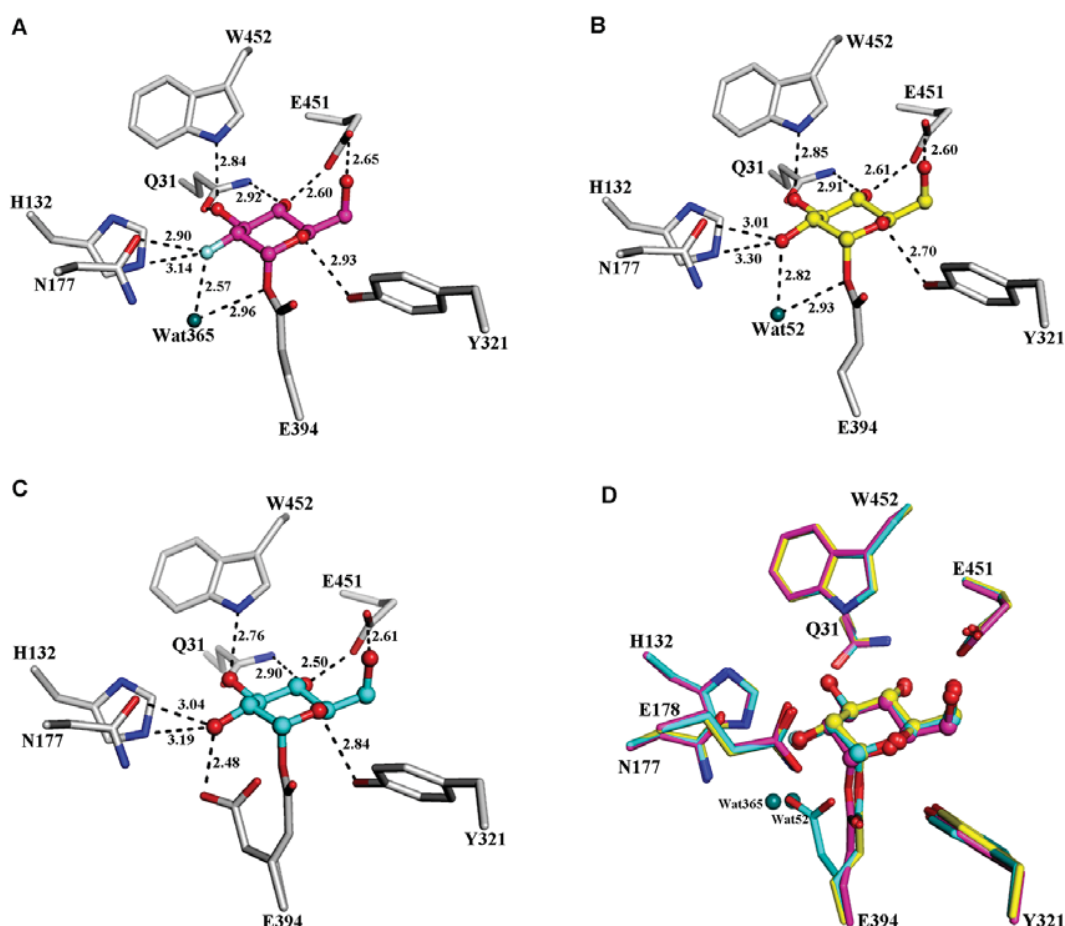


Figure 8. Comparison of the active site region of Os3BGlu6 bound with G2F (A) and those of the structures of Os3BGlu6 E178Q soaked with 4NPGlc (B) and Os3BGlu6 E178Q soaked with GA₄GE (C). The G2F and glucosyl moieties are represented as balls and sticks with carbon in pink for G2F, in yellow for Os3BGlu6 E178Q soaked with *p*NPGlc and cyan for Os3BGlu6 E178Q soaked with GA₄GE, oxygen in red and fluorine in pale cyan, (colors in the electronic version). The amino acid residues surrounding the ligands are represented by sticks with carbons in gray, nitrogens in dark blue and oxygens in red for all three structures. D. Superposition of the three structures in A, B and C, showing the nucleophile residues covalently bound to the glucosyl moiety at the -1 subsite of the Os3BGlu6E178Q structures soaked with 4NPGlc (yellow carbons in the electronic version) and GA₄GE (cyan carbons in the electronic version) and to the G2F moiety in Os3BGlu6 (3GNP, pink carbons in the electronic version). The unbound position of the nucleophile is represented by the alternative conformation of Os3BGlu6 E178Q soaked with GA₄GE. The molecules are represented as in A-C, except for the protein carbon colors.

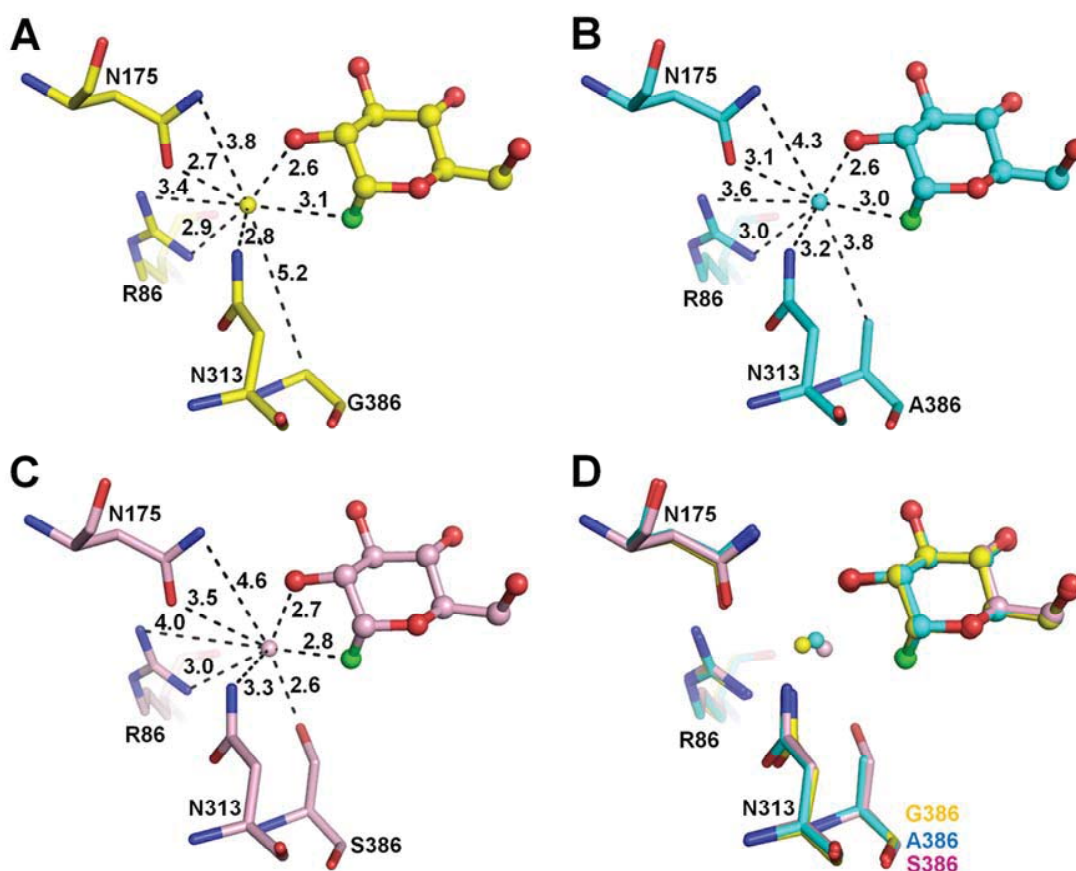


Figure 9. Active sites of the crystal structures of Os3BGlu7 E386G (A), E386A (B) and E386S (C) in complexes with α -fluoroglucoside. A fixed water found in each structure is shown as an unattached sphere for the oxygen molecule. The distances between the water molecule and the surrounding polar or mutated side chain atoms are shown as numbers next to the dotted lines connecting the atoms. D shows the superposition of the three structures, showing the positions of the atoms involved in hydrolysis near the missing nucleophile residue are very similar. The figure is essentially as presented in Wang et al. (2013).

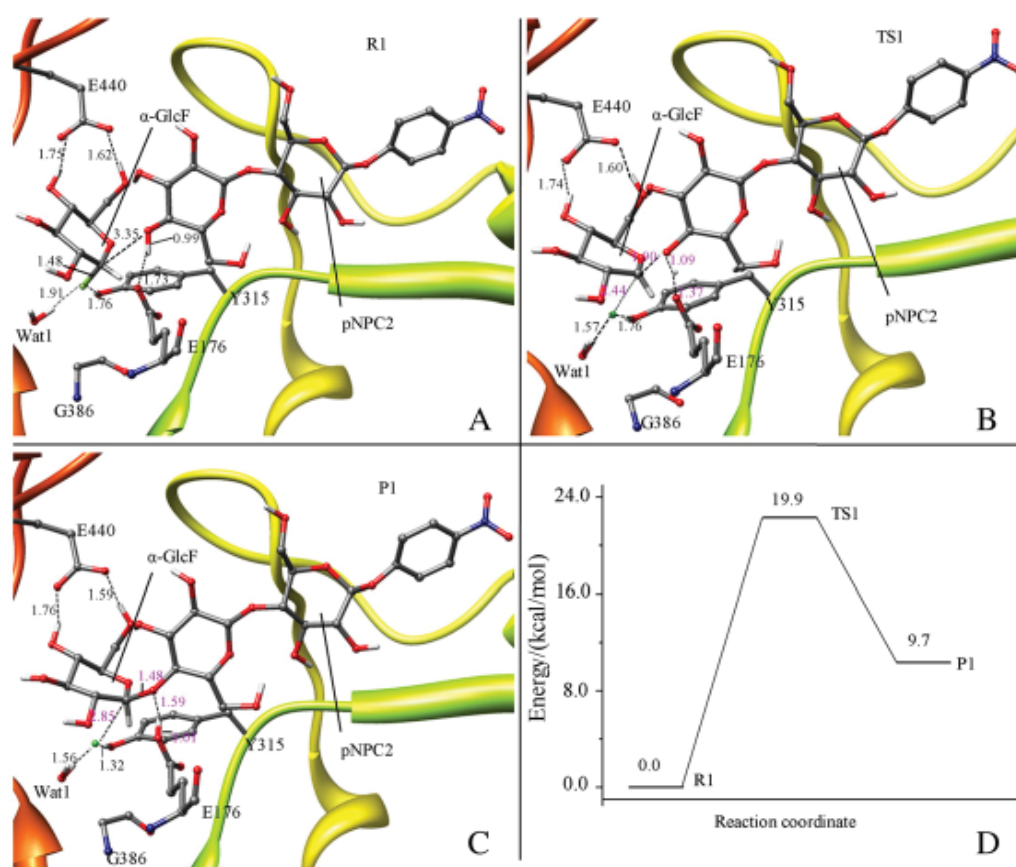


Figure 10. Optimized structures of the Os3BGlu7 E386G active site during the glycosynthase reaction in QM/MM simulations. A shows the computed substrate complex, including the α -fluoroglucoside from the crystal complex and a docked 4NP-cellobioside acceptor substrate. B shows the computed transition state configuration, and C shows the product at the completion of the reaction. The relative energies of the TS and product relative to the substrate complex are shown in D. Note that the calculated activation energy of 19.9 kcal/mol is lower than those computed for the Os3BGlu7 E386S (21.5 kcal/mol) and Os3BGlu7 E386A (21.9 kcal/mol) glycosynthases, in accordance with its higher reaction rate. The figure is from Wang et al. (2013).

7.6. Analysis of binding cleft residues binding cellooligosaccharides for hydrolysis and transglycosylation.

In this project, we characterized the activity and structure of rice Os3BGlu7 (BGlu1) E386G glycosynthase variants that had mutations of the putative cellooligosaccharide binding residues in the active site cleft, including the residues Asn245, Ser334 and Tyr341 in subsites +2, +3 and +4. We showed that the mutation of Tyr341 appeared to have larger effects on synthesis of long oligosaccharides by the glycosynthase than on hydrolysis of oligosaccharides by Os3BGlu7 (Pengthaisong et al., 2012a). The structure of the Os3BGlu7 E386G/Y341A mutant complexed with cellotetraose showed that the oligosaccharide bound in a different mode, with the glycosidic bond still in position for hydrolysis, suggesting a great plasticity in oligosaccharide binding for hydrolysis (Figure 11). The glucosyl residues in the +1 to +3 positions were moved to hydrogen bond to Gln183 on the opposite side of the cleft. To see whether this glutamine residue-binding mode was important to normal

hydrolysis and glycosynthase reactions, we have also mutated the Gln187 residue (Q187A). We found that this mutation had little effect on celooligosaccharide hydrolysis on its own, and also had little effect on β -glucosidase and glycosynthase activity when in combination with Y341A (Table 4). Inspection of the structures of Os3BGlu7 variants with the active site cleft mutations Y341A/Q187A in combination with either the E176Q or E386G inactivating mutations showed that the mutations with E386G still bound in a position similar to that seen in Y341A/E386G, while the Y341A or Y341A/Q187A mutations in the context of the E176Q showed movement of the cellotetraose in the opposite direction to bind in a position behind the normal position of the missing Tyr341 (Figure 12). We continue to work on identifying the alternative binding residues and determine their significance, as indicated in Table 4. It is notable that Arg178 and Trp337 mutations (R178A and W337A) have significant effects when made in combination with the Tyr341 mutation Y341A. This corresponds to the fact that in the variants with E176Q and Y341 mutations, the structures with cellotetraose showed the sugar chain moves to interact closely with Arg178 and for the nonreducing end glucosyl residue to stack on Trp337. Thus, the enzyme is able to adjust and form new binding sites to make up for the loss of oligosaccharide-binding residue functional groups.

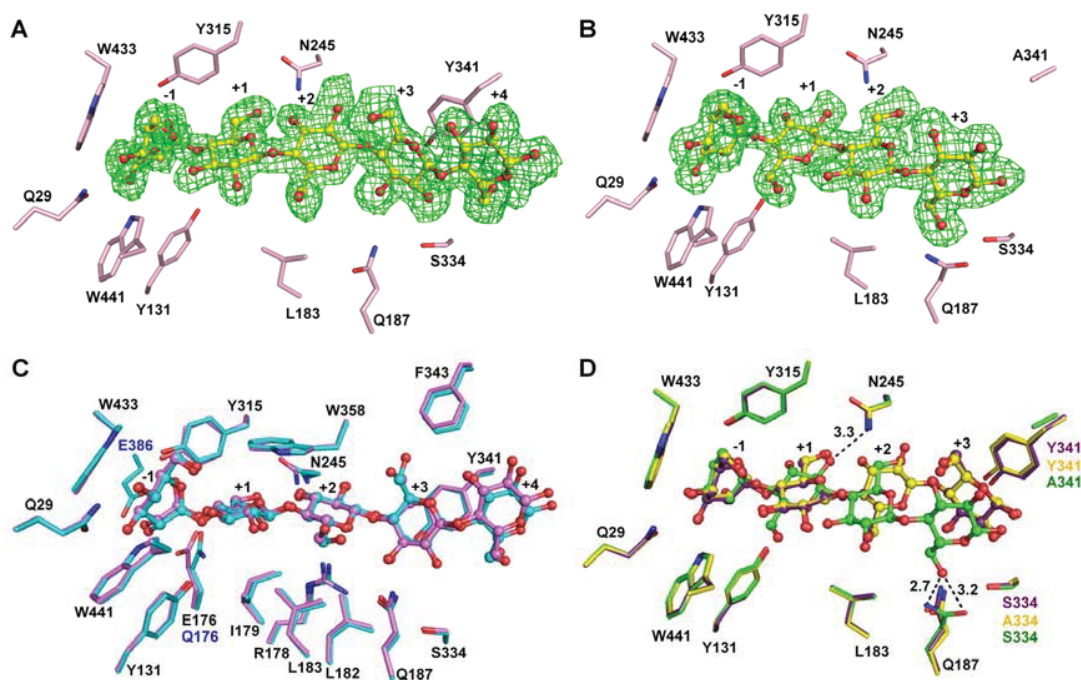


Figure 11. Cellooligosaccharide binding in the active site of Os3BGlu7 E386G glycosynthase and its active site cleft mutants. A. The electron density for the cellopentaose seen in the Os3BGlu7 E386G active site is shown in the context of the final refined structure for the cellopentaose and surrounding amino acid residues. B. The electron density for cellotetraose is shown, as in A. C. The structures of cellopentaose in the active sites of the Os3BGlu7 E176Q and E386G catalytic residue mutants are superimposed. The structures are very similar, but a subtle shift in the -1 subsite glucose results in a change in position and hydrogen bonding of Asn245 and slightly different reducing end residue positions between the two. In D, the structures of cellotetraose complexes of Os3BGlu7 E386G, E386G/S334A, and E386G/Y341A are superimposed to show the +1, +2, and +3 subsites are shifted in the E386G/Y341A mutant, so that these glucosyl residues are flipped and the fourth forms a hydrogen bond with Gln187.

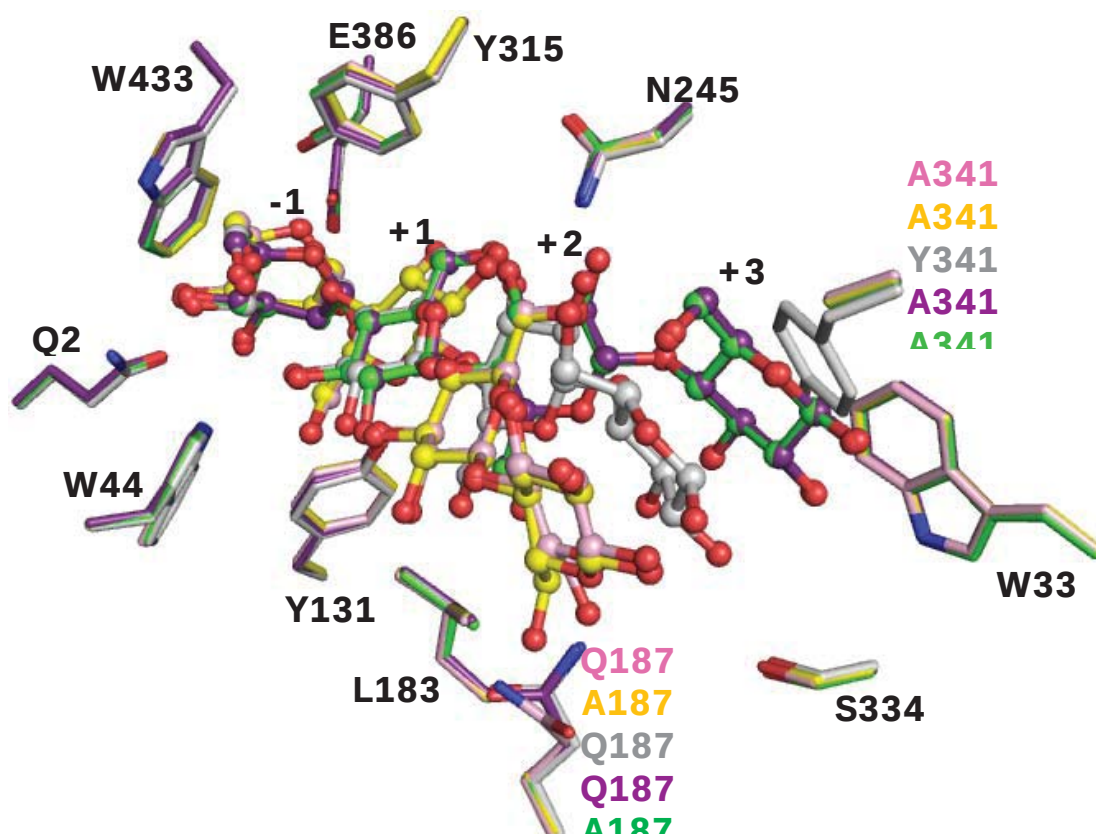


Figure 12: Comparison of Cellotetraose Binding in Os3BGlu7 variants with E386G or E176Q in combination with Y341A and Q187A mutations. The cellotetraose molecules are shown in ball and stick, while the surrounding amino acid residues are shown in stick representations. The central position of the cellotetraose is seen in E386G alone and is similar to that in E176Q alone. The lower two cellotetraose positions are seen in E386G/Y341A and E386G/Y341A/Q187A variants. The upper cellotetraose positions are seen in the E176Q/Y341A and E176Q/Y341A/Q187A variants. Thus, the subtle changes in binding in the -1 to +1 subsite between the E176Q and E386G mutants determine the alternative binding mode seen in the absence of Tyr341.

Table 4. Kinetic parameters for Os3BGlu7 and 10 variants for cellooligosaccharide hydrolysis.

Substrate	Kinetic parameters	Os3BGlu7	S334A	Y341A	Y341L	N245V	Q187A	Y341A/Q187A	N245V/Q187A	R178A	Y341A/R178A	Y341AW337A
Cellotriose	Km (mM)	0.50±0.05	0.62±0.05	0.97±0.07	0.97±0.05	11.18±0.79	0.48±0.04	0.75±0.06	15.15±0.85	1.91±0.94	6.41±0.48	2.12±0.16
	Vmax	17.06±0.64	22.09±0.72	28.85±0.79	25.41±0.49	39.4±1.0	14.81±0.42	21.69±0.6	37.10±0.84	20.7±1.78	37.93±1.07	30.35±0.80
	kcat (s-1)	15.64±0.59	20.25±0.66	26.45±0.72	23.29±0.45	36.13±0.92	13.58±0.39	19.88±0.55	34.01±0.77	18.98±1.63	34.77±0.98	27.82±0.73
	kcat/Km (s-1mM-1)	31.28	32.66	27.26	24.01	3.23	28.28	26.51	2.24	9.93	5.42	13.12
Cellobetraose	Km (mM)	0.37±0.04	0.52±0.04	0.70±0.05	0.54±0.03	5.30±0.23	0.42±0.03	0.57±0.04	6.66±0.42	1.03±0.07	3.87±0.27	1.09±0.05
	Vmax	25.81±1.18	31.77±1.33	38.76±1.08	28.94±0.48	112.9±1.7	26.15±1.05	32.23±0.77	104.71±2.77	47.50±1.32	40.50±1.03	37.01±0.60
	kcat (s-1)	23.66±1.08	29.12±1.22	35.53±0.99	26.53±0.44	103.5±1.6	23.97±0.96	29.54±0.71	95.98±2.54	43.54±1.21	37.13±0.94	33.93±0.55
	kcat/Km (s-1mM-1)	63.94	56.00	50.76	49.13	19.52	57.07	51.83	14.41	42.27	9.59	31.12
Cellopentaose	Km (mM)	0.28±0.02	0.36±0.03	0.44±0.02	0.36±0.02	3.83±0.18	0.28±0.03	0.43±0.03	4.46±0.43	0.59±0.03	3.80±0.34	1.18±0.06
	Vmax	29.99±1.26	34.7±1.7	35.2±0.64	27.94±0.39	151.7±3.5	28.11±1.84	27.43±0.62	128.25±5.81	49.76±0.96	27.80±1.14	33.97±0.63
	kcat (s-1)	27.49±1.16	31.81±1.53	32.3±0.59	25.61±0.36	139±3.2	25.77±1.69	25.14±0.57	117.56±5.32	45.61±0.88	25.48±1.05	31.14±0.58
	kcat/Km (s-1mM-1)	98.18	88.36	73.33	71.14	36.31	92.03	58.47	26.36	77.31	6.71	26.39

7.7. Investigation of structural differences between plant β -mannosidases and β -glucosidases.

In the previous grant, we tried to crystallize the barley HvBII β -mannosidase/ β -glucosidase without success, so we crystallized the closely related rice Os7BGlu26 β -mannosidase/ β -glucosidase instead and solved a preliminary structure. In the present work, we finalized the structure and solved a structure with a mannose molecule bound in the active site, as well as preliminary structures with the putative transition state analog inhibitors mannoimidazole, glucoimidazole, and isofagamine, which were identified from a larger set of putative inhibitors that we tested for inhibition against Os3BGlu7 β -glucosidase, Os7BGlu26 β -mannosidase and HvBII β -mannosidase. Similar structures are being generated with Os3BGlu7 β -glucosidase, which will be completed in the next project.

Initially, we solved the structure of the Os7BGlu26 β -mannosidase, based on a dataset at 2.20 Å resolution (Table 5). Attempts to soak the crystals of this protein with 2,4-dinitrophenyl β -D-2-deoxy-2-fluoromannoside and DNP2FG failed to generate covalent intermediates and the yellow color of the drops indicated that these putative inhibitors were cleaved by the enzyme. In fact, assays in solution showed these 2-fluoroglycosides did not significantly inhibit the enzyme and were rapidly turned over as substrates. We also attempted to soak with 100 mM D-mannose, and got electron density for mannose in the active site, which appeared to be bound to the HEPES usually found in the active site at the relatively low resolution of the structure (2.9 Å). Because it was unclear and later analysis suggested the mannose was a free sugar in the active site, we attempted to soak more crystals with 400 mM mannose and obtained a structure from a dataset cutoff at 2.45 Å (Table 5), which showed the free mannose and was the first structure of a product complex for a mannosidase. The crystals gave essentially isomorphous unit cells in the $P2_12_12_1$ space group with one protein molecule per asymmetric unit. Both structures were refined and the structures analyzed and compared to other plant family GH1 enzymes with known structures, since these are the first GH1 β -mannosidase structures (Tankrathok et al., In Press).

The fold of the structure of Os7BGlu26 is a classic TIM (β/α)₈-barrel, similar to other GH1 enzymes (Fig. 13A, Tankrathok et al., In Press). The structure placed the highly conserved E179 and E389 at the C-terminal ends of β -strands 4 and 7, respectively. These residues are positioned at the bottom of the active site cleft, as observed for the catalytic acid/base and nucleophile residues in other members of GH clan A (Jenkins et al., 1995; Henrissat et al., 1995). Further, nucleophilic rescue of mutants of the corresponding residues confirmed they are the catalytic residues in the closely related Os3BGlu7 β -D-glucosidase (Hommalai et al., 2007; Chuenchor et al., 2011). The four variable loops that have been reported to account for much of GH1 structural and functional diversity (Sanz-Aparicio *et al.*, 1998) were identified as loop A (A28-D68), loop B (E179-T209), loop C (H317-P366) and loop D (N390-D406) (Figs 13A). Although, no electron density was observed for the 14-residues from the N-terminal fusion tag and the five C-terminal residues of the Os7BGlu26 protein, five amino acid residues of the fusion protein linker region (A-A-P-F-T) and residues 1-478 of the mature Os7BGlu26 gave clear electron density for the structure. Two cis-peptide bonds were found between A194 and P195, and between W436 and S437, as seen in other plant GH1 enzymes (Barrett et al., 1995). The conserved disulfide bond found in nearly all plant GH1 enzymes was present on loop B between C198 and C201. The main chain torsion angles of the conserved active site tryptophan, Trp444,

fell in the outlier region of the Ramachandran plot in the model, but it is found in a similar outlier or borderline region in other GH1 enzymes (Czjzek et al., 2000; Chuenchor et al., 2008). Overall, the Ramachandran statistics were similar to those of other plant GH1 structures.

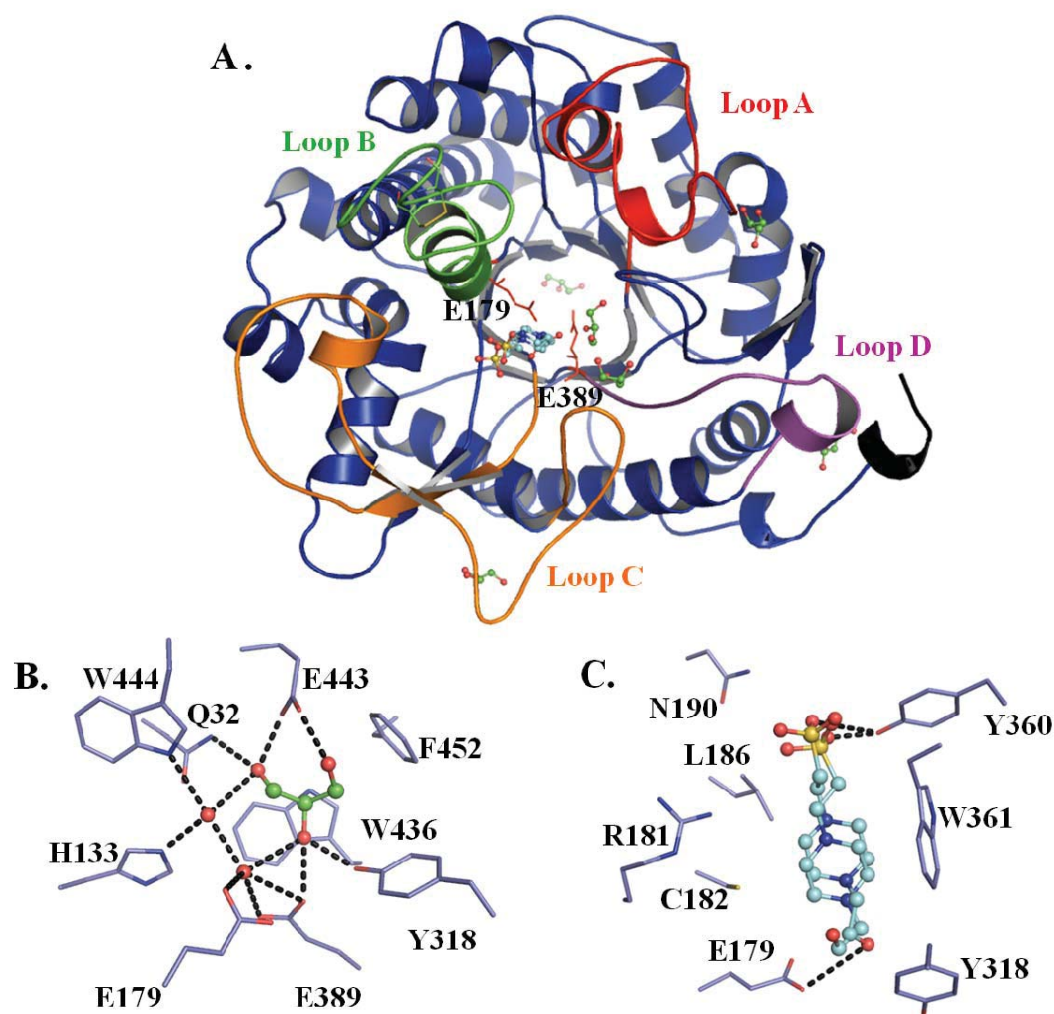


Figure 13. Structure of rice Os7BGlu26 β -D-mannosidase. A) Cartoon representation of the overall structure of rice Os7BGlu26. The catalytic nucleophile and acid/base residues are represented as sticks (E389 and E179), glycerol and HEPES are represented as ball-and-sticks, respectively. Loops A, B, C and D are indicated in red, green, orange and magenta, respectively, in the electronic version. (B) Amino acid residues around the -1 subsite, with contacts to a glycerol and water molecules occupying this site mediated through hydrogen bonds with distances between 2.4 and 3.1 Å, are indicated as dashed lines. (C) Amino acid residues interacting with the HEPES molecule in two alternate positions in the aglycon-binding region of the active site. The contacts through hydrogen bonds between the protein and HEPES are indicated as dashed lines. The figure is from Tankrathok et al. (In Press).

Table 5. Data-collection and processing statistics for Os7BGlu26 β -mannosidase structures.

	Native Os7BGlu26	Os7BGlu26/mannose complex
PDB code	4JHO	4JIE
Wavelength (Å)	1.00	1.00
Resolution range (Å)	30 – 2.20 (2.28 - 2.20)	30 – 2.45 (2.54 - 2.45)
Completeness (%)	99.9 (99.9)	95.8 (79.9)
Average redundancy per shell	5.5 (5.5)	6.1 (4.4)
$R_{\text{merge}}^{\dagger}$ (%)	8.8 (49.4)	14.9 (45.6)
$\{I/\sigma(I)\}$	18.0 (3.2)	11.9 (3.2)
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters (Å)	$a = 68.1, b = 71.7, c = 136.7$	$a = 68.0, b = 73.6, c = 134.0$
No. of unique reflections	34370	24116
No. of observed reflections	188040	147585
No. of molecules per ASU	1	1
R_{factor} (%)	17.6	15.3
$R_{\text{free}}^{\dagger\dagger}$ (%)	21.9	19.5
No. of protein atoms	3955	3960
No. of water molecules	305	234
No. of ligand atoms	None	12
No. of non solvent hetero atoms	51	51
r.m.s.d. bonds (Å)	0.010	0.010
r.m.s.d. angles (°)	1.155	1.212
Mean B -factor		
- Protein	22.9	23.7
- non solvent hetero atom	45.0	46.9
- solvent	32.8	30.8
- D-mannose	None	31.4
Ramachandran plot (%)		
Most favored	88.5	88.0
Allowed region	11.3	11.7
Outlier region	0.2	0.2

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$. Values in parentheses are for the outer shell.

$\dagger\dagger R_{\text{free}}$ represents the residual factor calculated from approximately 5% of the data that was not used in the refinement. Data are from Tankrathok et al. (In Press).

In the structure of Os7BGlu26 with D-mannose in the active site, the D-mannose fit the density best is a 1S_5 skew boat, which is the shape expected for a D-mannose product in the active site in mannose active enzymes from other families, if the deglycosylation step follows a trajectory the reverse of the 1S_5 skew boat to $B_{2,5}$ boat to 0S_2 skew boat proposed for the glycosylation step of hydrolysis in those enzymes (Figure 14 A). Since in the double displacement retaining mechanism, the enzymes go through the first glycosylation step to transfer the glucose to the enzyme nucleophile and the second deglycosylation step to transfer the glucose to a water (Rye and Withers, 2000), this deglycosylation step would be expected to be essentially the reverse of the glycosylation step. The distorted shape of the mannose was evidently supported by the set of strong, direct hydrogen bonds to the surrounding amino acid residues (Figure 14 B). When the apo and mannose complex structures were superimposed, no major changes were observed in the active site residue side chain positions upon binding the mannose (Figure 14 C). When the Os7BGlu26 structure with mannose was superimposed with the Os3BGlu7 structure with cellopentaose, the sugars in the -1 subsite were positioned very similarly, while the HEPES in the Os7BGlu26 structure was found to align with the glucosyl residues in the +1, +2 and +3 subsites of the Os3BGlu7 E176Q cellopentaose complex structure (Figure 14 D). Many of the residues surrounding the active site cleft were similar in the two enzymes, reflecting the fact that both enzymes hydrolyze β -1,4-linked oligosaccharides with similar kinetic properties (Kuntothom et al., 2009).

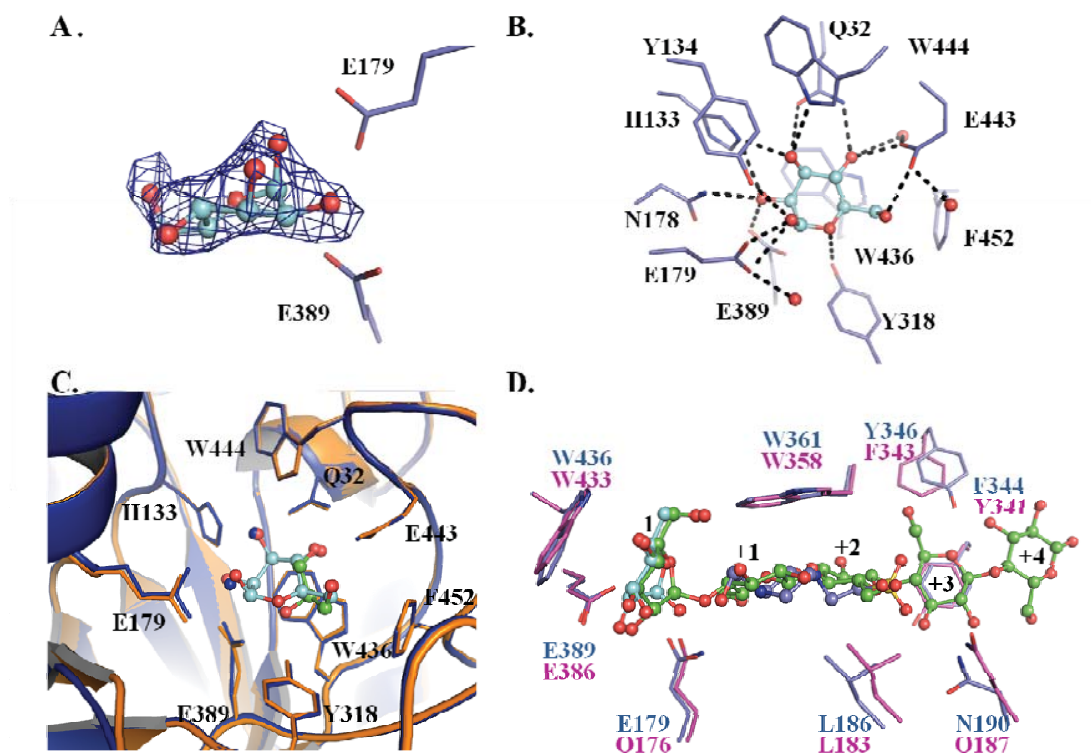


Figure 14. The structure of rice Os7BGlu26 in complex with β -D-mannose. A. An unbiased $F_o - F_c$ (OMIT) map of β -D-mannose is represented as a mesh contoured at 3σ . B. Amino acid residues interacting with β -D-mannose in the active site. The contacts through hydrogen bonds between the protein and β -D-mannose are indicated by dashed lines, while β -D-mannose is shown in ball-and-stick representation. The superimposition of the Os7BGlu26/ β -D-mannose complex (orange, in the electronic version) over the native Os7BGlu26 structure (blue in the electronic version) is shown in C. The superimposition of the Os7BGlu26/ β -D-mannose complex (carbons in blue in the electronic version) with the cellopentaose/Os3BGlu7 complex (carbons in violet in the electronic version) is shown in (D). β -D-mannoside, HEPES and cellopentaose are indicated in ball-and-stick representations. The glucosyl residue-binding subsites observed in the cellopentaose complex are marked by -1, +1, +2, +3, and +4.

To further explore the localization of substrates and surrounding residues in the active site, 3 approaches were taken: 1) the structures were superimposed with GH1 β -glucosidase structures complexed with substrates as described above; 2) the structures were superimposed with covalent intermediates of plant GH1 β -glucosidases (Os3BGlu6, Os3BGlu7 and Os4BGlu12 with G2F); and the 4NPGlc and 4NPMan substrates were computationally docked into the active site of the Os7BGlu26 structure. To avoid bias in assigning the sugar structures in the active site, the sugars were docked in four relatively low energy skew boat conformations, 1S_3 , 1S_5 , 0S_2 , and 3S_1 , and the energies of the lowest energy complexes compared (Figure 15). The lowest energy forms, 1S_3 Glc and 1S_5 Man, are shown in Figure 15. From these three methods, the residues surrounding the sugars in the -1 subsite were compared between β -glucosidase and β -mannosidase enzymes. It was noted that in the enzymes with significant mannosidase activities, a Tyr (Tyr134 in Os7BGlu26 and Tyr133 in Os3BGlu7) was found with its phenolic hydroxyl very close to the catalytic acid base, while this residue was replaced by Trp in Os3BGlu6 and Os4BGlu12, which have negligible β -mannosidase activities, and other GH1 β -glucosidases with known structures. The catalytic acid-base in Os7BGlu26 and Os3BGlu7 makes a short hydrogen bond to this group, while in Os3BGlu6, Os4BGlu12 and most other β -glucosidases with known structures, the catalytic acid-base is turned to hydrogen bond to a Thr in the active site cleft instead, which is in the position occupied by Cys182 in Os7BGlu26 and Ile179 in Os3BGlu7, while Val is found in this position in barley HvBII, Arabidopsis and tomato β -mannosidases. Thus, these residues around the catalytic acid-base of Os7BGlu26 were targeted for mutagenesis.

As shown in Table 6, conversion of Os7BGlu26 residues had differential effects on 4NPGlc and 4NPMan hydrolysis. Although mutation of the catalytic acid-base to Gln decreased the activity toward both substrates, it showed a slightly larger effect on 4NPMan than 4NPGlc. Conversion of Tyr134 to Trp increased the enzyme's k_{cat}/K_m toward both substrates, the change was much more substantial for the 4NPGlc, the K_m of which was lowered, compared to 4NPMan, for which the k_{cat} increased. This resulted in nearly equivalent k_{cat}/K_m ratios for 4NPMan and 4NPGlc in the Os7BGlu26 W134Y variant, although the maximum rate was still low for 4NPGlc. Converting Cys182 to Thr, allowing the acid/base to hydrogen bond to this position showed a decrease in k_{cat}/K_m for both glucoside and mannoside, but had a larger effect 4NPMan than 4NPGlc, so the ratio of β -mannosidase to β -glucosidase activity was reduced to approx. 2 from 3 in the wild type. The conversion of this residue to Ala had much smaller effects on k_{cat}/K_m , as predicted from the hypothesis that catalytic acid-base hydrogen bonding to the Thr was affecting the activity. Combining the W134Y and C182T mutations had nonadditive effects on the free energies of transition state binding ($\Delta\Delta G_{S^*mut}$), as expected for mutations of two residues that act on the same facet of catalysis (Mildvan et al., 1992). To see whether the removal of the hydrogen bonding group or the size of the aromatic group was the critical factor affecting the activity of the Y341W mutation, we also mutated Tyr341 to Phe. To our surprise, the mutation increased the k_{cat} and k_{cat}/K_m toward both substrates, and especially toward 4NPMan, so that the Os7BGlu26 enzyme is a much more effective β -mannosidase with a ratio of k_{cat}/K_m for 4NPMan over 4NPGlc of 35, compared to 3 in the wildtype. Thus, by modifying the residues around the catalytic acid/base, we were able to modify the relative β -mannosidase versus β -glucosidase activity.

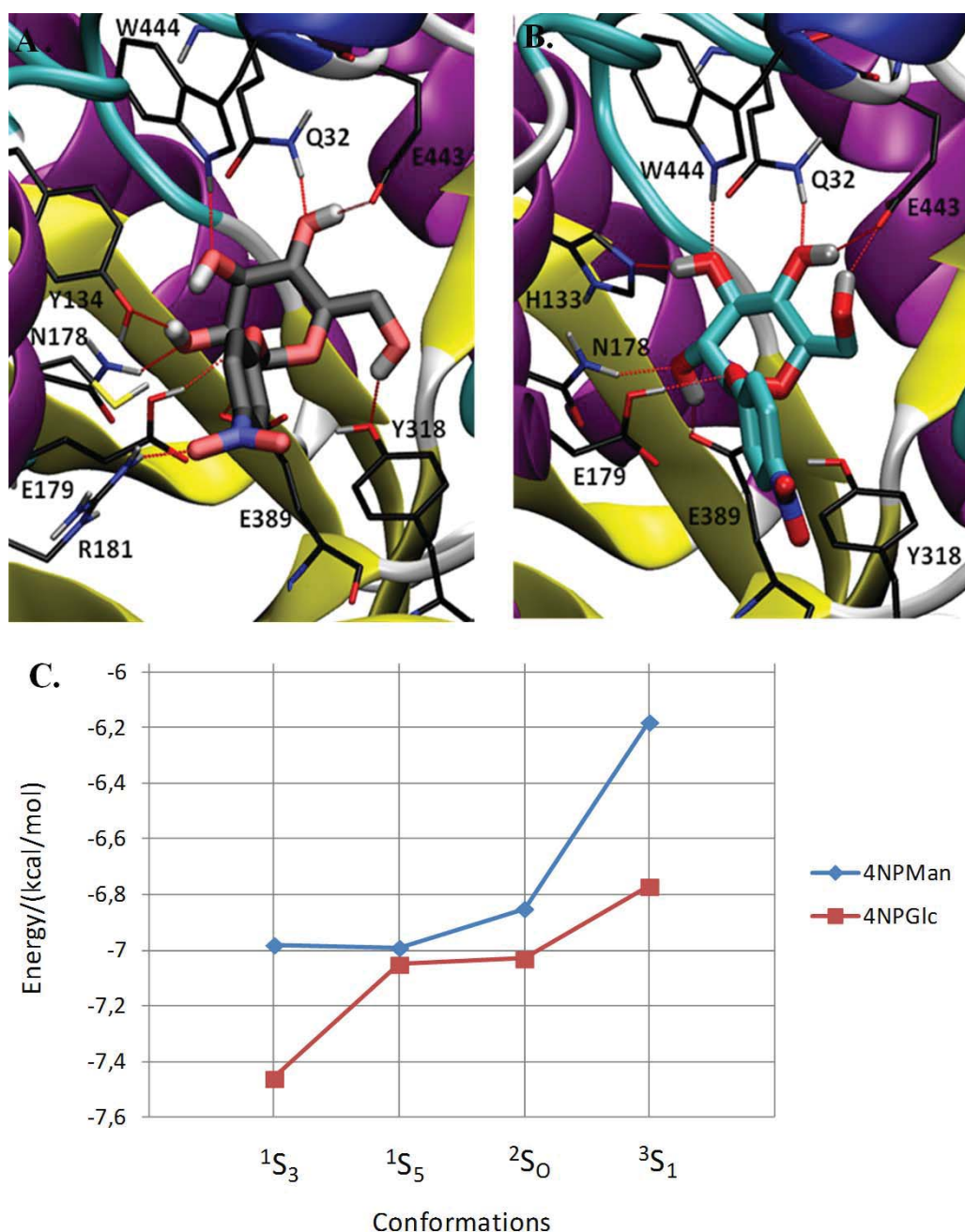


Figure 15. Computational docking of 4NPMann and 4NPGlc with different pyranose sugar conformations into the active site of the Os7BGlu26 β -mannosidase. A shows the structure of 4NPMann docked with a 1S_5 skew boat mannosyl moiety conformation, while B shows 4NPGlc docked with a 1S_3 glucosyl conformation. The distance between the carboxylic hydrogen atom of the catalytic acid/base residue (E179) and the glycosidic oxygen is 3.08 Å the glucoside and 3.58 Å for the mannoside. The distance between the O ϵ 1 of the nucleophile (E389) and the anomeric carbon is 2.98 Å for the glucoside and 3.0 Å for the mannoside. The binding energies of different conformers of 4NPMann and 4NPGlc in complex with Os7BGlu26 are shown in C. The figure was generated by Javier Iglesias-Fernández and Carme Rovira for Tankrathok et al. (In Press).

Table 6. Kinetic parameters of wild type Os7BGlu26 and its mutants using the 4NPMAN and 4NPGlc substrates.

	substrate	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (s ⁻¹ mM ⁻¹)	Mutant/wild type k_{cat}/K_m ratio	$\Delta\Delta G_{S^*mut}$ (kJmol ⁻¹)
Wild type	4NPMAN	0.48 ± 0.003	0.35 ± 0.004	0.714		
	4NPGlc	0.124 ± 0.002	0.029 ± 0.0001	0.237		
E179Q	4NPMAN	0.40 ± 0.003	0.008 ± 0.0004	0.02	0.028	+9.0
	4NPGlc	1.07 ± 0.02	0.0102 ± 0.0004	0.0095	0.040	+8.1
Y134W	4NPMAN	2.37 ± 0.04	2.4 ± 0.09	1.01	1.4	-0.9
	4NPGlc	0.032 ± 0.002	0.0306 ± 0.0006	0.959	4.0	-3.5
C182T	4NPMAN	12.4 ± 0.78	0.51 ± 0.03	0.043	0.060	+7.1
	4NPGlc	3.7 ± 0.14	0.078 ± 0.005	0.021	0.089	+6.1
Y134W/C182T	4NPMAN	11.56 ± 1.70	0.30 ± 0.03	0.026	0.036	+8.3
	4NPGlc	2.95 ± 0.17	0.058 ± 0.005	0.0198	0.084	+6.3
Y134F	4NPMAN	0.45 ± 0.02	12.5 ± 0.6	27.7	39	-9.2
	4NPGlc	0.167 ± 0.009	0.127 ± 0.003	0.759	3.2	-2.9
C182A	4NPMAN	2.46 ± 0.23	1.02 ± 0.07	0.414	0.58	+1.4
	4NPGlc	1.39 ± 0.11	0.242 ± 0.016	0.174	0.73	+0.8

$$\Delta\Delta G_{S^*mut} = -RT[\ln(k_{cat}/K_m)_{mutant} - \ln(k_{cat}/K_m)_{wildtype}] \text{ (Fersht } et al., 1987).$$

Since the structure of the Os7BGlu26 β -mannosidase and its complex with the mannose product did not give a complete picture of how β -D-glucosides and β -D-mannosides are distinguished by enzymes with different glycon specificities, we decided to investigate possible differences in the transition state. Thus, we collected the putative inhibitors shown in Figure 1 and tested their inhibition of Os7BGlu26 and HvBII β -D-mannosidases and Os3BGlu7 β -D-glucosidase, as shown in Table 7. Of those shown in Figure 1, only the inhibitors shown in Table 7 had significant inhibition at 100 μ M on the enzymes tested. As shown, the mannoimidazole was a much less effective inhibitor than the glucoimidazole for all three enzymes, even though the Os7BGlu26 and rHvBII enzymes prefer mannoside to glucoside. However, the glucoimidazoles had lower inhibition constants (greater energy of binding) in Os3BGlu7 β -glucosidase than in the β -mannosidases. Binding of these inhibitors to Os3BGlu7 was also verified by microcalorimetry, which showed that binding of glucoimidazole and isofagamine was driven by enthalpy, while the much weaker binding of mannoimidazole appeared to be entropy driven (data not shown).

Table 7. Inhibition of Os7BGlu26, Os3BGlu7 and rHvBII by putative transition state mimics.

	Os7BGlu26		Os3BGlu7		rHvBII	
	K_i (M)	ΔG (kJ/mol)	K_i (M)	ΔG (kJ/mol)	K_i (M)	ΔG (kJ/mol)
Mannoimidazole	13.2×10^{-6}	-28.3	3.8×10^{-6}	-31.4	10.6×10^{-6}	-28.9
Glucoimidazole	54.2×10^{-9}	-42.2	8.0×10^{-9}	-47.0	56.8×10^{-9}	-42.0
Phenethyl-Glucoimidazole	18.0×10^{-9}	-44.9	6.2×10^{-9}	-47.6	16.5×10^{-9}	-45.1
Isofagomine	2.97×10^{-6}	-32.1	0.58×10^{-6}	-36.2	0.10×10^{-6}	-40.6
Nojirimycin bisulfite	$>100 \times 10^{-6}$	>-23.2	17.3×10^{-6}	-27.6	5.05×10^{-6}	-30.7

The three inhibitors which had low inhibition constants and for which we had sufficient amounts were soaked into crystals of Os3BGlu7 β -glucosidase and Os7BGlu26 β -mannosidase. As shown in Figure 16, the crystal structure of mannoimidazole showed its ring was distorted between a 1S_5 skew boat and a $B_{2,5}$ boat, which is proposed to be the transition state conformation of mannosides. Glucoimidazole also had distortion to a 4E envelope conformation, which is close to the 4H_3 half chair transition state proposed for glucosides. When the structure of Os7BGlu26 in complex with isofagamine was solved, this inhibitor seemed to bind in a relaxed 4C_1 chair conformation. For comparison, we tried to solve the crystal structure of Os3BGlu7 β -glucosidase with the same inhibitors. We were able to solve the structure of Os3BGlu7 in complex with glucoimidazole, which also appeared to be close to a 4H_3 half-chair conformation, as predicted for glucosyl transition states. Similarly, the structure of the isofagamine also formed a 4C_1 relaxed chair, as in Os7BGlu26. When the structure of an Os3BGlu7 crystal soaked with

mannoimidazole was solved, the structure appeared to be glucoimidazole instead of mannoimidazole. When a second batch of mannoimidazole that was carefully synthesized was tested, it also gave a shape that was either unclear for the trajectory of the 2 hydroxyl (in molecule A of the asymmetric unit) or appeared to be glucoimidazole (in molecule B). Because the two inhibitors are synthesized together, we initially thought that this is a trace of glucoimidazole contaminant (estimated at 0.2% from the kinetic studies) was able to bind in the active site and outcompete the mannoimidazole present at about 500-fold higher concentration. Since mannoimidazole was clearly seen in the active site of Os7BGlu26 β -D-mannosidase, this shows a clear difference between these enzymes, despite the fact that they both bind glucoimidazole more tightly. However, it is still unclear whether this difference is really due to the presence of glucoimidazole at low concentration or a problem with the structure refinement of the mannoimidazole and an unusual shape that is not easily fit in this refinement process. We continue to work on these structures and to await for QM/MM and free energy landscape analysis from Prof. Carme Rovira and her student, which should be completed in the next project.

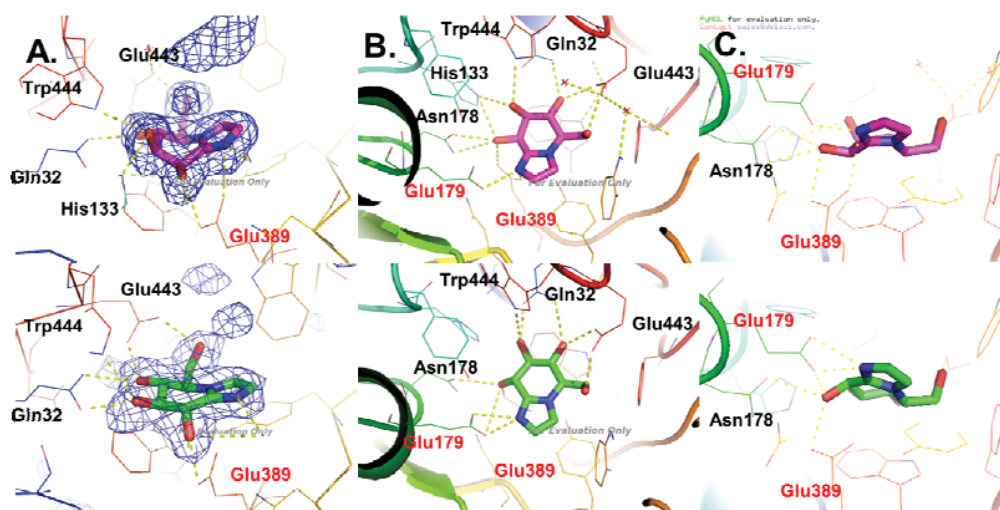


Figure 16. Shapes of glucoimidazole and mannoimidazole in the Os7BGlu26 active site. A. The electron density (2Fo-Fc maps) around the inhibitor position in the active sites for mannoimidazole (top) and glucoimidazole (bottom) calculated before insertion of the inhibitor models mapped into the active site with the final models. B. The models of the mannoimidazole (top) and glucoimidazole (bottom) in the active site from above the ring, showing the hydrogen bonding interactions with active site residues. C. The view of the inhibitors from the imidazole ring, demonstrating the $B_{2,5}$ boat shape of mannoimidazole (top) and 4E shape of glucoimidazole (bottom).

7.8. Structural and functional studies on barley and rice family GH3 enzymes

7.8.1. Effects of barley GH3 Exo I active site residues on substrate binding and hydrolysis.

In the previous project, we expressed barley HvExoI GH3 exoglucanase in *P. pastoris*, purified it, and crystallized it (Luang et al., 2010b), as well as making a series of active site mutations, including D95N, D95A, E220Q, E220A, E161A, E161Q, R158A, K206V, and R158A/E161A to test residues in the -1 subsite for

substrate binding and W286A and W434A in the +1 subsite. The expressed enzyme had very similar properties to that expressed in barley (Hrmova et al., 1996), and the crystals were isomorphous with those of native protein (Hrmova et al., 1998; Varghese et al., 1999). In this grant, we have continued this work by refining the crystal structures of the initial crystals, completing kinetic analysis of mutants and making more mutations, such as K206A and H207A and combination mutations of W286A/F and W434A/F in determining the -1 subsite substrate specificity, and these mutants were crystallized with β -1,2-, β -1,3-, β -1,4- and β -1,6- thio-linked oligosaccharides to see how they affect the binding of oligosaccharides with different linkages. We made a few more mutations of the Trp434 to Tyr (W434Y) and His (W434H) and found these to have small but significant activities.

To discern the roles of Trp286 and Trp434, the W286F and W434F mutants were tested, since the corresponding Ala mutant of Trp286 had no detectable activity and that of Trp434 had low activity. As seen in Table 8, the activities of these mutants were considerably higher, but still compromised compared to wild type HvExoI. Crystal structures were also determined for wild type HvExoI and the HvExoI W434A mutant in complexes with thio-linked analogues of laminaritrise, cellobiose, and gentiobiose. These structures showed that the without the Trp434, the catalytic acid/base, Glu491, lost a stabilizing hydrogen bond, thereby allowing it to take two alternate positions in the active site complexes with oligosaccharides, such as the complex with S-laminaritrise shown in Figure 17. The various oligosaccharides took different positions in the HvExoI W434A active site compared to the wildtype HvExoI, suggesting the substrates may bind less tightly in less optimal positions, accounting for the increased K_m and decreased k_{cat} for hydrolysis of laminaribiose with the W434A mutation.

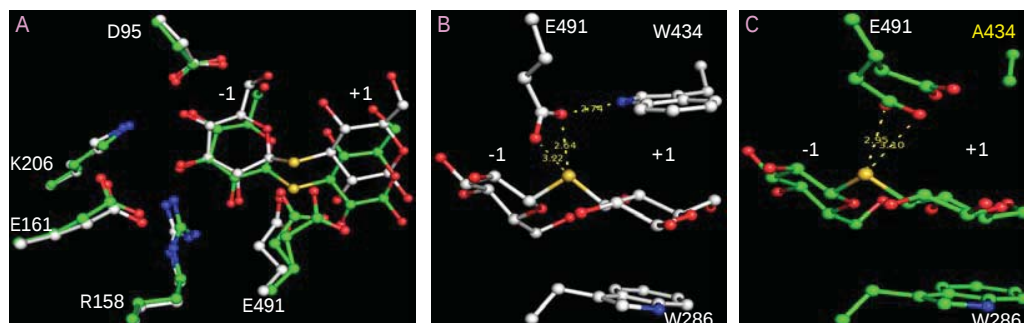


Figure 17. Structural comparison among wild type (grey) and Trp434Ala mutant (green) HvExoI in complex with S-laminaritrise. The superpositioned structures are shown in (A). In the active site of wild type HvExoI, Trp434 stabilizes the catalytic acid/base Glu491 and affects interactions between Glu491 and the S-glycosidic linkage (B). In the Trp434Ala variant, this stabilization is no longer possible and Glu491 occurs in two rotamers (C).

Table 8 Relative activities of HvExoI and its Glu220, Trp286 and Trp434 mutants.

		activity (umole/min/mg protein)					Relative activity (%)							
		glycosylated ExoI	deglycosyl ated ExoI	W434A	W434F	W286F	E220A	HvExoI	glycosylated ExoI	deglycosylated ExoI	W434A *	W434F *	W286F *	E220 A*
Polysaccharides Laminaran from <i>Laminaria digitata</i> Barley glucan Lichenans (<i>icelandic moss</i>)		54.4	65.3	nd	7.8	1.1	0.4	100.0	100.0	100.0	nd	14.4	2.0	0.7
		5.4	6.5	nd	nm	0.0	nd	10.0	10.0	10.0	nd	nd	0.05	nd
		0.7	nm	nd	nd	nd	nd	-	1.4	nm	nd	nd	nd	nd
		6.0	nm	nd	6.6	0.6	1.5	55.0	11.0	nm	nd	110.3	9.4	25.0
Oligosaccharides Sophorose (1-3)-beta-D linked oligosaccharides (2-7)	L2	20.8	nm	0.5	14.9	0.6	1.6	70.0	38.2	nm	2.4	71.5	2.8	7.7
	L3	42.1	nm	0.9	18.4	1.3	3.1	-	77.4	nm	2.1	43.7	3.1	7.4
	L4	39.1	nm	1.1	14.9	1.1	3.3	-	71.9	nm	2.8	38.2	2.9	8.4
	L5	33.8	nm	1.0	14.6	1.0	3.3	-	62.1	nm	3.0	43.2	3.0	9.8
	L6	46.9	nm	1.0	15.8	1.1	3.3	-	86.2	nm	2.1	33.6	2.3	7.0
	L7	43.9	nm	1.0	13.5	1.2	3.1	-	80.7	nm	2.3	30.8	2.7	7.1
	(1—4)-beta-D-linked oligosaccharides (2-6)	4.6	nm	nd	0.4	0.1	0.3	14.0	8.4	nm	nd	9.8	1.2	6.6
Gentiobiose	C2	8.8	nm	nd	1.2	0.1	0.4	-	16.3	nm	nd	13.8	1.4	4.5
	C4	8.6	nm	nd	1.9	0.1	0.5	-	15.8	nm	nd	22.2	1.5	5.8
	C5	8.4	nm	nd	1.7	0.1	0.3	-	15.4	nm	nd	20.2	1.4	3.6
	C6	9.1	nm	nd	1.8	0.1	0.6	-	16.8	nm	nd	19.6	1.1	6.6
		4.7	nm	nd	1.2	0.4	0.6	36.0	8.6	nm	nd	26.7	8.7	12.8
	Synthetic substrates 4NPGlc 4NPC2		12.4	11.0	6.1	5.0	0.9	1.6	10.0	22.8	16.8	49.2	40.2	7.1
		0.9	nm	nd	nd	nd	nd	-	1.7	nm	nd	nd	nd	nd

*relative activity compared to glycosylated ExoI wild type; nm=not measured; nd=not detected (not active)

7.8.2. Characterization of rice GH3 enzymes

In this project, we produced the OsExoI protein (derived from the Os03g0749500 gene locus and AK073110 cDNA) in recombinant *P. pastoris* yeast and the OsExoII protein (from the Os03g0749300 gene locus and NCBI Genbank accession AK065044 cDNA sequence and Os.59407 Unigene entry) and OsExoIII (from the Os01g0771900 locus) in recombinant *E. coli* and characterized the expressed enzymes. The OsExoI and OsExoII enzymes' pH (pH 5.5) and temperature (approx. 35 °C) optima were established and they were tested for hydrolysis of 4NP glycosides. As shown in Table 9, both hydrolyzed 4NPGlc much better than 4NP- α -L-arabinoside, 4NP- β -D-galactoside, 4NP- β -D-fucoside and 4NP- β -D-xyloside (4NPXyl), but the activity of OsExoI toward 4NPXyl was more significant, roughly 11% of its activity toward 4NPGlc. Upon checking hydrolysis of oligosaccharides and polysaccharides, OsExoI and OsExoII were shown to have similar relative activities for β -D-glucosyl-oligosaccharides of various linkages. Table 10 shows the kinetic parameters for OsExoII with various lengths of β -(1,3)- and β -(1,4)-linked oligosaccharides, but was much more efficient at hydrolyzing 4NPGlc and the polysaccharides laminarin and barley β -(1,3),(1,4)-glucan than oligosaccharides. Although OsExoI has not been fully characterized yet, due to it being more difficult to produce in *P. pastoris* than it is to produce OsExoII in *E. coli*, it did show a significantly better hydrolysis of the β -(1,3)-linked disaccharide laminaribiose than the β -(1,4)-linked disaccharide cellobiose, due to a much higher k_{cat} value with the former.

Although the cDNA encoding the mature protein for OsExoIII has been cloned into the pET32a vector and the protein expressed in Origami(DE3) *E. coli*, we have only shown that the expressed protein has activity to hydrolyze 4NPGlc, so further characterization is planned in the future.

Table 9. Relative activities of OsExoI and OsExoII against 4NP-glycosides and polysaccharides.

Substrate	Percent Relative Activity	
	OsExoI	OsExoII
4NP-glycosides ^a		
4NP- β -D-glucoside	100	100
4NP- β -D-fucoside	1.95	1.67
4NP- β -D-galactoside	2.0	2.3
4NP- β -D-xyloside	10.9	1.97
4NP- α -L-arabinoside	1.54	2.0
Polysaccharides ^b		
laminarin	100	100
barley (1,3)(1,4)- β -glucan	110	110
lichenan	61	57

^aAssays were done with 1 mM 4NP-glycosides and activity toward 4NPGlc was designated as 100% for each enzyme.

^bAssays were done with 1% polysaccharides by measuring glucose release and the activity toward laminarin was designated as 100% for each enzyme.

Table 10. Kinetic parameters for hydrolysis of oligosaccharides by OsExoI and OsExoII.

Substrate	Kinetics parameters		
		OsExoI	OsExoII
4NPGlc	K_m (mM)	2.0	0.57
	k_{cat} (s ⁻¹)	27.7	33.3
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	13.9	58.4
Cellobiose	K_m (mM)	1.2	0.4
	k_{cat} (s ⁻¹)	4.2	1.10
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	3.5	2.75
Cellotriose	K_m (mM)	nm	0.57
	k_{cat} (s ⁻¹)	nm	1.16
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	nm	2.03
Cellotetraose	K_m (mM)	nm	0.59
	k_{cat} (s ⁻¹)	nm	1.19
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	nm	2.01
Laminaribiose	K_m (mM)	1.3	0.45
	k_{cat} (s ⁻¹)	54.2	1.50
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	41.2	3.33
Laminaritriose	K_m (mM)	nm	0.62
	k_{cat} (s ⁻¹)	nm	1.63
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	nm	2.62
Laminaritetraose	K_m (mM)	nm	0.65
	k_{cat} (s ⁻¹)	nm	1.65
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	nm	2.53
Laminarin	K_m (mM)	nm	0.12
	k_{cat} (s ⁻¹)	nm	56.3
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	nm	469
Barley (1,3;1,4)- β -D-glucan	K_m (mM)	nm	0.15
	k_{cat} (s ⁻¹)	nm	76.5
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	nm	510

*nm means “not measured”.

8. Conclusions and Benefits

8.1. Scientific Conclusions and Benefits

In this project, we have increased the knowledge of how plant β -glucosidases, exoglucanases and β -mannosidases interact with their substrates and gain selectivity for one substrate over another. We have also shown that the determinants for specificity are subtle, but that single mutations can cause significant effects on relative activities toward substrates with different glycons or aglycons (Ketudat Cairns et al., 2012). Evolutionarily, this also suggests that relatively few amino acid changes are needed to change the selectivity from one substrate to another, so that global sequence similarity cannot always be trusted to indicate specific functions. On the other hand, the specificity for oligosaccharides or glycon may depend on the overall shape of the active site, as much or more than on the residues that line the active site and directly interact with the substrate.

The X-ray structure of Os4BGlu12, a rice β -D-glucosidase that acts on oligosaccharides and a large variety of glycosides, was determined and compared to other plant glycoside hydrolase family GH1 enzymes (Sansenya et al., 2011). The structures of its covalent complex with 2-deoxy-2-fluoroglucoside (G2F) and Michaelis complex with an as yet unreacted 2,4-dinitrophenyl β -D-2-deoxy-2-fluoroglucoside allowed comparison of the substrate-binding residues with other GH1 enzymes, to show the glycon-binding residues were similar, but aglycon-binding residues divergent. It was also found that the catalytic nucleophile of Os4BGlu12 bound to G2F was in a position similar to white mustard myrosinase, and significantly different from other β -D-glucosidases, which correlated to a significant thioglucosidase activity for this enzyme.

Characterization of the activity of the Os4BGlu12 enzyme prepared for crystallography from recombinant *E. coli* with the previous data for the activity of the fusion protein and the activities of Os4BGlu12 enzyme expressed in *P. pastoris* and purified from the plant indicated that removal of the fusion tag increased the activity of the enzyme, but it was still somewhat lower than enzymes purified from the rice and yeast for 4NP-glycosides (Himeno et al., 2013). In general, the relative activities from the different systems were similar, but with a few substrates, enzyme from one system would have significantly different relative activity. The Os4BGlu12 from all systems hydrolyzed helicin and the phytohormone glycosides SAG and TAG better than oligosaccharides, supporting a biological role in hydrolysis of these glycosides. We also established that Os3BGlu6 hydrolysis of a glucosyl ester, that of the gibberellin GA₄, appears to utilize the same mechanistic pathway, with the same glucosyl-enzyme intermediate seen in glucosides, and provided the structure of this intermediate in an acid-base mutant to clarify the interactions between the covalently linked glucose and surrounding residues (Hua et al., 2013).

A comparison of the active site cleft in the Os4BGlu12 structure with those of Os3BGlu6 and Os3BGlu7, showed that Os3BGlu6 has a bump in the active site cleft surface at Met251, which replaces the hydrogen-bonding residues Asn245 in Os3BGlu7 and His253 in Os4BGlu12. Replacing Met251 in Os3BGlu6 with Asn significantly increased binding of the third glucosyl residue of cellotriose, promoting its hydrolysis, while converting Os3BGlu7 Asn245 and Os4BGlu12 His253 to Met significantly decreased binding and hydrolysis of oligosaccharides (Sansenya et al., 2012). Thus, we were able to add or cripple a glucosyl residue binding site in the active site cleft based on observing the structures and comparing the activities.

We had previously observed that Os3BGlu7 E386G glycosynthase mutant catalyzes more rapid transglycosylation from α -fluoroglucoside to 4NP-cellooligosaccharides substrates than the Os3BGlu7 E386A and E386S mutants (Hommalai et al., 2007). Previously, it had been supposed that the replacement of the serine hydroxyl with a water molecule that is more flexibly positioned might account for this, but density for only one water was found in this part

of the active site of the structures in all three enzymes, suggesting more subtle differences affect the relative activities (Wang et al., 2013). QM/MM calculations starting from our structures were able to replicate the relative rates in terms of activation barriers for the three enzymes, and suggested that the positioning of the observed water is one of the key factors.

Further exploration of the ability of the Os3BGlu7 E386G glycosynthase to catalyze transfer reactions indicated that the enzyme can accept a broad range of donors and acceptors (Pengthaisong et al., 2012b). However, early studies were found to produce some products by wild type enzyme transglycosylation reactions, indicating a contamination with the wild type enzyme, which could be decreased to an acceptable level with a new construct that utilized a codon for the Gly that has two nucleotides different from that encoding Glu. The new construct was used for further analysis, which indicated the enzyme could transfer monosaccharide moieties to various sugar linkages, but only 1,4-linked chains were efficiently extended.

To verify the apparent glucosyl-binding residues in the outer subsites of Os3BGlu7 and see whether they were responsible for the synthesis of long chains by the Os3BGlu7 E386G glycosynthase, the residue that appeared to form stacking interactions with the fourth and fifth glucosyl residues, Tyr341, and a residue that seemed to contribute to the water-mediated hydrogen bonding network, Ser334, were mutated (Pengthaisong et al., 2012a). The Y341L and Y341A mutations seemed to cause only small effects on hydrolysis of oligosaccharides, but had a significant effect on the synthesis of long-chain oligosaccharides by the glycosynthase. The relatively small effect of the Y341A mutation seemed to be explained by the crystal structure of the Os3BGlu7 E386G/Y341A mutant in complex with cellotetraose, which showed that the 3 outer glucosyl residues were flipped and the fourth was hydrogen bonding to Gln187 on the opposite side of the active site. However, mutating Gln187 in the Q187A/Y341A mutant appeared to have little effect on hydrolysis or even the position of the oligosaccharide in the active site. The cellotetraose binding position appears to be dependent on the type of inactivating mutation as combining Y341A with E176Q resulted in binding of the outer glucosyl residues in the space created by removing the Y341 sidechain, where the sugars interacted with Arg178 and stacked on Trp337. Thus, the Os3BGlu7 enzyme shows a great plasticity in its cellooligosaccharide binding, such that when one interacting residue is removed, a new interaction can take its place. This may suggest the shape of the active site cleft and its overall mixture of polar, aromatic and nonpolar character may be the critical issue for cellooligosaccharide binding. Such information may be useful for engineering glucan-derived oligosaccharide hydrolyzing enzymes for biomass conversion and biofuel production in the future.

In terms of the relative activities of plant glycoside hydrolase family GH1 enzymes toward mannosides versus glucosides, we have identified the residues around the catalytic acid/base residue as having a profound effect on the hydrolysis of 4NPMann, as well as exploring the structural basis for substrate selection (Tankrathok et al., In Press). The structure of the Os7BGlu26 β -D-mannosidase had a somewhat narrower active site pocket than related Os3BGlu7 β -D-glucosidase and other plant β -D-glucosidases, suggesting the role of surrounding residues in the second and third shells in establishing the glycon specificity, since all the residues making direct contact with the glycon are conserved. The enzyme seems to use two different conformational trajectories and transition state shapes for hydrolysis of β -D-glucosides and β -D-mannosides, so we continue to explore the effect of transition state binding on the relative glycon specificities of related enzymes. Since the β -mannosidase-like enzymes are highly expressed in developing cereal grains, understanding how they hydrolyze the manno- and gluco- oligosaccharides may help in determining their role in development and whether they should be targeted to increase grain yields.

For glycoside hydrolase GH3 enzymes, we have collected a large amount of functional and structural data on barley HvExoI mutants and their complexes with various substrates and thio-oligosaccharide substrate analogs, in collaboration with Prof. Maria Hrmova in the University of Adelaide, Australia. These are helping to elucidate the basis for broad specificity of the HvExoI enzyme for oligosaccharides with different linkages and the role of the retained glucose product in catalysis and the mechanism for its exchange with a new substrate. Although these barley HvExoI studies have largely moved to Prof. Hrmova's laboratory with the move of Dr. Sukanya Luang from our group to hers, we continue to work on the rice GH3 exoglucanases. The rice OsExoII has been extensively characterized to show the kinetic parameters for its hydrolysis of oligosaccharides and the OsExoI enzyme is also in the process of kinetic characterization. We anticipate that the characterization of these enzymes will show whether they are likely to play the same roles as the barley enzymes or to have characteristic properties for rice that help to establish different kinetics of carbohydrate turnover in rice compared to barley. In addition, characterization of a third isoenzyme, OsExoIII, may yield more insights into glucan turnover in rice, which may impact on engineering rice development and the use of rice waste feedstocks in biomass conversion.

On the whole, we have generated a significant body of new knowledge on the substrate recognition of oligosaccharide-active exoglucanases, β -glucosidases, and β -mannosidases. This includes not only their action on oligosaccharides, but also on glycosides, including those of bioactive aglycons, such as phytohormones and flavonoids. This work will provide the basis for future studies to increase our understanding further and to engineer the GH1 and GH3 enzymes both for the release of glucose from carbohydrate feed stocks and for the production of glycosides with novel activities.

8.2. Other benefits.

This work has contributed to the publication of eight research papers and one invited review and to the training of five graduate students and one and a half postdoctoral fellows. It has also generated a large body of reserve data that will be completed in the new future and should provide the basis for another 4-5 international publications when supplemented with a few completing experiments. Indeed, we are in the process of completing one manuscript with Maria Hrmova's group, and others are anticipated soon.

8.3. Future benefits.

As noted above, our work has provided a basis for understanding oligosaccharide and other substrate binding that can guide further exploration of enzyme function and engineering. Since the world faces a shortage of fuels and food with the rapidly expanding world population, and concern about the release of carbon dioxide and other greenhouse gases causing global warming, our work serves to guide many scientists looking to improve biofuel and food production. Computational chemistry groups have used our structural models (Chuenchor et al., 2011) as the basis for simulating gluco-oligosaccharide hydrolysis, in the hopes of providing new information that can be used to engineer more efficient enzymes and conversion processes (Wang et al, 2011, 2013; Badiéyan et al., 2012). We anticipate this will continue in the future, and that more efforts to engineer carbohydrate active enzymes will take advantage of this information. For our part, we plan to utilize the information gained to apply the enzymes for the production of novel glycosides, as well as the use of glucan feed stocks for this purpose.

In addition to the work providing a large portion of the data for an additional 4-5 international papers, as mentioned above, it allows us to continue to explore both substrate binding and the application of glycosynthases and active site cleft mutants to synthesize

novel products in the next project. In addition, the students, postdoctoral fellows and research assistants trained in this work have already started to make their marks as researchers both within Thailand and abroad, so this training will continue to have impact in both related and unrelated investigations. Some have also become instructors, so their abilities to investigate scientific problems will help them to instill the scientific method and analytical thinking in new generations of science students. This is a major benefit of basic research that can feed into its appreciation and application in the future. Therefore, we anticipate that this project will continue to benefit Thailand and the world for some time to come.

9. References:

Badieyan S, Bevan DR, Zhang C. (2012) Probing the active site chemistry of β -glucosidases along the hydrolysis reaction pathway. *Biochemistry* 51, 8907-8918.

Barrett T, Suresh CG, Tolley SP, Dodson EJ, Hughes MA. (1995) The crystal structure of a cyanogenic β -glucosidase from white clover, a family 1 glycosyl hydrolase, *Structure* 5, 951-960.

Collaborative Computing Project Number 4. (1994) The CCP4 suite: programs for protein Crystallography. *Acta Crystallog. Sect. D* 50, 760-763.

Chen VB, Arendall WB III, Headd JJ, Keedy DA, Immomino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. (2010) Molprobity: all atom structure validation for macromolecular chromatography. *Acta Crystallog. Sect. D* 66, 12-21.

Chuenchor W, Pengthaisong S, Yuvaniyama J, Opassiri R, Svasti J and Ketudat Cairns JR. (2006) Purification, crystallization and preliminary X-ray analysis of rice BGlu1 β -glucosidase with and without 2-deoxy-2-fluoro- β -D-glucoside inhibitor *Acta Crystallog. Sect. F* 62, 798-801.

Chuenchor W, Pengthaisong S, Robinson RC, Yuvaniyama J, Oonanant W, Bevan DR, Esen A, Chen C-J, Opassiri R, Svasti J, Ketudat Cairns JR. (2008) Structural insights into rice BGlu1 β -glucosidase oligosaccharide hydrolysis and transglycosylation. *J. Mol. Biol.* 377, 1200-1215.

Chuenchor W, Pengthaisong S, Robinson RC, Yuvaniyama J, Svasti J, Ketudat Cairns JR. 2011. The structural basis of oligosaccharide binding by rice BGlu1 beta-glucosidase. *J. Struct. Biol.* 173, 169-179.

Czjzek M, M. Cicek M, V. Zamboni Z, W.P. Burmeister WP, D.R. Bevan DR, Henrissat B, A. Esen. (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 GJ, Ducros VMA, Varrot A, Zechel DL. (2003) Mapping the conformational itinerary of beta-glycosidases by X-ray crystallography. *Biochem. Soc. Trans.* 31, 523-527.

Emsley P, Cowtan K (2004) Coot-model-building tools for molecular graphics. *Acta Crystallog. Sect. D* 60, 2126-2132.

Fersht AR, Leatherbarrow RJ, Wells TNC. (1987) Structure-activity relationships in engineered proteins: analysis of use of binding energy by linear free energy relationships. *Biochemistry* 19, 6030-6038.

Henrissat B, Callebaut I, Fabrega S, Lehn P, Mornon JP, Davies G. (1995) Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases *Proc. Natl. Acad. Sci. USA* 92 (1995) 7090–7094.

Himeno N, Saburi W, Takeda R, Wakuta S, Matsuura H, Nabeta K, Sansenya S, Ketudat Cairns JR, Imai R, Mori H, Matsuda H. (2013) Identification of rice β -glucosidase with high hydrolytic activity towards salicylic β -glucoside. *Biosci. Biotechnol. Biochem.* 77, 934-939.

Hiromi K, Nitta Y, Namura C, Ono S. (1973) *Biochim. Biophys. Acta.* 302 362-375.

Hommalai G, Withers SG, Chuenchor W, Ketudat Cairns JR, Svasti J. (2007) Enzymatic synthesis of cello-oligosaccharides by mutated rice β -glucosidases. *Glycobiology* 17, 744-753.

Hrmova M, Harvey AJ, Wang J, Shirley NJ, Jones GP, Stone BA, Høj PB, Fincher GB. (1996) Barley β -D-glucan exohydrolases with β -D-glucosidase activity. Purification and determination of primary structure from a cDNA clone. *J. Biol. Chem.* 271, 5277–5286.

Hrmova M, Varghese JN, Høj PB, Fincher GB. (1998) Crystallization and preliminary X-ray analysis of beta-glucan exohydrolase isoenzyme ExoI from barley (*Hordeum vulgare*). *Acta Crystallogr. Sect. D* 54, 687-689.

Hua Y, Sansenya S, Saetang C, Wakuta S, Ketudat Cairns JR. (2013) Enzymatic and structural characterization of hydrolysis of gibberellin A4 glucosyl ester by a rice β -D-glucosidase. *Arch. Biochem. Biophys.* 537, 39-48.

Jenkins J, Lo Leggio L, Harris G, Pickersgill R. (1995) β -Glucosidases, β -galactosidases, family A cellulases, family F xylanases and two barley glucanases form a superfamily of enzymes with 8-fold β/α architecture and two conserved glutamates near the carboxyl-terminal ends of β strands four and seven. *FEBS Lett.* 362, 281-285.

Ketudat Cairns JR, Esen A. (2010) Beta-Glucosidases. *Cell Mol. Life Sci.* 67, 3389-3405.

Ketudat Cairns JR, Pengthaisong S, Luang S, Sansenya A, Tankrathok A, Svasti J. (2012) Protein-carbohydrate interactions leading to hydrolysis and transglycosylation in plant glycoside hydrolase family 1 enzymes. *J. Appl. Glycosci.* 59, 51-62.

Kuntothom T, Luang S, Harvey AJ, Fincher GB, Opasiri R, Hrmova M, Ketudat Cairns JR. (2009) Rice family GH1 glycosyl hydrolases with β -D-glucosidase and β -D-mannosidase activities. *Arch. Biochem. Biophys.* 491, 85-95.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.* 26, 283–291.

- Luang S, Hrmova M, Ketudat Cairns JR. (2010a) High-level expression of barley β -D-glucan exohydrolase HvExoI from a codon-optimized cDNA in *Pichia pastoris*. *Protein Express. Purific.* 73, 90-98.
- Luang S, Cairns JRK, Streltsov VA, Hrmova M. (2010b) Crystallisation of wild-type and variant forms of a recombinant plant enzyme β -D-glucan glucohydrolase from barley (*Hordeum vulgare* L.) and preliminary X-ray analysis. *Int. J. Mol. Sci.* 11, 2759-2769.
- Mildvan AS, Weber DJ, Kuliopulos A. (1992) *Arch. Biochem. Biophys.* 294, 327-349.
- Morris GM, Goodsell DS, Halliday, RS, Huey, R, Hart WE, Belew RK, Olson AJ. (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comp. Chem.* 19, 1639-1662.
- Murshudov GN, Lebedev A, Vagin AA, Wilson KS, Dodson EJ. (1999) Efficient anisotropic refinement of macromolecular structures using FFT. *Acta. Crystallogr. Sect. D* 55, 247–255.
- Opassiri R, Ketudat Cairns JR, Akiyama T, Wara-Aswapati O, Svasti J, and Esen A. (2003) Characterization of a rice β -glucosidase genes highly expressed in flower and germinating shoot *Plant Sci.* **165**, 627-638.
- Opassiri R, Hua Y, Wara-Aswapati O, Akiyama T, Svasti J, Esen A, and Ketudat Cairns JR. (2004) β -Glucosidase, exo- β -glucanase and pyridoxine transglucosylase activities of rice BGlu1. *Biochem. J.* **379**, 125-131.
- Opassiri R, Pomthong B, Okoksoong T, Akiyama T, Esen A, Ketudat Cairns JR. (2006) Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12 β -glucosidase *BMC Plant Biol.* **6**, 33.
- Opassiri R, Maneesan J, AkiyamaT, Pomthong B, Jin S, Kimura A, and Ketudat Cairns JR (2010) Rice Os4BGlu12 is a wound-induced β -glucosidase that hydrolyzes cell-wall- β -glucan-derived oligosaccharides and glycosides. *Plant Sci.* 179, 273-280.
- Otwinowski Z, Minor W. (1997) Processing of X-ray diffraction data collected in oscillation mode, in: C.W. Carter & R.M. Sweet (Eds.), *Meth. Enzymol.* Vol. 276 Academic Press, New York, pp 307–326.
- Pengthaisong S, Withers SG, Kuaprasert B, Svasti J, Ketudat Cairns JR. (2012a) The role of the oligosaccharide binding cleft of rice BGlu1 in hydrolysis of cellooligosaccharides and in their synthesis by rice BGlu1 glycosynthase. *Prot. Sci.* 21, 362-372.
- Pengthaisong S, Chen C-F, Withers SG, Kuaprasert B, Ketudat Cairns JR. (2012b) Rice BGlu1 glycosynthase and wild type transglycosylation activities distinguished by cyclophellitol inhibition. *Carbohydr. Res.* 352, 51-59.
- Rye, C.S. and Withers, S.G. (2000) Glycosidase mechanisms. *Curr. Opin. Chem. Biol.* **4**, 573-580.

- Sanz-Aparacio J, Hermoso JA, Martinez-Ripoll M, Lequerica JL, Polaina J. (1998) Crystal structure of β -glucosidase A from *Bacillus polymyxa*: insights into the catalytic activity in family 1 glycosyl hydrolases. *J. Mol. Biol.* 275, 491-502.
- Sansanya S, Opassiri R, Kuaprasert B, Chen C-J, Ketudat Cairns JR. (2011) The crystal structure of rice (*Oryza sativa* L.) Os4BGlu12, an oligosaccharide and tuberonic acid glucoside-hydrolyzing β -glucosidase with significant thioglucosidase activity. *Arch. Biochem. Biophys.* 510, 62-72.
- Sansanya S, Maneesan J, Ketudat Cairns JR. (2012) Exchanging a single amino acid residue generates or weakens a +2 cellooligosaccharide binding subsite in rice β -glucosidases. *Carbohydr. Res.* 351, 130-133.
- Seshadri S, Akiyama T, Opassiri R, Kuaprasert B, and Ketudat Cairns J. (2009) Structural and enzymatic characterization of Os3BGlu6, a rice β -glucosidase hydrolyzing hydrophobic glycosides and (1 $\rightarrow$ 3)- and (1 $\rightarrow$ 2)-linked disaccharides *Plant Physiol.* 151, 47-58.
- Tankrathok A, Iglesias-Fernández J, Luang S, Robinson R, Kimura A, Rovira C, Hrmova M, Ketudat Cairns J. (In Press) Structural analysis and insights into glycon specificity of the rice GH1 Os7BGlu26 β -D-mannosidase. *Acta Crystallog. Sect. D.*
- Toonkool P, Metheenukul P, Sujiwattanasat P, Paiboon P, Tongtubtim N, Ketudat-Cairns M, Ketudat-Cairns J, Svasti J. (2006) Expression and purification of dalcinase, a β -glucosidase from *Dalbergia cochinchinensis* Pierre, in yeast and bacterial hosts. *Protein Expression and Purification* 48, 195-204.
- Vagin A, Teplyakov A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* D30, 1022-1025.
- Varghese JN, Hrmova M, Fincher GB. (1999) Three-dimensional structure of a barley β -D-glucan exohydrolase, a family 3 glycosyl hydrolase. *Structure Fold. Des.* 7, 179-190.
- Wakuta S, Hamada S, Ito H, Matsuura H, Nabeta K, Matsui H. (2010). Identification of a β -glucosidase hydrolyzing tuberonic acid glucoside in rice (*Oryza sativa* L.). *Phytochemistry* 71, 1280-1288.
- Wakuta S, Hamada S, Ito H, Imai R, Mori H, Matsuura H, Nabeta K, Matsui H. 2011. Comparison of enzymatic properties and gene expression profiles of two tuberonic acid glucoside β -glucosidases from *Oryza sativa* L. *J. Appl. Glycosci.* 58, 67-70.
- Wang J, Hou Q, Dong L, Liu Y, Liu C. (2011) QM/MM studies on the glycosylation mechanism of rice BGlu1 β -glucosidase. *J. Molecul. Graph. Model.* 30, 148-152.
- Wang J, Hou Q, Sheng X, Gao J, Liu Y, Liu C. (2013) Theoretical study on the deglycosylation mechanism of rice BGlu1 β -glucosidase. *Int. J. Quant. Chem.* 113, 1071-1075.
- Wang J, Pengthaisong S, Ketudat Cairns JR, Liu Y. (2013) X-ray crystallography and QM/MM investigation on the oligosaccharide synthesis mechanism of rice BGlu1 glycosynthases. *Biochimica et Biophysica Acta Proteins and Proteomics.* 1834 (2), 536-545.

10. Output from grant project

10.1. International Journal Publications

- 10.1.1. Sansenya S, Opassiri R, Kuaprasert B, Chen C-J, Ketudat Cairns JR. (2011) The crystal structure of rice (*Oryza sativa* L.) Os4BGlu12, an oligosaccharide and tuberonic acid glucoside-hydrolyzing beta-glucosidase with significant thioglucohydrolase activity. *Archives of Biochemistry and Biophysics* **510**, 62-72. doi:10.1016/j.abb.2011.04.005 (ISI IF 2.935, 2011)
- 10.1.2. Pengthaisong S, Withers SG, Kuaprasert B, Svasti J, Ketudat Cairns JR. (2012) The role of the oligosaccharide binding cleft of rice BGlu1 in hydrolysis of cellobiosaccharides and in their synthesis by rice BGlu1 glycosynthase. *Protein Science* **21**, 362-372. doi: 10.1002/pro.2021 (ISI IF 2.735, 2012 JCR)
- 10.1.3. Sansenya S, Maneesan J, Ketudat Cairns JR. (2012) Exchanging a single amino acid residue generates or weakens a +2 cellobiosaccharide binding subsite in rice beta-glucosidases. *Carbohydrate Research*. **351**, 130-133. doi: 10.1016/j.carres.2012.01.010 (ISI IF 2.044, 2012 JCR)
- 10.1.4. Pengthaisong S, Chen C-F, Withers SG, Kuaprasert B, Ketudat Cairns JR. (2012) Rice BGlu1 glycosynthase and wild type transglycosylation activities distinguished by cyclophellitol inhibition *Carbohydrate Research* **352**(1), 51-59. doi: 10.1016/j.carres.2012.02.013 (ISI IF 2.044, 2011 JCR)
- 10.1.5. Ketudat Cairns JR, Pengthaisong S, Luang S, Sansenya A, Tankrathok A, Svasti J. (2012) Protein-carbohydrate interactions leading to hydrolysis and transglycosylation in plant glycoside hydrolase family 1 enzymes. *Journal of Applied Glycoscience* **59**, 51-62. doi:10.5458/jag.jag.JAG-2011_022 (no ISI listing)
- 10.1.6. Wang J, Pengthaisong S, Ketudat Cairns JR, Liu Y. (2013) X-ray crystallography and QM/MM investigation on the oligosaccharide synthesis mechanism of rice BGlu1 glycosynthases. *Biochimica et Biophysica Acta Proteins and Proteomics*. **1834** (2), 536-545. doi: 10.1016/j.bbapap.2012.11.003 (ISI IF 3.733, 2012 JCR).
- 10.1.7. Himeno N, Saburi W, Takeda R, Wakuta S, Matsuura H, Nabeta K, Sansenya S, Ketudat Cairns JR, Imai R, Mori H, Matsuda H. (2013) Identification of rice β -glucosidase with high hydrolytic activity towards salicylic β -glucoside. *Bioscience, Biotechnology and Biochemistry* **77** (5), 934-939. (ISI IF 1.269, 2012 JCR)
- 10.1.8. Hua Y, Sansenya S, Saetang C, Wakuta S, Ketudat Cairns JR. (2013) Enzymatic and structural characterization of hydrolysis of gibberellin A4 glucosyl ester by a rice β -D-glucosidase. *Archives of Biochemistry and Biophysics* **537**(1) 39-48. doi: 10.1016/j.abb.2013.06.005 (ISI IF 3.37, 2012 JCR)
- 10.1.9. Tankrathok A, Iglesias-Fernández J, Luang S, Robinson R, Kimura A, Rovira C, Hrmova M, Ketudat Cairns J. (In Press) Structural analysis and insights into glycon specificity of the rice GH1 Os7BGlu26 β -D-mannosidase. *Acta Crystallographica Section D* (ISI IF 14.103, 2012 JCR)

Cumulative Impact Factor: 32.233

10.2. International Meeting Presentations

- 10.2.1. Tankrathok A, Ketudat Cairns JR. 2011. Molecular basis for substrate preference in plant glycoside hydrolase family 1 β -D-mannosidase/ β -D-glucosidase isoenzymes. Third Conference of the Asia Pacific Protein Association, Shanghai University, Shanghai, China, 6-9 May, 2011. Student Oral and Poster.
- 10.2.2. Tankrathok, A., Ketudat Cairns, J.R., Hakki, Z., Williams S.J., Prasert, P., Pengthaisong, S., Luang, S., Hrmova, M., Kuntothom, T., Robinson, R.C., Kimura, A. 2012. Structural basis of glycon selection in plant β -D-mannosidases/ β -D-glucosidases. International Carbohydrate Symposium, Madrid, Spain, 23-27 July, 2012. Oral
- 10.2.3. Ketudat Cairns, J.R., 2011. Protein-carbohydrate interactions leading to hydrolysis and transglycosylation in plant enzymes belonging to glycoside hydrolase family 1. Japanese Applied Glycoscience Society Meeting on Alpha-Amylases and Related Enzymes. Hokkaido University, Sapporo, Hokkaido, Japan; 28-30 Sept., 2011. Special Invited Foreign Lecture.
- 10.2.4. Tankrathok, A., Prasert, P., Pengthaisong, S., Ketudat Cairns, J.R. 2011. Molecular basis for substrate preference in plant glycoside hydrolase family 1 β -D-mannosidase/ β -D-glucosidase isoenzymes. The Third Asia Pacific Protein Association Conference in conjunction with the Third Symposium of the Chinese Protein Society. Shanghai University, Shanghai, China. 6-9 May, 2011. Student talk & poster.
- 10.2.5. Svasti, J. Ketudat Cairns, J.R., Kongsaree, P. 2011. Structure-function relationships in Thai plant beta-glucosidases. The Third Asia Pacific Protein Association Conference in conjunction with the Third Symposium of the Chinese Protein Society. Shanghai University, Shanghai, China. 6-9 May, 2011. Invited lecture.
- 10.2.6. Ketudat Cairns JR, Tankrathok A, Pengthaisong S, Luang S, Prasert P, Hakki Z, Williams S, Hrmova M. 2012. Structural studies of hydrolysis of glucosides, mannosides and oligosaccharides by rice family GH1 glycoside hydrolases. 13th Federation of Asian and Oceanic Biochemists and Molecular Biologists Congress. BITEC Conference Center, Bangkok, Thailand, 25-29 November, 2012. Invited Lecture B-06.
- 10.2.7. Tankrathok A, Ketudat Cairns JR. 2012. Three-dimensional structure of the rice Os7BGlu26 beta-D-mannosidase in complex with a transition state mimic. 13th Federation of Asian and Oceanic Biochemists and Molecular Biologists Congress. BITEC Conference Center, Bangkok, Thailand, 25-29 November, 2012. Poster P-B-15.
- 10.2.8. Pengthaisong S, Cairns J.R. 2012. Investigation of cellooligosaccharide binding residues for transglucosylation of rice BGlu1 glycosynthase. 13th Federation of Asian and Oceanic Biochemists and Molecular Biologists Congress. BITEC Conference Center, Bangkok, Thailand, 25-29 November, 2012. Poster P-B-25.
- 10.2.9. Tankrathok A, Ketudat Cairns JR. 2013. Structural analysis of the rice GH1 Os7BGlu26 β -D-mannosidase: insights into glycon specificity. International Conference on Biomolecular Forms and Functions. A Celebration of 50 Years of the Ramachandran Map. Indian Institute of Science, Bangalore, India 8-11 January, 2013. Poster.

10.3. National Meeting Presentations

- 10.3.1. Tankrathok A, Ketudat Cairns JR. 2011. Structural insights into rice Os7BGlu26 β -mannosidase. The 6th International Symposium of the Protein Society of Thailand, Chulabhorn Research Institute, Bangkok, Thailand. 31 August-2 September, 2011. Oral Presentation 7.
- 10.3.2. Luang S, Ketudat-Cairns JR, Streltsov VA, Hrmova M. 2011. Two tryptophan residues control broad substrate specificity of barley β -D-glucan glucohydrolase. The 6th International Symposium of the Protein Society of Thailand, Chulabhorn Research Institute, Bangkok, Thailand. 31 August-2 September, 2011. Poster Presentation 14.
- 10.3.3. Pengthaisong S, Withers SG, Kuaprasert B, Ketudat Cairns JR. 2011. Structural analysis of the basis for oligosaccharide synthesis by rice BGlu1 glycosynthases. The 6th International Symposium of the Protein Society of Thailand, Chulabhorn Research Institute, Bangkok, Thailand. 31 August-2 September, 2011. Poster Presentation 43.
- 10.3.4. Sansenya S, Maneesan J, Ketudat Cairns JR. 2011. The effect of a single residue on +2 subsite binding of oligosaccharides by rice Os3BGlu6, Os3BGlu7, and Os4BGlu12 β -glucosidases. The 6th International Symposium of the Protein Society of Thailand, Chulabhorn Research Institute, Bangkok, Thailand. 31 August-2 September, 2011. Poster Presentation 55.
- 10.3.5. Prawisut A, Ketudat Cairns JR. Expression, purification, and characterization of rice (*Oryza sativa*) β -D-glucan glucanase. The 6th International Symposium of the Protein Society of Thailand, Chulabhorn Research Institute, Bangkok, Thailand. 31 August-2 September, 2011. Poster Presentation 60.
- 10.3.6. Pengthaisong, S., Wang J, Liu Y, Ketudat Cairns JR. 2012. X-ray crystallography and QM/MM investigation of the relative activities of rice BGlu1 glycosynthases. The 7th International Symposium of the Protein Society of Thailand. Chulabhorn Research Institute, Bangkok, Thailand. 29-31 August, 2012. Oral presentation 6.
- 10.3.7. Sansenya S, Wakuta S, Matsui H, Ketudat-Cairns JR. 2011. The structural basis of tuberonic acid glucoside binding by rice Os4BGlu12 β -glucosidase. The 7th International Symposium of the Protein Society of Thailand. Chulabhorn Research Institute, Bangkok, Thailand. 29-31 August, 2012. Poster Presentation 6.
- 10.3.8. Tankrathok A, Ketudat Cairns JR. 2012. Structural insights into glycon specificity of rice Os7BGlu26 β -mannosidase. The 7th International Symposium of the Protein Society of Thailand. Chulabhorn Research Institute, Bangkok, Thailand. 29-31 August, 2012. Poster Presentation 7.
- 10.3.9. Prawisut A, Ketudat Cairns JR. 2012. Expression, purification and characterization of rice (*Oryza sativa*) β -D-glucan exoglucanases. The 7th International Symposium of the Protein Society of Thailand. Chulabhorn Research Institute, Bangkok, Thailand. 29-31 August, 2012. Poster Presentation 30.
- 10.3.10. Prasert P, Ketudat Cairns JR. 2012. Crystals of BGlu1 E176Q with 4-nitrophenyl- β -D-mannoside and 2,4-dinitrophenyl-2-deoxy-2-fluoro-mannoside. The 7th International

Symposium of the Protein Society of Thailand. Chulabhorn Research Institute, Bangkok, Thailand. 29-31 August, 2012. Proceeding 3.

10.3.11. Pengthaisong S, Ketudat Cairns JR. 2013. Structural analysis of celooligosaccharide binding residues of rice BGlu1. The 8th International Symposium of the Protein Society of Thailand, Chulabhorn Research Institute, Bangkok, Thailand. 5-7 August, 2013. Poster Presentation 25.

10.3.12. Tankrathok A, Pengthaisong S, Ketudat Cairns JR. 2013. Structural, kinetic and thermodynamic studies of the rice BGlu1 beta-glucosidase in complex with transition state mimics. The 8th International Symposium of the Protein Society of Thailand, Chulabhorn Research Institute, Bangkok, Thailand. 5-7 August, 2013. Poster Presentation 31.

10.3.13. Prawisut A, Ketudat Cairns JR. 2013. Expression, purification, and characterization of rice (*Oryza sativa*) β -D-glucan exoglucanases. The 8th International Symposium of the Protein Society of Thailand, Chulabhorn Research Institute, Bangkok, Thailand. 5-7 August, 2013. Poster Presentation 42.

10.3.14. Sansenya S, Hua Y, Ketudat Cairns, JR. 2013. The covalent intermediate structure of rice Os3BGlu6 gibberellin glucosyl ester β -glucosidase E178Q with D-glucose. The 8th International Symposium of the Protein Society of Thailand, Chulabhorn Research Institute, Bangkok, Thailand. 5-7 August, 2013. Poster Presentation 47.

10.4. Graduate Students' Training

10.4.1. Ms. Salila Pengthaisong, Ph.D. 2012, Suranaree University of Technology, School of Biochemistry.

10.4.2. Mr. Sompong Sansenya, Ph.D. 2013, Suranaree University of Technology, School of Biochemistry.

10.4.3. Ms. Por-ngam Prasert, M.Sc. 2013, Suranaree University of Technology, School of Biochemistry.

10.4.4. Mr. Anupong Tankrathok, Studying for Ph.D. 2008-present, Suranaree University of Technology, School of Biochemistry.

10.4.5. Mr. Akkarawit Prawisut, Studying for Ph.D. 2008-present, Suranaree University of Technology, School of Biochemistry.

10.5. Journal covers

10.5.1. Cover of *Protein Science* for March, 2012.

10.6. Invited Lectures/Travel Awards

10.6.1. Dr. Ayako Suzuki and Dr. Yasuhito Takeda Travel Grant to present a Special Overseas Invited Lecture at the Japanese Applied Glycoscience Society's 16th Conference on α -Amylases and Related Enzymes in Hokkaido, Japan, September, 2011.

10.6.2. Invited Lecture at the Federation of Asian and Oceanic Biochemists and Molecular Biologists (FAOBMB) Symposium, Bangkok, Thailand, November, 2012.

10.7. International Protein Data Bank (PDB) Entries

- 10.7.1. 3PTK Apo Os4BGlu12 β -glucosidase structure.
- 10.7.2. 3PTM Os4BGlu12 β -glucosidase covalent complex with G2F
- 10.7.3. 3PTQ Os4BGlu12 β -glucosidase structure with DNP2FG
- 10.7.4. 3SCN BGlu1 (Os3BGlu7) E386G glycosynthase apo structure.
- 10.7.5. 3SCO BGlu1 (Os3BGlu7) E386G α -fluoroglucoside complex structure
- 10.7.6. 3SCP BGlu1 (Os3BGlu7) E386A glycosynthase apo structure.
- 10.7.7. 3SCQ BGlu1 (Os3BGlu7) E386A α -fluoroglucoside complex structure
- 10.7.8. 3SCR BGlu1 (Os3BGlu7) E386S glycosynthase apo structure.
- 10.7.9. 3SCS BGlu1 (Os3BGlu7) E386S α -fluoroglucoside complex structure
- 10.7.10. 3SCT BGlu1 (Os3BGlu7) E386G cellotetraose complex structure
- 10.7.11. 3SCU BGlu1 (Os3BGlu7) E386G cellopentaose complex structure
- 10.7.12. 3SCV BGlu1 (Os3BGlu7) E386G/S334A cellotetraose complex structure
- 10.7.13. 3SCW BGlu1 (Os3BGlu7) E386G/Y341A cellotetraose complex structure
- 10.7.14. 3WBA Os3BGlu6 E178Q 4NPGlc complex structure
- 10.7.15. 3WBE Os3BGlu6 E178Q GA₄Glc complex structure
- 10.7.16. 4JHO Os7BGlu26 apo structure
- 10.7.17. 4JIE Os7BGlu26 mannose complex structure

10.8. Collaborations

- 10.8.1. Prof. Dr. M.R. Jisnuson Svasti, Assoc. Prof. Dr. Jirundon Yuvaniyama, Assoc. Prof. Dr. Palangpol Kongsaree. Center for Protein Structure and Function, Faculty of Science, Mahidol University, and Laboratory of Biochemistry, Chulabhorn Research Institute, Bangkok, Thailand.
- 10.8.2. Dr. Prasart Kittakoop, Laboratory of Natural Products, Chulabhorn Research Institute, Bangkok, Thailand.
- 10.8.3. Dr. Buabarn Kuaprasert, Synchrotron Light Research Institute, Thailand.
- 10.8.4. Prof. Dr. Maria Hrmova, Australian Centre for Plant Functional Genomics, University of Adelaide, Waite Campus, Glen Osmond, SA, Australia.
- 10.8.5. Prof. Dr. Spencer Williams, University of Melbourne, School of Chemistry and Bio21 Molecular Science and Biotechnology Institute, Parkville, Victoria, Australia.
- 10.8.6. Prof. Dr. Stephen G. Withers, Department of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada.

- 10.8.7. Prof. Dr. Robert C. Robinson, Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, Singapore.
- 10.8.8. Prof. Dr. Atsuo Kimura, Graduate Faculty of Agriculture, Hokkaido University, Sapporo, Hokkaido, Japan.
- 10.8.9. Prof. Dr. Carme Rovira, Computer Simulation and Modeling Laboratory and Institut de Química Teòrica i Computacional (IQTUB), Parc Científic de Barcelona, and Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys, Barcelona, Spain.
- 10.8.10. Prof. Dr. Chun-Jung Chen, National Synchrotron Radiation Research Center, Hsinchu, Taiwan.
- 10.8.11. Prof. Dr. Hirokazu Matsui, Hokkaido University Research Faculty of Agriculture, Sapporo, Hokkaido, Japan.
- 10.8.12. Prof. Dr. Yongjun Liu Key Lab of Theoretical and Computational Chemistry in University of Shandong, School of Chemistry and Chemical Engineering, Shandong University, Jinan, Shandong 250100, China.



Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

The crystal structure of rice (*Oryza sativa* L.) Os4BGlu12, an oligosaccharide and tuberonic acid glucoside-hydrolyzing β -glucosidase with significant thioglucosidase activity

Sompong Sansenya^a, Rodjana Opassiri^a, Buabarn Kuaprasert^b, Chun-Jung Chen^c,
James R. Ketudat Cairns^{a,*}

^a Institute of Science, Schools of Biochemistry and Chemistry, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand

^b Synchrotron Light Research Institute (Public Organization), Nakhon Ratchasima 30000, Thailand

^c Life Science Group, Scientific Research Division, National Synchrotron Radiation Research Center, Hsinchu 30076, Taiwan

ARTICLE INFO

Article history:

Received 27 February 2011
and in revised form 8 April 2011
Available online 16 April 2011

Keywords:

Rice
Glycoside hydrolase
X-ray crystallography
Substrate specificity
Covalent intermediate
Michaelis complex

ABSTRACT

Rice Os4BGlu12, a glycoside hydrolase family 1 (GH1) β -glucosidase, hydrolyzes β -(1,4)-linked oligosaccharides of 3–6 glucosyl residues and the β -(1,3)-linked disaccharide laminaribiose, as well as certain glycosides. The crystal structures of apo Os4BGlu12, and its complexes with 2,4-dinitrophenyl-2-deoxyl-2-fluoroglucoside (DNP2FG) and 2-deoxy-2-fluoroglucose (G2F) were solved at 2.50, 2.45 and 2.40 Å resolution, respectively. The overall structure of rice Os4BGlu12 is typical of GH1 enzymes, but it contains an extra disulfide bridge in the loop B region. The glucose ring of the G2F in the covalent intermediate was found in a ⁴C₁ chair conformation, while that of the noncovalently bound DNP2FG had a ¹S₃ skew boat, consistent with hydrolysis via a ⁴H₃ half-chair transition state. The position of the catalytic nucleophile (Glu393) in the G2F structure was more similar to that of the *Sinapsis alba* myrosinase G2F complex than to that in covalent intermediates of other *O*-glucosidases, such as rice Os3BGlu6 and Os3BGlu7 β -glucosidases. This correlated with a significant thioglucosidase activity for Os4BGlu12, although with 200- to 1200-fold lower $k_{\text{cat}}/K_{\text{m}}$ values for *S*-glucosides than the comparable *O*-glucosides, while hydrolysis of *S*-glucosides was undetectable for Os3BGlu6 and Os3BGlu7.

© 2011 Elsevier Inc. All rights reserved.

Introduction

β -Glucosidases (3.2.1.21) catalyze the hydrolysis of glycosidic bonds between a nonreducing terminal β -D-glucosyl residue and an aglycon moiety, which may be an oligosaccharide or aryl or alkyl alcohol. According to the amino acid-sequence-based classification of Henrissat [1], β -glucosidases have been found in glycoside hydrolase (GH)¹ families GH1, GH3, GH5, GH9, GH30 and GH116 [2,3]. Of these, the greatest diversity of specificities has been reported in GH1, due in part to the large number of GH1 isoenzymes with different specificities found in plants [4,5]. In addition to β -D-glucosidases, GH1 also includes enzymes with β -D-phosphogalactosidase, β -D-phosphoglucosidase, β -D-thioglucosidase, β -D-man-

nosidase, β -D-fucosidase, β -D-galactosidase, 6-*O*-linked disaccharidase and hydroxyisourate hydrolase activities. Thus, GH1 serves as an excellent model system for investigating substrate specificity and the protein-carbohydrate and protein-aglycon interactions involved in determining this specificity.

The GH1 enzymes, along with GH5 and GH30, belong to the GH Clan A, all members of which have catalytic domains consisting of a (β/α)₈ or triose phosphate isomerase (TIM) barrel, which typically have the catalytic acid/base residing at the end of β -strand 4 and the nucleophile at the end of β -strand 7 [6,7]. These residues are found near the bottom of a cleft or crater-shaped active site at the carboxy-terminal end of the β -strands of the central barrel [2,3]. These enzymes act via a retaining double-displacement mechanism involving two carboxylate residues, as shown in Fig. 1 [8,9]. In the glycosylation step (the first reaction in Fig. 1), the catalytic acid protonates the aglycon leaving group as the nucleophile attacks the anomeric carbon to displace the aglycon and form a covalent intermediate. In the deglycosylation step (the second reaction in Fig. 1), an incoming water molecule is deprotonated by the catalytic acid/base and attacks the anomeric carbon to displace the enzyme from the glucose. Both steps are thought to involve the formation of oxocarbenium cation-like transition states. In plant thioglucosidases, such as the white mustard

* Corresponding author. Fax: +66 44 224185.

E-mail address: cairns@sut.ac.th (J.R. Ketudat Cairns).

¹ Abbreviations used: DNP2FG, dinitrophenyl 2-deoxy-2-fluoro- β -D-glucopyranoside; G2F, 2-fluoroglucopyranoside; GH, glycoside hydrolase; GH1, glycoside hydrolase family 1; GOL, glycerol; IMAC, immobilized metal affinity chromatography; Os3BGlu6, *Oryza sativa* β -glucosidase isoenzyme Os3BGlu6; Os3BGlu7, *O. sativa* β -glucosidase isoenzyme Os3BGlu7; Os4BGlu12, *O. sativa* β -glucosidase isoenzyme Os4BGlu12; pNP, para-nitrophenol; pNPGlc, pNP β -D-glucopyranoside; SbDhr1, *Sorghum bicolor* dhurrinase isoenzyme 1; ZmGlu1, *Zea mays* β -glucosidase isoenzyme Glu1.

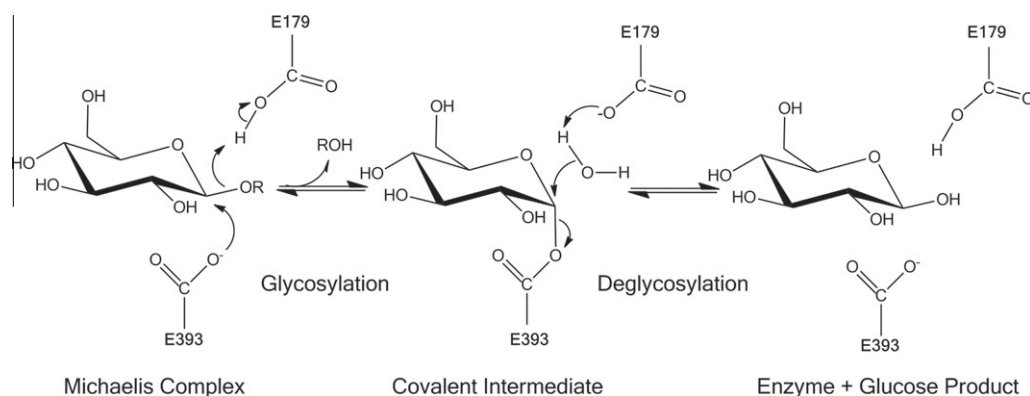


Fig. 1. Mechanism of retaining β -glycosidases, as seen in rice Os4BGlu12. Upon formation of the Michaelis complex between enzyme and substrate, the reaction proceeds via two steps, glycosylation to form the covalent intermediate, and deglycosylation to release the sugar from the protein. In Os4BGlu12, Glu179 serves as the catalytic acid/base, while Glu393 serves as the catalytic nucleophile.

(*Sinapis alba*) myrosinase, the catalytic acid/base is missing, since acid assistance is not necessary for cleavage of the *S*-glycosidic linkage and the base function is provided by ascorbate [10,11]. However, aphid myrosinase retains the acid/base in the conserved location [12].

Structural studies of the broad specificity bacterial *Sulfolobus solfataricus* β -glycosidase *S* and archaeal *Thermatoga maritima* β -glucosidase BglA in complexes with substrates, transition state analogues, and covalent intermediates have supported a conformational itinerary for the glucose that goes from a 1S_3 skew boat Michaelis complex to a 4H_3 half chair transition state to a relaxed 4C_1 chair in the covalent intermediate [13–15]. The observations of the 4C_1 chair shape of the covalent intermediate in white mustard myrosinase [10], plant β -glucosidases [16,17] and human cytoplasmic β -glucosidase [18], the 1S_3 skew boat in complexes of the sorghum (*Sorghum bicolor*) dhurrinase 1 mutant Glu189Asp with dhurrin and rice Os3BGlu6 with octyl-thiogluconide, and the 4H_3 half chair in glucotetraose in white mustard myrosinase [11] and maize (*Zea mays*) Glu1 β -glucosidase (ZmGlu1) [19] suggest this itinerary holds for eukaryotic enzymes as well. This is consistent with the fact that the glycon-binding residues are conserved in GH1 enzymes from all domains of living organisms.

Plant GH1 enzymes play a wide range of functions, including roles in defense, phytohormone activation, lignification, degradation of cell wall oligosaccharides, secondary metabolism and release of volatiles [3]. As such, regulation of many of the functions mentioned above dictates that these plant β -glucosidases should have at least some substrate selectivity. This is in contrast to many microbial and animal GH1 hydrolases, which tend to have rather broad aglycon specificities, in line with their release of glucose from nutritional sources. Thus, the structures of plant GH1 enzymes allow dissection of the interactions leading to aglycon specificity. Structures are available for eight plant GH1 enzymes with different specificities [9,10,16,17,19–23], the phylogenetic distribution of which is shown in Fig. 2. Aglycon specificity has been most thoroughly compared between ZmGlu1 and SbDhr1, since they share 70% amino-acid sequence identity and ZmGlu1 hydrolyzes a rather broad array of substrates, but not dhurrin, while SbDhr1 hydrolyzes dhurrin nearly exclusively [24]. Structural studies have shown that ZmGlu1 sandwiches its aglycon between the conserved Trp378 on one side and the nonconserved Phe198, Phe205, and Phe466 on the other side, while SbDhr1 interacts through the hydrophobic and aromatic stacking interactions with the corresponding Trp376, Val196, and Leu203, but also forms indirect hydrogen bonds between the phenolic hydroxyl of dhurrin and Ser462 and Asn259 [19]. Surprisingly, the aglycon-binding residues in the wheat and rye hydroxamic acid β -glucosidases,

which hydrolyze the same natural substrates as ZmGlu1, are not conserved with that of ZmGlu1, except for the conserved tryptophan [21,22]. These studies suggest that aglycon-binding is complex and the active site shape may play at least as critical a role as the surface residues.

Of the 34 GH1 genes in the rice genome that are likely to form active glycosidases [5], Os4BGlu12 showed the highest similarity to the cell wall associated β -glucosidase that had been previously characterized and shown to hydrolyze β -(1,3)- and β -(1,4)-linked gluco-oligosaccharides [25]. Os4BGlu12 expressed in recombinant *Escherichia coli* similarly hydrolyzed gluco-oligosaccharides, as well as certain glucosides with large rather apolar aglycons, such as apigenin 7-O- β -glucoside and deoxycorticosterone 21- β -glucoside [5,26]. It was also shown to be wounding induced and hydrolyze the β -(1,3),(1,4)-mixed linkage oligosaccharides released by a wounding induced β -glucanase. In contrast to rice Os3BGlu7 (also called rice BGlu 1) and its close relatives Os3BGlu8 and Os7BGlu26, which have clear increases in catalytic efficiencies (k_{cat}/K_m) as the lengths of β -1,4-linked gluco-oligosaccharides are increased from cellotriose to cellobiose [27,28], Os4BGlu12 shows only a slight increase in going from cellotriose to cellopentaose, and drops at cellobiose [26]. On the other hand, Os3BGlu6, which hydrolyzes β -(1,2)- and β -(1,3)-linked disaccharides and alkyl glycosides, barely hydrolyzes cello-oligosaccharides, with highest activity for cellobiose [17]. Recently, Os4BGlu13, the rice isoenzyme that is most closely related to Os4BGlu12, was identified as a tuberonic acid glucoside β -glucosidase (OsTAGG1) [29], and Os4BGlu12 was also found to act on this phytohormone [30].

In order to determine the basis for its unique behavior in binding and hydrolyzing oligosaccharides and apolar glycosides, we have determined the structures of Os4BGlu12 β -glucosidase and its noncovalent complex with 2,4-dinitrophenyl 2-deoxy-2-fluoroglucoside (DNP2FG) and covalent complex with 2-deoxy-2-fluoroglucoside (G2F).

Materials and methods

Expression, purification and crystallization of Os4BGlu12

Rice Os4BGlu12 β -glucosidase was expressed in *E. coli* strain Origami B(DE3) and purified by immobilized metal affinity chromatography (IMAC), cleavage of the N-terminal thioredoxin and His₆ tags with enterokinase, and removal of the fusion tag by IMAC, as described previously [31]. Crystals of Os4BGlu12 and its complexes were grown at 288 K by hanging drop vapor diffusion with a stock of 3.5 mg ml⁻¹ purified Os4BGlu12. The crystallization of

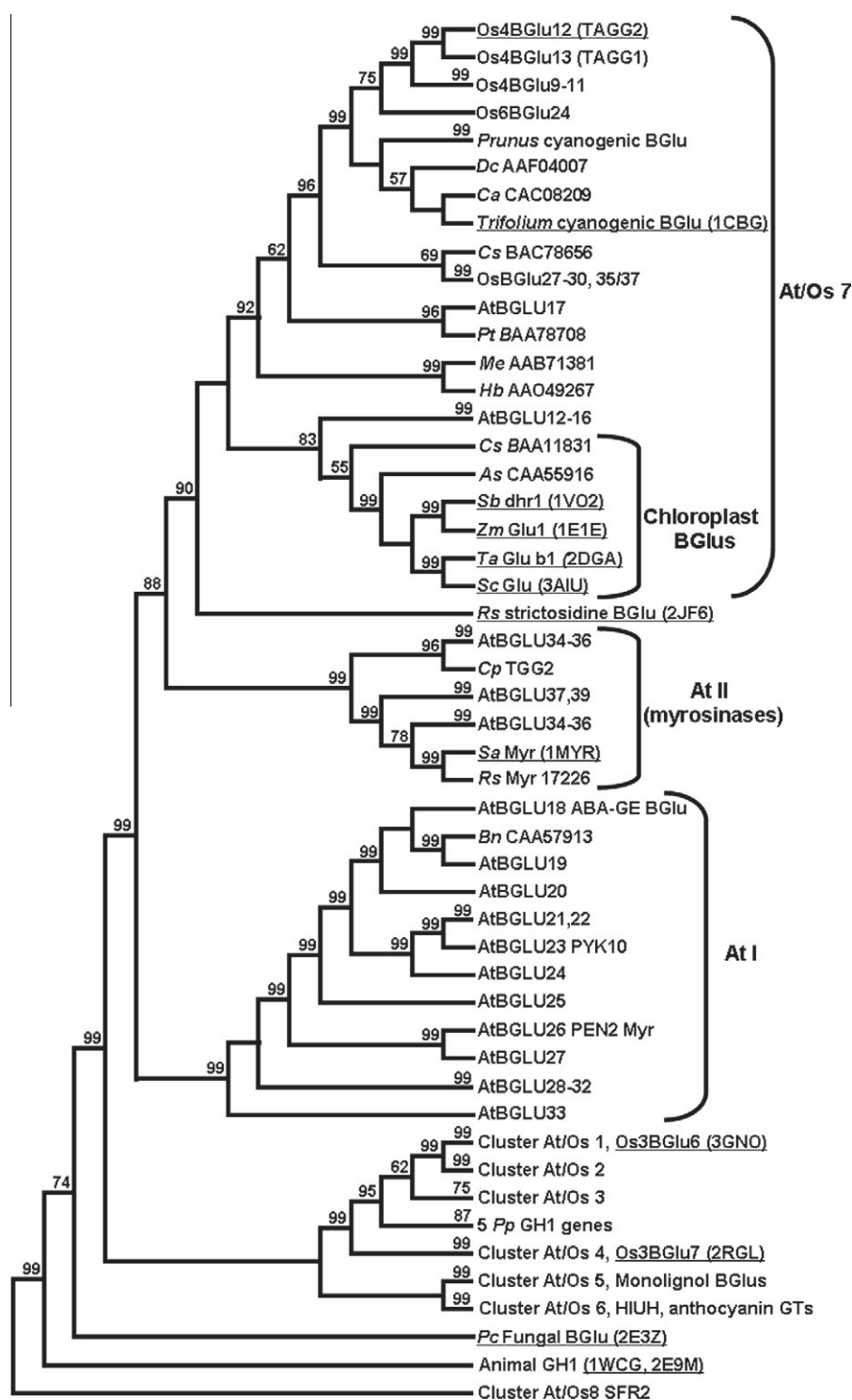


Fig. 2. Sequence-based phylogenetic tree of Plant GH1 Enzymes. The Minimum Evolution tree was computed with MEGA5 [53]. The optimal tree with the sum of branch length = 22.74 is shown. The confidence probability (multiplied by 100) that the interior branch length is greater than 0, as estimated using the bootstrap test (1000 replicates) is shown next to the branches for the branches with >50% support. The evolutionary distances were computed using the *p*-distance method [54] and are in the units of the number of amino acid differences per site. The ME tree was searched using the Close-Neighbor-Interchange (CNI) algorithm at a search level of 0. The initial tree was generated by the Neighbor-joining method. The analysis involved 131 amino acid sequences, aligned by the MUSCLE algorithm [55]. All ambiguous positions were removed for each sequence pair. There were a total of 907 positions in the final dataset. The SFR2 (Sensitive-to-Freezing 2) cluster (homologues from Arabidopsis, rice and the moss *Physcomitrella patens* (Pp)) serves to root the rest of the tree. Arabidopsis and rice sequences are named according to [4] and [5], respectively, papaya CPTGG2 myrosinase is named according to Ref. [52], proteins with known 3D structures are marked by representative PDB codes and underlined, and other sequences are given as their genus and species letters and Genbank accession numbers. The eight phylogenetic clusters containing both Arabidopsis and rice sequences (At/Os) and two crucifer-specific clusters, At I and At II (classic myrosinases) are designated according to [5].

apo enzyme and its complex with DNP2FG was reported previously [31]. For the Os4BGlu12 complex with DNP2FG, Os4BGlu12 was

crystallized in the presence of 2 mM DNP2FG (Sigma, St. Louis, MO, USA), and the crystal was soaked in the 23% (w/v) PEG 3350,

0.12 M Tris–HCl, pH 8.5, 0.27 M NaCl and 28% (v/v) glycerol cryoprotectant containing 10 mM DNP2FG before vitrification. Because this protocol resulted in the noncovalent complex, the co-crystal with G2F was soaked in the same cryoprotectant without the inhibitor for 15 s before flash vitrification in liquid nitrogen and X-ray diffraction in order to obtain the covalent glucosyl-enzyme intermediate.

Data collection

The crystals of Os4BGlu12 apo enzyme and the complex with DNP2FG diffracted X-rays to 2.50 and 2.45 Å, respectively, and belong to space group $P4_32_12$ with cell parameters and statistics for the two X-ray data sets, as described previously [28] and shown in Table 1. Diffraction data for the Os4BGlu12 complex with G2F was collected on an ADSC Quantum 315 CCD detector with 1.00 Å X-rays from beamline BL13B1 at the National Synchrotron Radiation Research Center (NSRRC), Taiwan. The data was indexed, integrated and scaled with HKL-2000 [32]. Data collection statistics are shown in Table 1.

Structure solution and refinement

The crystal structure of the Os4BGlu12 apo enzyme was solved by molecular replacement with the *MOLREP* program [33] with the structure of cyanogenic β -glucosidase from white clover [9] (PDB code 1CBG, residues 14–490; 63% identical to Os4BGlu12) as the search model. The crystals for the structures of Os4BGlu12 with DNP2FG and G2F were isomorphous with that for the apo Os4BGlu12 structure, allowing the structures to be solved by rigid body refinement with the native structure as a search model in the Refmac 5.0 program [34]. The three structures were also refined with

Refmac 5.0. The analysis of the electron density and model building were done in the Coot program [35]. The occupancies of the DNP2FG molecules binding in each of the active sites were refined to 0.8, by setting the occupancy at different values and refining the structures to determine which value gave the minimum values of the temperature factors (*B*-factors) of the ligand and R_{free} . The stereochemistry of the final model was checked with PROCHECK [36] and MolProbity [37]. The refinement and model statistics are given in Table 1. The structural superimposition and figures were produced with PyMol [38].

Kinetic analysis

p-Nitrophenyl β -D-glucopyranoside (pNPGlc), *n*-octyl- β -D-glucopyranoside, and *n*-heptyl- β -D-glucopyranoside were purchased from Sigma. *p*-Nitrophenyl- β -D-thioglucoside was provided by Ms. Yanling Hua (Center for Scientific and Technological Equipment, Suranaree University of Technology). All kinetic parameters were calculated from triplicate assays done at 30 °C, in 50 mM sodium acetate, pH 5.0. Activities on *p*-nitrophenol (pNP) glycosides were determined by measuring the released *p*-nitrophenolate, and activities on *n*-octyl- β -D-glucoside, *n*-heptyl- β -D-glucoside, and *n*-octyl- β -D-thioglucopyranoside were determined by measuring glucose released by the glucose oxidase/peroxidase-coupled assay, as described previously [39]. The initial velocities were determined at time points and enzyme concentrations (0.1 μ g per assay for *O*-glycosides and 1.0 μ g for *S*-glycosides) where the reaction rate was linear and the absorbance value was in the range of 0.1 to 1.0, and used to calculate the kinetic constants. Substrate concentrations over a range of approximately one-third to three times the apparent K_m were included in calculations of kinetic parameters. Kinetic parameters were calculated by nonlinear regression of

Table 1

Refinement statistics for the structures of Os4BGlu12 apoenzyme and its complexes with G2F and DNP2FG.

Dataset	Os4BGlu12	Os4BGlu12_G2F	Os4BGlu12_DNP2FG
PDB code	3PTK	3PTM	3PTQ
Beamline	BL13B1	BL13B1	BL13B1
Wavelength (Å)	1.00	1.00	1.00
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$
Unit-cell parameters (Å)	$a = b = 112.7$ $c = 182.8$	$a = b = 114.0$ $c = 184.1$	$a = b = 114.1$ $c = 184.5$
Resolution range (Å)	30–2.49	30–2.40	30–2.45
Resolution outer shell (Å)	2.56–2.49	2.49–2.40	2.51–2.45
No. unique reflections	39533	48334	43131
No. observed reflections	408534	329557	381316
Completeness (%)	100.0 (99.9)	99.6 (99.9)	100.0 (99.9)
Average redundancy per shell	9.7 (10.0)	6.8 (6.2)	8.4 (8.6)
$I/\sigma(I)$	19.6 (6.4)	17.5 (3.4)	21.7 (4.4)
$R_{\text{merge}}(\%)$	10.6 (41.7)	10.4 (53.7)	9.4 (49.9)
$R_{\text{factor}}(\%)$	20.3	19.6	20.9
$R_{\text{free}}(\%)$	24.8	23.4	25.1
No. of residues in protein	486 (2 molecules)	486 (2 molecules)	486 (2 molecules)
No. protein atoms	7720 (2 molecules)	7720 (2 molecules)	7720 (2 molecules)
No. ligand atoms	16 (Tris)	22 (G2F)	48 (DNP2FG)
No. other hetero atoms	1 (Zn)	37 (GOL, Zn)	31 (GOL, Zn)
No. waters	228	511	280
Mean <i>B</i> -factor			
Protein	32.47	28.31	33.18
Ligand	52.70 (Tris)	21.47 (G2F)	62.69 (DNP2FG)
Other hetero atoms	24.23 (Zn)	51.20 (GOL), 22.40 (Zn)	45.04 (GOL), 22.75 (Zn)
Waters	32.18	30.40	31.60
r.m.s. bond deviations (length)	0.009	0.006	0.010
r.m.s. angle deviations (degrees)	1.19	1.02	1.26
Ramachandran plot			
Residues in most favorable regions (%)	88.4	89.5	89.4
Residues in additional allowed regions (%)	10.5	9.6	9.4
Residues in generously allowed regions (%)	1.1	0.9	1.2
Residues in disallowed regions (%)	0	0	0

Numbers in parentheses are statistics for the outer resolution shell. Ramachandran values were determined from PROCHECK [35]. For the ligands, DNP2FG stands for dinitrophenyl 2-deoxy-2-fluoro- β -D-glucopyranoside, G2F for 2-fluoroglucopyranoside and GOL for glycerol.

the Michaelis–Menten curves with Grafit 5.0 (Erithacus Software, Horley, Surrey, UK).

Competitive K_i values for Tris–HCl were determined by incubating 0.5 μ g enzyme with 10 different concentrations of the Tris (0–200 mM) and 0.1, 0.3, 0.5, 1.0, 3.0 and 5.0 mM pNPGlc as the substrate, under the reaction conditions described above. A series of Lineweaver–Burk plots were drawn to verify the inhibition was competitive, then Dixon plots were used to calculate the inhibition constants [40].

Temperature stability assays

For comparison of thermostability with Os4BGlu12, free Os3BGlu7 was expressed from the pET32 plasmid and expressed in Origami(DE3) cells, and purified with removal of the fusion tag, as previously described [39,16]. In 50 μ L reactions, 0.5 μ g of Os4BGlu12 or Os3BGlu7 enzyme was incubated with or without 1 mM dithiothreitol (DTT) in 50 mM sodium acetate, pH 5.0, for 60 min at temperatures ranging from 0 to 60 °C. The enzyme was then cooled on ice and then assayed at 30 °C with 0.3 mM pNPGlc as substrate, as described above, for 10 min.

Structure accessions

The coordinate and structure factor files for the structures determined in this work were deposited in the Protein Data Bank (PDB) as accessions 3PTK (apo Os4BGlu12), 3PTM (complex with G2F), and 3PTQ (complex with DNP2FG).

Results and discussion

Model quality and overall structure

The crystal structures of apo Os4BGlu12, and Os4BGlu12 complexes with DNP2FG and G2F were solved by molecular replacement and refined at 2.50, 2.45 and 2.40 Å resolution, respectively. The crystals belonged to space group $P4_32_12$, with the unit-cell parameters $a = b = 112.7$ Å, $c = 182.8$ Å for the native structure, and with slightly larger unit cells for the isomorphous DNP2FG and G2F complexes (Table 1). As seen in Fig. 3, two protein molecules were found in the asymmetric unit with a Matthews coefficient (V_m) of $2.46 \text{ Å}^3 \text{ Da}^{-1}$ and solvent content of 50% for apo Os4BGlu12, while the crystals of Os4BGlu12 with DNP2FG and G2F had solvent contents of 54–55%, due to their slightly larger unit cell volumes. The refinement statistics are shown in Table 1.

In general, the electron density of almost all residues was seen clearly in the final models of all structures. The electron density for residues 9–486 were observed in the apo enzyme and the DNP2FG and G2F complex structures, but eight residues from Os4BGlu12, including Ala1, Tyr2, Asn3, Ser4, Ala5, Gly6, Glu7 and Pro8, and 19 amino acid residues remaining from the N-terminal fusion tag (Ala-Met-Ala-Asp-Ile-Thr-Ser-Leu-Tyr-Lys-Lys-Ala-Gly-Ser-Ala-Ala-Ala-Pro-Phe) were not visible. At the C-terminus, the side chains of Lys486 in all three structures were poorly defined due to poor electron densities. Residues Ser333, Asn334, Gly335, Leu336 and Asn337 in the loop C region had weak density, especially for molecule B, and the surface residues Arg76, Glu108, Glu228, Glu306 and Ile355 did not have defined electron densities for their side chains in all three structures. In fact, the poorly defined part of loop C sits at the entrance to the active site and was found to be cut between Asn334 and Gly335 in the protein purified from rice plants, the same position as proteolysis of OsTAGG1 [29,30], although no evidence of cleavage of this site was seen in the crystallized recombinant protein. Part of loop C was also found in two conformations in Os3BGlu7 β -glucosidase, and binding to

cellooligosaccharide substrates was found to stabilize one of these conformations [41]. So, it is possible that the unstructured nature of this loop reflects the lack of modification and that the proteolytic modification of this loop plays a role in substrate recognition.

The overall structure of Os4BGlu12 portrayed in Fig. 3 shows 478 residues for each monomer, starting at residue 9 after the predicted mature protein N-terminus and ending with residue 486 as the last well defined residue. The typical GH1 (β/α)₈-barrel structure shown in Fig. 3 placed the putative catalytic acid/base, Glu179, and nucleophile, Glu393, at the C-terminal ends of β -strands 4 and 7, respectively, on opposite sides of the glucose ring at the bottom of the active site, as expected for a GH A enzyme [6,7]. The substrate binding pocket of Os4BGlu12 is shaped primarily by the four extended variable loops connecting strands and helices at the C-terminal side of the barrel in Fig. 1: loop A (Ser25–Asp66, between β 1 and α 1), loop B (Glu179–Arg213, between β 4 and α 4), loop C (Trp321–Pro370, between β 6 and α 6) and loop D (Asn394–Asp412, between β 7 and α 7). The structures of monomers A and B in the asymmetric unit are nearly identical, with an RMSD of 0.06 Å over all atoms.

For structural comparison, cyanogenic β -glucosidase from *Trifolium repens* (1CBG) [9], rice Os3BGlu7 β -glucosidase (2RGM) and Os3BGlu6 (3GNO) [16,17] and myrosinase from *Sinapis alba* (1MYR) [10] were superimposed on the Os4BGlu12 structure, as shown in Supplementary Fig. 1. The 1CBG, 2RGM and 3GNO structures represent plant O-glucosidases, and 1MYR is a plant S-glycosidase. Os4BGlu12 showed the highest amino-acid sequence identity with 1CBG (63%), followed by 3GNO (53%), 2RGM (51%) and 1MYR (43%). The C_α atoms of molecule A from the 1CBG, 3GNO, 2RGL and 1MYR structures superimposed to rice Os4BGlu12 molecule A with RMSD of 0.61, 0.71, 0.77 and 0.78 Å for 444, 431, 428 and 413 C_α atoms, respectively. The main-chain trace of loop C of Os4BGlu12 differs significantly from that of other GH1 structures. This agrees with previous reports that this loop is less conserved [10]. Loop C and its neighboring loops are mainly responsible for formation of the aglycon-binding site and thus contribute to the substrate specificity among GH1 enzymes [10,12,16,21,42].

Enzyme stability

As shown in Fig. 3, the monomer structure of Os4BGlu12 contains two disulfide bridges between Cys198–Cys206, located in loop B, and Cys184–Cys219, connecting the α -helix at the beginning of loop B and α -helix 4 of the (β/α)₈ barrel, which follows loop B. The disulfide bond between Cys198–Cys206 is conserved among the known structures of plant plastid and secretory pathway-targeted GH1 enzymes [9,10,16,17,19,21,42]. Most of these GH1 proteins have only this one disulfide bridge, but *S. alba* myrosinase (1MYR) has another two disulfide bridges between its N-terminal region and helix 7 of the (β/α)₈-barrel (shown in Supplementary Fig. 1), which probably confer additional stability to the enzyme [10]. The conserved disulfide bond was also reported to be necessary for proper folding of maize ZmGlu [43]. It was hypothesized that the presence two disulfide bridges in the loop B region might help explain the fact that Os4BGlu12 has higher temperature stability than rice Os3BGlu7, which has only one disulfide bond [39]. Fig. 4 shows the activity levels for Os4BGlu12 (Fig. 4A) and Os3BGlu7 (Fig. 4B) after preincubation at temperatures of 0–90 °C for 60 min. The Os4BGlu12 activity remained stable after incubation at 20–50 °C for 60 min, whereas the activity dropped about 5% at 60 °C after 60 min. Inclusion of DTT in the enzyme solution resulted in decreases in the enzyme activity of 26% at 40 °C, and 80% at 60 °C. In contrast, Os3BGlu7 began to lose activity at 40 °C and was inactive at higher

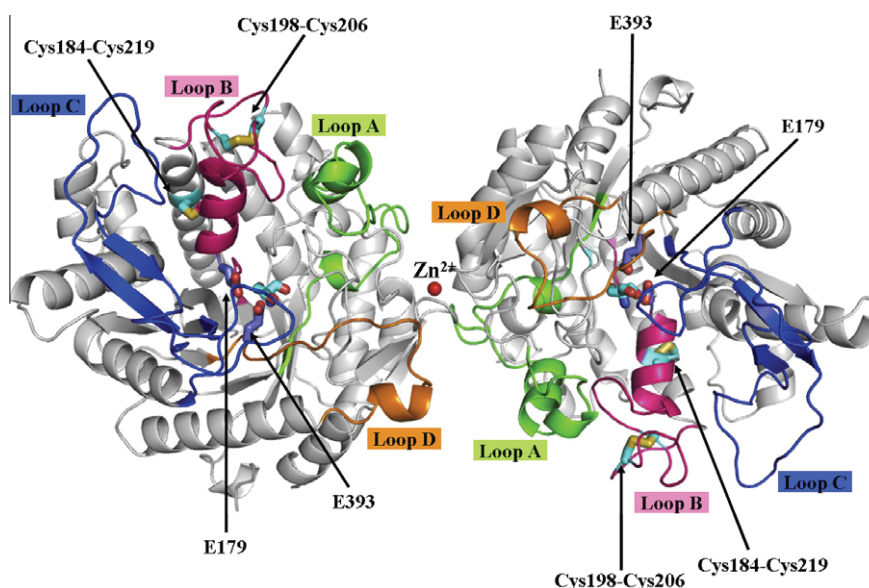


Fig. 3. Ribbon diagram representation of the asymmetric unit of the Os4BGlu12 apo protein structure. The catalytic acid/base, Glu179, and nucleophile, Glu393, are labeled and shown as sticks, colored in red for oxygen, purple for carbon in the online version. The four variable loops between the β -strands and α -helices of the $(\beta/\alpha)_8$ barrel are labeled and colored green for loop A, magenta for loop B, blue for loop C, and orange for loop D (including variable helices and strands that fall in these loop regions), in the online version. The Zn^{2+} ion between the A and B protein molecules in asymmetric unit is shown as a red sphere. Two molecules of Tris associated with the A and B molecules in the asymmetric unit of the Os4BGlu12 structure are represented by sticks with carbon in cyan, oxygen in red and nitrogen in blue. Two disulfide bridges located in the loop B region are represented by a thick bond (in gold with carbons in cyan). The Cys198-Cys206 disulfide bond is present in other structures of GH1 enzymes, while the Cys184-Cys219 bond is not observed in other GH1 enzymes but is located in the same loop region. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

temperatures, while inclusion of DTT resulted in a decrease in the enzyme activity of 25% at 10 °C and 90% at 30 °C suggesting that Os3BGlu7 stability was more sensitive to reduction of the disulfide bonds than Os4BGlu12 (Fig. 4A and B). Thus, although the conserved and nonconserved disulfide bonds may contribute slightly to Os4BGlu12 stability, other factors appear to explain its higher stability than Os3BGlu7.

Aside from the disulfide bonds, the structure of Os4BGlu12 contains several salt bridges and hydrogen bonds that contribute to its stability. Of the 25 arginine residues, 18 residues form direct salt bridges with other residues within the same monomer, while 15 of 30 lysine residues and 3 of 10 histidine residues form such salt bridges. Fifteen salt bridges are obtained from 27 aspartate residues, while 15 are formed by the 20 glutamate residues. For comparison, the structure of rice Os3BGlu7 (PDB entry 2RGM) has only 8 of 21 arginine residues, 5 of 25 lysine residues, 3 of 11 histidine residues, 14 of 24 aspartate residues, and 4 of 12 glutamate residues involved in intramolecular salt bridges. Thermophilic proteins have been observed to have more salt bridges than their mesophilic relatives [44], so the greater number of salt bridges in Os4BGlu12 may contribute to its greater stability than Os3BGlu7, along with other factors.

In the asymmetric unit of Os4BGlu12 crystals, His69 of both molecules and Asp66 of molecule A and Asp36 of molecule B chelate a Zn ion asymmetrically between the two protein molecules, as seen in Fig. 3, and shown in detail in Supplementary Fig. 2, but this stabilizing interaction is not expected to occur in solution. Os4BGlu12 is a monomeric protein in solution, as judged by gel-filtration and DLS (data not shown), and the maximum interface of 488 Å² between molecules A and B, 2.7% of the average monomer surface area of 18,307 Å², is predicted to not result in a dimer in solution by the PISA program [45]. The Zn^{2+} ion appears to have been a fortuitous contaminant in the purification solution or precipitant, which facilitated the contacts between the proteins in the crystal matrix, similar to our previous experience with Os3BGlu7 [16].

The binding of ligands in the active site

The active site of Os4BGlu12 is located at the bottom of an approximately 20 Å deep, slot-like pocket, with two walls approximately 8 Å apart at the narrowest point. As seen in Fig. 3, a Tris molecule was observed at the bottom of the active site of the apo Os4BGlu12 structure in two different conformations in molecules A and B of the asymmetric unit. To support this assignment, Tris was found to be a competitive inhibitor, although with a high K_i of 144 ± 0.1 mM. Because Tris was present at 0.1 M in the crystallization condition, it could be found in the active site of the apo enzyme of Os4BGlu12, despite its low affinity, but was not likely to compete effectively with specific ligands. Tris is commonly a weak inhibitor of β -glucosidases, and it was previously reported that the active site of the apo enzyme of Os3BGlu6 (PDB accession 3GNO) contained Tris [17], which was a component of the storage buffer of the protein and a competitive inhibitor with a K_i of 5.1 mM.

When Os4BGlu12 was crystallized with precipitant solution containing 2 mM DNP2FG then soaked in cryoprotectant containing 10 mM DNP2FG before cooling and X-ray diffraction, the anticipated covalent intermediate of α -2-deoxy-2-fluoroglucoside was not observed in the electron density. Instead, electron density map of Molecule A clearly revealed the presence of the 2-deoxy-2-fluoroglucose ring in a β -linkage to a weak density for the 2,4-dinitrophenol group, as shown in Fig. 5A and B. The whole ligand had a high average B-factor of 72 Å² at full occupancy and refined to a partial occupancy of 0.8 with the B-factor of 63 Å², in line with the transient nature of the Michaelis complex. The sugar ring was better defined than the nitrophenyl ring, and molecule B had poor density for the aglycone and higher B-factors, suggesting the active site may be partially occupied with covalently linked or free 2-F-glucose, with different occupancies for the two molecules. However, the glucosyl moiety fit the electron density better in the ¹S₃ skew-boat conformation previously observed for noncovalent β -glucosidase/substrate intermediates than in a relaxed ⁴C₁ chair

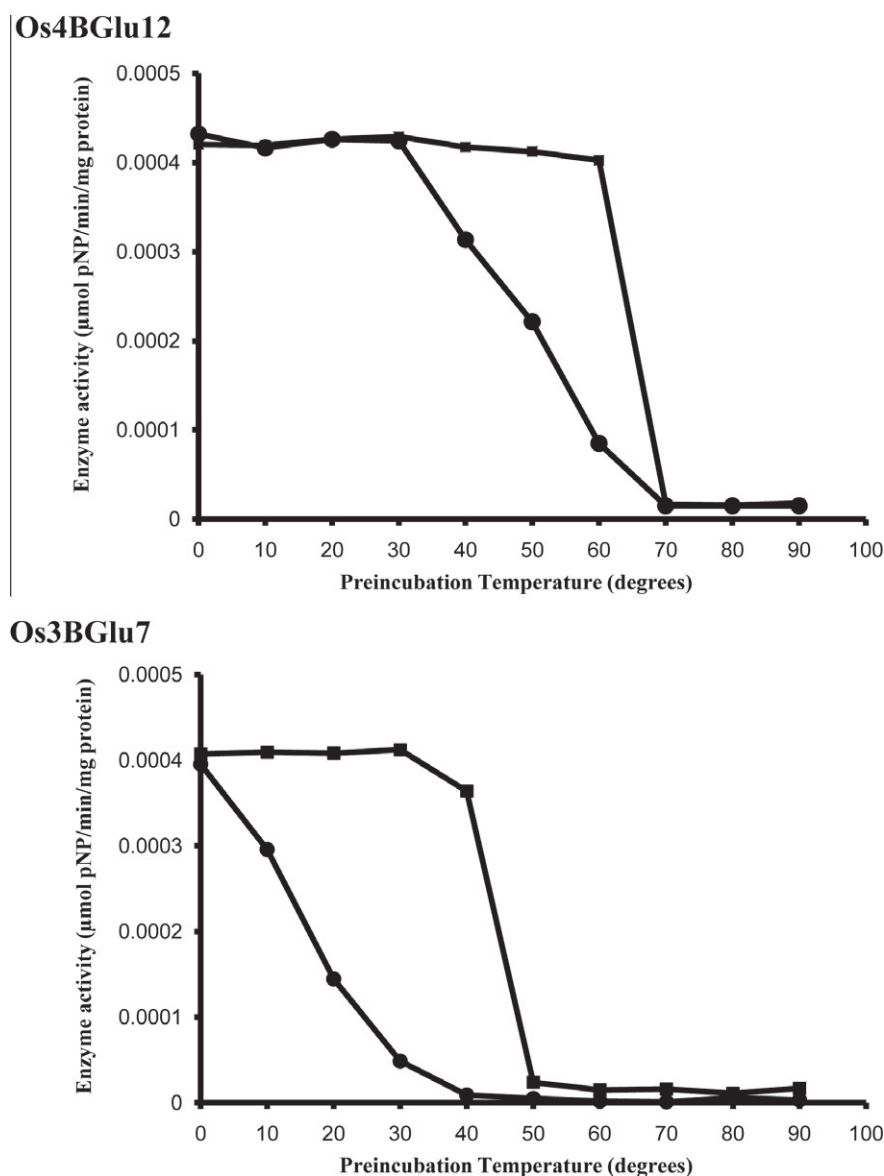


Fig. 4. Comparison of the thermostability of Os4BGlu12 and Os3BGlu7. The enzyme activities were measured in the standard assay (0.3 mM pNPGlc in 50 mM sodium acetate, pH 5.0, for 10 min) after incubation of enzyme with (●) and without (■) 1 mM of DTT for 1 h at the designated temperatures in 50 mM sodium acetate, pH 5.0.

conformation, which has generally been observed for covalent intermediates [10,14,16–18,41,42,46,47].

Previously, DNP2FG was used to show that the enzymatic mechanism of retaining glycosidases proceeds via a covalent glucosyl-enzyme intermediate, because the electron withdrawal by the 2-fluoro group slows both the glycosylation and deglycosylation steps (the first and second reactions of Fig. 1, respectively), but the use of a good leaving group (2,4-dinitrophenolate) as the aglycon allows the covalent intermediate with the nucleophile residue to form faster than hydrolysis [8]. However, it has been shown that acceptors for transglycosylation may allow rescue of the enzyme by accepting the 2-fluoroglucosyl moiety [47–49]. Therefore, the inclusion of 10 mM DNP2FG in the cryoprotectant might have allowed rescue of the enzyme, and resulted in the unhydrolyzed DNP2FG being seen in the active site. When the co-crystals were flash cooled after a 15 s incubation in cryoprotectant without DNP2FG, their diffraction yielded an electron diffraction map that showed the covalent intermediate of α -2-fluoroglucoside (G2F) bound to the O ϵ 2 of the catalytic nucleophile, Glu393, in a relaxed 4C_1 chair (Fig. 6A and B). Thus, the use of

different cryoprotectant conditions allowed solution of the structures of both the Michaelis complex of DNP2FG and the covalent intermediate with G2F, which supported the postulated conformational trajectory of the glucosyl residue through a 1S_3 skew boat to a 4H_3 half chair transition state to the 4C_1 chair covalent intermediate.

Michaelis complex of Os4BGlu12 with DNP2FG

Fig. 5A shows the Os4BGlu12 Michaelis complex with DNP2FG, in which the Glu393 O ϵ 1 was hydrogen bonded with O5 (3.09 Å) and within 3.2 Å of the anomeric carbon of the sugar ring. The O ϵ 2 of Glu393 was within hydrogen bond distance of the F2 (2.62 Å), suggesting it would likely hydrogen bond to the O2 atom of a normal glucosyl moiety. The distance between O ϵ 2 of the catalytic acid/base residue Glu179 and O1 of DNP2FG was 3.30 Å. Thus, the 1S_3 conformation of the Michaelis complex placed the substrate in position for nucleophilic attack and protonation by the catalytic acid. The glycon residue formed hydrogen bonds with 6 residues around the –1 subsite (Gln29, His133, Tyr322, Glu393, Gln449 and Trp450). In addition to the interactions described

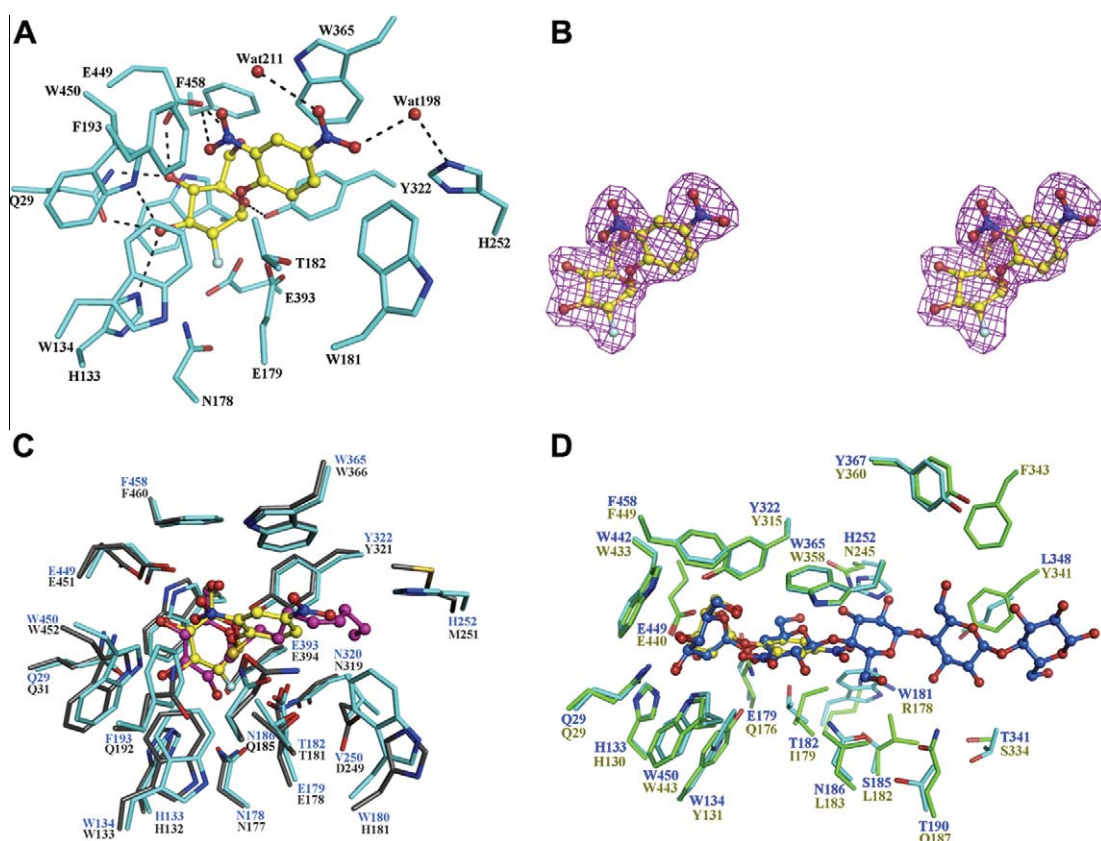


Fig. 5. Protein–ligand interactions in the active site of the Os4BGlu12 DNP2FG complex. (A) The amino acid residues surrounding DNP2FG are represented in stick representation, with carbon, nitrogen and oxygen atoms colored in cyan, blue and red, respectively. DNP2FG is drawn in ball and stick representation in the same colors, except carbon atoms are in yellow and fluorine in green. (B) Stereo view of the electron density ($2F_{\text{obs}} - F_{\text{calc}}$ map contoured at the 1σ level) of DNP2FG, calculated with a model from which DNP2FG was omitted. (C) Superimposition of Os4BGlu12_DNP2FG (cyan carbons) with Os3BGlu6 with *n*-octyl- β -D-thioglucopyranoside (gray carbons). The ligands are shown in ball and stick representations, with pink carbons for DNP2FG and yellow carbons for *n*-octyl- β -D-thioglucopyranoside. (D) Superimposition of the protein ligand binding of Os4BGlu12 with DNP2FG (cyan carbons) and Os3BGlu7 (3F5K) with cellopentaose (green carbons) and the ligand carbons drawn in yellow and blue for DNP2FG and cellopentaose, respectively.

above, O3 of the 2-F-glucose was hydrogen bonded with His133N ϵ 2 (2.88 Å), Glu29 O ϵ 1 (2.75 Å) and Trp450N ϵ 1 (2.82 Å). O4 hydrogen bonds with Gln29N ϵ 2 (2.71 Å) and Glu449 O ϵ 1 (2.54 Å), while O5 hydrogen bonds with Tyr322 (2.80 Å) and O6 with Glu449 O ϵ 2 (2.63 Å). These hydrogen bonds between the glucose and the active site residues have been observed for other GH1 β -glucosidase/substrate complexes [19]. Other conserved residues making up the glycon binding site include Trp134, Asn178, Asn320, Trp365, Trp442 and Phe458. The aglycon binding site is occupied by hydrophobic residues and a few polar residues. The 2,4-dinitrophenol group was bound by hydrophobic and aromatic stacking interactions with Trp181 and Phe193 on one side and Trp365 on the other. The polar residues found in the aglycon binding subsite include Thr182, Asn186 and Asn325 and Asn452. In GH1, Trp365 is a conserved residue, which corresponds to Trp358 in Os3BGlu7 and Trp366 in Os3BGlu6 (Fig. 5C and D).

The covalent intermediate of Os4BGlu12 with G2F

In the structure of Os4BGlu12 with a covalent bond between the anomeric carbon of G2F and the Glu393 nucleophile residue O ϵ 1, the 4C_1 chair conformation of the G2F ring was observed in subsite –1 (Fig. 6A and B). Aside from the covalent bonding of the nucleophile, the glucose is bound by hydrogen bonds similar to the Michaelis complex. Fig. 6B shows that the density was clear and Tyr322 was seen to have an alternative conformation in which it can hydrogen bond to the catalytic nucleophile in this complex, in addition to a conformation hydrogen bonding to O5, similar to

the noncovalent complex. The major structural differences between the apoenzyme and complex of many GH1 enzymes are the positions of nucleophilic residues. In the complex structures of the *P. polymyxa*, *S. solfataricus*, *T. maritima*, rice Os3BGlu6, rice Os3BGlu7, and human cytosolic β -glucosidases, the side-chain position of the nucleophilic residue has moved to link with anomeric carbon of G2F. For example, a displacement of C δ by 1.5–1.6 Å is observed when the rice Os3BGlu7 covalent complex, 2RGM (pink structure in Fig. 6C), is compared to the apoenzyme structure, 2RGL [16]. In these structures, the angles between the sugar anomeric carbon, O ϵ 1 and C δ of the nucleophilic residue, fall in the range of 113–117°. In Os4BGlu12, the C δ is only displaced 0.3 Å (molecule B) to 0.7 Å (molecule A), and the angle between the sugar anomeric carbon and Glu393 O ϵ 1 and C δ is strained (142°). This is similar to the *S. alba* myrosinase (shown in yellow in Fig. 6C), an S-glycosidase, where the displacement of C δ is only 0.2 Å in the covalent G2F complex (PDB accession 1E73) compared to the apo protein (1MYR) and the angle is similarly wide (138°) [11].

Thioglucosidase activity

The similarity of the covalent intermediate of Os4BGlu12 to that of *S. alba* myrosinase suggested that Os4BGlu12 might show mechanistic similarities to the myrosinase, so its ability to hydrolyze thioglucosides was tested. Os4BGlu12 could hydrolyze the S-glycosides *p*NP-thioglucoside and *n*-octyl- β -D-thioglucoside, as shown

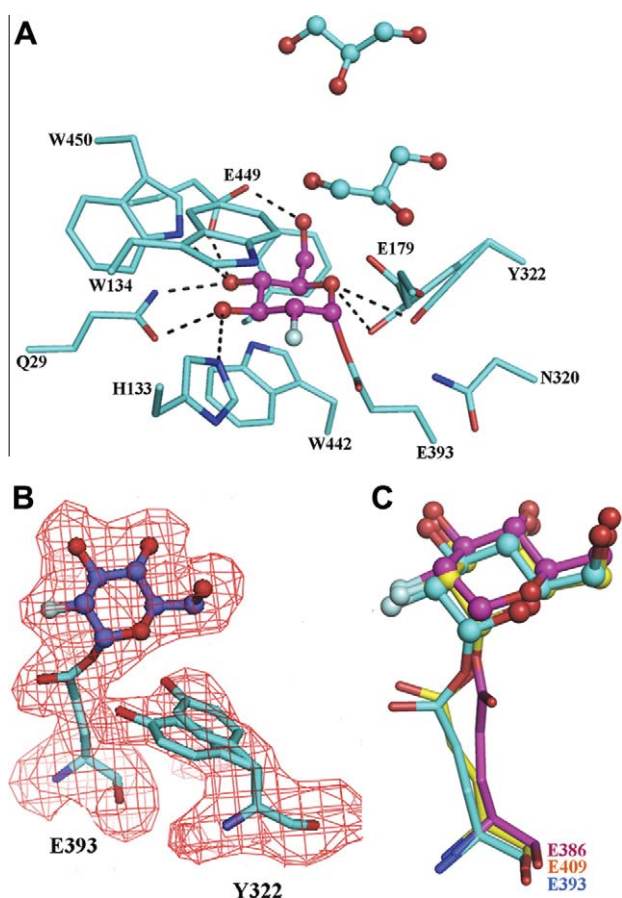


Fig. 6. The active site region of the covalent complex of Os4BGlu12 with G2F. (A) Protein ligand binding of the monomer A (Os4BGlu12/G2F complex) amino acids surrounding the ligand are represented as sticks with carbons colored cyan. The ligands are represented in ball and stick and include G2F and glycerol of the complex with carbons in pink and yellow, respectively. (B) $F_{\text{obs}} - F_{\text{calc}}$ electron density map of glycosyl-enzyme intermediate showing the two conformations of Tyr322 in the structure of Os4BGlu12 with G2F. The map is contoured at the 3σ level. (C) Superimposition of nucleophile and sugar residues of glycosyl-enzyme intermediates of the Os4BGlu12/G2F complex with those of *S. alba* myrosinase and rice Os3BGlu7. Carbons are shown in cyan for Os4BGlu12/G2F, in pink for rice Os3BGlu7 (2RGM) and in yellow for *S. alba* myrosinase (1E73). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

in Table 2. Os4BGlu12 β -glucosidase catalyzed the hydrolysis of *n*-octyl- β -D-thioglucopyranoside, although with about 1000-fold lower efficiency ($k_{\text{cat}}/K_{\text{m}}$ of $1.80 \times 10^{-2} \text{ mM}^{-1} \text{ s}^{-1}$) than it hydrolyzed *n*-octyl- β -D-glucopyranoside ($k_{\text{cat}}/K_{\text{m}}$ of $22.03 \text{ mM}^{-1} \text{ s}^{-1}$). Os4BGlu12 hydrolyzed pNP- β -D-thioglucoside about 200-fold less efficiently ($k_{\text{cat}}/K_{\text{m}}$ of $0.17 \text{ mM}^{-1} \text{ s}^{-1}$) than it hydrolyzed pNP- β -D-glucoside ($k_{\text{cat}}/K_{\text{m}}$ of $33 \text{ mM}^{-1} \text{ s}^{-1}$). The K_{m} values of *S*-glycosides were similar (1.61 mM for *n*-octyl- β -D-thioglucopyranoside and 1.64 mM for pNP-thioglucoside) but were higher than the corresponding *O*-glycosides by about 3-fold.

Table 2
Hydrolysis of *O*-glucosides and *S*-glucosides by Os4BGlu12.

Substrate	V_{max} ($\mu\text{mol min}^{-1} \text{ mg}^{-1} \text{ protein}$)	K_{m} (mM)	k_{cat}^a (s^{-1})	$k_{\text{cat}}/K_{\text{m}}$ ($\text{mM}^{-1} \text{ s}^{-1}$)
pNPG	20.1 ± 0.6	0.56 ± 0.03	18.5 ± 0.15	33
pNPSG	0.29 ± 0.007	1.64 ± 0.10	0.27 ± 0.015	0.17
<i>n</i> -Heptyl- β -D-glucoside	13.2 ± 0.13	0.91 ± 0.07	12.1 ± 0.05	13
<i>n</i> -Octyl- β -D-glucoside	10.7 ± 0.3	0.44 ± 0.015	9.8 ± 0.2	22
<i>n</i> -Octyl- β -D-thioglucopyranoside	$4.00 \times 10^{-2} \pm 0.003$	1.61 ± 0.10	$3.76 \times 10^{-2} \pm 0.004$	2.34×10^{-2}

^a The k_{cat} values given are apparent values calculated by division of the V_{max} by the molar amount of Os4BGlu12 protein.

Commercial sweet almond β -glucosidase has been shown to have thioglycohydrolase activities similar to Os4BGlu12, the K_{m} values for the *S*- and *O*-glycosidase are similar (0.9 mM for pNP- β -D-fucoside and 1.6 for pNP- β -D-thioglucoside), while the $k_{\text{cat}}/K_{\text{m}}$ value was about 1000-fold lower for the *S*-glycoside [50]. In contrast, Os3BGlu6 and Os3BGlu7 exhibited no significant hydrolysis of pNP- β -D-thioglucoside under the conditions of the assay (data not shown), although a slight hydrolysis of pNP- β -D-thioglucoside by Os3BGlu7 was recently reported at high concentrations of substrate and enzyme in NMR [51]. This is consistent with the fact that their covalent intermediate structures were different from myrosinase and Os4BGlu12, but the reason for this correlation remains unclear.

It has been suggested in the myrosinase literature that classic myrosinases are derived from the defense-related β -glucosidases, such as cyanogenic β -glucosidases [52]. Indeed, the minimum evolution-based phylogenetic tree in Fig. 2 shows the myrosinases group with cluster At/Os7, which contains the cyanogenic β -glucosidases, although in our previous analyses [5] and distance-based neighbor-joining and maximum likelihood trees of the same sequences failed to show a significance to this association (data not shown). The structural similarity of the covalent intermediate to that of *S. alba* myrosinase and the higher thioglucosidase activity of Os4BGlu12 and almond β -glucosidase [50] compared to the more distantly related Os3BGlu6 and Os3BGlu7 provide mechanistic support for this similarity.

Comparison of oligosaccharide binding by rice β -glucosidases

The three rice β -glucosidases for which structures have been determined, Os3BGlu6 (Fig. 5C), Os3BGlu7 (Fig. 5D) and Os4BGlu12, come from three distinct phylogenetic branches of plant GH1 enzymes, yet all hydrolyze oligosaccharides and glycosides, although with different preferences. Os3BGlu6 only hydrolyzes β -1,2- and β -(1,3)-linked disaccharides (sophorose and laminaribiose) efficiently, while Os3BGlu7 and Os4BGlu12 efficiently hydrolyze β -(1,4)-linked gluco-oligosaccharides (cellooligosaccharides) of 3–6 residues, as well as laminaribiose [26]. The superimposition of rice Os3BGlu6 and Os3BGlu7 found that Asn245 at the subsite + 2 for binding cellooligosaccharides in Os3BGlu7 (Fig. 5D) is replaced with Met251 in Os3BGlu6 (Fig. 5C) [17]. This Met251 appeared to block the third glucosyl residue of cellooligosaccharides from binding to Os3BGlu6 in the same position as in Os3BGlu7, and the Os3BGlu7 N245V mutation gave a $k_{\text{cat}}/K_{\text{m}}$ value for hydrolysis of cellotriose 15 times lower than wild type Os3BGlu7, due to a 42-fold increase in K_{m} [16]. These effects show the importance of Asn245 in Os3BGlu7 binding of the third glucosyl residue in cellooligosaccharides [41]. The corresponding residue in Os4BGlu12 is His252 (Fig. 5A, C and D), which also has the potential to hydrogen bond to the third glucosyl residue in glucooligosaccharides, thereby explaining the binding of Os4BGlu12 to cellotriose and longer oligosaccharides.

The efficiency of hydrolysis increases with the length of cellooligosaccharides from 3 to 6 for Os3BGlu7, but shows little change as the length increases with Os4BGlu12, which maintains a rather

high K_m of approximately 5 mM for all these cellooligosaccharides [26]. The distance from the deepest part of the active site slot to the surface entrance of Os4BGlu12 is approximately 20 Å, but the lack of increase in affinity for longer cellooligosaccharides suggests the outer parts of this cleft have little interaction with these cellooligosaccharides. Fig. 5D shows that superimposition of the structures of Os4BGlu12 with DNP2FG and Os3BGlu7 with cellopentaose (3F5K) [41] indicated that, while all residues in the –1 subsite of these two enzymes are the same, except for the Trp134 in Os4BGlu12 being replaced by Tyr131 in Os3BGlu7, their aglycon-binding sites show little similarity. Although both contain the conserved Trp at the +1 and +2 subsites (Trp 365 in Os4BGlu12, Trp 358 in Os3BGlu7), the aglycon subsite residues Trp181, Thr182, Ser185, Asn186, Thr190, Phe193, Phe254 and His252, in Os4BGlu12 are replaced by Arg178, Ile179, Leu182, Leu183, Gln187, Asn190, Asn245 and Tyr247, respectively, in Os3BGlu7. The Tyr341 residue, which stacks the residues in the +3 and +4 subsites in Os3BGlu7 [41] is replaced by Leu348 in Os4BGlu12, which may explain the relatively poor binding of cellotetraose and longer oligosaccharides in Os4BGlu12 compared to Os3BGlu7 [26].

Os4BGlu12 had a 2-fold higher K_m for *n*-heptyl-β-D-glucopyranoside than for *n*-octyl-β-D-glucopyranoside, suggesting that Os4BGlu12 may prefer glycosides with more extensive apolar regions. This is similar to Os3BGlu6, which showed a similar preference for octyl over heptyl glucoside, but Os4BGlu12 is over 10 times more efficient at hydrolyzing both substrates than Os3BGlu6. Os4BGlu12 was previously noted to efficiently hydrolyze the large, hydrophobic deoxycorticosteroid glucoside ($k_{cat}/K_m = 20.1 \text{ mM}^{-1} \text{ s}^{-1}$) [26]. This seems to reflect the presence of various nonpolar groups in the active site, as mentioned above, although Os3BGlu6 and Os4BGlu12 show little difference in the overall polarity of the aglycone binding site, as seen in Fig. 5C.

Conclusion

In summary, the overall structure of Os4BGlu12 exhibits the (β/α)₈ barrel fold characteristic of GH1 enzymes, but contains two disulfide bridges (non-conserved Cys184–Cys219 and conserved Cys198–Cys206), located in loop B, while most plant GH1 β-glucosidases only have the conserved one. However, the extra disulfide bond did not appear to explain the higher stability of Os4BGlu12, compared to Os3BGlu7 produced in the same system. Both the Michaelis complex with DNP2FG (Fig. 1 left structure and Fig. 5) and the subsequent covalent intermediate with G2F (Fig. 1 middle structure and Fig. 6) could be observed by including or omitting DNP2FG from the cryoprotectant, respectively. In the complex structures, the surrounding conserved amino acid residues and the glucosyl group formed hydrogen bonds similar to those seen in other GH1 enzymes. The position of the catalytic nucleophile, Glu393, in the Os4BGlu12/G2F complex was similar to that of the nucleophile in the *S. alba* myrosinase, an S-glycosidase, consistent with its catalysis of the hydrolysis of S-glycosides. The apparent Os4BGlu12 aglycon-binding site included both aromatic and polar amino acids, accounting for its ability to hydrolyze glycosides with a range of hydrophobic and polar aglycons. The aglycon-binding cleft of Os4BGlu12 appears to be less adapted for binding long cellooligosaccharides than Os3BGlu7, so it represents a broad specificity, generalist β-glucosidase. Os4BGlu12 should serve as a good model for understanding the binding of large apolar aglycons, such as the phytohormone tuberonic acid [30].

Acknowledgments

Phimonphan Chuankhayan and the protein X-ray crystallography beamline staff of the National Synchrotron Radiation Research

Center (NSRRC) Hsinchu, Taiwan are thanked for assistance in data collection and useful comments. The authors are grateful to Yanling Hua for providing *p*-nitrophenyl-thio-β-D-glucoside and to Salila Pengthaisong for providing Os3BGlu7 protein. Funding was provided by the Synchrotron Light Research Institute (Thailand), which also provided an in-house X-ray diffractometer usage for preliminary studies, the Thailand Research Fund [Grants BRG5380017 and MRG5280249] and Suranaree University of Technology. Data collection was supported by the NSRRC, a national user facility supported by the National Science Council of Taiwan, ROC. The Synchrotron Radiation Protein Crystallography Facility is supported by the National Research Program for Genomic Medicine.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.abb.2011.04.005.

References

- [1] B. Henrissat, *Biochem. J.* 280 (1991) 309–316.
- [2] B.L. Cantarel, P.M. Coutinho, C. Rancurel, T. Thomas Bernard, V. Vincent Lombard, B. Henrissat, *Nucleic Acids Res.* 37 (2009) 233–238.
- [3] J.R. Ketudat Cairns, A. Esen, *Cell. Mol. Life Sci.* 67 (2010) 3389–3405.
- [4] Z. Xu, L.L. Escamilla-Treviño, L. Zeng, M. Lalgondar, D.R. Bevan, B.S.J. Winkel, A. Mohamed, C.L. Cheng, M.C. Shih, J.E. Poulton, A. Esen, *Plant Mol. Biol.* 55 (2004) 343–367.
- [5] R. Opassiri, B. Pomthong, T. Onkkoosong, T. Akiyama, A. Esen, J.R. Ketudat Cairns, *BMC Plant Biol.* 6 (2006) 33.
- [6] J. Jenkins, L.L. Leggio, G. Harris, R. Pickersgill, *FEBS Lett.* 362 (1995) 281–285.
- [7] B. Henrissat, I. Callebaut, S. Fabrega, P. Lehn, J.P. Mornon, G. Davies, *Proc. Natl. Acad. Sci. USA* 92 (1995) 7090–7094.
- [8] S.G. Withers, R.A.J. Warren, I.P. Street, K. Rupitz, J.B. Kempton, R. Aebersold, *J. Am. Chem. Soc.* 112 (1990) 5887–5889.
- [9] T. Barrett, C.G. Suresh, S.P. Tolley, E.J. Dodson, M.A. Hughes, *Structure* 5 (1995) 951–960.
- [10] W.P. Burmeister, S. Cottaz, H. Driguez, R. Iori, S. Palmieri, B. Henrissat, *Structure* 5 (1997) 663–675.
- [11] W.P. Burmeister, S. Cottaz, P. Rollin, A. Vasella, B. Henrissat, *J. Biol. Chem.* 275 (2000) 30385–30393.
- [12] H. Husebye, S. Arzt, W.P. Burmeister, F.V. Haertel, A. Brandt, J.T. Rossiter, A.M. Bones, *Insect. Biochem. Mol. Biol.* 35 (2005) 1311–1320.
- [13] G.J. Davies, V.M.-A. Ducros, A. Varrot, D.L. Zechel, *Biochem. Soc. Trans.* 31 (2003) 523–527.
- [14] D.L. Zechel, A.B. Boraston, T. Gloster, C.M. Boraston, J.M. Macdonald, D.M. Tilbrook, R.V. Stick, G.J. Davies, *J. Am. Chem. Soc.* 125 (2003) 14313–14323.
- [15] T.M. Gloster, S. Roberts, V.M. Ducros, G. Perugini, M. Rossi, R. Hoos, M. Moracci, A. Vasella, G.J. Davies, *Biochemistry* 43 (2004) 6101–6109.
- [16] W. Chuenchor, S. Pengthaisong, R.C. Robinson, J. Yuvaniyama, W. Oonanant, D.R. Bevan, A. Esen, C.J. Chen, R. Opassiri, J. Svasti, J.R. Ketudat Cairns, *J. Mol. Biol.* 377 (2008) 1200–1215.
- [17] S. Seshadri, T. Akiyama, R. Opassiri, B. Kuaprasert, J.R. Ketudat Cairns, *Plant Physiol.* 151 (2009) 47–58.
- [18] J. Noguchi, Y. Hayashi, Y. Baba, O. Zkino, M. Kimura, *Biochem. Biophys. Res. Comm.* 374 (2008) 549–552.
- [19] L. Verdoucq, J. Morinière, D.R. Bevan, A. Esen, A. Vasella, B. Henrissat, M. Czjzek, *J. Biol. Chem.* 279 (2004) 31796–31803.
- [20] M. Czjzek, M. Cicek, V. Zamboni, W.P. Burmeister, D.R. Bevan, B. Henrissat, A. Esen, *Proc. Natl. Acad. Sci. USA* 97 (2000) 13555–13560.
- [21] M. Sue, K. Yamazaki, S. Yajima, T. Nomura, T. Matsukawa, H. Iwamura, T. Miyamoto, *Plant Physiol.* 141 (2006) 1237–1247.
- [22] M. Sue, C. Nakamura, T. Miyamoto, S. Yajima, *Plant Sci.* 180 (2011) 268–275.
- [23] L. Barleben, S. Panjikar, M. Ruppert, J. Koepke, J. Stöckigt, *Plant Cell* 19 (2007) 2886–2897.
- [24] M. Cicek, A. Esen, *Biotechnol. Bioeng.* 63 (1999) 392–400.
- [25] T. Akiyama, H. Kaku, N. Shibuya, *Phytochemistry* 48 (1998) 49–54.
- [26] R. Opassiri, J. Maneesan, T. Akiyama, B. Pomthong, S. Jin, A. Kimura, J.R. Ketudat Cairns, *Plant Sci.* 179 (2010) 273–280.
- [27] R. Opassiri, Y. Hua, O. Wara-Aswapati, T. Akiyama, J. Svasti, A. Esen, J.R. Ketudat Cairns, *Biochem. J.* 379 (2004) 125–131.
- [28] T. Kuntothom, S. Luang, A.J. Harvey, G.B. Fincher, R. Opassiri, M. Hrmova, J.R. Ketudat Cairns, *Arch. Biochem. Biophys.* 491 (2009) 85–95.
- [29] S. Wakuta, S. Hamada, H. Ito, H. Matsuura, K. Nabeta, H. Matsui, *Phytochemistry* 71 (2010) 1280–1288.
- [30] S. Wakuta, S. Hamada, H. Ito, R. Imai, H. Mori, H. Matsuura, K. Nabeta, H. Matsui, *J. Appl. Glycosci.* 58 (2011) 67–70.
- [31] S. Sansenya, J.R. Ketudat Cairns, R. Opassiri, *Acta Crystallogr. F66* (2010) 320–323.

- [32] Z. Otwinowski, W. Minor, *Meth. Enzymol.* Academic Press, New York 276 (1997) 307–326.
- [33] A. Vagin, A. Teplyakov, *J. Appl. Crystallogr.* 30 (1997) 1022–1025.
- [34] G.N. Murdushov, A.A. Vagin, A. Lebedev, K.S. Wilson, E.J. Dodson, *Acta Crystallogr. D55* (1999) 247–255.
- [35] P. Emsley, K. Cowtan, *Acta Crystallogr. D60* (2004) 2126–2132.
- [36] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, *J. Appl. Crystallogr.* 26 (1993) 283–291.
- [37] I.W. Davis, A. Leaver-Fay, V.B. Chen, J.N. Block, G.J. Kapral, X. Wang, L.W. Murray, W.B. Arendall, J. Snoeyink, J.S. Richardson, D.C. Richardson, *Nucleic Acids Res.* 35 (2007) 375–383.
- [38] W.L. Delano, The Pymol Molecular Graphics System (2002) (Online). Available from: <<http://www.pymol.org/>>.
- [39] R. Opasiri, J.R. Ketudat-Cairns, T. Akiyama, O. Wara-Aswapati, J. Svasti, A. Esen, *Plant Sci.* 165 (2003) 627–638.
- [40] M. Dixon, *Biochem. J.* 55 (1953) 170–171.
- [41] W. Chuenchor, S. Pengthaisong, R.C. Robinson, J. Yuvaniyama, J. Svasti, *J. Struct. Biol.* 173 (2011) 169–179.
- [42] M. Czjzek, M. Cicek, V. Zamboni, W.P. Burmeister, D.R. Bevan, B. Henrissat, A. Esen, *Biochem. J.* 354 (2001) 37–46.
- [43] V. Rotrekl, E. Nejedla, I. Kucera, F. Abdallah, K. Palme, B. Brzobohaty, *Eur. J. Biochem.* 266 (1999) 1056–1065.
- [44] J.J. Tanner, R.M. Hecht, K.L. Krause, *Biochemistry* 35 (1996) 2597–2609.
- [45] E. Krissinel, K. Henrick, *J. Mol. Biol.* 372 (2007) 774–797.
- [46] P. Isorna, J. Polaina, L. Latorre-García, F.J. Cañada, B. González, J. Sanz-Aparicio, *J. Mol. Biol.* 371 (2007) 1204–1218.
- [47] C. Malet, A. Planas, *FEBS Lett.* 440 (1998) 208–212.
- [48] M. Okuyama, H. Mori, K. Watanabe, A. Kimura, S. Chiba, *Biosci. Biotechnol. Biochem.* 66 (2002) 928–933.
- [49] G. Hommalai, S.G. Withers, W. Chuenchor, J.R. Ketudat Cairns, J. Svasti, *Glycobiology* 17 (2007) 744–753.
- [50] H. Shen, L.D. Byers, *Biochem. Biophys. Res. Commun.* 362 (2007) 717–720.
- [51] T. Kuntothom, M. Raab, I. Tvaroska, S. Fort, S. Pengthaisong, F.J. Cañada, L. Calle, J. Jimenez-Barbero, J.R. Ketudat Cairns, M. Hrmova, *Biochemistry* 49 (2010) 8779–8793.
- [52] M. Wang, D. Lee, X. Sun, Y.J. Zhu, H. Nong, J. Zhang, *Plant Sci.* 177 (2009) 716–723.
- [53] K. Tamura, J. Dudley, M. Nei, S. Kumar, *Mol. Biol. Evol.* 24 (2007) 1596–1599.
- [54] M. Nei, S. Kumar, *Molecular Evolution and Phylogenetics*, Oxford University Press, New York, 2000.
- [55] R.C. Edgar, *Nucleic Acids Res.* 32 (2004) 1792–1797.

The role of the oligosaccharide binding cleft of rice BGlu1 in hydrolysis of cellooligosaccharides and in their synthesis by rice BGlu1 glycosynthase

Salila Pengthaisong,¹ Stephen G. Withers,² Buabarn Kuaprasert,³ Jisnusun Svasti,⁴ and James R. Ketudat Cairns^{1*}

¹Schools of Biochemistry and Chemistry, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand

²Centre for High-throughput Biology (CHiBi) and Department of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1

³Research Facility Division, Synchrotron Light Research Institute (Public Organization), Nakhon Ratchasima 30000, Thailand

⁴Department of Biochemistry and Center of Excellence in Protein Structure and Function, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

Received 13 November 2011; Revised 26 December 2011; Accepted 27 December 2011

DOI: 10.1002/pro.2021

Published online 11 January 2012 proteinscience.org

Abstract: Rice BGlu1 β -glucosidase nucleophile mutant E386G is a glycosynthase that can synthesize *p*-nitrophenyl (*p*NP)-cellooligosaccharides of up to 11 residues. The X-ray crystal structures of the E386G glycosynthase with and without α -glucosyl fluoride were solved and the α -glucosyl fluoride complex was found to contain an ordered water molecule near the position of the nucleophile of the BGlu1 native structure, which is likely to stabilize the departing fluoride. The structures of E386G glycosynthase in complexes with cellotetraose and cellopentaose confirmed that the side chains of N245, S334, and Y341 interact with glucosyl residues in cellooligosaccharide binding subsites +2, +3, and +4. Mutants in which these residues were replaced in BGlu1 β -glucosidase hydrolyzed cellotetraose and cellopentaose with $k_{\text{cat}}/K_{\text{m}}$ values similar to those of the wild type enzyme. However, the Y341A, Y341L, and N245V mutants of the E386G glycosynthase synthesize shorter *p*NP-cellooligosaccharides than do the E386G glycosynthase and its S334A mutant, suggesting that Y341 and N245 play important roles in the synthesis of long oligosaccharides. X-ray structural studies revealed that cellotetraose binds to the Y341A mutant of the glycosynthase in a very different, alternative mode not seen in complexes with the E386G glycosynthase, possibly explaining the similar hydrolysis, but poorer synthesis of longer oligosaccharides by Y341 mutants.

Keywords: glycosynthase; oligosaccharide synthesis; rice; transglucosylation; X-ray crystallography

Abbreviations: α -GlcF, α -glucosyl fluoride; DP, degree of polymerization; ESI-MS, electrospray ionization-mass spectrometry; G2F, 2-deoxy-2-fluoro- α -D-glucosyl moiety; Glc, glucose; IMAC, immobilized metal affinity chromatography; LC-MS, liquid chromatography-mass spectrometry; *p*NP, *para*-nitrophenyl; *p*NPC2, *p*NP- β -cellobioside; *p*NPC3, *p*NP- β -cellotrioside; *p*NPC4, *p*NP- β -cellotetraoside; *p*NPC5, *p*NP- β -cellopentaoside; *p*NPC6, *p*NP- β -cellohexaoside; *p*NPC7, *p*NP- β -celloheptaoside; *p*NPC8, *p*NP- β -cellooctaoside; *p*NPC9, *p*NP- β -cellonanaoside; *p*NP_{Glc}, *p*NP- β -D-glucopyranoside.

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Synchrotron Light Research Institute (Public Organization) of Thailand; Grant number: GS-51-D03; Grant sponsor: Thailand Research Fund; Grant number: BRG5380017; Grant sponsors: Suranaree University of Technology National Research University Project of the Commission on Higher Education of Thailand; Diffraction data collection at the NSRRC was supported by the National Research Program for Genomic Medicine and National Research Council of Taiwan, ROC.

*Correspondence to: James R. Ketudat Cairns, Schools of Biochemistry and Chemistry, Institute of Science, Suranaree University of Technology, 111 University Avenue, Muang District, Nakhon Ratchasima 30000, Thailand. E-mail: cairns@sut.ac.th

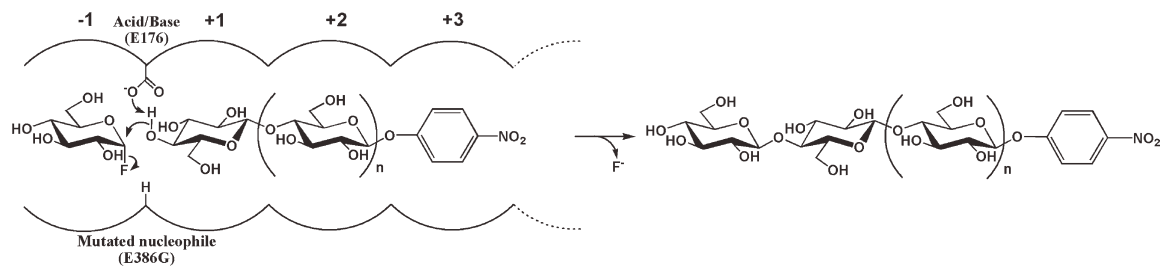


Figure 1. Mechanism of rice BGlul E386G glycosynthase.

Introduction

Glycosynthases are mutants of glycosidases in which the catalytic nucleophile of a retaining glycosidase has been replaced. They can be used for the synthesis of specific oligosaccharides in quite attractive yields. The catalytic nucleophile of these mutants is typically replaced by a nonnucleophilic residue such that the mutant cannot form a reactive glycosyl-enzyme intermediate for hydrolysis or transglycosylation.¹ However, when a glycosyl fluoride of the opposite anomeric configuration to that of the parent substrate is present as a glycosyl donor, the mutant enzymes are able to transfer the glycosyl moiety to acceptor alcohols without hydrolysis of the products² (Fig. 1). Moreover, a glycosynthase was recently derived from an inverting glycosidase, which ordinarily hydrolyzes the glycosidic bond by a single displacement mechanism.³ While high yields and selectivities are key features of the glycosynthase reaction, mutagenesis can also play an important role in the improvement and alteration of the catalytic activities of glycosynthases.^{4–9}

Exoglycosynthases, glycosynthases derived from exoglycosidases, have moderate substrate specificity and regioselectivity and can synthesize short chain oligosaccharides (di-, tri-, and tetra-oligosaccharides) that have various glycosidic linkages.^{10–12} Endoglycosynthases, which are derived from endoglycosidases, generally have high regioselectivity and can catalyze the synthesis of specific glycosidic linkages. They can synthesize longer oligosaccharides than exoglycosynthases because they have long glycone-binding sites that accommodate longer glycosyl donor substrates.^{13–16}

A glycoside hydrolase family 1 β -D-glucosidase from rice (BGlul, systematically called Os3BGlul7) has been expressed in *Escherichia coli* and characterized.¹⁷ Rice BGlul hydrolyzes β -1,3- and β -1,4-linked oligosaccharides and pyridoxine 5'-O- β -D-glucoside, and also has high transglycosylation activity to produce these same compounds as products.¹⁸ Moreover, kinetic subsite analysis of cellooligosaccharide hydrolysis indicated that rice BGlul has subsites for binding of at least six β -1,4-linked D-glucosyl residues. Glycine, alanine, and serine mutants of rice BGlul nucleophile, E386, have been gener-

ated and the E386G mutant, acting as glycosynthase, catalyzed the most rapid accumulation of transglycosylation products.¹⁹ The unique property of this enzyme compared to previous glycosynthases derived from exoglycosidases is that it could synthesize much longer chains (of at least 11 β -1,4-linked glucosyl residues) than do other exoglycosidase-derived glycosynthases. This was hypothesized to be a consequence of an extended active site in rice BGlul, as represented by the outer subsites in Figure 1. The X-ray crystal structures of wild type BGlul and of its E176Q acid-base mutant in complexes with oligosaccharides revealed that the glucosyl residues were primarily bound by aromatic-sugar stacking interactions and water-mediated hydrogen bonds with several residues along a long active site cleft, except for the nonreducing terminal glucosyl residue that was distorted through strong hydrogen bonding interactions with the surrounding amino acids.^{20,21}

To assess the role of the long oligosaccharide-binding cleft of rice BGlul E386G glycosynthase in synthesis of long oligosaccharides, its structures alone and in complexes with the donor substrate α -glucosyl fluoride or the products cellotetraose and cellopentaose were determined by X-ray crystallography. The importance of the interactions in the long oligosaccharide binding cleft for hydrolysis and glycosynthase production of long cellooligosaccharides was investigated by site-directed mutagenesis of residues in this cleft in the wild type BGlul and BGlul E386G glycosynthase. Enzymatic and structural analysis of these mutants showed flexibility in the binding of cellooligosaccharides for hydrolysis, while synthesis of long cellooligosaccharides by the glycosynthase appeared more sensitive to loss of an outer glucosyl-residue binding amino acid.

Results and Discussion

3D structures of BGlul E386G glycosynthase complexed with substrates

To assess the structural basis for the efficient glycosynthase activity of BGlul E386G, its structures alone and in complex with α -glucosyl fluoride (α -GlcF) were determined by X-ray crystallography.

Table I. Data Collection Statistics

	E386G	E386G_α-GlcF	E386G_C4	E386G_C5	S334A_C4	Y341A_C4
PDB code	3SCN	3SCO	3SCT	3SCU	3SCV	3SCW
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters (Å, °)	$a = 79.4$ $b = 100.8$ $c = 127.4$ $\alpha = \beta = \gamma = 90^\circ$	$a = 79.9$ $b = 100.8$ $c = 127.1$ $\alpha = \beta = \gamma = 90^\circ$	$a = 79.3$ $b = 101.0$ $c = 127.2$ $\alpha = \beta = \gamma = 90^\circ$	$a = 79.5$ $b = 100.8$ $c = 127.2$ $\alpha = \beta = \gamma = 90^\circ$	$a = 79.4$ $b = 101.5$ $c = 127.3$ $\alpha = \beta = \gamma = 90^\circ$	$a = 79.3$ $b = 101.3$ $c = 127.4$ $\alpha = \beta = \gamma = 90^\circ$
No. of molecules per ASU	2	2	2	2	2	2
Resolution range (Å)	30–2.20 (2.28–2.20)	30–1.95 (2.02–1.95)	30–1.60 (1.66–1.60)	30–1.58 (1.64–1.58)	30–2.10 (2.18–2.10)	30–1.90 (1.97–1.90)
No. of unique observations	51,498	73,721	134,069	139,830	60,078	80,727
Redundancy	5.6 (5.4)	6.8 (4.8)	4.7 (4.6)	6.0 (6.0)	7.3 (7.2)	7.3 (7.3)
Completeness (%)	97.5 (92.6)	97.8 (92.9)	99.1 (97.3)	100.0 (100.0)	99.9 (100.0)	99.3 (98.9)
$I/\sigma(I)$	13.9 (4.4)	21.4 (3.9)	19.0 (3.5)	23.7 (3.9)	13.0 (4.3)	16.7 (4.6)
R_{sym} (%) ^a	11.4 (46.3)	7.9 (39.5)	6.8 (45.8)	7.1 (45.5)	15.0 (46.9)	10.8 (44.3)

Numbers in parentheses are outer shell parameters.

^a $R_{\text{sym}} = \sum_{\text{hkl}} \sum_i |I_i(\text{hkl}) - \langle I(\text{hkl}) \rangle| / \sum_{\text{hkl}} \sum_i I_i(\text{hkl})$.

Crystals of the apo BGlu1 E386G glycosynthase and of its complexes were isomorphous with wild type BGlu1 crystals.²⁰ Diffraction data statistics and model parameters for the structures are summarized in Tables I and II.

Structures of the E386G glycosynthase, both cocrystallized and soaked with 10 mM *p*NP-β-cellobioside (*p*NPC2), an acceptor substrate, as well as with both 10 mM α-GlcF and 10 mM *p*NPC2 were determined, but *p*NPC2 was not observed in the active site. Nonetheless, α-GlcF was observed in the active site of ternary complex crystals and, if the crystals were soaked for longer periods, the crystal mosaicities generally increased and the crystals broke down after soaking for ~3 h, suggesting they may have produced long oligosaccharides that pushed the crystal matrix apart.

The α-GlcF complex structure is the first description of an α-GlcF donor substrate complex with a glycosynthase. The α-GlcF binds to the glycosynthase in a relaxed ⁴C₁ chair conformation stacked onto the indole ring of W433 at the -1 subsite in the same position as the 2-deoxy-2-fluoro-α-D-glucosyl moiety (G2F) in its covalent complex with wild type BGlu1 (PDB: 2RGM)²⁰ and made all the hydrogen bonds previously observed for that complex. The active site residues R86, N175, and N313, as well as the fluorine and O2 of the α-GlcF substrate interact with a single water molecule [Fig. 2(A,B)] that is also found in the structure of the covalent intermediate of BGlu1 with G2F, at the same position as the nucleophile carboxylate oxygen of apo wild type BGlu1 [Fig. 2(C), PDB: 2RGL].²⁰

The BGlu1 E386G glycosynthase activity was threefold higher than that of E386S and 19-fold higher than E386A glycosynthases.¹⁹ It was previously speculated that the high activity of the glycine glycosynthase of *Agrobacterium* sp. β-glucosidase⁴ may be due to binding of a water in a position similar to that of the serine glycosynthase hydroxyl group. However, there was no evidence for a water molecule in this position in the BGlu1 E386G structures. The position of the one water molecule found near the α-GlcF ligand was conserved with that of a water observed to stabilize the covalent complex of the BGlu1 with 2-deoxy-2-fluoroglucose and was similar to that of the water in the structure of Man26A β-mannanase E320G glycosynthase complexed with mannobiose.²³ The BGlu1 E386G mutant is likely an efficient glycosynthase due to the flexible positioning of this water molecule and surrounding protein atoms, as well as other dynamic issues that cannot be fully understood from the current static X-ray crystallographic structural model of the initial Michaelis complex structure.²⁴

Removal of either the catalytic base that activates the nucleophilic water or the hydroxyl that coordinates binding of the same water molecule to

Table II. Refinement Statistics

	E386G	E386G_ α -GlcF	E386G_C4	E386G_C5	S334A_C4	Y341A_C4
PDB code	3SCN	3SCO	3SCT	3SCU	3SCV	3SCW
Resolution range (Å)	30–2.20	30–1.95	30–1.60	30–1.58	30–2.10	30–1.90
No. of amino-acid residues	944	944	944	944	944	944
No. of protein atoms	7608	7608	7608	7608	7606	7594
No. of water molecules	651	833	912	922	781	902
Refined carbohydrate	None	α -GlcF	Cellotetraose	Cellopentaose	Cellotetraose	Cellotetraose
No. of carbohydrate atoms	None	24	135	112	135	135
No. of other hetero atoms	35	47	35	35	35	35
R_{factor} (%) ^a	19.8	17.2	19.1	17.4	17.4	17.0
R_{free} (%) ^b	24.0	20.7	20.8	19.2	20.6	19.8
Ramachandran statistics ^c						
Most favored regions (%)	88.7	89.8	89.1	89.4	90.0	89.3
Additionally allowed regions (%)	11.3	10.2	10.9	10.6	10.0	10.7
Outlier regions (%)	0.0	0.0	0.0	0.0	0.0	0.0
R.m.s.d. from ideality						
Bond distances (Å)	0.011	0.013	0.008	0.010	0.010	0.010
Bond angle (°)	1.35	1.39	1.36	1.38	1.33	1.32
Mean B factors (Å ²)						
All protein atoms	16.0	16.6	13.5	11.3	10.6	11.1
Waters	25.1	28.5	27.6	24.9	21.3	25.5
Hetero atoms	35.5	33.0	24.3	19.9	27.5	26.1
Carbohydrate atoms	None	14.0	42.9	37.6	36.1	30.3
Subsite −1 (A/B)	None	14.0/13.9	30.5/31.2	21.8/21.9	28.5/33.0	18.6/19.5
Subsite +1 (A/B)	None	None	39.9/39.6	33.0/34.8	35.4/37.6	21.8/23.9
Subsite +2 (A/B)	None	None	42.5/42.1	38.8/41.1	40.1/40.7	27.0/27.4
Subsite +3 (A/B)	None	None	45.7/46.1	42.5/44.0	43.5/44.3	40.3/40.9
Subsite +4 (A/B)	None	None	None	46.9/48.2	None	None

^a $R_{\text{factor}} = (\sum |F_o| - |F_c|) / \sum |F_o|$.^b Based on 5% of the unique observations not included in the refinement.^c Ramachandran values were determined from PROCHECK.²²

disrupt hydrolysis but allow transglycosylation, can produce a glycosynthase from an inverting glycosidase.^{3,6} The prominent position of the water stabilizing the α -GlcF suggests that modulating binding of a critical water molecule may be important for production of an efficient glycosynthase from a retaining glycosidase, as well.

Structures of BGlul E386G glycosynthase complexed with celooligosaccharides

The oligosaccharides cellotetraose and cellopentaose, which could serve as glycosynthase acceptors or products, were soaked into the E386G crystals and the structures were determined, with the X-ray diffraction statistics shown in Tables I and II. The observed electron densities, shown in Figure 3(A), and Supporting Information Figures S1 and S2, indicated that these oligosaccharides bound with the nonreducing residue in the −1 site as glycosynthase products or hydrolase substrates, as previously seen for the E176Q acid–base mutant,²¹ rather than in the position of acceptor substrates. To obtain sufficient ligand occupancies in the active site, the crystals were soaked in cryoprotectant saturated with the cellotetraose or cellopentaose, rather than the 2 mM cellotetraose and 1 mM cellopentaose used for the E176Q mutant. This resulted in the binding of

an extracellotetraose on the surface of the protein between Molecule A of one asymmetric unit and Molecule B of another asymmetric unit, but this is unlikely to be of any biological significance.

The comparisons of the structures of BGlul E386G in complex with cellotetraose and cellopentaose with those of the BGlul E176Q mutant in complex with cellotetraose and cellopentaose (PDB: 3F5J and 3F5K)²¹ in Figure 3(B) and Supporting Information Figure S3(A) reveal very similar positions of the amino acid residues, except the E176 (the acid/base, which is mutated to glutamine in the BGlul E176Q complex structures) and N245 side chain positions were shifted slightly. Likewise the Glc1, Glc2, Glc3, and Glc4 glucosyl residues from the nonreducing ends of cellotetraose and cellopentaose at subsites −1, +1, +2, and +3 in the active site of BGlul E386G adopt similar positions in the two complexes, although Glc5 in the active site of BGlul E386G had lower electron density than that in the cellopentaose complex with BGlul E176Q, and a slightly different position [Fig. 3(B)].

Hydrogen bonding interactions between the glucosyl residues of cellotetraose or cellopentaose, ordered water molecules and amino acid residues in the active site of BGlul E386G are shown in Figure 3(C). The hydrogen bonds and aromatic-sugar

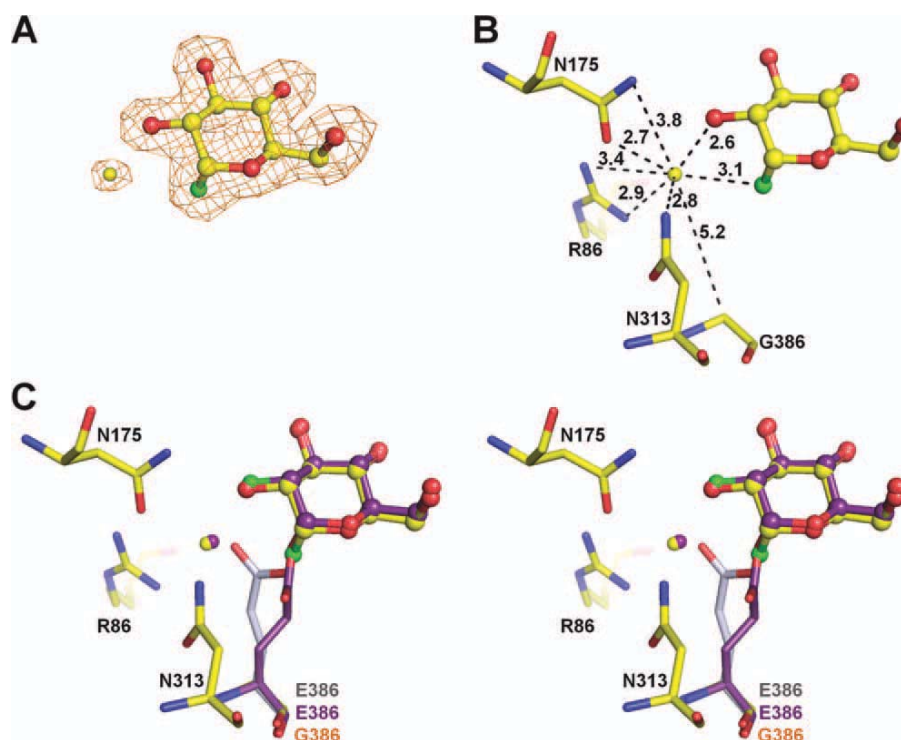


Figure 2. Comparison of active sites of rice BGlu1 E386G glycosynthase and free BGlu1 and its 2-fluoroglucosyl enzyme intermediate. (A) The electron densities of α -GlcF and water in the active sites of E386G (yellow carbons). The electron density ($F_o - F_c$) maps were calculated with all heteroatoms omitted from the active site, and are shown in brown mesh at $+3\sigma$. An active site water molecule is seen in the wild type covalent intermediate as well as in the glycosynthase mutant complex with α -GlcF. This water is bound at the same location as the catalytic nucleophile in the apo wild type enzyme. The superimposition of rice BGlu1 E386G (yellow carbons) and apo BGlu1 (gray carbons) with the structure of the 2-fluoroglucosyl enzyme intermediate (purple carbons) is shown in wall-eyed stereo in (C). The α -GlcF is represented as balls and sticks and the water molecule is represented as a ball. Oxygen is shown in red, nitrogen in blue, and fluorine in green. Distances (in Å) are shown from the one water molecule to the surrounding polar groups and the mutated residue. No other water molecules were observed in this part of the active site.

stacking interactions of the Glc3, Glc4, and Glc5 residues at subsites +2, +3, and +4 are similar to those in the BGlu1 E176Q complex with cellopentaose. The slight shifts in the positions of the oligosaccharide sugar moiety and of the N245 side chain allowed the N245 N δ to form hydrogen bonds with Glc2 O6, at subsite +1, and Glc3 O2, at subsite +2, rather than the bonds to Glc3 O2 and O3 observed in the E176Q complexes.²¹ This shift reflects an inherent plasticity in oligosaccharide binding by this enzyme, as demonstrated by the hydrogen bonding of N245 to sugars in either the +1 subsite (in BGlu1 E176Q with laminaribiose) or the +2 subsite (in BGlu1 E176Q with cellopentaose and cellotetraose),²¹ or both (in the BGlu1 E386G structures presented here).

Other residues that make obvious interactions with the oligosaccharides include S334 and Y341, residues that interact primarily with Glc4 and Glc5 at the +3 and +4 subsites. The Y341 residue interacts via a water-mediated hydrogen bond to the glycosidic oxygen between subsite +2 and +3 and aromatic-sugar stacking interactions at subsites +3 and +4,

while the side chain and α -carbonyl of S334 form water-mediated hydrogen bonds at subsite +3.

Oligosaccharide-binding residue mutants

Based on the observed interactions between S334 and Y341 and the oligosaccharides, the BGlu1 mutants S334A, Y341A, and Y341L were generated to investigate how they would affect hydrolysis and synthesis of cellobiosaccharides. The activity of the N245V mutant, which was previously shown to have higher K_m values for *p*NP- β -D-glucopyranoside (*p*NPGlc), cellobiose, and cellotriose than does wild type BGlu1,²¹ was also investigated with longer oligosaccharides. Only slightly higher values of K_m and slightly lower catalytic efficiency values (k_{cat}/K_m) for cellotriose, cellotetraose, and cellopentaose were seen for the BGlu1 S334A, Y341A, and Y341L mutants compared to wild type BGlu1 (Table III). The BGlu1 N245V mutant had a 10-fold lower k_{cat}/K_m value than wild type for cellotriose, but only threefold lower for cellotetraose and cellopentaose. Kinetic subsite analysis by the method of Hiromi *et al.*²⁵ (with the assumption of an

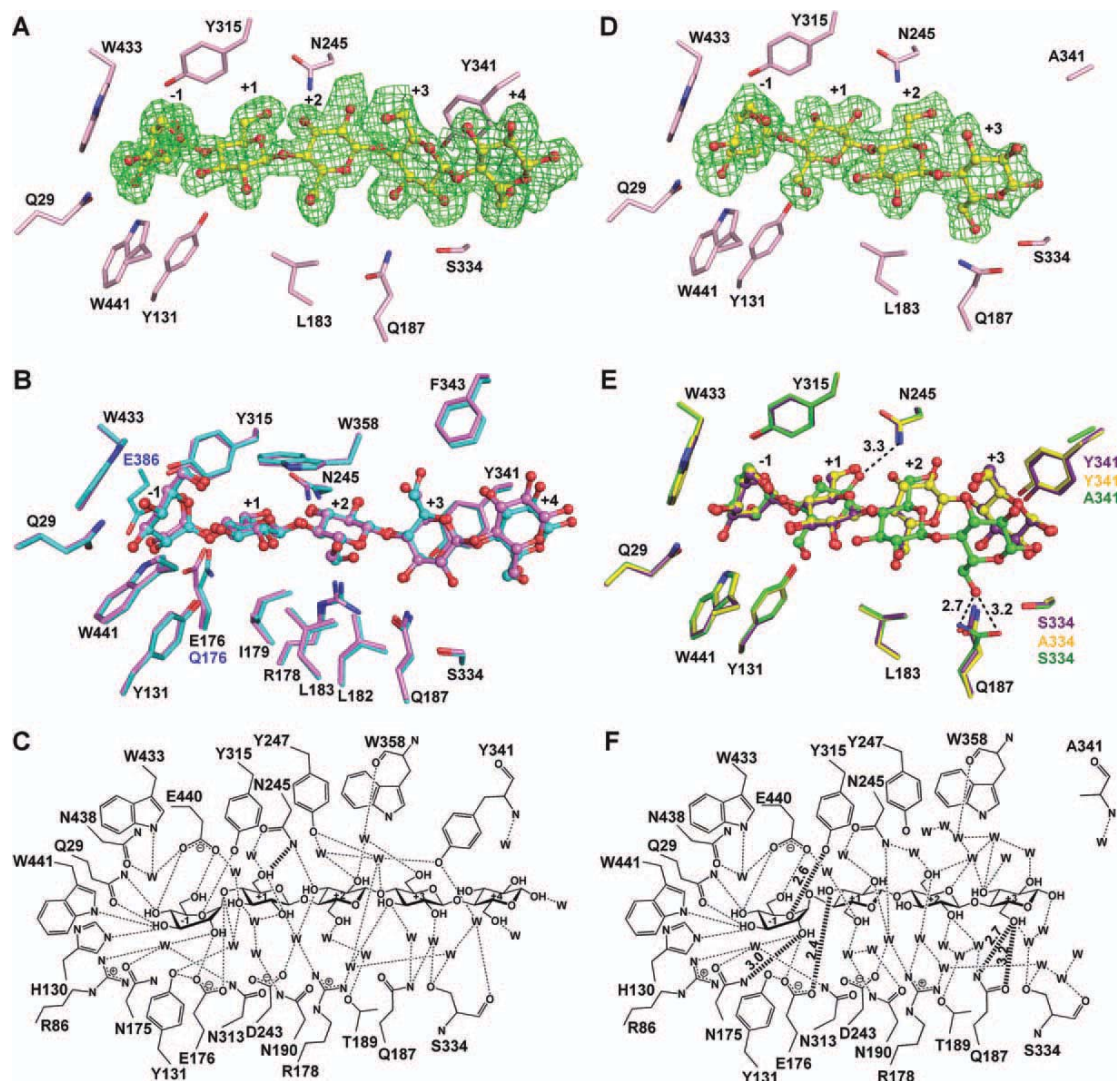


Figure 3. Cellooligosaccharide binding in the active site of BGlu1 E386G and BGlu1 E386G/Y341A. (A) The $2F_o - F_c$ electron density in the active site of BGlu1 E386G with cellopentaose. The $2F_o - F_c$ map was calculated with heteroatoms in the active sites, and are shown in green mesh at $+1\sigma$. The ligands from the final structures are shown in ball and stick representation. The carbons are colored in yellow and oxygen is shown in red. The interacting amino acids in the active sites are shown with carbons in pink. (B) Superimposition of the active sites of the BGlu1 E386G glycosynthase complex with cellopentaose (pink) and the BGlu1 E176Q acid-base mutant complex with cellopentaose (cyan). Superimposition of these structures with the cellotetraose complexes can be seen in Supporting Information Figure S3(A). (C) Diagram of the strong hydrogen bonds between the glucose residues of cellopentaose, surrounding amino acid residues and the network of water molecules (W) in the active site of BGlu1 E386G. Hydrogen bonds with measured distances of 3.2 Å or less are shown as dashed lines to indicate the hydrophilic interactions. The direct hydrogen bond between N245 and Glc2, which is not seen in the complexes of cellotetraose and cellopentaose with BGlu1 E176Q, is shown as a darker dashed line. Hydrogen bonding for the cellotetraose complex was similar, as shown in Supporting Information Figure S3(B). (D) The $2F_o - F_c$ electron density in the active site of BGlu1 E386G/Y341A with cellotetraose. (E) Superimposition of the active sites of BGlu1 E386G (violet carbons), BGlu1 E386G/S334A (yellow carbons) and BGlu1 E386G/Y341A (green carbons) in complexes with cellotetraose. The direct hydrogen bonds between the glucosyl moieties in the +1 to +3 subsites of the active site of BGlu1 E386G/Y341A were identified as O6 of Glc4 with Gln187 N ϵ and O ϵ and a weak interaction of O3 of Glc2 with Asn245 N δ , which are shown as dashed lines. The ligands are represented as balls and sticks. (F) Diagram of the strong hydrogen bonds in active site of the complex of the E386G/Y341A mutant with cellotetraose, showing the flip of the glucosyl residues in subsites +1 to +3 that results in a completely different network. The direct hydrogen bonds with Q187 and shortened hydrogen bonds from N175 and Y315 are shown in darkened dashed lines with their distances in Å marked on the lines.

intrinsic k_{cat}) verified that the S334A, Y341A, and Y341L mutations had little effect on the apparent affinity at the +3 and +4 subsites. Even the very low

efficiency of the N245V mutant in the hydrolysis of cellotriose appeared to be compensated by binding subsequent glucosyl residues in cellotetraose and

Table III. Kinetic Parameters of BGlu1 Wild Type and Mutants for Hydrolysis of Cellotriose, Cellotetraose, and Cellopentaose

Substrate	Kinetic parameters	BGlu1	S334A	Y341A	Y341L	N245V
Cellotriose	K_m (mM)	0.50 ± 0.05	0.62 ± 0.05	0.97 ± 0.07	0.97 ± 0.05	11.2 ± 0.8
	k_{cat} (s^{-1})	15.6 ± 0.6	20.3 ± 0.7	26.5 ± 0.7	23.3 ± 0.5	36.1 ± 0.9
	k_{cat}/K_m ($M^{-1} s^{-1}$)	31,000	33,000	27,000	24,000	3,230
	K_m (mM)	0.37 ± 0.04	0.52 ± 0.04	0.70 ± 0.05	0.54 ± 0.03	5.3 ± 0.2
Cellotetraose	k_{cat} (s^{-1})	23.7 ± 1.1	29.1 ± 1.2	35.5 ± 1.0	26.5 ± 0.4	103.5 ± 1.6
	k_{cat}/K_m ($M^{-1} s^{-1}$)	64,000	56,000	58,000	49,000	20,000
	Subsite affinity	1.8	1.4	1.6	1.8	4.5
	(subsite +3, $kJ mol^{-1}$)					
Cellopentaose	K_m (mM)	0.28 ± 0.02	0.36 ± 0.03	0.44 ± 0.02	0.36 ± 0.02	3.83 ± 0.18
	k_{cat} (s^{-1})	27.5 ± 1.2	31.8 ± 1.5	32.3 ± 0.6	25.6 ± 0.4	139 ± 3
	k_{cat}/K_m ($M^{-1} s^{-1}$)	98,000	88,000	73,000	71,000	36,300
	Subsite affinity	1.1	1.1	0.93	0.93	1.6
	(subsite +4, $kJ mol^{-1}$)					

Subsite affinities were calculated by the method of Hiromi et al.²⁵

$A_i = -\Delta G = RT \ln[(k_{cat}/K_m)_i / (k_{cat}/K_m)_{n-1}]$ as a means of estimating the effect of each successive sugar residue.

cellopentaose, giving the mutant a high apparent subsite +3 affinity.

Transglycosylation

Previously, we hypothesized that the long celooligosaccharide-binding cleft of rice BGlu1 allowed its glycosynthase to synthesize longer oligosaccharides than do other exoglycosidase-based glycosynthases.¹⁹ Because the mutations generated were in the oligosaccharide binding site and had small effects on celooligosaccharide hydrolysis, the same mutations of E386G glycosynthase were tested to see whether they affected the production of long oligosaccharides in transglycosylation reactions with α -GlcF donor and pNPC2 acceptor. When the E386G glycosynthase and its four mutants, S334A, Y341A, Y341L, and N245V, were compared with α -GlcF donor and pNPC2 acceptor at molar ratios of 1:2, 1:1, 2:1, and 5:1 transparent product solutions resulted for E386G/Y341A, E386G/Y341L, and E386G/N245V, but the reactions of 2:1 donor:acceptor with E386G and 5:1 donor to acceptor with E386G and E386G/S334A resulted in precipitated products (data not shown). These insoluble products were previously shown to comprise celooligosaccharides of up to 11 glucosyl residues for the E386G glycosynthase.¹⁹

The soluble products were monitored by thin-layer chromatography (TLC, Fig. 4 and Supporting Information Fig. S4) and HPLC (Supporting Information Table SI and Supporting Information Fig. S5). The S334A mutant has similar glycosynthase activity to E386G and synthesized long pNP-oligosaccharide products up to pNP- β -cellonanaoside (pNPC9), as judged by LC-MS analysis (Supporting Information Fig. S5). In TLC, the longer products appeared as a dark spot at the origin. The N245V mutant synthesizes small but significant amounts of pNP- β -cellopentaoside (pNPC5) for all reactions and longer pNP-celooligosaccharides for the reaction at

the 5:1 donor:acceptor ratio, based on TLC (Supporting Information Fig. S4), but pNP-celloheptaoside (pNPC7) was the longest product detected in the soluble products of the 1:1 ratio reaction (Supporting

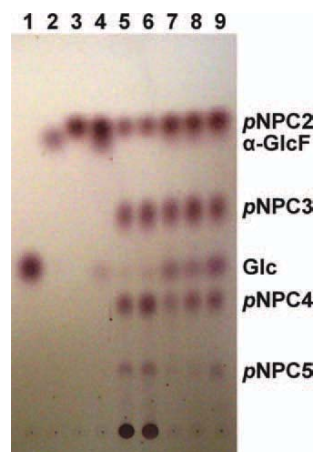


Figure 4. TLC detection of products of transglucosylation reactions between α -GlcF donor and pNPC2 acceptor at a molar ratio of 1:1. The reactions were incubated with E386G and E386G mutants in NH_4HCO_3 (pH 7.0) at 30°C for 16 h. The soluble products were separated by TLC, using EtOAc-MeOH-water (7:2.5:1) as a developing solvent and detected by exposure to 10% sulfuric acid in ethanol followed by charring to detect carbohydrates. Lane 1, glucose; Lane 2, α -GlcF; Lane 3, pNPC2; Lane 4, control reaction without enzyme; Lanes 5–9, reactions with glycosynthase enzymes: Lane 5, E386G; Lane 6, E386G/S334A; Lane 7, E386G/Y341A; Lane 8, E386G/Y341L; and Lane 9, E386G/N245V. The TLC detection of the products of the reactions between α -GlcF donor and pNPC2 acceptor at molar ratios of 1:2, 2:1, and 5:1 are shown in Supporting Information Figure S4. Based on the LC-MS analysis (Supporting Information Table SI and Supporting Information Fig. S5) and Hommalai et al.,¹⁹ the products were identified as pNPC3, pNPC4, pNPC5 and longer pNP-celooligosaccharides. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Information Table SI). However, the Y341A and Y341L mutants synthesized only very small amounts of longer products (*p*NPC5) unless the donor to acceptor ratio was increased to 5:1, although a tiny peak of *p*NPC6 could be detected by LC-MS of the 1:1 ratio reactions (Supporting Information Table SI). Small and similar amounts of the hydrolysis product *p*NPGLc were also seen in the HPLC of E386G glycosynthase and its mutants, due to ~0.1% BGlu1 E386 β -glucosidase contamination in the expression system (Pengthaisong *et al.*, unpublished). Thus, the BGlu1 E386G/Y341A, E386G/Y341L, and E386G/N245V glycosynthases synthesize shorter *p*NP-cellooligosaccharides than does the unmodified E386G glycosynthase and the E386G/S334A mutant glycosynthase under the same conditions.

Another hydrolysis product, glucose, which is produced from both α -GlcF and *p*NP-oligosaccharides, was more evident in the aglycone binding cleft mutants (Fig. 4). Because only a small amount of *p*NPGLc was produced from *p*NPC2 hydrolysis in each case (<1% of the total *p*NP-oligosaccharides, Supporting Information Table SI), and the binding site mutants of the hydrolase are slightly less efficient at release of Glc from *p*NP-oligosaccharides, this is unlikely to result from hydrolysis of *p*NP oligosaccharides. Therefore, the extra glucose in the mutants is likely a result of the less efficient acceptor binding, resulting in an increased partitioning of the reaction to hydrolysis of α -GlcF relative to transglycosylation.

Structures of oligosaccharide-binding site mutants

The relatively small effects of mutation of Y341 on the hydrolysis of oligosaccharides was initially surprising, particularly for those with glucose residues that would stack on its aromatic ring at subsites +3 and +4. To explain the effects of these mutations, the X-ray crystal structures of E386G glycosynthase mutants complexed with cellotetraose and cellopentaose were investigated by X-ray crystallography. Cellotetraose was observed in the active site of the BGlu1 E386G/S334A and E386G/Y341A mutants. The glucosyl residues of cellotetraose and amino acid residues in the active site of the structure of the BGlu1 E386G/S334A glycosynthase mutant were similar to those in the cellotetraose complex of BGlu1 E386G [Fig. 3(E)], while the Glc2, Glc3, and Glc4 residues at subsites +1, +2, and +3 in the structure of the Y341A glycosynthase mutant were flipped nearly 180° compared to the structure of BGlu1 E386G with cellotetraose [Fig. 3(D–F)]. The Glc3 and Glc4 residues in the active site of the BGlu1 E386G/Y341A mutant moved from their positions in the active site of BGlu1 E386G, so that O6 of Glc4 directly hydrogen bonds to Q187 N ϵ at the new subsite +3, while N245 may now form only a weak interaction with Glc2 O3 (3.3 Å), rather than

the hydrogen bonds with Glc2 and Glc3 observed in the E386G complex described earlier.

The alternative binding mode observed for cello-tetraose in the structures of the Y341A glycosynthase mutant appears to allow the hydrolytic reaction to occur, since the -1 subsite is filled correctly. However, synthesis of oligosaccharides by glycosynthase requires the oligosaccharide acceptor to bind with the nonreducing residue in the weakly binding +1 subsite, where the shift in position may be more destabilizing. The fact that only the initial binding mode is observed in BGlu1 E176Q and E386G crystals without additional mutations suggests that it is the most stable binding mode. However, the small effects of the Y341 mutants on hydrolysis suggest that the energy of binding in the second mode observed in the BGlu1 E386G/Y341A mutant complex with cellotetraose may be only slightly less favorable than that in the previously observed mode. This suggests that the wild type BGlu1 can bind cel-looligosaccharides in more than one position, which may allow efficient funneling of these substrates into the active site.

Conclusion

In conclusion, the crystal structure of the BGlu1 E386G glycosynthase with α -GlcF showed a stabilizing water molecule at the position of the nucleophile of apo BGlu1 β -glucosidase, the flexible positioning of which is likely the key to high activity in the E386G glycosynthase. Plasticity of the active site also appears to be important for binding oligosaccharides, since N245 can make direct hydrogen bonds in either the +1 or +2 subsite, and cellotetraose glucosyl residues may bind with their faces in different orientations. The greater effect of loss of Y341 on glycosynthase activity suggests that the binding of acceptors is not as flexible, so the effects of mutation of aglycone-binding residues on hydrolysis and transfer reactions may not be strictly comparable.

Materials and Methods

Mutation of rice BGlu1 and BGlu1 E386G

The mutations for BGlu1 N245V and the BGlu1 E386G (previously designated E414G, based on its position in the BGlu1 precursor protein) have been described previously.^{17,19} The mutations S334A, Y341A, Y341L, E386G/S334A, E386G/Y341A, E386G/Y341L, and E386G/N245V were made in the rice *bglu1* and rice *bglu1* E386G cDNA^{17,19} with the QuikChange site-directed mutagenesis kit (Stratagene, La Jolla, CA). The primers used were as follows: 5'-G ACA CCG ACG AGT TAC GCA GCC GAT TGG CAG-3' and its reverse complement for S334A; 5'-C GAT TGG CAG GTT ACC GCT GTT

TTT GCG AAA AAC GGC-3' and its reverse complement for Y341A; 5'-CC GAT TGG CAG GTT ACC CTT GTT TTT GCG AAA AAC GGC-3' and its reverse complement for Y341L; and 5'-AA GTT GGA ATA GTT CTG GAC TTC GTA TGG TAT GAA GCT CTT TCC AAC TC-3' and its reverse complement for N245V (the mutated nucleotides are underlined).

Protein expression, purification, and crystallization

The recombinant proteins of BGlu1, its glycosynthase mutants and the mutants described above were expressed and purified as previously described for wild type BGlu1.^{17,20} The purified protein concentration was determined by measuring absorbance at 280 nm. Extinction coefficients ϵ_{280} of 113,560 M⁻¹ cm⁻¹ for most of the BGlu1 proteins and ϵ_{280} of 112,280 M⁻¹ cm⁻¹ for BGlu1 Y341A and Y341L were calculated by the method of Gill and von Hippel.²⁶

The BGlu1 E386G and its mutants were crystallized with and without ligands, including 10 mM α -GlcF, 10 mM pNPC2, both 10 mM α -GlcF and 10 mM pNPC2, 2 mM cellotetraose and 2 mM cellopentaose by hanging drop vapor diffusion with microseeding, optimized in 16–26%, polyethylene glycol monomethyl ether (PEG MME) 5000, 0.12–0.26 M (NH₄)₂SO₄, and 2–5 mg mL⁻¹ rice BGlu1 in 0.1M MES (pH 6.7) at 288 K, as previously described.²⁰ Before flash cooling in liquid nitrogen, the crystals with 10 mM α -GlcF, 10 mM pNPC2, and both 10 mM α -GlcF and 10 mM pNPC2 were soaked in cryo-solution (18% (v/v) glycerol in precipitant solution) containing the same concentrations of ligands for 1–5 min, while the crystals with cellotetraose and cellopentaose were soaked in cryo solution saturated with the ligands (~75–100 mM cellotetraose or 25 mM cellopentaose).

Data collection and processing

Preliminarily, diffraction was tested in a nitrogen cryostream from a 700 series Cryostream Cooler (Oxford Cryosystems, Oxford, England), using a Cu K α rotating anode source mounted on a MicroSTAR generator operating at 45 kV and 60 mA connected to Rayonix SX-165 CCD detector at the Synchrotron Light Research Institute (SLRI, Nakhon Ratchasima, Thailand). Datasets were collected with 1.0 Å wavelength X-rays and an ADSC Quantum 315 CCD detector on the BL13B1 beamline at the National Synchrotron Radiation Research Center (NSRRC in Hsinchu, Taiwan). The crystals were maintained at 105 K during data collection with a nitrogen cold stream (Oxford Instruments). Data were processed and scaled with the HKL-2000 package.²⁷

Structure solution and refinement

The crystals of glycosynthases were isomorphous with wild type BGlu1 crystals,²⁰ allowing the structures to be solved by rigid body refinement with the free BGlu1 structure (PDB: 2RGL) in REFMAC5²⁸ with the two molecules in the asymmetric unit refined as independent domains. The refinement was executed with REFMAC5 with tight noncrystallographic symmetry (NCS) restraints and model building with Coot.²⁹ Water molecules were added with the Coot and ARP/wARP programs in the CCP4 suite. Glucosyl residues were built into the electron densities in the shapes that fit the densities best (⁴C₁ relaxed chairs or ¹S₃ skew boats) and refined. The refined sugar residue coordinates were assigned their final conformation designation according to their Cremer-Pople parameters,³⁰ calculated by the Cremer-Pople parameter calculator of Dr. Shinya Fushinobu (University of Tokyo, <http://www.ric.hi-ho.ne.jp/asfushi/>). The occupancy of cellopentaose binding in the active site of E386G was refined to 0.7 by setting at different values and refining to find the occupancy that minimized the values of the temperature factors (*B*-factors) of the ligand and the free residual factor (*R*_{free}). The final models were analyzed with PROCHECK²² and MolProbity.³¹ The figures of protein structures were generated in PyMol (Schrödinger LC).

Oligosaccharide hydrolysis

Wild type and mutant BGlu1 enzymes were purified by immobilized metal affinity chromatography (IMAC), enterokinase digestion and IMAC, as previously described.²⁰ The activities of the enzymes toward cellotriose, cellotetraose, and cellopentaose were assayed in 50 mM sodium acetate buffer (pH 5.0) at 30°C, as previously described.¹⁷ The kinetic parameters *V*_{max} and *K*_m were calculated by nonlinear regression of Michaelis–Menten plots with the Grafit 5.0 computer program (Erithacus Software, Horley, UK) and divided by the protein concentrations to determine the apparent *k*_{cat} and *k*_{cat}/*K*_m. Relative subsite affinities were determined by the method of Hiromi *et al.*²⁵

Oligosaccharide synthesis by BGlu1 glycosynthase enzymes

The BGlu1 E386G glycosynthase and its four mutants, BGlu1 E386G/S334A, BGlu1 E386G/Y341A, BGlu1 E386G/Y341L, and BGlu1 E386G/N245V, were purified as described above and the buffer was exchanged with 50 mM phosphate buffer (pH 6.0). The proteins were incubated with α -GlcF donor and pNPC2 acceptor in 150 mM ammonium bicarbonate buffer (pH 7.0) at 30°C for 16 h.¹⁹ The reaction mixture was centrifuged at 13,000g for 5 min. The enzymes in the supernatant were removed

by centrifugal filtration (Microcon YM-10, Millipore, USA) and the soluble products were analyzed by electrospray ionization-mass spectrometry (ESI-MS) and monitored by TLC (silica gel 60 F₂₅₄, Merck, Germany) using 7:2.5:1 ethyl acetate (EtOAc)-methanol (MeOH)-water as solvent. Plates were visualized under ultraviolet (UV) light and by exposure to 10% sulfuric acid in ethanol followed by charring. Five microliters of the products from the reactions with 1:1 molar ratios of α -GlcF donor:pNPC2 acceptor were loaded onto a ZORBAX carbohydrate column (4.6 mm $\times$ 250 mm, Agilent, USA) connected to an Agilent 1100 series LC-MS. The column was eluted with a linear gradient from 90 to 50% acetonitrile in water over 30 min at a flow rate of 1 mL min⁻¹. The eluted peaks were detected at 300 nm with a UV-visible diode array detector and the product masses determined by mass spectrometry.

Acknowledgments

The E386G glycosynthase vector was originally generated by Greengrai Hommalai and Watchalee Chuenchor and the authors are grateful to these investigators whose ideas contributed to the project. Portions of this research were conducted at the National Synchrotron Radiation Research Center.

References

- Wang Q, Graham RW, Trimbur D, Warren RAJ, Withers SG (1994) Changing enzymatic reaction mechanism by mutagenesis, conversion of a retaining glucosidase to an inverting enzyme. *J Am Chem Soc* 116: 11594–11595.
- Williams SJ, Withers SG (2000) Glycosyl fluorides in enzymatic reactions. *Carbohydr Res* 327:27–46.
- Honda YW, Kitaoka M (2006) The first glycosynthase derived from an inverting glycoside hydrolase. *J Biol Chem* 281:1426–1431.
- Mayer C, Zechel DL, Reid SP, Warren RAJ, Withers SG (2000) The E358S mutant of *Agrobacterium* sp. β -glucosidase is a greatly improved glycosynthase. *FEBS Lett* 466:40–44.
- Kim YW, Lee SS, Warren RAJ, Withers SG (2004) Directed evolution of a glycosynthase from *Agrobacterium* sp. increases its catalytic activity dramatically and expands its substrate repertoire. *J Biol Chem* 279: 42787–42793.
- Honda Y, Fushinobu S, Hidaka M, Wakagi T, Shoun H, Taniguchi H, Kitaoka M (2008) Alternative strategy for converting an inverting glycoside hydrolase into a glycosynthase. *Glycobiology* 18:325–330.
- Wilkinson SM, Liew CW, Mackay JP, Salleh HM, Withers SG, McLeod MD (2008) *Escherichia coli* glucuronosynthase, an engineered enzyme for the synthesis of beta-glucuronides. *Org Lett* 10:1585–1588.
- Addington T, Calisto B, Alfonso-Prieto M, Rovira C, Fita I, Planas A (2011) Re-engineering specificity in 1,3-1,4- β -glucanase to accept branched xyloglucan substrates. *Proteins* 79:365–375.
- Cobucci-Ponzano B, Strazzulli A, Rossi M, Moracci M (2011) Glycosynthases in biocatalysis. *Adv Synth Catal* 353:2284–2300.
- Perugino G, Trinccone A, Rossi M, Moracci M (2004) Oligosaccharide synthesis by glycosynthases. *Trends Biotechnol* 22:31–37.
- Faijes M, Saura-Valls M, Pérez X, Conti M, Planas A (2006) Acceptor dependent regioselectivity of glycosynthase reactions by *Streptomyces* E383A β -glucosidase. *Carbohydr Res* 341:2055–2065.
- Müllegger J, Chen H, Chan WY, Reid SP, Jahn M, Warren RAJ, Salleh HM, Withers SG (2006) Thermostable glycosynthases and thioglycoligases derived from *Thermotoga maritima* β -glucuronidase. *Chem Biochem* 7: 1028–1030.
- Fort S, Boyer V, Greffe L, Davies GJ, Moroz O, Christensen L, Schülein M, Cottaz S, Driguez H (2000) Highly efficient synthesis of β -(1,4)-oligo- and -polysaccharides using a mutant cellulase. *J Am Chem Soc* 122:5429–5437.
- Fairweather JK, Faijes M, Driguez H, Planas A (2002) Specificity studies of *Bacillus* 1,3-1,4- β -glucanases and application to glycosynthase-catalyzed transglycosylation. *Chem Biochem* 3:866–873.
- Hrmova M, Imai T, Rutten SJ, Fairweathers JK, Pelosi L, Bulone V, Driguez H, Fincher GB (2002) Mutated barley (1,3)- β -D-glucan endohydrolases synthesize crystalline (1,3)- β -D-glucans. *J Biol Chem* 277: 30102–30111.
- Kim YW, Chen H, Withers SG (2005) Enzymatic transglucosylation of xylose using a glycosynthase. *Carbohydr Res* 340:2735–2741.
- Opassiri R, Ketudat Cairns JR, Akiyama T, Wara-Aswapati O, Svasti J, 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, Svasti J, Esen A, Ketudat Cairns JR (2004) β -glucosidase, exo- β -glucanase and pyridoxine transglucosylase activities of rice BGlu1. *Biochem J* 379:125–131.
- Hommalai G, Withers SG, Chuenchor W, Ketudat Cairns JR, Svasti J (2007) Enzymatic synthesis of cello-oligosaccharides by mutated rice β -glucosidases. *Glycobiology* 17:744–753.
- Chuenchor W, Pengthaisong S, Robinson RC, Yuvaniyama J, Oonant W, Bevan DR, Esen A, Chen CJ, Opassiri R, Svasti J, Ketudat Cairns JR (2008) Structural insights into rice BGlu1 beta-glucosidase oligosaccharide hydrolysis and transglycosylation. *J Mol Biol* 377:1200–1215.
- Chuenchor W, Pengthaisong S, Robinson RC, Yuvaniyama J, Svasti J, Ketudat Cairns JR (2011) The structural basis of oligosaccharide binding by rice BGlu1 beta-glucosidase. *J Struct Biol* 173:169–179.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK, a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26:283–291.
- Jahn M, Stoll D, Warren RAJ, Szabó L, Singh P, Gilbert HJ, Ducros VM, Davies GJ, Withers SG (2003) Expansion of the glycosynthase repertoire to produce defined manno-oligosaccharides. *Chem Commun* 12: 1327–1329.
- Ducros VM, Tarling CA, Zechel DL, Brzozowski AM, Frandsen TP, von Ossowski I, Schülein M, Withers SG, Davies GJ (2003) Anatomy of glycosynthesis, structure and kinetics of the *Humicola insolens* Cel7B E197A and E197S glycosynthase mutants. *Chem Biol* 10: 619–628.
- Hiromi K, Nitta Y, Namura C, Ono S (1973) Subsite affinities of glucoamylase, examination of the validity

- of the subsite theory. *Biochim Biophys Acta* 302: 362–375.
26. Gill SC, von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* 182:319–326.
 27. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
 28. Murshudov GN, Lebedev A, Vagin AA, Wilson KS, Dodson EJ (1999) Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr D* 55:247–255.
 29. Emsley P, Cowtan K (2004) Coot, model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.
 30. Cremer D, Pople JA (1975) A general definition of ring puckering coordinates. *J Am Chem Soc* 97: 1354–1358.
 31. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB III, Snoeyink J, Richardson JS, Richardson DC (2007) MolProbity, all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35: W375–W383.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at SciVerse ScienceDirect

Carbohydrate Research

journal homepage: www.elsevier.com/locate/carres

Note

Exchanging a single amino acid residue generates or weakens a +2 cellooligosaccharide binding subsite in rice β -glucosidases

Sompong Sansenya^a, Janjira Maneesan^{a,†}, James R. Ketudat Cairns^{a,b,*}^a Schools of Biochemistry and Chemistry, Institute of Science, Suranaree University of Technology, 111 University Avenue, Nakhon Ratchasima 30000, Thailand^b Laboratory of Biochemistry, Chulabhorn Research Institute, Vibhavadee Rangsit Road, Bangkok 10210, Thailand

ARTICLE INFO

Article history:

Received 31 October 2011

Received in revised form 8 January 2012

Accepted 18 January 2012

Available online 28 January 2012

Keywords:

Rice
Glycoside hydrolase
Oligosaccharide
Mutagenesis
Kinetics

ABSTRACT

Os3BGlu6, Os3BGlu7, and Os4BGlu12 are rice glycoside hydrolase family 1 β -glucosidases, the structures of which have been solved by X-ray crystallography. In complex structures, Os3BGlu7 residue Asn245 hydrogen bonds to the second sugar in the +1 subsite for laminaribiose and the third sugar in the +2 subsite for cellotetraose and cellopentaose. The corresponding Os3BGlu6 residue, Met251, appears to block the binding of cellooligosaccharides at the +2 subsite, whereas His252 in this position in Os4BGlu12 could hydrogen bond to oligosaccharides. Mutation of Os3BGlu6 Met251 to Asn resulted in a 15-fold increased k_{cat}/K_m value for hydrolysis of laminaribiose compared to wild type Os3BGlu6 and 9 to 24-fold increases for cellooligosaccharides with degrees of polymerization (DP) of 2–5. On the other hand, mutation of Os3BGlu7 Asn245 to Met decreased the k_{cat}/K_m of hydrolysis by 6.5-fold for laminaribiose and 17 to 30-fold for cellooligosaccharides with DP >2, while mutation of Os4BGlu12 His252 to Met decreased the corresponding k_{cat}/K_m values 2 to 6-fold.

© 2012 Elsevier Ltd. All rights reserved.

In plants, glycosyl hydrolase family 1 (GH1) β -glucosidases act in many processes, such as chemical defense against herbivores, hydrolysis of cell wall-derived oligosaccharides, phytohormone regulation and lignification.^{1,2} In order to carry out their various functions in plants, GH1 enzymes must have diverse specificities. The basis for these diverse substrate specificities in GH1 has been probed by many structural and functional studies, especially in the monocot plastid β -glucosidases maize (*Zea mays*) Glu1 (ZmGlu1), sorghum (*Sorghum bicolor*) dhurrinase 1 (SbDhr1), wheat (*Triticum aestivum*) TaGlu1b, and rye (*Secale cereale*).^{3–8} In general these studies identified nonconserved binding cleft residues that interact with the specific substrates, but these residues varied in different enzymes, even among those with similar substrate specificities. Both surface residues and underlying residues that shape the active site determine aglycone-binding, as also appears to be the case for glycone specificity.⁹

The 40 rice (*Oryza sativa*)¹⁰ and 48 Arabidopsis (*Arabidopsis thaliana*)¹¹ genes found in genomic sequences have been divided into eight protein-sequence-based phylogenetic clusters that

contain both rice and Arabidopsis proteins.¹⁰ To date, only a few rice GH1 β -glucosidase isoenzymes have been characterized for their possible functions, which include hydrolyzing gibberellin glucoconjugates,¹² cell wall-derived oligosaccharides,¹³ and tuberonic acid β -glucoside.¹⁴ Rice GH1 enzymes characterized at the molecular and structural level include Os3BGlu6, Os3BGlu7 (also called rice BGlu1), and Os4BGlu12, which all hydrolyze oligosaccharides and glycosides and share 51–53% sequence identity but represent separate phylogenetic clusters.^{10,15–18}

Rice Os3BGlu7 and the other members of its phylogenetic cluster that have been characterized hydrolyze cellooligosaccharides with increasing efficiency as the degree of polymerization (DP) increases up to six.^{18–20} Os4BGlu12 shows a similar trend, but with a lower dependency on the DP of cellooligosaccharides with DP greater than 2. Os4BGlu12 has very low efficiency in hydrolysis of cellobiose, but 40-fold higher efficiency for cellotriose and a slight increase in the k_{cat}/K_m value as the DP of cellooligosaccharides rises from 3 to 5.¹⁶ In contrast, the Os3BGlu6 enzyme efficiently hydrolyzes the β -(1,3)-linked disaccharide laminaribiose, but hydrolyzes β -(1,4)-linked oligosaccharides very poorly.¹⁷

Native and complex structures of rice Os3BGlu6, Os3BGlu7, and Os4BGlu12 β -glucosidases have been reported.^{17,21–23} These include covalent intermediates of all three enzymes with 2-fluoroglucofuranose, oligosaccharide-bound complexes of an inactive Os3BGlu7 mutant, and substrate-like inhibitor complexes with Os3BGlu6 and Os4BGlu12. These structures are all typical (β/α)-barrels with the active site at the C-terminal ends of the internal

Abbreviations: DP, degree of polymerization; GH1, glycoside hydrolase family 1; Glc, glucose.

* Corresponding author. Tel.: +66 44 224304; fax: +66 44 224185.

E-mail addresses: sompong_biochem@yahoo.co.th (S. Sansenya), jmaneesan@hotmail.com (J. Maneesan), cairns@sut.ac.th (J.R. Ketudat Cairns).

[†] Present address: Research Faculty of Agriculture, Hokkaido University, Sapporo 060-8589, Japan.

β -strands, as seen for all GH1 structures.²⁴ The amino acid residues surrounding the nonreducing end glucosyl residue in the -1 subsite of the three enzymes are all conserved, while their aglycone binding subsites are not.

One important position for aglycone binding is that of Os3BGlu7 Asn245, which was seen to make hydrogen bonds to the glucosyl residue in subsite +2 in cellobiosaccharides and to that in subsite +1 in laminaribiose.²² A comparison of the cellobiosaccharide complex structures with Os3BGlu6 shows that the Met251 at the position corresponding to Os3BGlu7 Asn245 extends into the active site cleft in a position that would sterically block the binding of β -(1,4)-linked oligosaccharides in the same position in Os3BGlu6 (Fig. 1A).^{17,22} In contrast Os4BGlu12 has His252 in this position, which may be able to form hydrogen bonds to the cellobiosaccharides.²³ To assess the effects of the residue corresponding to Os3BGlu6 Met251 and Os3BGlu7 Asn245, the kinetic constants for hydrolysis of the β -(1,3)-linked disaccharide laminaribiose and β -(1,4)-linked cellobiosaccharides (DP 2–6) were determined for Os3BGlu6 M251N, Os3BGlu7 N245M, and Os4BGlu12 H252M, and compared to the wild type enzymes, as shown in Table 1.

It was expected that Met251, which extends into the Os3BGlu6 active site cleft would interfere with binding of cellobiosaccharides in the same mode as in Os3BGlu7, since it comes within 0.9 Å of O2 of the third glucosyl residue of cellobiosaccharide or cellopentaose, when the Os3BGlu6 structure is superimposed on structures of those complexes, as illustrated in Figure 1A. When cellobiosaccharide was docked into the active site, the nonreducing glucosyl residue in the +2 subsite turned away from this position in Os3BGlu6, but the subsite -1 and +1 glucosyl residues were in

similar positions to those in cellobiosaccharides in the Os3BGlu7 active site (Fig. 1B and C). Replacement of Met251 in Os3BGlu6 with the hydrogen-bonding Asn of Os3BGlu7 afforded the opportunity to introduce a new glucosyl residue-binding subsite, while relieving the steric block. In fact, Os3BGlu6 M251N showed a 10 to 24-fold increase in the k_{cat}/K_m values for all the substrates compared with the wild type enzyme. The Os3BGlu6 M251N mutation increased the efficiency of hydrolysis (k_{cat}/K_m) for laminaribiose and cellobiosaccharide by 15 to 16-fold. The greatest change was seen for cellobiosaccharide, where the 24-fold increase in k_{cat}/K_m corresponds to an 8 kJ/mol improvement in the binding of the transition state ($\Delta\Delta G_{S\rightarrow mut} = -8$ kJ/mol). The k_{cat}/K_m of cellobiosaccharide was threefold that of cellobiosaccharide, but that of cellobiosaccharide was within error of that of cellobiosaccharide, while that of cellobiosaccharide was lower.

Thus, the Asn251 in the Os3BGlu6 mutant can carry out a similar role to Asn245 in Os3BGlu7 in stabilizing β -(1,3)- and β -(1,4)-linked gluco-disaccharide binding and produce a third glucosyl residue binding site. The fact that cellobiosaccharide and longer gluco-oligosaccharides were hydrolyzed with similar or lower k_{cat}/K_m values suggests that only the +2 subsite was created by the mutation and no further binding sites are accessible for longer cellobiosaccharides.

It was hypothesized that replacement of Os3BGlu7 Asn245 with Met might provide a steric block to cellobiosaccharide binding, similar to the role of Met251 in Os3BGlu6. The Os3BGlu7 N245M mutant showed a 6.5-fold drop in the k_{cat}/K_m value for laminaribiose and 17 to 30-fold drops in those for cellobiosaccharides with DP of 3–6 compared to wild type. Os3BGlu7 N245M did show an increase in k_{cat}/K_m values with increasing DP from 2 to 6, similar to the wild type enzyme, but the largest decrease in k_{cat}/K_m

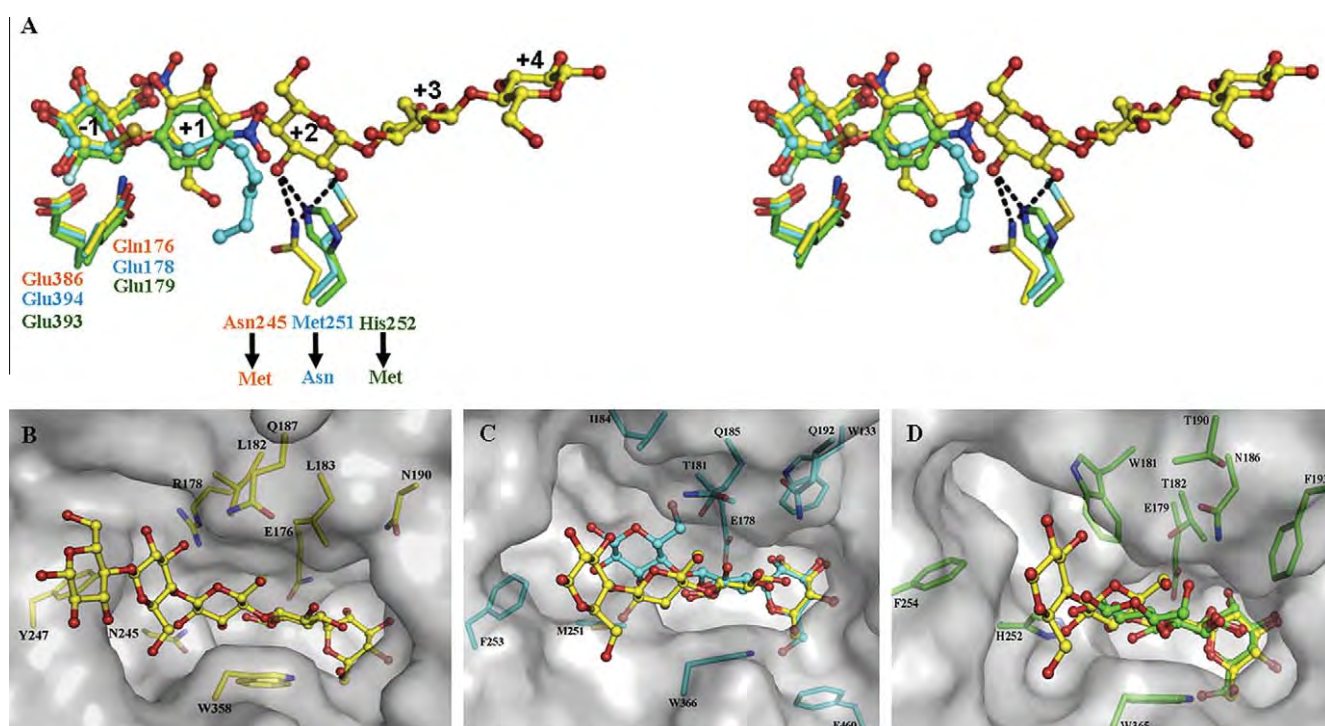


Figure 1. Comparison of the structures of three rice β -glucosidase in complexes with substrates or inhibitors. (A) Stereo view of the ligands in the active sites of the superimposed structures of the complexes of Os3BGlu7 E176Q with cellopentaose (PDB ID: 3F5K, yellow carbons),²² Os3BGlu6 with *n*-octyl- β -D-thioglucopyranoside (PDB ID: 3GNP, cyan carbons),¹⁷ and Os4BGlu12 with 2,4-dinitrophenyl 2-deoxy-2-fluoro- β -D-glucopyranoside (PDB ID: 3PTQ).²³ The acid/base and nucleophile residues are shown, as well as the residues at the positions mutated in this work, with the mutations designated by the arrows. (B) The active site of the Os3BGlu7 E176Q complex with cellopentaose showing the surface and underlying residues that interact with the ligand. (C) The active site of Os3BGlu6 with cellobiosaccharide docked into the active site, and the cellobiosaccharide from the superimposed complex of Os3BGlu7 (PDB ID: 3FKJ) shown for comparison. (D) The active site of Os4BGlu12 with cellobiosaccharide docked into it and cellobiosaccharide from the superimposed complex of Os3BGlu7 with cellobiosaccharide shown for comparison. In all frames, the amino acids are represented by sticks and ligands by balls and sticks with oxygen in red, sulfur in dark yellow and nitrogen in bright blue, carbon in yellow for Os3BGlu7 with cellopentaose or cellobiosaccharide, cyan for Os3BGlu6 with *n*-octyl- β -D-thioglucopyranoside and green for Os4BGlu12 with 2,4-dinitrophenyl 2-deoxy-2-fluoro- β -D-glucopyranoside. Docking of cellobiosaccharide in C and D was done with Autodock 4.0. The figure was generated with Pymol (Schrödinger LLC).

Table 1

Kinetic constants of Os3BGlu6, Os3BGlu6M251N, Os3BGlu7 Os3BGlu7N245M and Os4BGlu12, Os4BGlu12H252M for hydrolysis of oligosaccharides

Substrate	Kinetic parameters	Os3BGlu6	Os3BGlu6 M251N	Os3BGlu7 ^b	Os3BGlu7N245M	Os4BGlu12 ^c	Os4BGlu12 H252M
Laminaribiose (β -1,3-linked Glc ₂)	K_m (mM)	8.7 $\pm$ 0.07	3.5 $\pm$ 0.2	2.05 $\pm$ 0.10	4.5 $\pm$ 0.3	5.1 $\pm$ 0.2	7.2 $\pm$ 0.4
	k_{cat} (s ⁻¹)	1.64 $\pm$ 0.06	9.9 $\pm$ 0.5	32 $\pm$ 3	9.6 $\pm$ 0.3	23 $\pm$ 1	10.6 $\pm$ 0.3
	k_{cat}/K_m (M ⁻¹ s ⁻¹)	180	2830	15,700	2410	4500	1460
	$\Delta\Delta G_{S=mut}$ (kJ mol ⁻¹) ^d		-6.9		+4.7		+1.3
Cellobiose	K_m (mM)	17.5 $\pm$ 0.3	14.3 $\pm$ 1.2	31.5 $\pm$ 1.6	46 $\pm$ 3	27 $\pm$ 2	22.7 $\pm$ 1.9
	k_{cat} (s ⁻¹)	0.27 $\pm$ 0.004	1.97 $\pm$ 0.09	1.52 $\pm$ 0.13	1.18 $\pm$ 0.09	5.9 $\pm$ 0.4	2.42 $\pm$ 0.11
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	15	140	50	30	220	110
	$\Delta\Delta G_{S=mut}$ (kJ mol ⁻¹) ^d		-5.6		+1.3		+1.7
Cellotriose	K_m (mM)	25.1 $\pm$ 0.8	3.2 $\pm$ 0.2	0.72 $\pm$ 0.002	1.51 $\pm$ 0.10	5.0 $\pm$ 0.5	11.2 $\pm$ 0.4
	k_{cat} (s ⁻¹)	0.41 $\pm$ 0.014	1.32 $\pm$ 0.12	18.1 $\pm$ 0.4	2.16 $\pm$ 0.13	44.3 $\pm$ 0.3	41.9 $\pm$ 0.13
	k_{cat}/K_m (M ⁻¹ s ⁻¹)	17	410	25,400	1430	8820	3580
	$\Delta\Delta G_{S=mut}$ (kJ mol ⁻¹) ^d		-8.0		+7.2		+2.3
Cellotetraose	K_m (mM)	20.3 $\pm$ 0.3	3.9 $\pm$ 0.2	0.28 $\pm$ 0.01	1.78 $\pm$ 0.16	5.9 $\pm$ 0.5	23.4 $\pm$ 0.16
	k_{cat} (s ⁻¹)	0.54 $\pm$ 0.004	1.62 $\pm$ 0.10	17.3 $\pm$ 0.6	5.5 $\pm$ 0.2	89 $\pm$ 3	87.8 $\pm$ 0.2
	k_{cat}/K_m (M ⁻¹ s ⁻¹)	27	420	61,000	3100	15,000	5760
	$\Delta\Delta G_{S=mut}$ (kJ mol ⁻¹) ^d		-6.9		+7.5		+2.4
Cellopentaose	K_m (mM)	27.0 $\pm$ 0.5	4.6 $\pm$ 0.4	0.24 $\pm$ 0.01	1.32 $\pm$ 0.11	5.5 $\pm$ 0.4	20.25 $\pm$ 0.11
	k_{cat} (s ⁻¹)	0.25 $\pm$ 0.003	0.62 $\pm$ 0.04	16.9 $\pm$ 0.1	6.7 $\pm$ 0.5	105 $\pm$ 9	62.1 $\pm$ 0.5
	k_{cat}/K_m (M ⁻¹ s ⁻¹)	9	140	72,000	5050	19,000	3070
	$\Delta\Delta G_{S=mut}$ (kJ mol ⁻¹) ^d		-6.9		+6.7		+4.6
Cellohexaose	K_m (mM)	ND ^a	ND	0.11 $\pm$ 0.01	1.52 $\pm$ 0.08	ND	ND
	k_{cat} (s ⁻¹)	ND	ND	16.9 $\pm$ 0.3	7.73 $\pm$ 0.4	ND	ND
	k_{cat}/K_m (mM ⁻¹ s ⁻¹)	ND	ND	153,000	5100	ND	ND
	$\Delta\Delta G_{S=mut}$ (kJ mol ⁻¹) ^d				+8.6		

^a ND means not determined.^b Kinetics constants were taken from Ref. 18.^c Kinetic constants for wild type Os4BGlu12 were taken from Ref. 16.^d The change in Gibbs free energy for transition state binding by the mutants compared to the wildtype enzyme was calculated as $\Delta\Delta G_{S=mut} = -RT[\ln(k_{cat}/K_m)_{mutant} - \ln(k_{cat}/K_m)_{wildtype}]$, so a negative value indicates stronger binding by the mutant, while a positive value indicates poorer transition state binding.²⁷

compared to wild type was seen at cellohexaose. The decreases in k_{cat}/K_m for cellobiose and cellotriose were similar to the 2-fold and 15-fold drops for these substrates in the Os3BGlu7 N245V mutant, which was not expected to introduce a steric block.²¹ This suggests that the introduced Met may not extend into the active site cleft in Os3BGlu7 as much as in Os3BGlu6, or that the substrate is able to adjust its position without further loss of energy. We recently observed that mutation of Os3BGlu7 Tyr341, which interacts with glucosyl residues in the +2 to +4 subsites,²² resulted in cellotetraose binding in an alternative productive position, which interacted less with Asn245.²⁵ So, the steric influence of the Met could cause cellotriose and longer oligosaccharides to bind in this alternative position.

When the structure of Os4BGlu12 was superimposed on the Os3BGlu7 cellopentaose and laminaribiose complex structures, His252Nε2 of Os4BGlu12 appeared to be within range to form hydrogen bonds with the glucosyl residues in the +1 and +2 subsites. Docking of cellotriose into Os4BGlu12 placed the 3 glucosyl residues in nearly the same positions in the lowest energy conformation, including a possible hydrogen bond from O3 of the glucosyl residue in subsite +2 to His252 (Fig. 1D). Mutation of Os4BGlu12 His252 to Met decreased the k_{cat}/K_m threefold for laminaribiose and 2 to 6-fold for cellooligosaccharides with DP of 2–5. The relatively minor effect on cellotriose hydrolysis ($\Delta\Delta G_{S=mut}$ of 2.3 kJ/mol), compared to the corresponding mutation in Os3BGlu7 ($\Delta\Delta G_{S=mut}$ of 7.2), shows weaker interaction between His252 and the oligosaccharide. This suggests that the position of binding of cellooligosaccharides by Os4BGlu12 is similar to, but not the same as, that seen in Os3BGlu7, at least at the +2 subsite.

In conclusion, the residue corresponding to Asn245 in Os3BGlu7 is critical to binding of oligosaccharides by the divergent rice GH1 β -glucosidases Os3BGlu6, Os3BGlu7, and Os4BGlu12. Conversion of the corresponding Met251 to Asn, created a new cellooligosaccharide binding subsite for Os3BGlu6 and allowed it to hydrolyze cellooligosaccharides with DP 2–5 at much higher rates, while

the inverse replacements of Os3BGlu7 Asn245 and Os4BGlu12 His252 with Met decreased hydrolysis of cellooligosaccharides. In each case, the single amino acid residue change did not convert the oligosaccharide hydrolysis properties of the acceptor β -glucosidase to that of the source enzyme for that residue type, but it did take a substantial step in that direction. Therefore, the structure-based engineering of one active site cleft residue can dramatically alter the oligosaccharide specificity of plant GH1 enzymes.

1. Experimental

1.1. Site-directed mutagenesis

The mutagenesis to produce cDNA encoding Os3BGlu6 M251N, Os3BGlu7 N254M, and Os4BGlu12 H252M was performed with the QuikChange[®] Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA, USA). The oligonucleotide primers used for the mutagenesis were 5'-GCTTGGGATAGCGTTCGACGTGAATTGGTTCGAGCCG-3' and its reverse complement for Os3BGlu6 M251N; 5'-GAGTTGGAAA GAGCTTCATACCACATGAAGTCCAGAACTATTCCAAC-3' and its reverse complement for Os3BGlu7 N245M; and 5'-GATTGAATAA CTCTGGTCTCGATGTGGTTTGTCCCTTCTCC-3' and its reverse complement for Os4BGlu12 H252M. The previously published pET32 expression vectors were used as the templates^{10,15,17} and the full cDNA inserts were sequenced to verify that only the desired mutations were present in the mutated plasmids.

1.2. Expression and purification of enzymes

The wild type Os3BGlu6 and the Os3BGlu6 M251N, Os3BGlu7 N245M, and Os4BGlu12 H252M mutant proteins were expressed in *Escherichia coli* as thioredoxin-fusion proteins, as previously described for the wild type proteins.^{10,15,17} The three mutant enzymes were purified by immobilized metal affinity chromatography

(IMAC), followed by removal of the N-terminal thioredoxin and His₆ tags with enterokinase for Os3BGlu7N245M and Os4BGlu12H252M and TEV protease for Os3BGlu6, and removal of the fusion tag by IMAC, as previously described.^{17,21,26}

1.3. Kinetic studies

The release of glucose from laminaribiose and cellooligosaccharides with DP of 2–6 by the enzymes was monitored by glucose oxidase assay, as previously described,¹⁵ to determine the kinetic parameters. The enzyme concentrations used were 2 µg of wild type Os3BGlu6 and 0.3–0.5 µg of mutant Os3BGlu6, Os3BGlu7, and Os4BGlu12. Times of reactions giving initial velocities (v_0) were determined by measuring the activity at 3–6 different time points. Substrate concentrations over a range of approximately 0.2 to 5 times the apparent K_m were used to determine the kinetic parameters. The apparent kinetic parameters including k_{cat} , K_m , k_{cat}/K_m were calculated by nonlinear regression of Michaelis–Menten plots with the Grafit 5.0 computer program (Erithacus Software, Horley, UK). Changes in transition state binding energies in the mutations ($\Delta\Delta G_{S*mut}$) were calculated as described by Ferst et al.:²⁷

$$\Delta\Delta G_{S*mut} = -RT[\ln(k_{cat}/K_m)_{mutant} - \ln(k_{cat}/K_m)_{wildtype}].$$

1.4. Docking studies

Docking trials were run in Autodock 4.0 using the Lamarckian genetic algorithm,²⁸ with cellotriose with the nonreducing sugar in the ¹S₃ skew boat configuration, as previously described for Os3BGlu7.²¹ Ten docking trials were run under default parameters from Autodock 4.0 for each of Os3BGlu6 and Os4BGlu12. Due to lack of water in the simulation, these trials were not expected to give accurate energies, but simply to show the feasibility of binding in positions similar to cellooligosaccharides in crystal structures of Os3BGlu7.

Acknowledgments

Supriya Seshadri is acknowledged for initial characterization of rice Os3BGlu6 and Rodjana Opassiri and Jisnusun Svasti are thanked for insightful comments during the work. This work was supported by the Thailand Research Fund grant BRG5380017, Suranaree University of Technology and the National Research University Project of the Thai Commission on Higher Education.

References

- Cantarel, B. L.; Coutinho, P. M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. *Nucleic Acids Res.* **2009**, *37D*, 233–238.
- Ketudat Cairns, J. R.; Esen, A. *Cell. Mol. Life Sci.* **2010**, *67*, 3389–3405.
- Verdougq, L.; Czjzek, M.; Moriniere, J.; Bevan, D. R.; Esen, A. *J. Biol. Chem.* **2003**, *278*, 25055–25062.
- Czjzek, M.; Cicek, M.; Zamboni, V.; Bevan, D. R.; Henrissat, B.; Esen, A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13555–13560.
- Czjzek, M.; Cicek, M.; Zamboni, V.; Burmeister, W. P.; Bevan, D. R.; Henrissat, B.; Esen, A. *Biochem. J.* **2001**, *354*, 37–46.
- Zouhar, J.; Vevodova, J.; Marek, J.; Damborsky, J.; Su, X. D.; Brzobohatý, B. *Plant Physiol.* **2001**, *127*, 973–985.
- Verdougq, L.; Moriniere, J.; Bevan, D. R.; Esen, A.; Vasella, A.; Henrissat, B.; Czjzek, M. *J. Biol. Chem.* **2004**, *279*, 31796–31803.
- Sue, M.; Nakamura, C.; Miyamoto, T.; Yajima, S. *Plant Sci.* **2011**, *180*, 268–275.
- Mendonça, L. M. F.; Marana, S. R. *Biochim. Biophys. Acta* **2011**, *1814*, 1616–1623.
- Opassiri, R.; Pomthong, B.; Onkoksoong, T.; Akiyama, T.; Esen, A.; Ketudat Cairns, J. R. *BMC Plant Biol.* **2006**, *6*, 33.
- Xu, Z.; Escamilla-Trevino, L. L.; Zeng, L.; Lalgonar, M.; Bevan, D. R.; Winkel, B. S. J.; Mohamed, A.; Cheng, C. L.; Shih, M. C.; Poulton, J. E.; Esen, A. *Plant Mol. Biol.* **2004**, *55*, 343–367.
- Schlieman, W. J. *Plant Physiol.* **1984**, *116*, 123–132.
- Akiyama, T.; Kaku, H.; Shibuya, N. *Phytochemistry* **1998**, *48*, 49–54.
- Wakuta, S.; Hamada, S.; Ito, H.; Matsuura, H.; Nabeta, K.; Matsui, H. *Phytochemistry* **2010**, *71*, 1280–1288.
- Opassiri, R.; Ketudat-Cairns, J. R.; Akiyama, T.; Wara-Aswapati, O.; Svasti, J.; Esen, A. *Plant Sci.* **2003**, *165*, 627–638.
- Opassiri, R.; Maneesan, J.; Akiyama, T.; Pomthong, B.; Jin, S.; Kimura, A.; Ketudat Cairns, J. R. *Plant Sci.* **2010**, *179*, 273–280.
- Seshadri, S.; Akiyama, T.; Opassiri, R.; Kuaprasert, B.; Ketudat Cairns, J. R. *Plant Physiol.* **2009**, *151*, 47–58.
- Opassiri, R.; Hua, Y.; Wara-Aswapati, O.; Akiyama, T.; Svasti, J.; Esen, A.; Ketudat Cairns, J. R. *Biochem. J.* **2004**, *379*, 125–131.
- Hrmova, M.; MacGregor, E. A.; Biely, P.; Stewart, R. J.; Fincher, G. B. *J. Biol. Chem.* **1998**, *273*, 11134–11143.
- Kuntothom, T.; Luang, S.; Harvey, A. J.; Fincher, G. B.; Opassiri, R.; Hrmova, M.; Ketudat Cairns, J. R. *Arch. Biochem. Biophys.* **2009**, *491*, 85–95.
- Chuenchor, W.; Pengthaisong, S.; Robinson, R. C.; Yuvaniyama, J.; Oonant, W.; Bevan, D. R.; Esen, A.; Chen, C. J.; Opassiri, R.; Svasti, J.; Ketudat Cairns, J. R. *J. Mol. Biol.* **2008**, *377*, 1200–1215.
- Chuenchor, W.; Pengthaisong, S.; Robinson, R. C.; Yuvaniyama, J.; Svasti, J.; Ketudat Cairns, J. R. *J. Struct. Biol.* **2011**, *173*, 169–179.
- Sansenya, S.; Opassiri, R.; Kuaprasert, B.; Chen, C. J.; Ketudat Cairns, J. R. *Arch. Biochem. Biophys.* **2011**, *510*, 62–72.
- Barrett, T.; Suresh, C. G.; Tolley, S. P.; Dodson, E. J.; Hughes, M. A. *Structure* **1995**, *3*, 951–960.
- Pengthaisong, S.; Withers, S. G.; Kuaprasert, B.; Svasti, J.; Ketudat Cairns, J. R. *Protein Sci.*, in press. doi:10.1002/pro.2021.
- Sansenya, S.; Ketudat Cairns, J. R.; Opassiri, R. *Acta Crystallogr., Sect. F* **2010**, *66*, 320–323.
- Fersht, A. R.; Leatherbarrow, R. J.; Wells, T. N. C. *Biochemistry* **1987**, *26*, 6030–6038.
- Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. *J. Comput. Chem.* **1998**, *19*, 1639–1662.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at SciVerse ScienceDirect

Carbohydrate Research

journal homepage: www.elsevier.com/locate/carres

Rice BGlu1 glycosynthase and wild type transglycosylation activities distinguished by cyclophellitol inhibition

Salila Pengthaisong^a, Chi-Fan Chen^b, Stephen G. Withers^b, Buabarn Kuaprasert^c,
James R. Ketudat Cairns^{a,*}

^aSchools of Biochemistry and Chemistry, Institute of Science, Suranaree University of Technology, 111 University Avenue, Muang District, Nakhon Ratchasima 30000, Thailand

^bCentre for High-throughput Biology (CHiBi) and Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, British Columbia, Canada V6T 1Z1

^cResearch Facility Division, Synchrotron Light Research Institute (Public Organization), 111 University Avenue, Muang District, Nakhon Ratchasima 30000, Thailand

ARTICLE INFO

Article history:

Received 1 December 2011

Received in revised form 6 February 2012

Accepted 10 February 2012

Available online 21 February 2012

Keywords:

Glycosylation

Glycosynthase

Oligosaccharide synthesis

Rice

Transglucosylation

 β -Glucosidase

ABSTRACT

The rice BGlu1 β -D-glucosidase nucleophile mutant E386G is a glycosynthase that catalyzes the synthesis of cellooligosaccharides from α -D-glucopyranosyl fluoride (GlcF) donor and *p*-nitrophenyl (pNP) cellobioside (Glc2-pNP) or cello-oligosaccharide acceptors. When activity with other donors and acceptors was tested, the initial enzyme preparation cleaved pNP- β -D-glucopyranoside (Glc-pNP) and pNP- β -D-fucopyranoside (Fuc-pNP) to pNP and glucose and fucose, suggesting contamination with wild type BGlu1 β -glucosidase. The products from reaction of GlcF and Fuc-pNP included Fuc- β -(1 $\rightarrow$ 3)-Fuc-pNP, Glc- β -(1 $\rightarrow$ 3)-Fuc-pNP, and Fuc- β -(1 $\rightarrow$ 4)-Glc- β -(1 $\rightarrow$ 3)-Fuc-pNP, suggesting the presence of both wild type BGlu1 and its glycosynthase. Inhibition of the BGlu1 β -glucosidase activity within this preparation by cyclophellitol confirmed that the E386G glycosynthase preparation was contaminated with wild type BGlu1. Rice BGlu1 E386G-2, generated from a new construct designed to minimize back-mutation, showed glycosynthase activity without wild type hydrolytic or transglycosylation activity. E386G-2 catalyzed transfer of glycosyl residues from GlcF, α -L-arabinosyl fluoride, α -D-fucosyl fluoride, α -D-galactosyl fluoride, α -D-mannosyl fluoride, and α -D-xylosyl fluoride donors to Glc2-pNP acceptor. The synthetic products from the reactions of α -fucosyl fluoride and α -mannosyl fluoride donors were confirmed to result from addition of a β -(1 $\rightarrow$ 4)-linked glycosyl residue. Moreover, the E386G glycosynthase transferred glucose from GlcF donor to glucose, cellobiose, Glc-pNP, Fuc-pNP, pNP- β -D-galactopyranoside, and pNP- β -D-xylopyranoside acceptors, but little to pNP- β -D-mannopyranoside. Production of longer oligosaccharides occurred most readily on acceptors with an equatorial 4-OH. Elimination of wild type contamination thereby allowed a clear assessment of BGlu1 E386G glycosynthase catalytic abilities.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Glycosynthases are nucleophile mutants of glycosidases that effect synthesis of oligosaccharides and glycoconjugates from glycosyl fluoride donors and suitable acceptors without hydrolysis of the products (Fig. 1).^{1–3} Glycosynthases derived from endoglycosidases use oligosaccharyl fluorides as donor substrates and have high regioselectivity.^{3–8} Glycosynthases derived from exoglycosidases are typically capable of efficient glycoside assembly with monosaccharyl fluoride donors, but sometimes with lower substrate specificity and regioselectivity.^{3,9,10}

A glycoside hydrolase family 1 β -D-glucosidase from rice (BGlu1, systematically called Os3BGlu7) has a broad substrate specificity, with particularly high hydrolytic activity toward β -(1 $\rightarrow$ 3)- and β -(1 $\rightarrow$ 4)-linked oligosaccharides and pyridoxine 5'-O- β -D-glucoside.^{11–13} It also has high transglucosylation activity

using 4-nitrophenyl β -D-glucoside (Glc-pNP) as donor for glucosyl transfer to pyridoxine or Glc-pNP acceptors, resulting in the synthesis of pyridoxine 5'-O- β -D-glucoside and pNP gluco-oligosaccharides. Rice BGlu1 has previously been mutated to form a glycosynthase (Fig. 1C) that can synthesize pNP-cellooligosaccharides of at least 11 β -(1 $\rightarrow$ 4)-linked glucosyl residues.¹⁴ In that study, it was found that the E386G mutation (designated E414G in the previous study, based on the nucleophile position in the BGlu1 precursor) resulted in much higher glycosynthase activity than the E386A or E386S mutations. The ability to synthesize long cellooligosaccharides was attributed to their binding in a long oligosaccharide binding cleft, which has recently been verified in the X-ray crystal structures of BGlu1 acid-base and glycosynthase mutants.^{15,16}

In this paper we expand our exploration of the glycosynthase activity of the BGlu1 E386G enzyme, in particular, we explore the ability of the enzyme to utilize alternate glycosyl fluoride donors and acceptors, thereby broadening the spectrum of di- and trisaccharides produced by this glycosynthase. Wild type

* Corresponding author. Tel.: +66 44 224304; fax: +66 44 224185.

E-mail address: cairns@sut.ac.th (J.R. Ketudat Cairns).

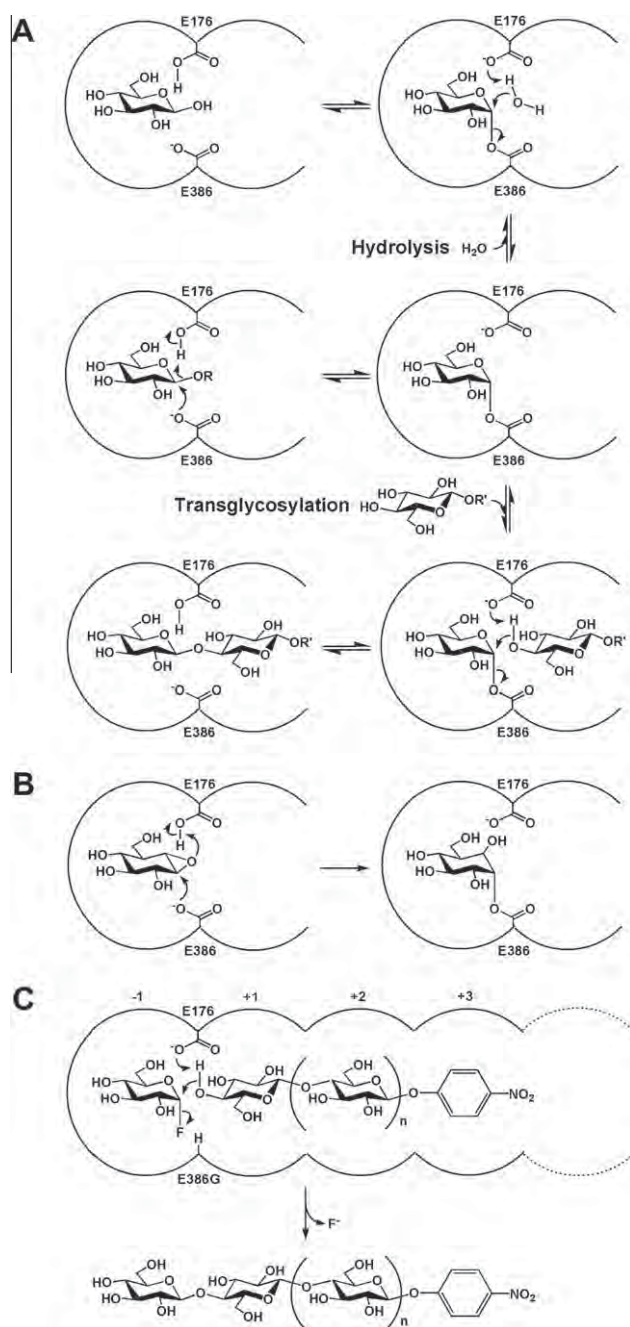


Figure 1. Reaction mechanisms of hydrolysis and transglycosylation catalyzed by rice BGlul1 β-glucosidase (A); covalent intermediate trapping of rice BGlul1 by cyclophellitol (B), and rice BGlul1 E386G glycosynthase (C). Note that cyclophellitol cannot react with BGlul1 E386G due to its lack of a catalytic nucleophile.

BGlul1 contamination of BGlul1 E386G was confirmed and eliminated by inhibition with cyclophellitol, thereby allowing the actual glycosynthase products to be identified.

2. Results and discussion

2.1. Evidence of wild type BGlul1 contamination in the BGlul1 E386G preparation

As is the case with many GH1 glycosidases,¹⁷ the wild type BGlul1 β-D-glucosidase has relatively broad substrate specificity, since it can hydrolyze 4-nitrophenyl β-D-glucopyranoside

(Glc-pNP), pNP β-D-fucopyranoside (Fuc-pNP), pNP β-D-galactopyranoside (Gal-pNP), pNP β-D-mannopyranoside (Man-pNP), pNP β-D-xylopyranoside (Xyl-pNP), and pNP α-L-arabinopyranoside, as well as β-(1→2), β-(1→3), β-(1→4), and β-(1→6) glycosidic linkages.^{11,12} This suggests that BGlul1 glycosynthase might also utilize other glycosyl fluoride donors and thus make mixed products.

When the acceptor specificity of the BGlul1 E386G glycosynthase preparation was investigated, both transglycosylation and hydrolysis products were detected. The Glc-pNP and Fuc-pNP acceptors were cleaved to pNP, glucose and fucose, while Glc-pNP was produced from Glc2-pNP, which suggested wild type contamination of the BGlul1 E386G glycosynthase preparation. Since new column packings were used for all glycosynthase preparations, cross-contamination during purification was unlikely. It thus appeared that a mutation in the Origami(DE3) *Escherichia coli* strain being used for expression may have led to a higher than normal reversion rate, possibly by misreading on the ribosome or reversion mutations in a few cells in the expression culture.^{18,19} This suggested that some of the products formed may have been generated by a combination of glycosynthase and wild type transglycosylation activities. Characterization of purified products provided further evidence for the combined action of BGlul1 wild type and E386G glycosynthase, as described below.

Mass spectrometric and NMR characterizations identified the purified products of reaction of α-glucosyl fluoride (GlcF) and Fuc-pNP catalyzed by the BGlul1 E386G preparation as Fuc-β-(1→3)-Fuc-pNP, Glc-β-(1→3)-Fuc-pNP, and Fuc-β-(1→4)-Glc-β-(1→3)-Fuc-pNP (Table 1, products 1, 2, and 3, respectively). Since glycosynthase mutants of BGlul1 show very little pNP-glycoside hydrolysis,¹⁴ the nonreducing end fucosyl moieties in products of Fuc-β-(1→3)-Fuc-pNP and Fuc-β-(1→4)-Glc-β-(1→3)-Fuc-pNP from reaction of GlcF and Fuc-pNP may have come from transglycosylation from Fuc-pNP by wild type BGlul1 contamination, as diagrammed in Figure 1A.

2.2. Cyclophellitol inhibition

The presence of wild type BGlul1 contamination in the E386G preparation was confirmed by inhibition of the wild type enzyme with the covalent inhibitor cyclophellitol. As shown in Figure 1B, cyclophellitol inactivates retaining β-D-glucosidases by forming a covalent adduct with the catalytic nucleophile. This is a stable adduct that does not undergo reactivation by transglycosylation.²⁰ Figure 2 shows that inactivation of wild type rice BGlul1 by cyclophellitol follows pseudo-first order kinetics of inactivation, similar to previous reports for other β-glucosidases.^{20,21} Cyclophellitol acts as an irreversible inactivator of the rice BGlul1 β-glucosidase with a K_i of 5.7 μM and a k_i of 0.35 min⁻¹ (Fig. 2B and C).

The E386G preparation was thus pre-incubated with cyclophellitol for 2 h, which resulted in complete inhibition of wild type BGlul1 activity (Fig. 2A), before testing the transglycosylation of GlcF onto Fuc-pNP. The E386G glycosynthase preparation synthesized two products in the range of pNP-disaccharides and a few trisaccharides in the reaction without cyclophellitol pre-incubation (Fig. 3 lane EG -) and a single product, corresponding to Glc-Fuc-pNP, after pre-incubation with 2 mM cyclophellitol (Fig. 3 lane EG +).

As expected, untreated wild type BGlul1 β-glucosidase synthesized transglycosylation products, while that pre-incubated with cyclophellitol did not (Fig. 3, lanes BG). Moreover, uninhibited wild type BGlul1 synthesized only one product in the pNP-disaccharide range: Fuc-Fuc-pNP, which was also seen with the untreated E386G preparation, but not when it was preincubated with cyclophellitol. This result confirms that the original E386G glycosynthase was contaminated with wild type BGlul1, and that, as would be expected, the wild type BGlul1 does not transfer from GlcF, suggesting

Table 1
Structures of products of glycosynthase reactions of BGlu1 E386G

#	Donor	Acceptor	Product
1			(1)*
			β -D-Fuc-(1 $\rightarrow$ 3)- β -D-Fuc-pNP
			(2)
2			(3)**
			β -D-Fuc-(1 $\rightarrow$ 4)- β -D-Glc-(1 $\rightarrow$ 3)- β -D-Fuc-pNP
			(4)
3			(5)
			β -D-Fuc-(1 $\rightarrow$ 4)- β -D-Glc-(1 $\rightarrow$ 4)- β -D-Glc-pNP

* Formed by wild type BGlu1 β -glucosidase contaminant.

** Formed by successive action of glycosynthase and wild type BGlu1 contaminant.

it is unlikely to have contributed to the production of hetero-disaccharide products where the added glycosyl residues were derived from an α -glycosyl fluoride donor.

The cyclophellitol inhibition results show that the combination of wild type BGlu1 and E386G glycosynthase seen in the contaminated prep was likely responsible for synthesis of the novel pNP-oligosaccharide Fuc- β -(1 $\rightarrow$ 4)-Glc- β -(1 $\rightarrow$ 3)-Fuc-pNP, which could not be produced by either enzyme alone. This points to the importance of testing for wild type hydrolase activity when assessing glycosynthase products, since the transglycosylation products arising from small amounts of contaminating hydrolase can be significant, given the large amount of glycosynthase needed for the transfer reactions.

2.3. Verification of linkages of glycosynthase products

As noted above, the wild type BGlu1 β -glucosidase could not transfer glycosyl residues from α -glycosyl fluoride donors, so any products generated from such transfers reflect the glycosynthase regiospecificity, even when derived from the mixture with wild type in the E386G preparation. It was previously demonstrated that reaction of BGlu1E386G with GlcF and Glc2-pNP specifically yields β -(1 $\rightarrow$ 4)-linked oligosaccharides.¹⁴ The MS and NMR analyses described above identified a product generated from transfer of a single glucosyl residue onto Fuc-pNP as β -D-Glc-(1 $\rightarrow$ 3)- β -D-Fuc-pNP (Table 1, product 2). Similarly, in reaction products with Glc2-pNP as acceptor, a single transfer product from α -D-mannosyl fluoride (ManF) was identified as Man- β -(1 $\rightarrow$ 4)-Glc- β -(1 $\rightarrow$ 4)-Glc-pNP (Table 1, product 4) and the corresponding product with α -D-fucosyl fluoride (FucF) donor was Fuc- β -(1 $\rightarrow$ 4)-Glc- β -(1 $\rightarrow$ 4)-Glc-pNP (Table 1, product 5). Thus, in each case, β -(1 $\rightarrow$ 4)-linkages

were preferred on Glc residues with equatorial 4-OH groups, but the glycosyl residue was added onto the equatorial 3-OH of Fuc residues, where the 4-OH is axial.

Streptomyces β -glucosidase glycosynthase is regiospecific for production of β -(1 $\rightarrow$ 3)-linkages to pNP-monosaccharide acceptors and can use Fuc-pNP and Gal-pNP acceptors,⁹ while *Agrobacterium* β -glucosidase has a strong preference for formation of β -(1 $\rightarrow$ 4)-linkages.^{22–25} As such, rice BGlu1 appears to be more similar to *Streptomyces* β -glucosidase glycosynthase in its activity toward these acceptors.

2.4. Transglycosylation activity of E386G-2

A new construct for the production of BGlu1 E386G was made by changing the codon used from GAA (Glu) to GGC instead of GGG to give Gly at the nucleophile position. Use of this double-changed codon should decrease the probability of misreading as Glu (GAG) by the ribosome or of a reversion mutation in the expression culture. The protein expressed from this mutation (E386G-2) indeed exhibited 33-fold lower Glc-pNP hydrolysis activity compared to the original E386G glycosynthase (the Glc-pNP hydrolysis activities of BGlu1, BGlu1 E386G, and BGlu1 E386G-2 were 1900, 1.1, and 0.033 nmol min⁻¹ mg⁻¹ protein, respectively). When the E386G-2 glycosynthase preparation was used to catalyze the reaction between Fuc-pNP and GlcF, only a single product, corresponding to Glc-Fuc-pNP, was detected after pre-incubation either without or with 2 mM cyclophellitol (Fig. 3, lanes EG2 – and +).

Since no wild type hydrolysis and transglycosylation activities were detected in the E386G-2 preparation under the conditions of the standard reaction, this sample was used to test the BGlu1

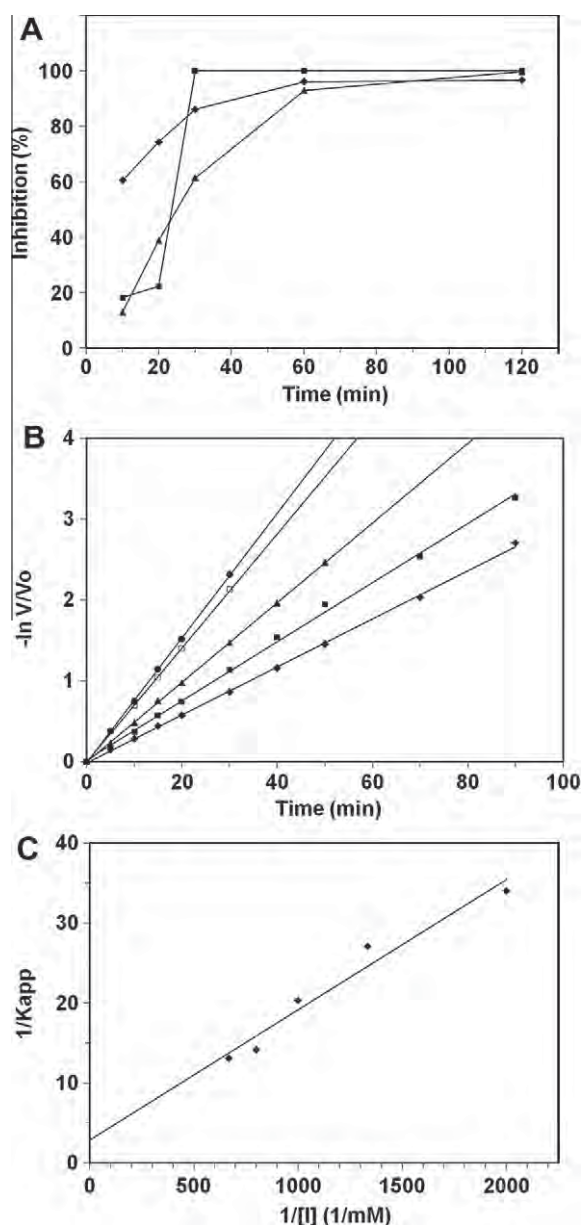


Figure 2. Cyclophellitol inactivation of rice BGlul. (A) Time course for incubation of the enzymes with cyclophellitol. The enzyme solutions of 0.2 μM BGlul (▲), 40 μM E386G (◆), and 40 μM E386G-2 (■) were pre-incubated with 1 μM cyclophellitol for different times and then assayed for release of pNP from Glc-pNP. (B) Semi-logarithmic plot of wild type BGlul β-glucosidase activity versus time at various concentrations of cyclophellitol, (●), 0.5 μM; (□), 0.75 μM; (▲), 1 μM; (■), 1.25 μM; (◆), 1.5 μM. (C) Replot of the inverse of the apparent first order rate constants (k_{app}) from B versus the inverse of the inhibitor concentration, $1/[I]$.

E386G glycosynthase activity with various donors and acceptors. The E386G-2 glycosynthase successfully transferred a range of donors (α -L-arabinosyl fluoride (AraF), FucF, α -D-galactosyl fluoride (GalF), GlcF, ManF, and α -D-xylosyl fluoride (XylF)) onto pNP-cellobiose to produce trisaccharides, and a faint spot of tetrasaccharide was seen with FucF donor (Fig. 4).

The E386G-2 glycosynthase could also use glucose, cellobiose, Fuc-pNP, Gal-pNP, Glc-pNP, Xyl-pNP, and Glc2-pNP as acceptors (Fig. 4). pNP-Oligosaccharides resulting from addition of more than one Glc could only be detected by TLC with Glc2-pNP as acceptor. Longer oligosaccharides were, however, synthesized by E386G-2 using GlcF donor when Glc-pNP, Man-pNP, Xyl-pNP, or Glc2-pNP

acceptors were employed, similar to what had been seen with glycosynthases from *Agrobacterium* β-glucosidase.^{22,23} By contrast, the *Streptomyces* E383A glycosynthase could only synthesize disaccharides when using GlcF and pNP-monosaccharide acceptors.⁹ When using Fuc-pNP as acceptor E386G-2 synthesized only Glc-Fuc-pNP disaccharide and a tiny amount of Glc-Glc-Fuc-pNP trisaccharide, and showed no hydrolytic activity toward acceptors, verifying that the longer Fuc-pNP oligosaccharides seen with the original E386G prep resulted from the wild type BGlul transglycosylation activity. The pNP-disaccharide product spot seen when Glc-pNP was used as an acceptor was very faint (Fig. 4B, lane pG), explaining why we failed to detect it in our previous work.¹⁴ This spot had a mobility slightly lower than Glc2-pNP, suggesting it might be pNP-gentiobiose, but this identity was not confirmed due to the small amount produced. Similarly, the product spots obtained with cellobiose as acceptor were very faint, though at least two products could be detected by sulfuric acid charring (Fig. 4D, lane C2).

Surprisingly, a clear spot of laminaribiose was observed when glucose was used as an acceptor, and this spot was also seen in the other reactions, likely due to the presence of glucose from hydrolysis of GlcF. The production of laminaribiose is in line with the fact that laminaribiose is a much better hydrolytic substrate for the wild-type BGlul β-glucosidase than are cellobiose or other disaccharides¹² and consistent with structural studies in complexes with laminaribiose.¹⁵ The enzyme also catalyzed autocondensation of GlcF, as seen by the smear above glucose in the TLC and LC-MS peaks with masses of 379 m/z ($[M+^{35}Cl]^-$).

Transglycosylation yields for reaction of BGlul E386G-2 with GlcF and a variety of acceptors are shown in Table 2, and those with a variety of donors and Glc2-pNP acceptor are shown in Table 3. Relatively high yields were obtained with FucF and ManF donors, but only a low yield with XylF. The absence of a substituent at C-5 apparently affects binding at the -1 subsite, although the activity was higher than that of non-evolved glycosynthases derived from *Agrobacterium* β-glucosidase^{22–24} and *Streptomyces* E383A,⁹ for which XylF was not a donor substrate. Moreover, E386G-2 uses AraF as a donor, something which has not previously been reported for a GH1 glycosynthase. The E386G-2 glycosynthase gave low yields with Fuc-pNP, Gal-pNP, Glc-pNP, and Xyl-pNP acceptors, and GlcF donor, compared to the *Streptomyces* E383A.⁹

Although Gal-pNP and Fuc-pNP acceptors gave the highest yields for disaccharide production, only very small amounts of trisaccharides and no tetrasaccharides were detected on LC-MS (Table 2). In contrast, Glc-pNP gave higher production of trisaccharide and tetrasaccharide products were also seen for Glc-pNP, Xyl-pNP, and Man-pNP, which all have equatorial 4-OH groups. This agrees with the previous report that the BGlul E386G glycosynthase appears to require a β-(1→4)-linked acceptor, such as Glc2-pNP or cellotriose, before it can efficiently transfer successive glucosyl residues onto a β-(1→4)-linked chain with high regioselectivity.¹⁴

With Glc2-pNP as acceptor, the BGlul E386G-2 glycosynthase gave lower yields of trisaccharide products (26% with GlcF), but higher yields of tetrasaccharide (16%) and longer products, compared to *Agrobacterium* E358A (79% trisaccharide and 13% tetrasaccharide under comparable conditions) and E358S (88% trisaccharide, 7% tetrasaccharide) and *Streptomyces* E383A glycosynthase (80% trisaccharide and 3% tetrasaccharide). As noted previously, the unique feature of rice BGlul glycosynthases is their ability to produce long oligosaccharides,¹⁴ so that BGlul E386G also produced cellopentaose in 11% yield. At donor to acceptor ratios above 1:1, the yields could not be assessed by HPLC with GlcF donor and Glc2-pNP acceptor, since they resulted in precipitated products. Thus, BGlul E386G is more useful for synthesis of longer oligosaccharides, while the bacterial enzymes are better for synthesis of di- and tri-saccharides.

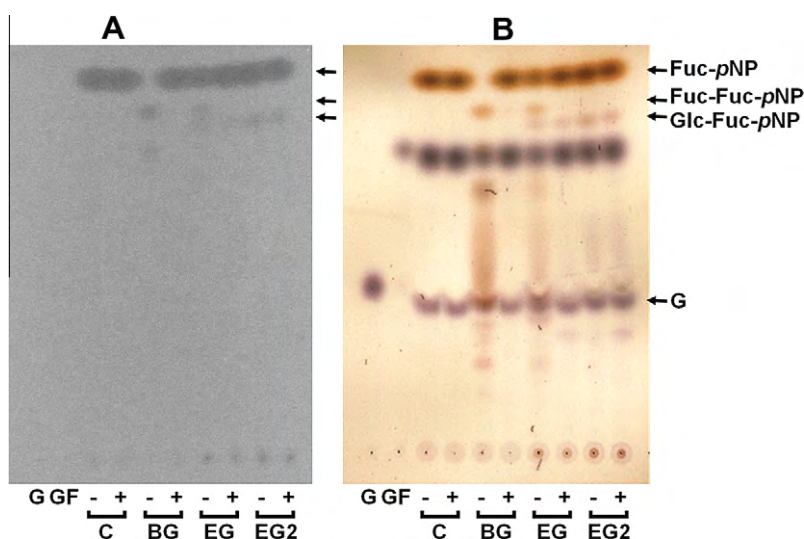


Figure 3. Effect of cyclophellitol inhibition on transglycosylation products of rice BGLu1, BGLu1 E386G, and BGLu1 E386G-2 with Fuc-pNP and GlcF. The 0.2 μ M BGLu1, 40 μ M E386G, and 40 μ M E386G-2 were pre-incubated without (–) or with (+) 2 mM cyclophellitol for 2 h, then transglycosylation reactions with 10 mM Fuc-pNP and 10 mM α -GlcF in 150 mM NH_4HCO_3 , pH 7.0, were incubated at 30 °C for 16 h. Silica gel TLC analysis was performed with EtOAc–MeOH–water (7:2.5:1) as a developing solvent, and products detected under UV light (A) or by dabbing with 10% sulfuric acid in ethanol followed by charring (B). Lane G, glucose standard; GF, GlcF; C, control of 10 mM Fuc-pNP and 10 mM GlcF without enzyme; BG, BGLu1; EG, BGLu1 E386G; EG2, BGLu1 E386G-2.

2.5. Conclusion

Pure rice BGLu1 E386G glycosynthase transfers glycosyl residues from a variety of α -glycosyl fluoride donors, in the reactivity order Glc > Fuc > Ara > Man > Gal > Xyl. Likewise it transfers glucosyl residues from GlcF to several different acceptors, with a preference for formation of 1–4 linkages. Only equatorial hydroxyls on the ring appear to be glycosylated, so if the C-4 hydroxyl is axial, the monosaccharide is added to the C-3 hydroxyl. However the products formed in this way do not appear to serve as acceptors for further transfer reactions. Care must be taken to avoid contamination by wild type enzyme arising either from laboratory procedures or translational misincorporation, since otherwise, initially formed glycosynthase products can be hydrolyzed, and unexpected transglycosylation products formed.

3. Experimental

3.1. Materials

Cello-oligosaccharides (DP 3–6) were purchased from Seikagaku Kogyo Co. (Tokyo, Japan). Glucose, cellobiose, Glc2-pNP, Glc-pNP, Fuc-pNP, Gal-pNP, Man-pNP, and Xyl-pNP were purchased from Sigma Chemical Co. (St. Louis, MO). GlcF, AraF, FucF, GalF, ManF, and XylF were synthesized by previously described methods.²⁶ Cyclophellitol was a generous gift from Kah-Yee Li and Professor Herman Overkleeft, University of Leiden.

3.2. Protein expression and purification

The recombinant BGLu1 and BGLu1 E386G proteins were expressed in *E. coli* strain Origami(DE3) and purified by immobilized metal affinity chromatography (IMAC), enterokinase digestion, and subtractive IMAC, as previously described.^{11,14,27} The buffer was exchanged with 50 mM phosphate buffer, pH 6.0, and concentrated in an Amicon ultra-centrifugal filter (Ultracel-10K, Millipore, USA). The purified protein concentration was determined by measuring its absorbance at 280 nm. An extinction coefficient ϵ_{280} of 113560 $\text{M}^{-1} \text{cm}^{-1}$, calculated from the amino acid composition by the method of Gill and von Hippel,²⁸ was used to calculate the BGLu1 enzyme concentrations.

3.3. Transglycosylation activity of E386G with various donors

The E386G glycosynthase was incubated at 40 μ M with 10 mM of each of the donors: AraF, FucF, GalF, ManF, or XylF and 10 mM Glc2-pNP acceptor in 150 mM ammonium bicarbonate buffer, pH 7.0, at 30 °C for 16 h.¹⁴ Products were analyzed by TLC (Silica Gel 60 F₂₅₄, aluminum-backed, Merck) with ethyl acetate (EtOAc)–methanol (MeOH)–water (7:2.5:1) as solvent. Plates were visualized under ultraviolet (UV) light and by exposure to 10% sulfuric acid in ethanol followed by charring. Samples (5 μ L) were diluted in 200 μ L of 20% MeOH, 0.1% formic acid, and 5 μ L of the solution was injected onto an ESI-MS (Waters, Alliance HPLC, Micromass ZQ ESCi, USA). The quadrupole mass analyzer was scanned over a range of 100–1000 m/z .

3.4. Transglycosylation by BGLu1 E386G to various acceptors

The BGLu1 E386G glycosynthase was incubated at a concentration of 40 μ M with 10 mM GlcF donor and 10 mM acceptor: glucose, cellobiose, Glc-pNP, Fuc-pNP, Gal-pNP, Man-pNP, or Xyl-pNP, in 150 mM ammonium bicarbonate, pH 7.0, at 30 °C for 16 h. The products were analyzed by TLC with EtOAc–MeOH–water (7:2.5:1 as solvent for glucose and pNP-oligosaccharides, and 2:1:1 for cellobiose). The products were further verified by ESI-MS, as described above. Samples (3 μ L) of the products from the reactions with 1:1 molar ratios of GlcF: Glc-pNP or GlcF: Fuc-pNP were loaded onto a ZORBAX carbohydrate column (4.6 mm $\times$ 250 mm, Agilent, USA) connected to an Agilent 1100 series LC–MSD. The column was eluted with a linear gradient from 90% to 60% acetonitrile in water over 30 min at a flow rate of 0.8 mL/min. The eluted peaks were detected by UV absorbance at 300 nm. Reaction yields were determined by integration of the product peaks within the HPLC profile and the molecular masses of the eluted products were determined by mass spectrometry.

3.5. Determination of linkages within glycosynthase products

Oligosaccharide products from reactions of FucF and Glc2-pNP, ManF and Glc2-pNP, and GlcF and Fuc-pNP were synthesized with E386G. Reaction mixtures containing FucF or ManF (14.6 mg, 4 equiv) and Glc2-pNP (9.2 mg, 1 equiv); GlcF (48 mg, 1 equiv)

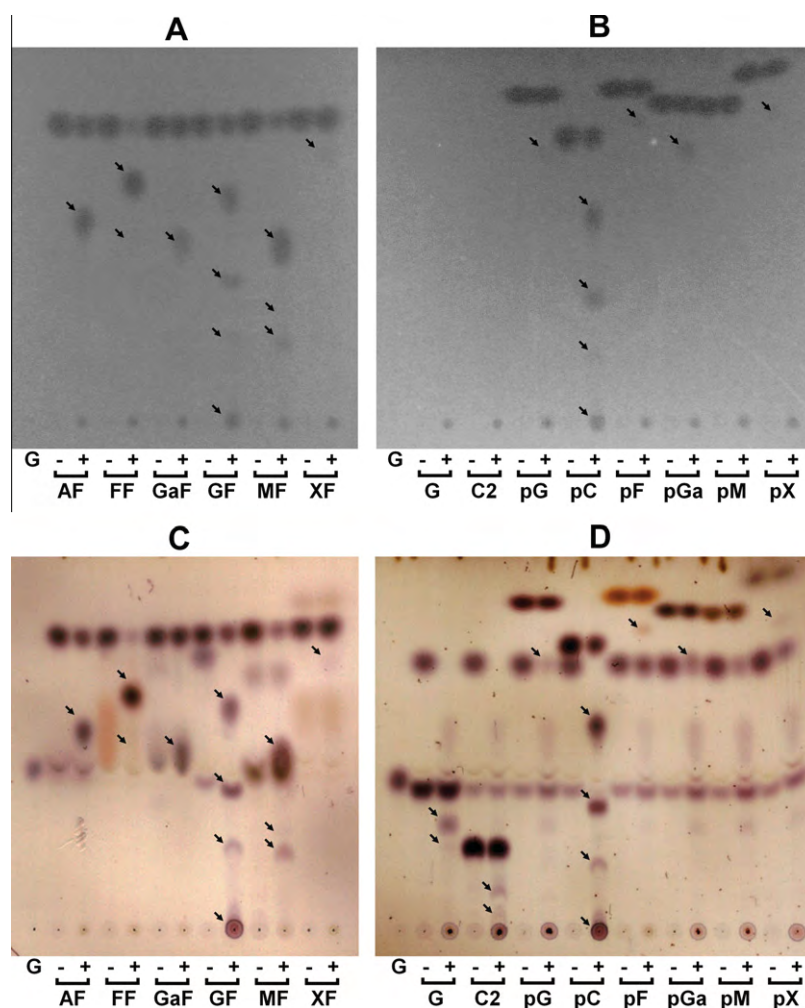


Figure 4. TLC detection of transglycosylation catalyzed by BGlu1 E386G-2 with various donors or acceptors. The E386G-2 was incubated with 10 mM Glc2-*p*-NP and 10 mM donors (A and C) or 10 mM GlcF and 10 mM acceptors (B and D) in 150 mM NH_4HCO_3 , pH 7.0, at 30 °C for 16 h. The reactions were performed without (–) or with (+) enzyme. The products were separated by silica gel TLC with EtOAc–MeOH–water (7:2.5:1) as a developing solvent and detected under UV light (A and B) or by exposure to 10% sulfuric acid in ethanol, followed by charring (C and D). G is glucose standard. A and C show reactions with various donors: AF, AraF; FF, FucF; GaF, GalF; GF, GlcF; MF, ManF; and XF, XylF. B and D show reactions with various acceptors: G, glucose; C2, cellobiose; pG, Glc-*p*-NP; pC, Glc2-*p*-NP; pF, Fuc-*p*-NP; pGa, Gal-*p*-NP; pM, Man-*p*-NP; and pX, Xyl-*p*-NP.

and Fuc-*p*-NP (75 mg, 1 equiv) and 40 μM of the appropriate enzyme were incubated at 30 °C in 150 mM ammonium bicarbonate, pH 7.0, (2–10 ml total reaction volume) for 24–48 h. The reaction mixtures were filtered and loaded onto a reverse-phase column (Sep-Pak *tc*18, sorbent weight 2 g, Waters), which was eluted with step gradients of 0–30% methanol in water. Eluted fractions were checked by TLC, as described above. Fractions containing the compounds were pooled, dried on a rotary evaporator, and further purified by TLC. Silica gel containing each compound was scratched from the TLC plate and eluted with 30% MeOH, then the methanol was evaporated and the sample solution was dried by lyophilization. Masses were determined by ESI-MS.

A mixture of pyridine and acetic anhydride (5 ml, $V_{\text{pyridine}}:V_{\text{acetic anhydride}} = 3:2$) was added to the purified glycoside products at room temperature and the reaction mixture was stirred overnight. Ice water was added to quench the unreacted acetic anhydride and the reaction was stirred for another 10 min. After evaporation of the solvent, the residue was dissolved in ethyl acetate and washed successively with ice water, 1 M HCl, NaHCO_3 , and brine, and dried with MgSO_4 . The organic solvent was evaporated in vacuo, and the resulting residue was purified by flash column chromatography on silica gel using ethyl acetate/petroleum ether = 1:1 as eluent.

All ^1H and ^1H COSY nuclear magnetic resonance (NMR) spectra were recorded on a Bruker Avance-400 inv spectrometer

(400 MHz) and chemical shifts were reported in δ units (ppm) and were calibrated based on the deuterium solvent chemical shift (CDCl_3). Multiplicity of signals is described as follows: br-broad, s-singlet; d-doublet; t-triplet; m-multiplet.

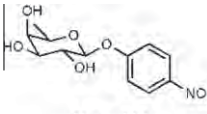
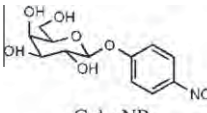
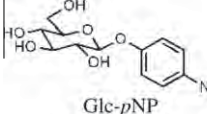
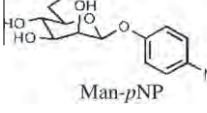
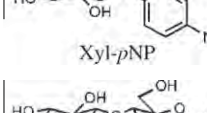
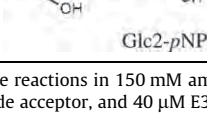
3.5.1. NMR spectra

3.5.1.1. 4-Nitrophenyl (2,3,4-tri-*O*-acetyl- α -*D*-fucopyranosyl)-(1 $\rightarrow$ 3)-2,4-di-*O*-acetyl- β -*D*-fucopyranoside (1). ^1H NMR (400 MHz, CDCl_3): δ 8.21 (m, 2H, Ar), 7.05 (m, 2H, Ar), 5.50 (dd, 1H, $J_{1,2}$ 8.1, $J_{2,3}$ 9.9 Hz, H-2), 5.35 (br d, 1H, $J_{3,4}$ 3.4 Hz, H-4), 5.20 (br d, 1H, $J_{3',4'}$ 3.3 Hz, H-4'), 5.11 (dd, 1H, $J_{1',2'}$ 7.9, $J_{2',3'}$ 10.7 Hz, H-2'), 5.07 (d, 1H, $J_{1,2}$ 8.1 Hz, H-1), 4.95 (dd, 1H, $J_{2',3'}$ 10.7, $J_{3',4'}$ 3.4 Hz, H-3'), 4.55 (d, 1H, $J_{1',2'}$ 7.9 Hz, H-1'), 3.98 (dd, 1H, $J_{2,3}$ 9.9, $J_{3,4}$ 3.4 Hz, H-3), 3.94 (m, 1H, H-5'), 3.76 (m, 1H, H-5), 2.20–1.98 (5s, 15H, 5 $\times$ CH_3CO), 1.30 (d, 3H, H6'a, H6'b, H6'c), 1.20 (d, 3H, H6a, H6b, H6c). ESI-MS: m/z 454 $[\text{M}+\text{Na}]^+$.

3.5.1.2. 4-Nitrophenyl (2,3,4,6-tetra-*O*-acetyl- α -*D*-glucopyranosyl)-(1 $\rightarrow$ 3)-2,4-di-*O*-acetyl- β -*D*-fucopyranoside (2). ^1H NMR (400 MHz, CDCl_3): δ 8.21 (m, 2H, Ar), 7.06 (m, 2H, Ar), 5.46 (dd, 1H, $J_{1,2}$ 8.0, $J_{2,3}$ 10.0 Hz, H-2), 5.34 (br d, 1H, $J_{3,4}$ 3.4 Hz, H-4), 5.17 (dd, 1H, $J_{3',4'}$ 9.3, $J_{4',5'}$ 9.0 Hz, H-4'), 5.12 (dd, 1H, $J_{2',3'}$ 9.0, $J_{3',4'}$ 9.3 Hz, H-3'), 5.08 (d, 1H, $J_{1,2}$ 8.0 Hz, H-1), 4.93 (dd, 1H, $J_{1',2'}$ 7.8, $J_{2',3'}$ 9.0 Hz, H-2') 4.65 (d, 1H, $J_{1',2'}$ 7.8 Hz, H-1'), 4.33 (dd, 1H, $J_{5',6'a}$

Table 2

Products of glycosynthase reactions of BGlu1 E386G-2 glycosynthase using α -glucosyl fluoride donor and different acceptors^a

#	Acceptor	pNP-Oligosaccharide product (% yield)								Total yield
		Di	Tri	Tetra	Penta	Hexa	Hepta	Octa	Nana	
4b c	 Fuc-pNP	4.2 7.3 (482)	0.040 0.074 (644)	—	—	—	—	—	—	4.24 7.37
5b c	 Gal-pNP	9.1 11 (498)	0.15 0.21 (660)	—	—	—	—	—	—	9.25 11.2
6b c	 Glc-pNP	3.0 3.4 (498)	0.49 0.55 (660)	0.15 0.14 (822)	0.12 0.13 (984)	—	—	—	—	3.76 4.22
7b c	 Man-pNP	0.062 0.068 (498)	0.035 0.032 (660)	0.046 0.039 (822)	—	—	—	—	—	0.143 0.139
8b c	 Xyl-pNP	3.1 4.0 (468)	0.1 0.17 (630)	0.052 0.051 (792)	—	—	—	—	—	3.25 4.22
9b	 Glc2-pNP		26 (660)	16 (822)	11 (984)	3.5 (1146)	0.47 (1308)	0.03 (1471)	0.002 (1633)	57.0

^a Preparative reactions in 150 mM ammonium bicarbonate buffer.

pNP-glycoside acceptor, and 40 μ M E386G-2 concentration for 24 h. The relative percents are in terms of peak area per total 300 nm absorbance in pNP-glycoside products separated by HPLC on a ZORBAX carbohydrate column. The molecular mass [Mass+³⁵Cl][−] of each eluted compound was confirmed by mass spectrometry as shown in parenthesis.

^b pH 7.0, 30 °C, at a 1:1 molar ratio.

^c 3:1 molar ratios of GlcF donor.

2.6, $J_{6'a,6'b}$ 12.3 Hz, H-6'a), 4.15 (dd, 1H, $J_{5',6'b}$ 3.8, $J_{6'a,6'b}$ 12.3 Hz, H-6'b), 3.97–3.92 (m, 2H, H-3, H-5), 3.69–3.65 (m, 2H, H-5'), 2.18–2.00 (6s, 18H, 6 $\times$ CH₃CO), 1.25 (d, 3H, H6a, H6b, H6c). ESI-MS: m/z 470 [M+Na]⁺.

3.5.1.3. 4-Nitrophenyl [(2,3,4-tri-O-acetyl- β -D-fucopyranosyl)-(1 $\rightarrow$ 4)-O-2,3,6-tri-O-acetyl- β -D-glucopyranosyl]-(1 $\rightarrow$ 3)-O-2,4-di-O-acetyl- β -D-fucopyranoside (3).

¹H NMR (400 MHz, CDCl₃): δ 8.21 (m, 2H, Ar), 7.06 (m, 2H, Ar), 5.48 (dd, 1H, $J_{1,2}$ 8.0, $J_{2,3}$ 10.0 Hz, H-2), 5.33 (br d, 1H, $J_{3,4}$ 3.0 Hz, H-4), 5.20 (br d, 1H, $J_{3',4'}$ 3.3 Hz, H-4''), 5.17–5.03 (m, 4H, H-1, H-3', H-2'', H-3''), 4.89 (dd, 1H, $J_{1',2'}$ 7.9, $J_{2',3'}$ 8.8 Hz, H-2'), 4.58 (d, 1H, $J_{1',2'}$ 7.9 Hz, H-1'), 4.45 (d, 1H, $J_{1'',2''}$ 8.2 Hz, H-1''), 4.30 (dd, 1H, $J_{5,6'a}$ 3.0, $J_{6'a,6'b}$ 12.1 Hz, H-6'a), 4.14 (dd, 1H, $J_{5',6'b}$ 3.9, $J_{6'a,6'b}$ 12.1 Hz, H-6'b), 3.98–3.91 (m, 2H, H-3, H-5), 3.73–3.63 (m, 3H, H-4', H-5', H-5''), 2.19–2.00 (8s, 24H, 8 $\times$ CH₃CO), 1.35 (d, 3H, H6a, H6b, H6c), 1.25 (d, 3H, H6'a, H6'b, H6'c). ESI-MS: m/z 616 [M+Na]⁺.

3.5.1.4. 4-Nitrophenyl [(2,3,4,6-tetra-O-acetyl- β -D-mannopyranosyl)-(1 $\rightarrow$ 4)-O-2,3,6-tri-O-acetyl- β -D-glucopyranosyl]-(1 $\rightarrow$ 4)-O-2,3,6-tri-O-acetyl- β -D-glucopyranoside (4).

¹H NMR (400 MHz, CDCl₃): δ 8.21 (m, 2H, Ar), 7.06 (m, 2H, Ar), 5.40 (br d, 1H, $J_{2'',3''}$ 3.0, H-2''), 5.29 (dd, 1H, $J_{2,3}$ 8.5, $J_{3,4}$ 8.5 Hz, H-3), 5.21 (dd, 1H, $J_{1,2}$ 7.3, $J_{2,3}$ 8.5 Hz, H-2), 5.20 (dd, 1H, $J_{3',4'}$ 10.0, $J_{4'',5''}$ 8.5 Hz, H-4''), 5.17 (d, 1H, $J_{1,2}$ 7.3 Hz, H-1), 5.15 (dd, 1H, $J_{2',3'}$ 9.4, $J_{3',4'}$ 9.7 Hz, H-3'), 5.01 (dd, 1H, $J_{2'',3''}$ 3.0, $J_{3',4'}$ 10.0 Hz, H-3''), 4.88 (dd, 1H, $J_{1',2'}$ 7.9, $J_{2',3'}$ 9.4 Hz, H-2'), 4.64 (br s, 1H, H-1''), 4.55 (dd,

1H, $J_{5,6'a}$ 1.8, $J_{6'a,6'b}$ 11.8 Hz, H-6'a), 4.52 (d, 1H, $J_{1',2'}$ 7.9 Hz, H-1'), 4.34 (dd, 1H, $J_{6'a,6'b}$ 12.45, $J_{5',6'a}$ 5.2 Hz, H-6'a), 4.31–4.28 (m, 2H, H-6'a, H-6'b), 4.12 (dd, 1H, $J_{6'a,6'b}$ 12.5, $J_{5',6'b}$ 2.1 Hz, H-6'b), 4.11 (dd, 1H, $J_{6'a,6'b}$ 11.8, $J_{5,6'b}$ 4.9 Hz, H-6b), 3.91–3.81 (m, 3H, H-4, H-4', H-5), 3.64–3.58 (m, 2H, H-5', H-5''), 2.16–1.99 (10s, 30H, 10 $\times$ CH₃CO). ESI-MS: m/z 648 [M+Na]⁺.

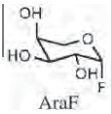
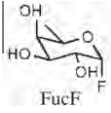
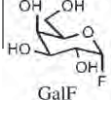
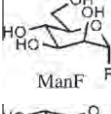
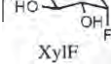
3.5.1.5. 4-Nitrophenyl [(2,3,4-tri-O-acetyl- β -D-fucopyranosyl)-(1 $\rightarrow$ 4)-O-2,3,6-tri-O-acetyl- β -D-glucopyranosyl]-(1 $\rightarrow$ 4)-O-2,3,6-tri-O-acetyl- β -D-glucopyranoside (5).

¹H NMR (400 MHz, CDCl₃): δ 8.21 (m, 2H, Ar), 7.05 (m, 2H, Ar), 5.29 (dd, 1H, $J_{2,3}$ 8.5, $J_{3,4}$ 8.2 Hz, H-3), 5.23–5.15 (m, 4H, H-2, H-3', H4'', H-1) 5.08 (dd, 1H, $J_{1''2''}$ 8.0, $J_{2''3''}$ 10.3 Hz, H-2''), 4.94 (dd, 1H, $J_{2',3'}$ 10.3, $J_{3',4'}$ 3.4 Hz, H-3''), 4.87 (dd, 1H, $J_{1',2'}$ 8.1, $J_{2',3'}$ 9.5 Hz, H-2'), 4.55 (m, 1H, H-6a), 4.53 (d, 1H, $J_{1',2'}$ 8.1 Hz, H-1'), 4.43 (d, 1H, $J_{1''2''}$ 8.0 Hz, H-1''), 4.41 (m, 1H, H-6'a), 4.16–4.14 (m, 2H, H-6b, H-6'b), 3.88 (dd, 1H, $J_{3,4}$ 8.2, $J_{4,5}$ 10.1 Hz, H-4), 3.81 (dd, 1H, $J_{3',4'}$ 9.5, $J_{4',5'}$ 9.4 Hz, H-4'), 3.91–3.78 (m, 1H, H-5), 3.76–3.71 (m, 1H, H-5''), 3.65–3.59 (m, 1H, H-5'), 2.20–1.90 (9s, 27H, 9 $\times$ CH₃CO), 1.20 (d, 3H, Me). ESI-MS: m/z 632 [M+Na]⁺.

3.6. Generation of E386G-2 mutant of rice BGlu1

BGlu1 E386G-2 was constructed by changing the DNA codon for GAA to GGC by QuikChange mutagenesis (Stratagene, La Jolla, CA, USA) with the primers 5'-G ACA GTC GTC ATA ACT GGC AAC GGA ATG GAT CAA C-3' and 5'-G TTG ATC CAT TCC GTT GCC AGT TAT

Table 3
Products of glycosynthase reactions of BGlu1 E386G-2 glycosynthase using different donors and Glc2-pNP acceptor^a

#	Donor	pNP-Oligosaccharide product (% yield)		
		Tri	Tetra	Total yield
10b c	 AraF	25 59 (630)	0.051 0.072 (762)	25.1 59.1
11b c	 FucF	49 98 (644)	0.011 1.4 (790)	49.0 99.4
12b c	 GalF	15 42 (660)	0.013 0.018 (822)	15.0 42.0
13b c	 ManF	19 71 (660)	0.21 7.9 (822)	19.2 78.9
14b c	 XylF	1.2 3.1 (630)	0.0091 0.013 (762)	1.21 3.11

^a Preparative reactions in 150 mM ammonium bicarbonate buffer.

^b pH 7.0, 30 °C, at a 1:1.

^c 3:1 molar ratios of glycosyl fluoride donor:Glc2-pNP acceptor, and 40 μM E386G-2 concentration for 24 h. The relative percents are in terms of peak area per total 300 nm absorbance in pNP-glycoside products separated by HPLC on a ZORBAX carbohydrate column. The molecular mass [Mass+³⁵Cl][−] of each eluted compound was confirmed by mass spectrometry, as shown by the mass in parenthesis (in amu).

GAC GAC TGT C-3' and the pET32a/BGlu1 expression vector¹¹ as template. The transglucosylation activity of the new mutation was compared to the E386G (codon: GGG) toward various donors of AraF, FucF, GalF, GlcF, ManF, and XylF with Glc2-pNP acceptor, and GlcF donor with various acceptors of glucose, cellobiose, Glc-pNP, Glc2-pNP, Fuc-pNP, Gal-pNP, Man-pNP, and Xyl-pNP, as described above. Samples (10 μl) of the products from the reactions with 1:1 and 3:1 molar ratios of donor:pNP acceptor were analyzed by HPLC on a ZORBAX carbohydrate column as described above.

3.7. Cyclophellitol inhibition

The cyclophellitol inhibition of BGlu1 wild type, E386G, and E386G-2 activities was monitored by analysis of residual enzyme activity as a function of time. Solutions of 0.2 μM BGlu1, 40 μM E386G, and 40 μM E386G-2 in 50 mM sodium acetate, pH 5.0, were pre-incubated with 1 μM cyclophellitol at 30 °C for different times and Glc-pNP hydrolysis was assayed as previously described.¹¹

Inactivation of BGlu1 by cyclophellitol was performed by incubating the BGlu1 in 50 mM sodium acetate, pH 5.0, at 30 °C in the presence of the cyclophellitol concentrations given in Figure 2B. Aliquots (10 μl) were removed at appropriate time intervals and assayed for hydrolysis of Glc-pNP in the standard assay.

The β-glucosidase activity of glycosynthase mutants was inactivated by pre-incubating 0.2 μM BGlu1, 40 μM E386G, or 40 μM E386G-2 with 2 mM cyclophellitol for 2 h, and then transglycosylation activity was tested with 10 mM Fuc-pNP and 10 mM GlcF in 150 mM NH₄HCO₃, pH 7.0, at 30 °C for 24 h. The products were analyzed by TLC, using EtOAc–MeOH–water (7:2.5:1) as solvent, as described above.

Acknowledgements

The authors are thankful to Grienggrai Hommalai and Jisnusun Svasti for initiation of this work and for sharing preliminary data

on galactoside production with BGlu1 E386S glycosynthase. We also thank Kah-Yee Li and Professor Herman Overkleeft, University of Leiden, for their generous gift of cyclophellitol. This work was supported by a Ph.D. scholarship to SP by the Synchrotron Light Research Institute (Public Organization) of Thailand [GS-51-D03], Suranaree University of Technology National Research University Project from the Commission on Higher Education of Thailand, Grant number BRG5380017 from the Thailand Research Fund and the Natural Sciences and Engineering Research Council of Canada through Discovery and Strategic Grants and a CREATE training Grant.

References

- Williams, S. J.; Withers, S. G. *Carbohydr. Res.* **2000**, 327, 27–46.
- Williams, S. J.; Withers, S. G. *Aust. J. Chem.* **2002**, 55, 3–12.
- Perugino, G.; Trincone, A.; Rossi, M.; Moracci, M. *Trends Biotechnol.* **2004**, 22, 31–37.
- Fort, S.; Boyer, V.; Greffe, L.; Davies, G. J.; Moroz, O.; Christiansen, L.; Schüle, M.; Cottaz, S.; Driguez, H. *J. Am. Chem. Soc.* **2000**, 122, 5429–5437.
- Fairweather, J. K.; Fajjes, M.; Driguez, H.; Planas, A. *Chem. Biochem.* **2002**, 3, 866–873.
- Hrmova, M.; Imai, T.; Rutten, S. J.; Fairweathers, J. K.; Pelosi, L.; Bulone, V.; Driguez, H.; Fincher, G. B. *J. Biol. Chem.* **2002**, 277, 30102–30111.
- Kim, Y. W.; Chen, H.; Withers, S. G. *Carbohydr. Res.* **2005**, 279, 42787–42793.
- Hancock, S. M.; Vaughan, M. D.; Withers, S. G. *Curr. Opin. Chem. Biol.* **2006**, 10, 509–519.
- Fajjes, M.; Saura-Valls, M.; Pérez, X.; Conti, M.; Planas, A. *Carbohydr. Res.* **2006**, 341, 2055–2065.
- Müllegger, J.; Chen, H.; Chan, W. Y.; Reid, S. P.; Jahn, M.; Warren, R. A. J.; Salleh, H. M.; Withers, S. G. *Chem. Biochem.* **2006**, 7, 1028–1030.
- Opassiri, R.; Ketudat Cairns, J. R.; Akiyama, T.; Wara-Aswapati, O.; Svasti, J.; Esen, A. *Plant Sci.* **2003**, 165, 627–638.
- Opassiri, R.; Hua, Y.; Wara-Aswapati, O.; Akiyama, T.; Svasti, J.; Esen, A.; Ketudat Cairns, J. R. *Biochem. J.* **2004**, 379, 125–131.
- Opassiri, R.; Pomthong, B.; Onkoksoong, T.; Akiyama, T.; Esen, A.; Ketudat Cairns, J. R. *BMC Plant Biol.* **2006**, 33.
- Hommalai, G.; Withers, S. G.; Chuenchor, W.; Ketudat Cairns, J. R.; Svasti, J. *Glycobiology* **2007**, 17, 744–753.
- Chuenchor, W.; Pengthaisong, S.; Robinson, R. C.; Yuvaniyama, J.; Svasti, J.; Ketudat Cairns, J. R. *J. Struct. Biol.* **2011**, 173, 169–179.

16. Pengthaisong, S.; Withers, S. G.; Kuaprasert, B.; Svasti, J.; KetudatCairns, J. R. *Protein Sci.* **2012**, 21, 362–372.
17. Day, A. G.; Withers, S. G. *Biochem. Cell Biol.* **1986**, 64, 914–922.
18. Parker, J. *Microbiol. Rev.* **1989**, 53, 273–298.
19. Kramer, E. B.; Farabaugh, P. J. *RNA* **2007**, 13, 87–96.
20. Gloster, T. M.; Madsen, R.; Davies, G. J. *Org. Biomol. Chem.* **2007**, 5, 444–446.
21. Withers, S. G.; Umezawa, K. *Biochem. Biophys. Res. Commun.* **1991**, 177, 532–537.
22. Mackenzie, L. F.; Wang, Q.; Warren, R. A. J.; Withers, S. G. *J. Am. Chem. Soc.* **1998**, 120, 5583–5584.
23. Mayer, C.; Zechel, D. L.; Reid, S. P.; Warren, R. A.; Withers, S. G. *FEBS Lett.* **2000**, 466, 40–44.
24. Kim, Y. W.; Lee, S. S.; Warren, R. A.; Withers, S. G. *J. Biol. Chem.* **2004**, 279, 42787–42793.
25. Prade, H.; Mackenzie, L. F.; Withers, S. G. *Carbohydr. Res.* **1998**, 305, 371–381.
26. Yokoyama, M. *Carbohydr. Res.* **2000**, 327, 5–14.
27. Chuenchor, W.; Pengthaisong, S.; Robinson, R. C.; Yuvaniyama, J.; Oonant, W.; Bevan, D. R.; Esen, A.; Chen, C. J.; Opassiri, R.; Svasti, J.; Ketudat Cairns, J. R. *J. Mol. Biol.* **2008**, 377, 1200–1215.
28. Gill, S. C.; von Hippel, P. H. *Anal. Biochem.* **1989**, 182, 319–326.

Contents lists available at [SciVerse ScienceDirect](http://www.elsevier.com/locate/bbapap)

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

X-ray crystallography and QM/MM investigation on the oligosaccharide synthesis mechanism of rice BGlu1 glycosynthases

Jinhu Wang^{a,b}, Salila Pengthaisong^c, James R. Ketudat Cairns^{c,*}, Yongjun Liu^{a,**}

^a Key Lab of Theoretical and Computational Chemistry in University of Shandong, School of Chemistry and Chemical Engineering, Shandong University, Jinan, Shandong 250100, China

^b College of Chemistry Chemical Engineering and Material Science, Zaozhuang University, Zaozhuang, Shandong 277160, China

^c School of Biochemistry and Chemistry, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand

ARTICLE INFO

Article history:

Received 19 August 2012

Received in revised form 7 November 2012

Accepted 13 November 2012

Available online 19 November 2012

Keywords:

Oligosaccharide synthesis

Rice BGlu1 E386 mutant

QM/MM

Energy barrier

rice BGlu1 E386G E386A, E386S

ABSTRACT

Nucleophile mutants of retaining β -glycosidase can act as glycosynthases to efficiently catalyze the synthesis of oligosaccharides. Previous studies proved that rice BGlu1 mutants E386G, E386S and E386A catalyze the oligosaccharide synthesis with different rates. The E386G mutant gave the fastest transglucosylation rate, which was approximately 3- and 19-fold faster than those of E386S and E386A. To account for the differences of their activities, in this paper, the X-ray crystal structures of BGlu1 mutants E386S and E386A were solved and compared with that of E386G mutant. However, they show quite similar active sites, which implies that their activities cannot be elucidated from the crystal structures alone. Therefore, a combined quantum mechanical/molecular mechanical (QM/MM) calculations were further performed. Our calculations reveal that the catalytic reaction follows a single-step mechanism, i.e., the extraction of proton by the acid/base, E176, and the formation of glycosidic bond are concerted. The energy barriers are calculated to be 19.9, 21.5 and 21.9 kcal/mol for the mutants of E386G, E386S and E386A, respectively, which is consistent with the order of their experimental relative activities. But based on the calculated activation energies, 1.1 kcal/mol energy difference may translate to nearly 100 fold rate difference. Although the rate limiting step in these mutants has not been established, considering the size of the product and the nature of the active site, it is likely that the product release, rather than chemistry, is rate limiting in these oligosaccharides synthesis catalyzed by BGlu1 mutants.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

The synthesis of carbohydrates is challenging since there are several functional groups and chiral centers presenting on the monosaccharides [1,2]. Although chemical methods can be used to synthesize oligosaccharides, their analogues, and glycoconjugates with exceptional flexibility, stereochemical control required during the synthesis process must be achieved by sequential protection and deprotection of the functional groups. These complex procedures often result in low yields of oligosaccharides. Enzymatic oligosaccharide synthesis approaches provide alternative tools to obtain the desired oligosaccharide by the use of substrate specificity and stereoselectivity of glycosidases and glycosyltransferases.

Enzymatic oligosaccharide synthesis involves two main classes of carbohydrate enzymes: glycosidases and glycosyltransferases. The

oligosaccharide synthesis reactions promoted by glycosidases may be achieved either with a large excess of acceptors (thermodynamically controlled synthesis) or with activated glycosyl donors (kinetically controlled transglycosylation) [3,4]. However, the typical retaining glycosidases usually give poor synthetic yields (10–40%) due to hydrolysis of substrates and products. Glycosyltransferases can catalyze the glycosyl transfer from a saccharide donor to the acceptor with strict regioselectivity and stereospecificity without hydrolysis [5]. However, although glycosyltransferases have been used for oligosaccharide synthesis, their poor availability and high cost of substrates limit their exploitation.

In 1998, Withers and colleagues discovered that a mutant of the retaining *Agrobacterium* sp. β -glucosidase (E.C. 3.2.1.21) in which the catalytic nucleophile was mutated to a nonnucleophilic amino acid residue (Glu358Ala) could be used for transglycosylation reactions [6]. This enzyme, denoted as a glycosynthase, could efficiently catalyze the synthesis of oligosaccharides with high yields (>80%) without detectable hydrolysis. Since then, engineering a retaining glycosidase into a glycosynthase by the mutation of the catalytic nucleophile to a smaller nonnucleophilic residue, to form a hydrolytically inactive enzyme has become a popular strategy for synthesizing various oligosaccharides [7]. When the glycosyl fluorides of inverted configuration relative to

Abbreviations: QM/MM, quantum mechanical/molecular mechanical; α -GlcF, α -glucosyl fluoride; pNPC2, pNP-cellobioside; RMSD, root-mean-square deviation; MD, molecular dynamics; HB, hydrogen bond; R, reactant; TS, transition state; P, product

* Corresponding author. Tel.: +66 44 224304; fax: +66 44 224185.

** Corresponding author. Tel.: +86 531 88564464; fax: +86 531 88365576.

E-mail addresses: cairns@sut.ac.th (J.R.K. Cairns), yongjunliu_1@sdu.edu.cn (Y. Liu).

the natural substrates are used, these hydrolytically inactive enzymes can be used to synthesize oligosaccharides on a large scale in the presence of suitable glycosyl acceptor, giving products that accumulate in a very high yield.

Catalytic nucleophile mutants of retaining glycosidases have been used as glycosynthases to synthesize various glycosides and oligosaccharides [8–10]. Although glycosynthases were originally derived from retaining β -glycosidases, their sources were recently expanded to retaining α -glycosidases and inverting glycosidases [11,12]. Furthermore, the catalytic activity and substrate specificity of the glycosynthase has been improved greatly by additional mutations of their catalytic sites, while the regioselectivity of glycosidic linkages could be enhanced by changing the stereochemistry of the acceptor substrate [13]. Thus, glycosynthase technology has become a versatile tool for the synthesis of novel glycosides and oligosaccharides [14,15]. So far, a variety of novel glycosynthases have been derived from different glycosidase families, including endoglycosidases and exoglycosidases [16–21].

Rice (*Oryza sativa*) BGlu1 β -glucosidase, belonging to the glycoside hydrolase family 1 enzyme, shows hydrolysis activities towards β -(1,4)- and short β -(1,3)-linked gluco-oligosaccharides [22–24], and glycosides such as pNP- β -glycoside. The hydrolysis mechanism of rice BGlu1 β -glucosidase has been studied by the combined quantum mechanical/molecular mechanical (QM/MM) method. As the energy barriers were 15.7 kcal/mol for the glycosylation step [25] and 21.4 kcal/mol for the deglycosylation step [26], the hydrolysis rate may proceed quickly. Therefore, although the wild-type BGlu1 enzyme displays transglucosylation activity on oligosaccharide and glycoside substrates, the yields of products are relatively low and the newly formed products are subsequently hydrolyzed at the same time [22,27].

Because the aglycone-binding site of rice BGlu1 β -glucosidase was relatively long, the ability of its glycosynthase to synthesize the long oligomeric saccharides was explored [28]. Hommalai et al. reported that the mutations of rice BGlu1 β -glucosidase at glutamate residue 386 (the residue was numbered according to its position in the BGlu1 precursor, as E414 in that paper) destroyed the hydrolytic activities of the enzyme. When the catalytic activities of three candidate mutants, E386G, E386S and E386A, were compared with α -glucosyl fluoride (α -GlcF) as donor and pNP-cellobioside (pNPC2) as acceptor, all these mutants displayed transglucosylation activities to synthesize mixed length oligosaccharides (see Fig. 1). This enzymatic oligosaccharide synthesis reaction was proposed to be a concerted single-step mechanism [6]. The acid/base residue abstracts a proton from one hydroxyl group of substrate pNPC2, as this deprotonated hydroxyl group attacks on the anomeric carbon of α -GlcF, and then the fluorine atom departs. The BGlu1 E386G mutant gave the fastest transglucosylation rate, which was approximately 3- and 19-fold faster than those of E386S and E386A, respectively, and gave yields of up to 70–80% insoluble products.

Recently, the structure of the BGlu1 E386G glycosynthase mutant alone and in complexes with α -GlcF, cellotetraose or cellopentaose were reported [29]. However, those structures could not explain the relative activities of the BGlu1 E386G, E386S and E386A glycosynthases. Therefore, in this paper we solved the structures of the E386S and E386A glycosynthases complexed with α -GlcF for comparison and explored the dynamic issues contributing to these relative activities by the QM/MM method.

Using the Gaussian frequency data, we also determined the kinetic isotope effect (KIE) of the proton transfer process in the E386 mutants based on two expressions [30]. One expression is the semi-classical Eyring equation, where the KIE is given as

$$\left(\frac{k_H}{k_D}\right)_s = \exp\left(-\frac{(G_H - G_H^R) - (G_D - G_D^R)}{RT}\right) \quad (1)$$

The other one is a simple “quantum correction” equation with the semiclassical $(k_H/k_D)_s$ multiplying the Wigner quantum correction Q^{corr} factor

$$\left(\frac{k_H}{k_D}\right)_w = Q^{\text{corr}} \left(\frac{k_H}{k_D}\right)_s \quad (2)$$

$$Q^{\text{corr}} = \frac{1 + \frac{u_t^2}{24}}{1 + \frac{u_t'^2}{24}} \quad (3)$$

$$u_t = \frac{h\nu_H}{k_B T} \quad (4)$$

$$u_t' = \frac{h\nu_D'}{k_B T} \quad (5)$$

where ν is the imaginary frequency of the transition state.

2. Methods

2.1. Protein X-ray crystal structure determination

The recombinant proteins of BGlu1 E386A and BGlu1 E386S [28] were expressed and purified as previously described for wild type BGlu1 [22,27]. Crystallization of BGlu1 E386S and E386A without and with 10 mM α -GlcF (kindly provided by Prof. Stephen G. Withers) was optimized by hanging drop vapor diffusion with microseeding around the conditions used for crystallization of BGlu1 [22,27], varying the concentrations of polyethylene glycol monomethyl ether (PEG MME) 5000 over the range of 16–26%, $(\text{NH}_4)_2\text{SO}_4$ between 0.12 and 0.26 M, and protein between 2 and 6 mg/ml in 0.1 M MES, pH 6.7, at 288 K. Before flash cooling in liquid nitrogen, the crystals with 10 mM α -GlcF were soaked

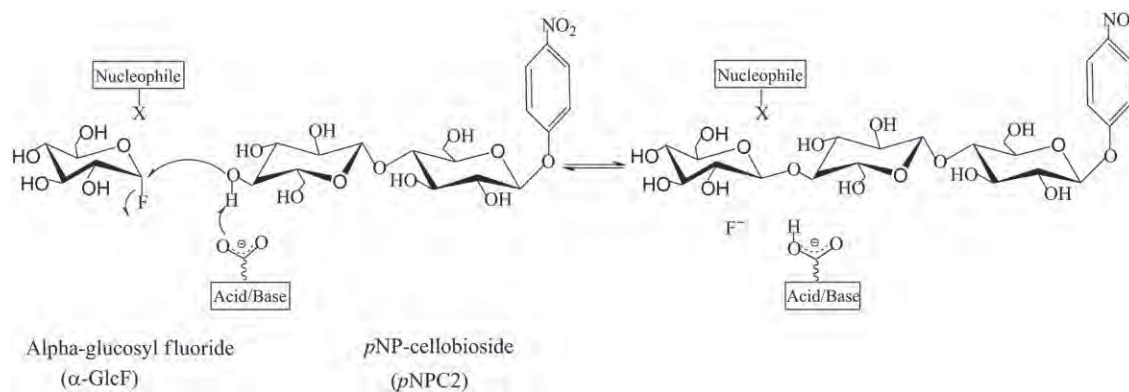


Fig. 1. Proposed oligosaccharide synthesis mechanism of glycosynthase derived from a nucleophile mutant of retaining glycosidases. X represents the mutated nucleophile residue.

in cryo solution (18% (v/v) glycerol in precipitant solution) containing 10 mM α -GlcF for 1–5 min.

Preliminary diffraction experiments were done with a Cu K α rotating anode X-ray source mounted on a MicroSTAR generator operating at 45 kV and 60 mA connected to Rayonix SX-165 CCD detector at the Synchrotron Light Research Institute (Public Organization) (SLRI, Nakhon Ratchasima, Thailand). Promising crystals were used to diffract 1.0 Å wavelength X-rays on the BL13B1 beamline at the National Synchrotron Radiation Research Center (NSRRC in Hsinchu, Taiwan), and reflections were recorded with an ADSC Quantum 315 CCD detector. The crystals were maintained at 105 K during diffraction with a nitrogen cold stream (Oxford Instruments). Data were processed and scaled with the HKL-2000 package [31]. Structures were solved by molecular replacement with wild type BGlu1 (PDB code: 2RGL), refined, and validated as previously described for BGlu1 E386G and its complex with α -GlcF [29].

2.2. Automated docking setup

Recently, we published the structure of the E386G mutant and its complex with α -GlcF (PDB code: 3SCO) [29]. Since the crystal of rice BGlu1 E386G mutant was only complexed with the donor α -GlcF, and the transglucosylation activities were detected under the acceptor pNPC2, the pNPC2 was docked into the binding pocket using the Autodock 4.0 program [32]. Before docking, the pNPC2 substrate was optimized at the B3LYP/6-31 + G(d) level with the Gaussian 03 package [33]. When docking, the grid scale was set as 60 Å × 60 Å × 60 Å based on a grid module, with a spacing of 0.375 Å between the grid points. Gasteiger charges [34] were set for both the ligand and protein. Fifty independent docking runs were performed. In the whole calculation, the protein was kept rigid, and all the torsional bonds of the ligand were kept free. Based on a root-mean-square deviation (RMSD) criterion of 10 Å, the docking results were clustered. Finally, the conformation with the most cluster members and the lowest protein-ligand interaction energy was chosen as the bioactive structure.

The models of the newly obtained BGlu1 E386A and BGlu1 E386S were constructed the same as the crystal structure of E386G mutant mentioned above, in which the acceptor substrate pNPC2 was also docked into these active sites. The parameter settings were consistent with those of E386G mutant. As the derived docking conformations of the E386S and E386A mutants were similar with that of E386G mutant, only the obtained docking structure of E386G mutant is shown (see Fig. 2).

2.3. Computational model

The obtained docking ternary complex structures of BGlu1 E386G, E386S and E386A glycosynthases were used as the initial structures in molecular dynamics (MD) simulations. For the glycosynthase mechanism of these mutants, the catalytic acid/base (E176) acts as a base to extract a proton from the substrate, and was modeled in its deprotonated state. The protonation/deprotonation states of other ionizable residues were altered on the basis of PROPKA method [35,36]. The Mulliken charge parameters of the substrates (α -GlcF and pNPC2) were achieved based on the QM (at the B3LYP/6-31 + G(d) level) calculations. All the hydrogen atoms were added with the HBUILD facility of the CHARMM package [37]. The crystallographic water molecules were kept in their original positions and an extra 4,648 water molecules were used to solvate each of the three systems with a 37.9 Å water sphere centered on E176. Nine Cl[−] ions were also added at random positions to neutralize each system. To equilibrate the prepared systems, a minimization followed by a 1000 ps MD simulation was performed with the CHARMM22 force field [38]. To confirm the equilibrium of the initial model, the RMSD curves for the mutants of E386G, E386S and E386A have been checked, which are given in supporting information (Fig. S1–S3).

During the subsequent QM/MM calculations, the QM region included residues E176, Y315, E440, the mutated nucleophile residue X386, the substrates α -GlcF, pNPC2 (terminal glucosyl residue) and one order water Wat1, where X referred to Gly, Ser and Ala for the E386G, E386S and E386A mutants, respectively (see Fig. 3). Fig. 3 shows the labels to indicate the serial numbers of relevant atoms in the substrate rings. The total number of QM atoms were 85, 86 and 88 for the E386G, E386S and E386A mutants, respectively. The remainder of the enzyme and waters was set as the MM region. Any residue with at least one atom within 10 Å of E176 (including the QM region and part of the MM region) was kept loose, while the remaining part was kept frozen. In the geometry optimizations, the QM region was treated with quantum mechanics by the Turbomole module [39] and the MM part with molecular mechanics under the CHARMM22 force field by the DL-POLY program [40]. For the whole QM/MM optimizations, calculations were performed at the B3LYP/6-31G(d,p)//CHARMM22 level. The electrostatic interactions between the QM and MM regions were described by the standard electronic embedding scheme [41]. To avoid the hyperpolarization effect during the QM treatment, the MM atomic partial charges were incorporated into the one-electron Hamiltonian of the QM calculation. The charge shift model with hydrogen linked atoms was used to simulate covalent bonds across the QM/MM boundary [42]. The ChemShell package [43], incorporating

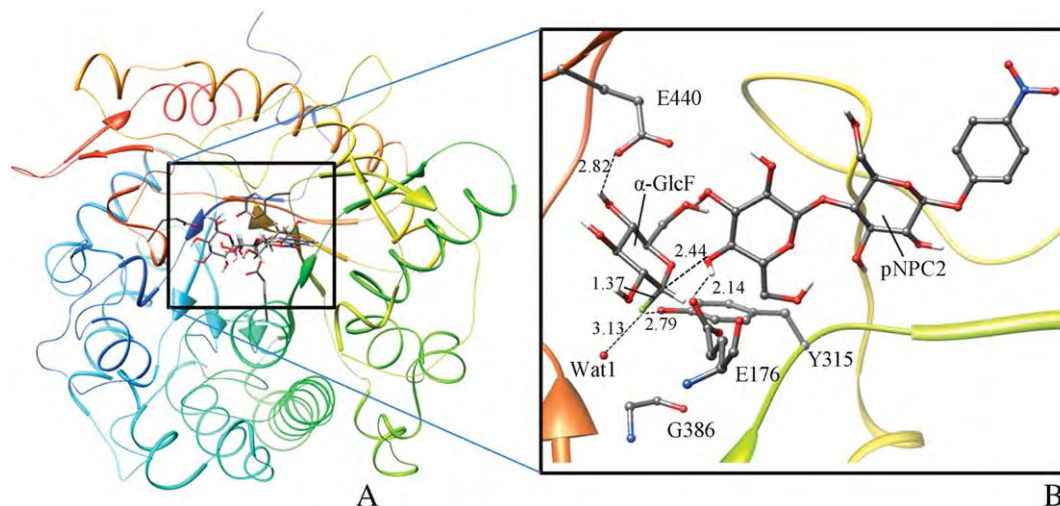


Fig. 2. (A) The overall obtained docking structure of rice BGlu1 E386G mutant (PDB code: 3SCO). (B) The corresponding residues in the active site.

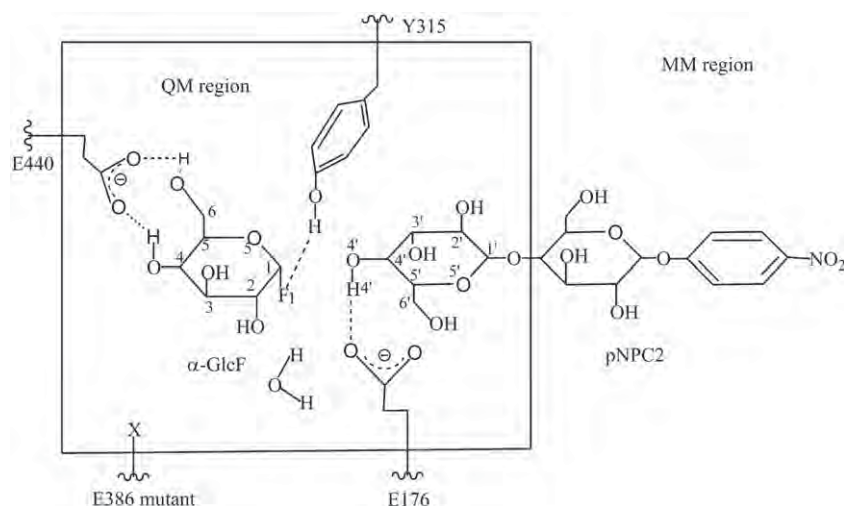


Fig. 3. The selected quantum mechanics (QM) region in QM/MM calculations; X refers to the mutated nucleophile residue in the active site.

the Turbomole and DL-POLY programs, were used to perform the QM/MM calculations. A geometry optimizer of hybrid delocalized internal coordinates (HDLC) [44] was adopted for the geometry optimizations. Stationary points were searched by the quasi-Newton limited memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm [45,46], and transition states with the algorithm of partitioned rational function optimization (P-RFO) [47,48]. The vibrational frequencies of the QM region atoms were determined at the same level by using the Gaussian 03 package [33]. After all the stationary points were located, high level single point electronic energy calculations were performed at a larger basis set 6-31++G(d,p) to obtain accurate energies.

3. Results and discussion

3.1. 3D Structures of three glycosynthases in complexes with α -GlcF

To assess the structural basis for the differing glycosynthase efficiencies of BGlul E386G, E386S and E386A, the structures of the E386S and E386G glycosynthases and those of their complexes with α -GlcF were determined by X-ray crystallography and compared with the corresponding structure of E386G. Crystals of the apo glycosynthase enzymes and of their complexes were isomorphous with wild type BGlul crystals [22,27]. Diffraction data parameters are shown in Table S1 and model parameters for the structures of the glycosynthases are summarized in Table 1. The resolution limits of 1.85 Å for BGlul E386S and 2.10 Å for E386A and free residual (R_{free}) values of 20.5% for E386S and 20.6% for E386A in complexes with α -GlcF are comparable to the previously reported values for the corresponding complex of E386G of 1.95 Å and 20.7% [29]. However, the temperature (B) factors reported for the E386G protein (16.6 Å²) and α -GlcF (14.0 Å²) in the E386G complex are significantly higher than those in the E386S (9.5 and 5.7 Å², respectively) and E386A (10.8 Å² and 8.2 Å², respectively) complexes. These differences may support the supposition of a greater flexibility in the E386G protein and its complex, although we cannot exclude other factors contributing to the higher crystal structure disorder.

The overall structures of the three BGlul glycosynthase mutants and the positions of residues and α -GlcF in the active site are very similar. In each structure, the α -GlcF binds to the glycosynthases in a relaxed ⁴C₁ chair conformation stacked onto the indole ring of W433 at the −1 subsite, in the same position as the 2-deoxy-2-fluoro- α -D-glucosyl moiety (G2F) in its covalent complex with BGlul (PDB code: 2RGM) [22,27], with the same hydrogen bonds as observed for that complex. As shown in Fig. 4, the active site residue N313, as well as the fluorine and O2 of α -GlcF, interact with a single ordered water molecule in all

three mutant complexes. An ordered water molecule is also found in the same position in the structure of the covalent intermediate of BGlul with G2F, and this is also the position of the nucleophile carbonyl oxygen of apo wild type BGlul (PDB code: 2RGL) [22,27]. As seen in Fig. 4, the water is closer to the fluorine in the E386S structure, while it is generally closer to the hydrogen-bonding side chains in the BGlul E386G structure and is in an intermediate position in the BGlul E386A structure, but these differences may not be significant.

Although the determination of the structures of the BGlul E386G, E386A and E386S glycosynthases in apo and α -GlcF-bound forms is the most thorough structural investigation of one enzyme's glycosynthase mutants to date, the structures alone cannot easily explain the relative activities of BGlul E386G, E386S and E386A. The previous idea that a Gly nucleophile may be a better catalyst than a Ser nucleophile due to

Table 1
Structure refinement statistics for E386A and E386S X-ray structures.

	E386S	E386S_α-GlcF	E386A	E386A_α-GlcF
PDB code	3SCR	3SCS	3SCP	3SCQ
Resolution range (Å)	23–1.80	25–1.85	30–2.10	25–2.10
No. of amino-acid residues	944	944	944	944
No. of protein atoms	7612	7612	7610	7610
No. of water molecules	865	870	760	782
Refined carbohydrate	None	α-GlcF	None	α-GlcF
No. of carbohydrate atoms	None	24 ^c	None	24 ^c
No. of other hetero atoms	47	47	47	35
R_{factor} (%) ^a	17.8	17.5	17.8	17.7
R_{free} (%) ^b	20.8	20.5	20.8	20.6
Ramachandran statistics ^c				
Most favored regions (%)	89.2	90.0	89.2	89.9
Additionally allowed regions (%)	10.8	9.9	10.8	9.9
Outlier regions (%)	0.0	0.1	0.0	0.2
R.m.s.d. from ideality				
Bond distances (Å)	0.013	0.013	0.013	0.010
Bond angle (°)	1.40	1.38	1.40	1.30
Estimated coordinate error (Å) ^d	0.12	0.12	0.17	0.17
Mean B factors (Å ²)				
All protein atoms	11.5	9.5	12.6	10.8
Waters	24.5	20.6	24.7	20.9
Hetero atoms	24.6	21.8	33.9	24.9
Carbohydrate atoms	None	5.7	None	8.2
Subsite −1 α-GlcF (A/B)	None	5.8/5.6	None	7.3/9.1

^a $R_{\text{factor}} = (\sum |F_o| - |F_c|) / \sum |F_o|$.

^b Based on 5% of the unique observations not included in the refinement.

^c Ramachandran values were determined from PROCHECK [51].

^d Based on R_{free} values, as calculated by Refmac 5.5.0110 [52].

^e Twelve per protein molecule in the asymmetric unit.

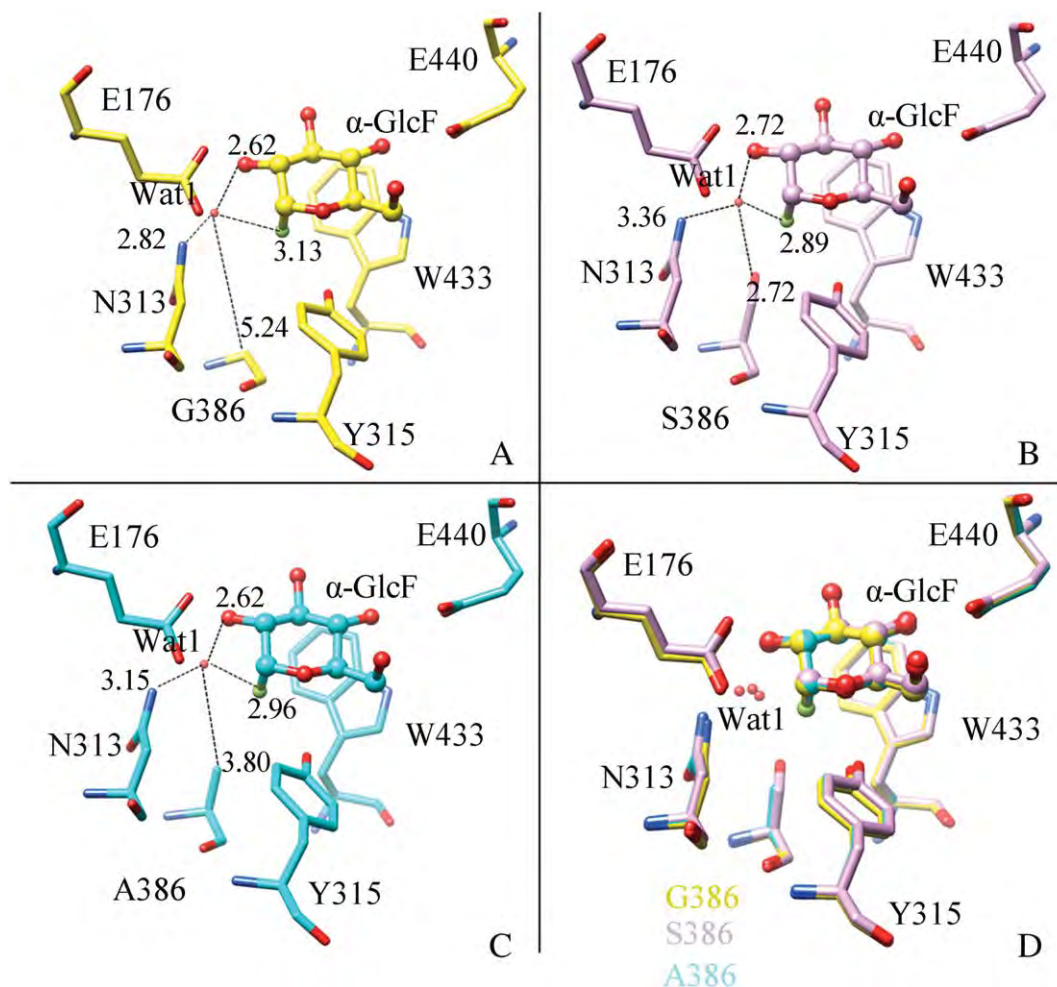


Fig. 4. Comparison of the active sites of rice BGlu1 E386G, E386A and E386S glycosynthases. An active site water molecule is seen in the wild type covalent intermediate as well as in all three glycosynthase mutant complexes with α -GlcF. This water occupies a very similar position in BGlu1 E386G [29], BGlu1 E386A (B) and BGlu1 E386S (C), with the superimposition of all three of these rice BGlu1 glycosynthases being shown in (D). The α -GlcF is represented as balls and sticks, the surrounding amino acid residues are shown as sticks, and the water molecule is represented as a ball. Oxygen is shown in red, nitrogen in blue, and fluorine in green. Distances (in Angstroms) are shown from the one water molecule to the surrounding polar groups and the mutated sidechains. No other water molecules were observed in this part of the active site in any of the structures. The figures of protein structures were generated in PyMol (Schrödinger LC).

a water taking the place of the Ser hydroxyl but with more optimal geometry [49], since only the one water molecule was observed in this area in both structures. As described above, this water molecule appears in a critical position to stabilize the leaving group in all three structures, and S386 points toward it to stabilize its position in the E386S mutant, which could be used to explain the higher activity of E386S than E386A. However, this gives little insight into the high relative activity of the E386G mutant. As noted by Ducros et al. [50], X-ray structures, such as those described here of initial Michaelis complex structures, do not account for changes in the ternary complex with the acceptor and at the transition state, which may be more critical for determining the relative activities. Therefore, a combined quantum mechanical/molecular mechanical (QM/MM) calculations were performed to account for the difference of the relative activities of the reactions.

3.2. Docking structure

Since no acceptor density was seen in crystals containing α -GlcF and pNPC2 (data not shown), docking was used to produce the ternary complex structures. The overall structure of rice BGlu1 E386G mutant in complex with α -GlcF and docked pNPC2 is shown in Fig. 2A, and the pocket residues are given in Fig. 2B. The donor α -GlcF is fixed by two hydrogen bonds (HBs), with its fluorine atom (F1, as labeled in Fig. 3) to the phenolic hydroxyl of residue Y315 (2.79 Å)

and the C4-OH group to the E440 carboxyl (2.82 Å). The important crystal water molecule Wat1 lies near the F1 atom with a distance of 3.13 Å. The acceptor pNPC2 locates in an appropriate position to facilitate the reaction, and forms one strong HB with the acid/base residue E176 with a distance of 2.14 Å. Furthermore, the distance between C1 atom of α -GlcF and O4' of C4'-OH group of pNPC2 is close enough for the formation of the glycosidic bond with a distance of 2.44 Å. Thus, the docking structure appears reasonable for the following QM/MM calculations.

3.3. Analyses of reaction pathways

For the mutants of E386G, E386S and E386A, the optimized geometries of the reactants, transition states and products obtained at the 6-31G(d,p) level are shown in Figs. 5–7, respectively. The bond distances which are formed or broken in each state are colored magenta. Furthermore, the single point energies were calculated at the 6-31++G(d,p) level, and the obtained energy profiles for each mutant are shown in the corresponding figures.

3.3.1. E386G mutant

Fig. 5 gives the structures of reactant (R1), transition state (TS1), product (P1) and energy profile of the oligosaccharide synthesis process catalyzed by E386G Mutant.

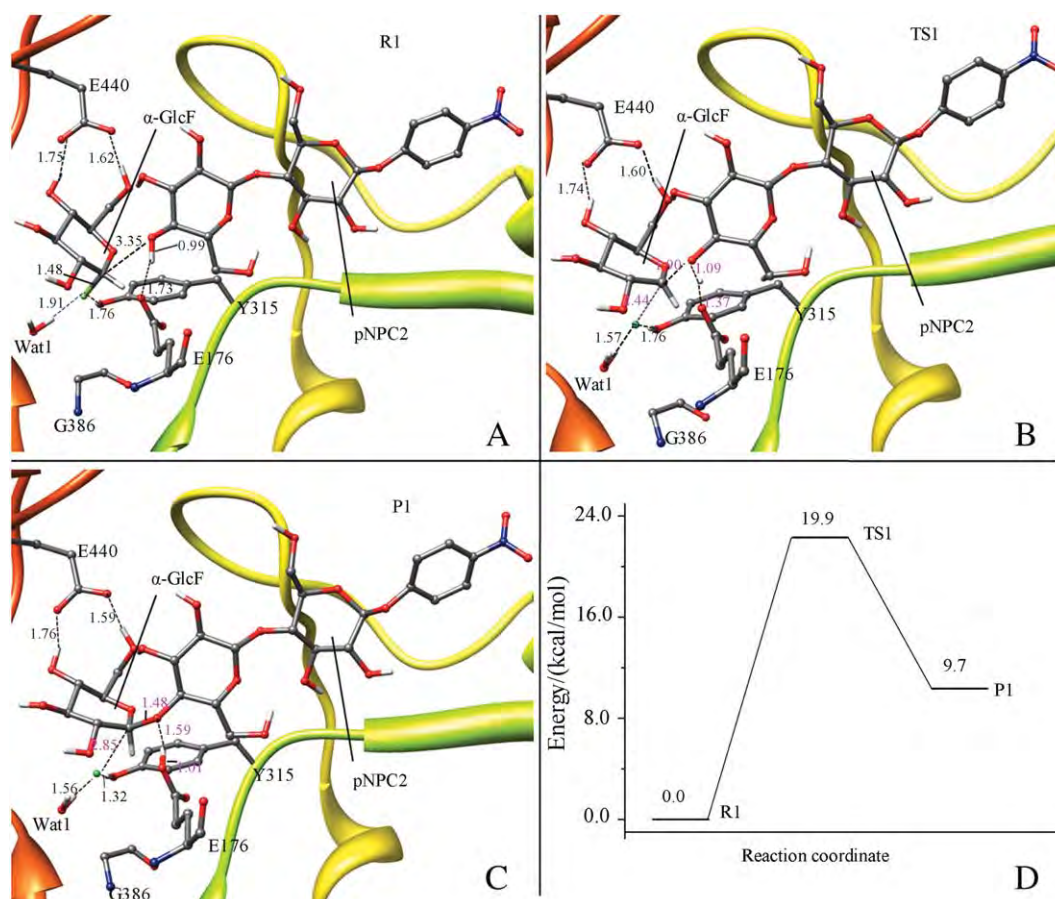


Fig. 5. Optimized structures of Reactant (A), Transition state (B) and Product (C) of glycosynthase reaction and energy profile (D) for the E386G mutant. The distances of newly formed or broke bonds are indicated in magenta.

In **R1** (Fig. 5A), both of the donor α -GlcF and acceptor *p*NPC2 are located in favorable positions for the reaction to occur. Compared with the docking structure in Fig. 2B, the donor α -GlcF is still hydrogen bonded with residues E440 and Y315, but the interactions of these HBs have been strengthened. For example, the distance between F1 atom and hydroxyl group of Y315 decreases to 1.76 Å. A new HB forms between E440 and C6–OH of α -GlcF with a distance of 1.62 Å. Furthermore, the crystal water Wat1 comes closer to F1 atom, establishing a new HB (1.91 vs 3.13 Å in Fig. 2B). The acceptor *p*NPC2 moves a little further away from α -GlcF, but it still remains at a reasonable distance (3.35 Å) for the formation of glycosidic bond. In addition, the HB length between the C4'–OH group and side chain of acid/base E176 decreases from 2.14 Å in Fig. 2B to 1.73 Å, which is feasible for the nucleophilic attack of E176.

At **TS1** (Fig. 5B), the *p*NPC2 O4'–H4' bond has almost broken (bond length distance of 1.09 Å). The proton H4' is much closer to the acid/base residue E176 with a distance of 1.37 Å. As E176 extracts the proton from the *p*NPC2 C4'–OH group, the interaction between the donor and acceptor is also strengthened. The distance between the C1 and O4' atoms decreases to 1.90 Å (3.35 Å in **R1**). The leaving F1 atom moves away from α -GlcF with a distance of 2.44 Å and forms two strong HBs with water Wat1 and Y315.

In the structure of **P1** (Fig. 5C), the O4'–H4' bond in *p*NPC2 has been further elongated to 1.59 Å and the H4' atom has formed a covalent H–O bond with E176 (1.01 Å). The glycosidic bond between C1 and O4' has formed with a distance of 1.48 Å. The departed F1 atom has moved far away from C1 (2.85 Å), but still forms two strong HBs with Wat1 and Y315 with distances of 1.56 and 1.32 Å, respectively. Obviously, the water Wat1 and residue Y315 play an important

role in stabilizing the departed F1 atom. Fig. 5D shows that the energy barrier is 19.9 kcal/mol, and the reaction is endothermic with the reaction energy of 9.7 kcal/mol.

3.3.2. E386S mutant

Fig. 6 gives the optimized structures of reactant (**R2**), transition state (**TS2**), product (**P2**) and energy profile of E386S mutant enzyme.

In **R2** (Fig. 6A), the binding modes of substrates α -GlcF and *p*NPC2 are similar with that of **R1** in Fig. 5A. The side chain of E176 still forms one strong HB with C4'–OH group of *p*NPC2, while the α -GlcF is H-bonded with E440, Y315 and water Wat1. The main difference is the newly formed HB between F1 atom and side chain of S386 (the distance of which is 2.42 Å). By comparing **R2** with **R1**, we find that though S386 has established the HB interaction with the donor, its side chain makes the active site crowd, which may be unfavorable for the departure of the leaving F1 atom.

Fig. 6B shows the transition state **TS2**. The *p*NPC2 O4'–H4' covalent bond has been weakened, but it remains bonded to O4' (1.03 vs 0.99 Å) in **R2**. With the weakness of the O4'–H4' bond, the interaction between the donor and acceptor has been strengthened: the distance between the C1 and O4' atoms decreases from 3.55 Å in **R2** to 2.17 Å. Thus, the glycosidic bond C1–O4' has been partially formed. In this structure, as the donor comes closer to the acceptor, the C1–F1 bond length increases to 2.62 Å, indicating this bond also becomes much weakened. The leaving F1 atom is hydrogen bonded with Wat1, S386 and Y315 with distances of 1.63, 1.68 and 1.57 Å, respectively. It can be seen that the leaving F1 atom locates in a crowd space, which may hinder the departure of leaving group and thereby influence the activity of the E386S mutant.

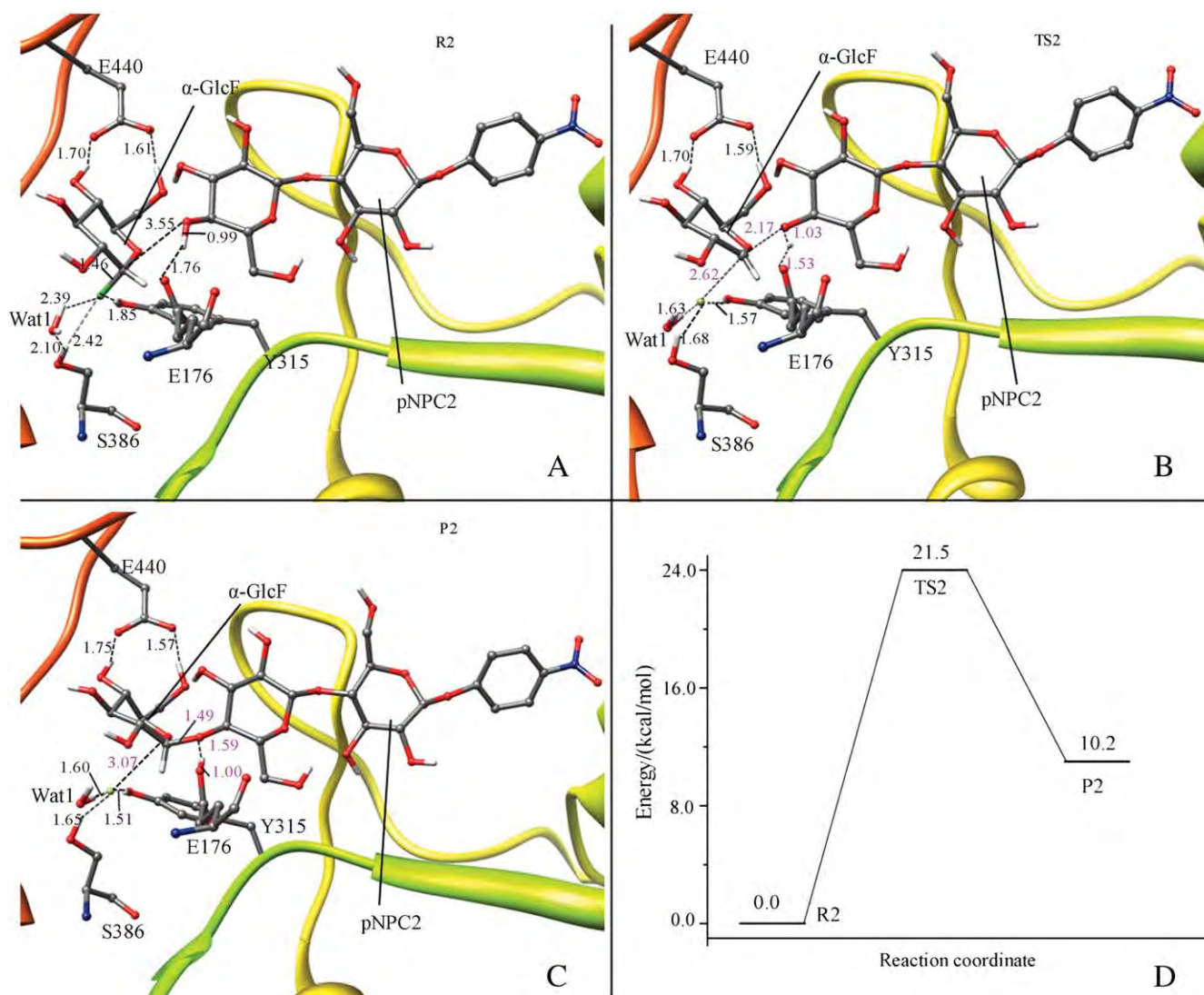


Fig. 6. Optimized structures of Reactant (A), Transition state (B) and Product (C) of glycosynthase reaction and energy profile (D) for the E386S mutant. The distances of newly formed or broke bonds are indicated in magenta.

In the product, **P2** (Fig. 6C), the acid/base E176 has extracted the H4' proton from the C4'-OH group of pNPC2 forming a new H-O bond (distance of which is 1.00 Å). The donor of α -GlcF has connected with pNPC2, establishing a new glycosidic bond of 1.49 Å. At the same time, the leaving F1 atom moves further away (3.07 Å from the C1 atom) but still interacts with the surrounding residues (Y315 and S386) and water Wat1. All these indicate that the oligosaccharide synthesis process is accomplished at this state. Fig. 6D shows that the energy barrier is 21.5 kcal/mol, which is higher than that of E386G mutant (19.9 kcal/mol). The possible reason might be that the side chain of S386 is larger than that of G386, making the active site crowded, which restricts the freedom of the water Wat1 to interact optimally with the departing fluoride and blocks the departure of leaving group (F1 atom).

3.3.3. E386A Mutant

Fig. 7 gives the structures of reactant (**R3**), transition state (**TS3**), product (**P3**) and energy profile of the E386A mutant enzyme.

In the reactant **R3** (Fig. 7A), the side chain of A386 is a methyl group, which is bigger than that of residue G386 in E386G. The sterically hindered methyl group occupies the active site, which is relatively unfavorable for the binding of α -GlcF. The distances between

the F1 atom and the surrounding Wat1 and Y315 are 2.01 and 1.81 Å, respectively, which are longer than those of the distances in **R1** (lengths of which are 1.91 and 1.76 Å, respectively). Besides, the extra methyl group also pushes the donor away from the acceptor: the distance between C1 and O4' atoms is increased from 3.35 Å in Fig. 5A to 3.41 Å in Fig. 7A.

In **TS3** (Fig. 7B), the pNPC2 O4'-H4' bond has been partially broken (distance of which is 1.07 vs. 0.99 Å in **R3**) and the proton H4' has partially moved to the carboxyl oxygen of E176 (1.41 vs. 1.73 Å in **R3**). With the weakness of the O4'-H4' bond, the C1 atom of α -GlcF has come closer to the O4' atom of pNPC2 with a distance of 1.93 Å. Therefore, the interaction between the donor and acceptor has been strengthened and the glycosidic bond of C1-O4' has been partially established. With the formation of the C1-O4' bond, the C1-F1 bond becomes weakened, and its bond length increases to 2.52 Å. Additionally, the leaving group F1 atom is still hydrogen bonded with Wat1 and Y315 with distances of 1.60 and 1.42 Å, respectively, which are favorable for the departure of F1 atom.

In product **P3** (Fig. 7C), the pNPC2 O4'-H4' covalent bond has broken with a long distance of 1.56 Å, and the H4' atom has formed a covalent H-O bond with the side chain of E176 (distance of 1.01 Å). The glycosidic bond C1-O4' is completely established (length of 1.48 Å).

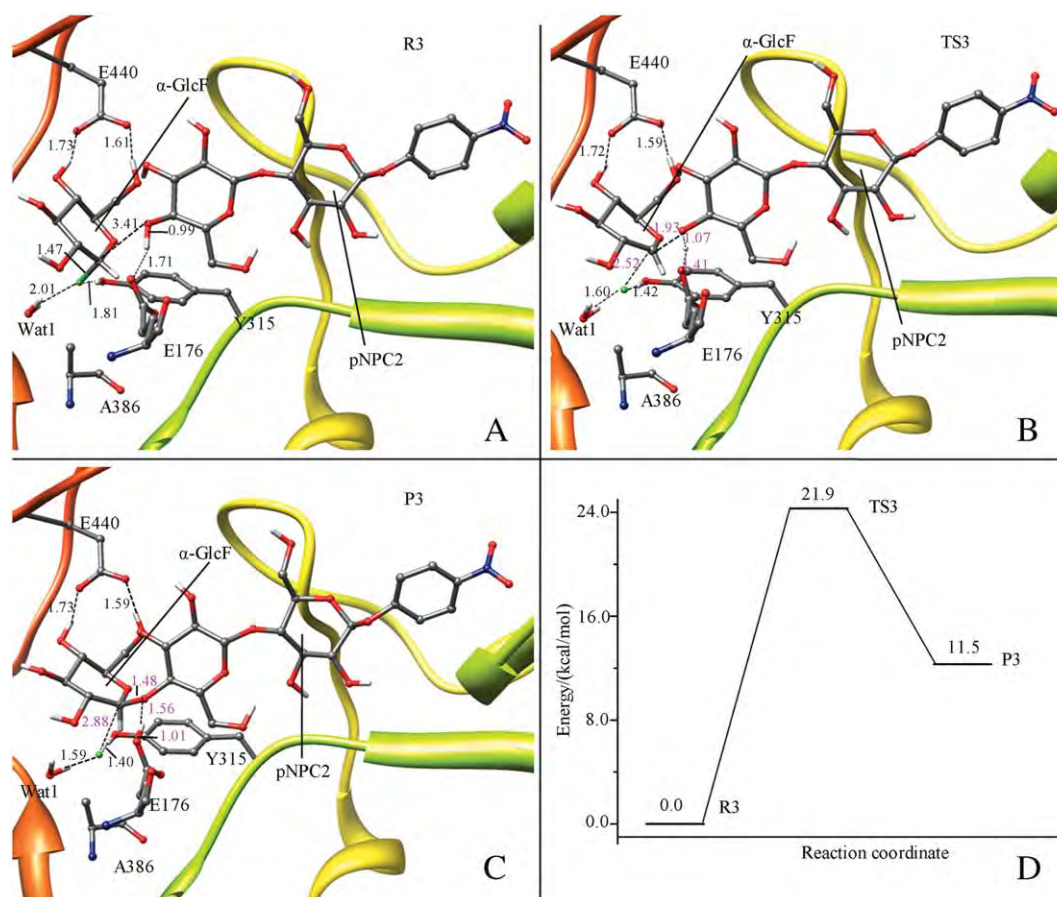


Fig. 7. Optimized structures of Reactant (A), Transition state (B) and Product (C) of glycosynthase reaction and energy profile (D) for the E386A mutant. The distances of newly formed or broke bonds are indicated in magenta.

at this state. Since the residue A386 has a methyl group on its side chain, it has no HB interaction with the leaving F1 atom, which is different from residue S386 in the E386S mutant. Similar to the structure of **P2**, the departed F1 atom also forms two strong HBs with Wat1 and Y315 with distances of 1.59 and 1.40 Å, respectively, indicating that the water Wat1 and Y315 are important in stabilization of the leaving group. Owing to the steric hindrance of A386, the activity of this mutant is lower than that of the previous two mutants, which can be reflected from the energy barrier. Fig. 7D shows that the energy barrier of this reaction is 21.9 kcal/mol, while for E386G this barrier is 19.9 kcal/mol and for E386S is 21.5 kcal/mol. The order of the energy barriers is consistent with the experimental relative activities of these enzymes [28]: the experimental k_{cat} value of Gly mutant is about 6 per minute, and 2 per minute for Ser, and about 0.3 per minute for Ala. But based on calculated activation energies (19.9, 21.5, and 21.9 kcal/mol for Gly, Ser and Ala mutants, respectively), 1.1 kcal/mol energy difference translates to more than 100 fold rate difference with Gly mutant. Although the rate limiting step in these mutants has not been established, Shoham has suggested that product release is likely a rate limiting step in some glycosynthases [53]. Considering the size of the product and the nature of the active site, it is likely that the product release, rather than chemistry, is rate limiting in BGlut1 mutants. Glycine may well be the fastest as it presents little steric hindrance for the departure of the trisaccharide from the active site.

The crystal structures of E386G, E386S and E386A (Fig. 4) shows that the mutant structures look very similar and one critical water molecule (Wat1) appears in the active site. The calculations were also performed with the absence of Wat1 in the QM regions at the same level. The

corresponding optimized structures (reactants, transition states and products) and energy profiles were shown in the Supporting Information (Fig. S4–S6). The geometries of these structures are similar with those structures with the presence of the water Wat1 in the QM regions (see Figs. 5–7). Differences can be seen in the energy barriers. In the absence of Wat1, the energy barrier is calculated to be 22.4 kcal/mol for catalytic reaction of E386G mutant (Fig. S4), 25.7 kcal/mol for E386S (Fig. S5) and 28.2 kcal/mol for E386A (Fig. S6). We can see that all the energy barrier are higher than those of corresponding mutants with considering the water Wat1 in the QM regions, but they are all consistent with the order of activities [28]. The lower barriers indicate that this crystal water performs an assist role in the catalytic reaction of three mutants. Obviously, the hydrolysis reactions of E386 mutants are endothermic, the extra energy may be abstracted from the surroundings.

3.3.4. Kinetic Isotope Effect

To approximately evaluate the nuclear quantum dynamical effect on the proton transfer in E386 mutants, we calculated the semiclassical kinetic isotope effect (KIE) based on the QM region defined in the above QM/MM calculation [54,55]. The required free energy of activation and vibrational frequencies were determined at the B3LYP/6-31 G(d,p) level by using the Gaussian 03 package [33]. Detailed formulations of the calculations are shown in the Supporting Information. Both the semiclassical KIE values and the derived Wigner-corrected KIE values for the proton-transfer processes are listed in Table 2. It exhibits that the proton-transfer processes of all the three E386 mutants give the small KIE and Wigner tunneling corrections,

Table 2

Predicted imaginary frequencies for the transition states and KIE values for corresponding proton-transfer processes.

TS	Frequencies (cm ⁻¹)		H/D KIE	
	H	D	KIE ^{semi}	KIE ^{Wigner}
TS1 (E386G)	−300.6	−292.4	2.39	2.41
TS2 (E386S)	−109.2	−109.1	1.40	1.40
TS3 (E386A)	−229.3	−227.2	1.87	1.87

meaning that the quantum mechanical tunneling effect is insignificant for E386 mutants.

4. Conclusion

The oligosaccharide synthesis mechanism catalyzed by rice BGLu1 E386 mutants (E386G, E386S and E386A) has been studied by crystallographic and QM/MM approaches. The structures of these glycosynthases were quite similar, so the computational approach was necessary to account for the dynamic issues leading to the differences in their activities. In the dynamic model, the extraction of the proton H4' (pNPC2) by the acid/base E176, the formation of the glycosidic bond (C1-O4') and the departure of the leaving group (F1 atom) are concerted in a single step. Our calculations indicate that the energy barriers are influenced by the mutated residues in the active sites. The energy barrier was calculated to be 19.9 kcal/mol for E386G mutant, 21.5 kcal/mol for E386S mutant and 21.9 kcal/mol for E386A mutant, which was consistent with the order of their experimental relative activities. The lower energy barrier of the glycine mutant than the serine mutant might be due to the greater freedom of the water Wat1 to interact optimally with the departing fluoride in the glycine mutant and the steric constraints on this departure imposed by the Ser hydroxymethyl side chain. We also demonstrate that, the presence of Wat1 is able to assist the catalytic reaction for each mutant. Although the rate limiting step in these mutants has not been established, considering the size of the product and the nature of the active site, it is likely that the product release, rather than chemistry, is rate limiting in BGLu1 mutants.

Accession Numbers

Coordinates and structure factors have been deposited in the Protein Data-bank with accession codes 3SCR, 3SCS, 3SCP and 3SCQ.

Acknowledgments

This work was supported by the Natural Science Foundation of China (21173129), the Thailand Research Fund (grant BRG380017) and the Commission on Higher Education National Research University Project to Suranaree University of Technology, which supported SP as a postdoctoral fellow.

Appendix A Supplementary data

Diffraction data parameters for E386S and E386A mutants (Table S1); The RMSD curves for the mutants of E386G, E386S and E386A (Fig. S1–3); The optimized structures (reactants, transition states and products) and energy profiles with the presence of Wat1 in the MM regions (Fig. S4–6). This material is available free of charge via the Internet. Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.bbapap.2012.11.003>.

References

- [1] P. Sears, C.H. Wong, Toward automated synthesis of oligosaccharides and glycoproteins, *Science* 291 (2001) 2344–2350.
- [2] C.R. Bertozzi, L.L. Kiessling, Chemical glycobiology, *Science* 291 (2001) 2357–2364.

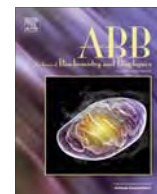
- [3] D. Koshland Jr., Stereochemistry and the mechanism of enzymatic reactions, *Biol. Rev.* 28 (1953) 416–436.
- [4] K.M. Koeller, C.H. Wong, Emerging themes in medicinal glycoscience, *Nat. Biotechnol.* 18 (2000) 835–841.
- [5] M.M. Palcic, Biocatalytic synthesis of oligosaccharides, *Curr. Opin. Biotechnol.* 10 (1999) 616–624.
- [6] L.F. Mackenzie, Q. Wang, R.A.J. Warren, S.G. Withers, Glycosynthases: mutant glycosidases for oligosaccharide synthesis, *J. Am. Chem. Soc.* 120 (1998) 5583–5584.
- [7] S.J. Williams, S.G. Withers, Glycosyl fluorides in enzymatic reactions, *Carbohydr. Res.* 327 (2000) 27–46.
- [8] M. Moracci, A. Trincone, M. Rossi, Glycosynthases: new enzymes for oligosaccharide synthesis, *J. Mol. Catal. B: Enzym.* 11 (2001) 155–163.
- [9] S.J. Williams, S.G. Withers, Glycosynthases: mutant glycosidases for glycoside synthesis, *Aust. J. Chem.* 55 (2002) 3–12.
- [10] G. Perugini, B. Cobucci-Ponzano, M. Rossi, M. Moracci, Recent advances in the oligosaccharide synthesis promoted by catalytically engineered glycosidases, *Adv. Synth. Catal.* 347 (2005) 941–950.
- [11] M. Okuyama, H. Mori, K. Watanabe, A. Kimura, S. Chiba, α -Glucosidase mutant catalyzes “ α -glycosynthase”-type reaction, *Biosci. Biotechnol. Biochem.* 66 (2002) 928–933.
- [12] Y. Honda, M. Kitaoka, The first glycosynthase derived from an inverting glycoside hydrolase, *J. Biol. Chem.* 281 (2006) 1426.
- [13] R.V. Stick, K.A. Stubbs, A.G. Watts, Modifying the regioselectivity of glycosynthase reactions through changes in the acceptor, *Aust. J. Chem.* 57 (2004) 779–786.
- [14] Y.W. Kim, S.S. Lee, R.A.J. Warren, S.G. Withers, Directed evolution of a glycosynthase from *Agrobacterium* sp. increases its catalytic activity dramatically and expands its substrate repertoire, *J. Biol. Chem.* 279 (2004) 42787–42793.
- [15] Y.W. Kim, H. Chen, S.G. Withers, Enzymatic transglycosylation of xylose using a glycosynthase, *Carbohydr. Res.* 340 (2005) 2735–2741.
- [16] A. Trincone, G. Perugini, M. Rossi, M. Moracci, A novel thermophilic glycosynthase that effects branching glycosylation, *Bioorg. Med. Chem. Lett.* 10 (2000) 365–368.
- [17] O. Nashiru, D.L. Zechel, D. Stoll, T. Mohammadzadeh, R.A.J. Warren, S.G. Withers, β -Mannosynthase: synthesis of β -mannosides with a mutant β -mannosidase, *Angew. Chem.* 113 (2001) 431–434.
- [18] D.L. Jakeman, S.G. Withers, On expanding the repertoire of glycosynthases: mutant beta-galactosidases forming beta-(1, 6)-linkages, *Can. J. Chem.* 80 (2002) 866–870.
- [19] J. Drone, H. Feng, C. Tellier, L. Hoffmann, V. Tran, C. Rabiller, M. Dion, Thermophilus glycosynthases for the efficient synthesis of galactosyl and glucosyl β -(1,3)-glycosides, *Eur. J. Org. Chem.* 2005 (2005) 1977–1983.
- [20] M. Fajjes, M. Saura-Valls, X. Pérez, M. Conti, A. Planas, Acceptor-dependent regioselectivity of glycosynthase reactions by *Streptomyces* E383A β -glucosidase, *Carbohydr. Res.* 341 (2006) 2055–2065.
- [21] J. Mülleger, H.M. Chen, W.Y. Chan, S.P. Reid, M. Jahn, R.A.J. Warren, H.M. Salleh, S.G. Withers, Thermostable glycosynthases and thioglycosylases derived from *Thermotoga maritima* β -glucuronidase, *ChemBioChem* 7 (2006) 1028–1030.
- [22] R. Opassiri, J.R.K. Cairns, T. Akiyama, O. Wara-Aswapati, J. Svasti, A. Esen, Characterization of a rice beta-glucosidase highly expressed in flower and germinating shoot, *Plant Sci.* 165 (2003) 627–638.
- [23] R. Opassiri, Y.L. Hua, O. Wara-Aswapati, T. Akiyama, J. Svasti, A. Esen, J.R.K. Cairns, beta-glucosidase, exo-beta-glucanase and pyridoxine transglucosylase activities of rice BGLu1, *Biochem. J.* 379 (2004) 125–131.
- [24] W. Chuenchor, S. Pengthaisong, R.C. Robinson, J. Yuvaniyama, J. Svasti, J.R.K. Cairns, The structural basis of oligosaccharide binding by rice BGLu1 beta-glucosidase, *J. Struct. Biol.* 173 (2011) 169–179.
- [25] J.H. Wang, Q.Q. Hou, L.H. Dong, Y.J. Liu, C.B. Liu, QM/MM studies on the glycosylation mechanism of rice BGLu1 beta-glucosidase, *J. Mol. Graph. Model.* 30 (2011) 148–152.
- [26] J.H. Wang, Q.Q. Hou, X. Sheng, J. Gao, Y.J. Liu, C.B. Liu, Theoretical study on the deglycosylation mechanism of rice BGLu1 β -glucosidase, *Int. J. Quantum Chem.* (2012), <http://dx.doi.org/10.1002/qua.24131>.
- [27] W. Chuenchor, S. Pengthaisong, R.C. Robinson, J. Yuvaniyama, W. Oonant, D.R. Bevan, A. Esen, C.J. Chen, R. Opassiri, J. Svasti, J.R.K. Cairns, Structural insights into rice BGLu1 beta-glucosidase oligosaccharide hydrolysis and transglycosylation, *J. Mol. Biol.* 377 (2008) 1200–1215.
- [28] G. Hommalai, S.G. Withers, W. Chuenchor, J.R.K. Cairns, J. Svasti, Enzymatic synthesis of cello-oligosaccharides by rice BGLu1 β -glucosidase glycosynthase mutants, *Glycobiology* 17 (2007) 744–753.
- [29] S. Pengthaisong, S.G. Withers, B. Kuaprasert, J. Svasti, J.R. Ketudat Cairns, The role of the oligosaccharide binding cleft of rice BGLu1 in hydrolysis of cellobiosaccharides and in their synthesis by rice BGLu1 glycosynthase, *Protein Sci.* 21 (2012) 362–372.
- [30] L. Melander, W.H. Saunders, Reaction rates of isotopic molecules, Robert E. Krieger Publishing Company, Malabar, FL, 1987. (Chapter 2).
- [31] Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol.* 276 (1997) 307–326.
- [32] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, *J. Comput. Chem.* 19 (1998) 1639–1662.
- [33] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery Jr., T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, H.P. Hratchian, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui,

- A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez, J.A. Pople, Gaussian 03, Revision C. 02, Gaussian, Inc., Wallingford CT, 2004.
- [34] J. Gasteiger, M. Marsili, Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges, *Tetrahedron* 36 (1980) 3219–3228.
- [35] H. Li, A.D. Robertson, J.H. Jensen, Very fast empirical prediction and rationalization of protein pKa values, *Proteins* 61 (2005) 704–721.
- [36] D.C. Bas, D.M. Rogers, J.H. Jensen, Very fast prediction and rationalization of pKa values for protein–ligand complexes, *Proteins* 73 (2008) 765–783.
- [37] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, CHARMM: a program for macromolecular energy, minimization, and dynamics calculations, *J. Comput. Chem.* 4 (1983) 187–217.
- [38] A.D. MacKerell Jr., D. Bashford, M. Bellot, R.L. Dunbrack, J.D. Evanseck Jr, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B* 102 (1998) 3586–3616.
- [39] R. Ahlrichs, M. Bar, M. Haser, H. Horn, C. Kolmel, Electronic structure calculations on workstation computers: the program system TURBOMOLE, *Chem. Phys. Lett.* 162 (1989) 165–169.
- [40] W. Smith, T.R. Forester, DL_POLY_2.0: a general-purpose parallel molecular dynamics simulation package, *J. Mol. Graph. Model.* 14 (1996) 136–141.
- [41] D. Bakowies, W. Thiel, Hybrid models for combined quantum mechanical and molecular mechanical approaches, *J. Phys. Chem.* 100 (1996) 10580–10594.
- [42] A.H. de Vries, P. Sherwood, S.J. Collins, A.M. Rigby, M. Rigutto, G.J. Kramer, Zeolite structure and reactivity by combined quantum-chemical–classical calculations, *J. Phys. Chem. B* 103 (1999) 6133–6141.
- [43] P. Sherwood, A.H. de Vries, M.F. Guest, G. Schreckenbach, C.R.A. Catlow, S.A. French, A.A. Sokol, S.T. Bromley, W. Thiel, A.J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S.C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjovoll, A. Fahmi, A. Schafer, C. Lennartz, QUASI: a general purpose implementation of the QM/MM approach and its application to problems in catalysis, *J. Mol. Struct. (THEOCHEM)* 632 (2003) 1–28.
- [44] S.R. Billeter, A.J. Turner, W. Thiel, Linear scaling geometry optimisation and transition state search in hybrid delocalised internal coordinates, *Phys. Chem. Chem. Phys.* 2 (2000) 2177–2186.
- [45] J. Nocedal, Updating quasi-Newton matrices with limited storage, *Math. Comput.* 35 (1980) 773–782.
- [46] D.C. Liu, J. Nocedal, On the limited memory BFGS method for large scale optimization, *Math. Program.* 45 (1989) 503–528.
- [47] A. Banerjee, N. Adams, J. Simons, R. Shepard, Search for stationary points on surfaces, *J. Phys. Chem.* 89 (1985) 52–57.
- [48] J. Baker, An algorithm for the location of transition states, *J. Comput. Chem.* 7 (1986) 385–395.
- [49] C. Mayer, D.L. Zechel, S.P. Reid, R.A.J. Warren, S.G. Withers, The E358S mutant of *Agrobacterium* sp. β -glucosidase is a greatly improved glycosynthase, *FEBS Lett.* 466 (2000) 40–44.
- [50] V. Ducros, C.A. Tarling, D.L. Zechel, A.M. Brzozowski, T.P. Frandsen, I. von Ossowski, M. Schülein, S.G. Withers, G.J. Davies, Anatomy of glycosynthesis: structure and kinetics of the *Humicola insolens* Cel7B E197A and E197S glycosynthase mutants, *Chem. Biol.* 10 (2003) 619–628.
- [51] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr.* 26 (1993) 283–291.
- [52] G.N. Murshudov, A.A. Vagin, A. Lebedev, K.S. Wilson, E.J. Dodson, Efficient anisotropic refinement of macromolecular structures using FFT, *Acta Crystallogr. D* 55 (1999) 247–255.
- [53] A. Ben-David, G. Shoham, Y. Shoham, A universal screening assay for glycosynthases: directed evolution of glycosynthase XynB2(E335G) suggests a general path to enhance activity, *Chem. Biol.* 15 (2008) 546–551.
- [54] R.B. Wu, H.J. Xie, Z.X. Cao, Y.R. Mo, Combined quantum mechanics/molecular mechanics study on the reversible isomerization of glucose and fructose catalyzed by *pyrococcus furiosus* phosphoglucose isomerase, *J. Am. Chem. Soc.* 130 (2008) 7022–7031.
- [55] Y. Wang, D. Kumar, C.L. Yang, K.L. Han, S. Shaik, Theoretical study of N-demethylation of substituted *N,N*-dimethylanilines by cytochrome P450: the mechanistic significance of kinetic isotope effect profiles, *J. Phys. Chem. B* 111 (2007) 7700–7710.



Contents lists available at SciVerse ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbiEnzymatic and structural characterization of hydrolysis of gibberellin A4 glucosyl ester by a rice β -D-glucosidaseYanling Hua^{a,b}, Sompong Sansenya^a, Chiraporn Saetang^a, Shinji Wakuta^{a,1}, James R. Ketudat Cairns^{a,*}^a School of Biochemistry and Chemistry, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand^b The Center for Scientific and Technological Equipment, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand

ARTICLE INFO

Article history:

Received 15 May 2013

and in revised form 11 June 2013

Available online 26 June 2013

Keywords:

 β -Glucosidase

Gibberellin glucosyl ester

Covalent intermediate

X-ray crystal structure

Transglycosylation

ABSTRACT

In order to identify a rice gibberellin ester β -D-glucosidase, gibberellin A4 β -D-glucosyl ester (GA₄-GE) was synthesized and used to screen rice β -glucosidases. Os3BGlu6 was found to have the highest hydrolysis activity to GA₄-GE among five recombinantly expressed rice glycoside hydrolase family GH1 enzymes from different phylogenetic clusters. The kinetic parameters of Os3BGlu6 and its mutants E178Q, E178A, E394D, E394Q and M251N for hydrolysis of *p*-nitrophenyl β -D-glucopyranoside (pNPGlc) and GA₄-GE confirmed the roles of the catalytic acid/base and nucleophile for hydrolysis of both substrates and suggested M251 contributes to binding hydrophobic aglycones. The activities of the Os3BGlu6 E178Q and E178A acid/base mutants were rescued by azide, which they transglucosylate to produce β -D-glucopyranosyl azide, in a pH-dependent manner, while acetate also rescued Os3BGlu6 E178A at low pH. High concentrations of sodium azide (200–400 mM) inhibited Os3BGlu6 E178Q but not Os3BGlu6 E178A. The structures of Os3BGlu6 E178Q crystallized with either GA₄-GE or pNPGlc had a native α -D-glucosyl moiety covalently linked to the catalytic nucleophile, E394, which showed the hydrogen bonding to the 2-hydroxyl in the covalent intermediate. These data suggest that a GH1 β -glucosidase uses the same retaining catalytic mechanism to hydrolyze 1-O-acyl glucose ester and glucoside.

© 2013 Elsevier Inc. All rights reserved.

Introduction

Beta-glucosidases (β -D-glucopyranoside glucohydrolases, E.C. 3.2.1.21) are enzymes that hydrolyze the β -O-glycosidic bond at the anomeric carbon of a nonreducing terminal D-glucosyl moiety to release D-glucose from glycosides and oligosaccharides. These enzymes are found in essentially 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 [8] by releasing glucose blocking groups from their inactive glucosides in plants. To achieve specificity for these various functions, different β -glucosidases must bind to distinct aglycones with a wide variety of structures, in addition to the glucose of the substrate.

β -Glucosidases have been classified into glycoside hydrolase (GH)² families GH1, GH3, GH5, GH9, GH30, and GH116, based on their amino acid sequences and structural similarity, while other β -glucosidases remain to be classified [9–12]. Most characterized plant β -glucosidases belong to GH1, which falls in GH Clan A, as do GH5 and GH30. GH Clan A enzymes have (β/α)₈-barrel structures with catalytic acid/base and nucleophile on the ends of strands 4 and 7 of the barrel, respectively. The mechanism by which GH1 enzymes recognize and hydrolyze substrates with different specificities remains an area of intense study.

Among β -glucosidase functions in plants, the function of phytohormone activation has attracted many peoples' interest. Phytohormone glucosyl conjugates are ubiquitous in plants, including those of gibberellic acids (GAs), abscisic acid (ABA), cytokinins,

* Corresponding author. Address: Institute of Science, Suranaree University of Technology, 111 University Avenue, Nakhon Ratchasima 30000, Thailand. Fax: +66 44 224185.

E-mail address: cairns@sut.ac.th (J.R. Ketudat Cairns).

¹ Present address: Research Faculty of Agriculture, Hokkaido University, Kita-ku, Sapporo 062-8555, Japan.

² Abbreviations used: A₄₀₅, absorbance at 405 nm; ABA, abscisic acid; ABA-GE, ABA β -D-glucosyl ester; API-ES, atmospheric pressure ionization-electrospray; DAD, diode array detector; DMSO, dimethyl sulfoxide; DNP2FG, 2,4-dinitrophenyl-2-deoxy-2-fluoro- β -D-glucopyranoside; EtOAc, ethyl acetate; GA, gibberellic acid(s); GA₄-GE, gibberellin A4 β -D-glucosyl ester; G2F, 2-deoxy-2-fluoroglucoside; GH, glycoside hydrolase; IAA, indole acetic acid; IAA-GE, IAA β -D-glucosyl ester; IMAC, immobilized metal affinity chromatography; MeOH, Methanol; MES, 2-(N-morpholino)ethanesulfonic acid; NaOAc, sodium acetate; PEG5000MME, polyethylene glycol 5000 monomethyl ester; pNPG, *p*-nitrophenol; pNPGlc, *p*-nitrophenyl β -D-glucopyranoside; TEV, tobacco etch virus (protease).

auxins, and jasmonic acid derivatives, and β -glucosidases that hydrolyze these conjugates have been found in plants [4,13–20].

ABA and the auxin indole-3-acetic acid (IAA) have both been found as inactive 1-*O*-acyl glucose esters in plants. ABA is critical for plant growth, development, and adaptation to various stress conditions. Plants have to adjust ABA levels constantly to respond to changing physiological and environmental conditions [14]. Abscissic acid glucose ester (ABA-GE) is a biologically inactive form that constitutes a reserve form of ABA. Two *Arabidopsis* β -glucosidase (AtBG1 and AtBG2) were found to hydrolyze ABA-GE to form free ABA and to be essential to proper response to drought stress and delay in seed germination, despite the fact that other ABA-GE hydrolyzing enzymes were detected [14,17]. IAA glucosyl ester (IAA-GE) is also an inactive, stored form that can also be hydrolyzed by IAA-glucose hydrolase to release active IAA [15].

Both 1-*O*-acyl glucosyl esters and glucosides of GAs are found in plants. GAs promote germination, shoot elongation, and flower development, among their many functions [21]. It has been suggested that GA glucosyl esters are deactivated GAs that can be enzymatically reconverted to active GAs, thus serving as a reserve form of biologically active GAs [22]. After [^{13}C]GA₂₀- β -D-glucosyl ester was injected into light-grown maize seedlings, the metabolites, [^{13}C]GA₂₀, [^{13}C]GA₂₉, [^{13}C]GA₂₀-13-*O*-glucoside, [^{13}C]GA₂₉-2-*O*-glucoside, [^{13}C]GA₈ and [^{13}C]GA₈-2-*O*-glucoside were identified in the extracts of the seedlings made 24 h after the injection [23]. This showed that the endogenous hydrolysis of the introduced conjugate and its reconjugation led to the three new glucosides. In rice, [^3H]GA₁, [^3H]GA₂, [^3H]GA₃₄, the glucosides of [^3H]GA₂, [^3H]GA₄, [^3H]GA₈ and [^3H]GA₃₄, and the glucosyl ester of [^3H]GA₄ (GA₄-GE) have been found after application of [^3H]GA₄ to cell suspension cultures of *Oryza sativa* cv. nipponbare [24].

β -Glucosidases have been proposed to hydrolyze the GA conjugates to active GAs, but the molecular identification of these enzymes and investigation of their modes of action have yet to be reported. Schliemann [13] reported that β -glucosidases extracted from mature rice seeds and seedlings have different hydrolytic activities toward GA₈-2-*O*-glucoside, GA₃-3-*O*-glucoside and 1-*O*-GA₃-glucosyl ester, but he did not purify and characterize the β -glucosidases. Furthermore, the substrate specificity of β -glucosidase to GA conjugates and how β -glucosidase binds to GA conjugates has yet to be reported. In this study, we identified a rice GA₄-GE β -D-glucosidase and characterized its mechanism of glucosyl ester hydrolysis by comparing hydrolysis and transglycosylation activities of wild type and acid/base mutant enzymes with GA₄-GE. To better understand the ester-enzyme interaction, structures of mutant enzymes soaked with GA₄-GE and pNPGlc were also determined.

Materials and methods

Synthesis of GA₄-GE

GA₄-GE was synthesized from GA₄ (Jiangsu Fengyuan Bioengineering Co.Ltd, P.R. China) following the method of Hiraga, et al. [25]. The acetylated and deacetylated GA₄-GE were obtained with 43.7% and 40.5% yields, respectively. The synthesized acetylated and deacetylated GA₄-GE structures were confirmed by NMR spectra on a 300 MHz NMR spectrometer with a Varian 300 ID/PFG probe at a frequency of 299.986 MHz (Unity INOVA, Varian, USA). Deuterated chloroform (CDCl₃) and dimethyl sulfoxide-*d*₆ (DMSO-*d*₆) were used as solvents for acetylated and deacetylated GA₄-GE, respectively. The ^1H NMR was consistent with the published data for GA₄-GE [25].

The identity of the deacetylated GA₄-GE was also confirmed from its mass spectrum (data not shown). In the positive mode,

we detected [M+Na]⁺ at *m/z* 517.1, [M+H-H₂O]⁺ at *m/z* 477.2, and [M+H-Glc]⁺ at *m/z* 315.5.

Expression of rice GH1 enzymes to screen for GA₄-GE hydrolysis

Five GH1 enzymes that have been expressed in our laboratory, Os3BGlu6 [26], Os3BGlu7 [27], Os4BGlu12 [28], Os4BGlu18 (S. Baiya et al., unpublished) and Os9BGlu31 [29] were expressed as N-terminal thioredoxin fusion proteins in *Escherichia coli* and purified by immobilized metal affinity chromatography, as previously described. The purified fusion proteins were tested for the hydrolysis activity to pNPGlc and GA₄-GE according to the method described below.

Site-directed mutagenesis of Os3BGlu6

Mutagenesis of the pET32/Os3BGlu6 expression vector [26] to create the Os3BGlu6E178A, Os3BGlu6E178Q, Os3BGlu6E394D and Os3BGlu6E394Q mutations was performed with the QuikChange[®] Site-Directed Mutagenesis Kit (Stratagene, Agilent Corp., La Jolla, CA, USA) according to the supplier's instructions. The following oligonucleotides were used for mutagenesis: for E178A, 5'-GATCACGCTCAACGCGCCGCACACGGTG3' and its reverse complement; for E178Q, 5'-GGATCAGCTCAACCAACCGCACACGGTGGC3' and its reverse complement; for E394D, 5'-CCAGTGTACATCACTGATAACGGGATGGATGACAGC3' and its reverse complement; and for E394Q, 5'-CCACCAGTGTACATCACTCAGAACGGGATGGATGACAGC3' and its reverse complement. The cDNA were confirmed to include the desired mutations and be free of additional mutations by automated DNA sequencing (Macrogen Corp., Rep. of Korea).

Recombinant expression and purification of the mutants of Os3BGlu6

The wild type rice Os3BGlu6 and its mutants M251N, E178Q, E178A, E394D and E394Q were expressed in *E. coli* strain Origami (DE3) as fusion proteins with N-terminal thioredoxin and His₆ tags as described previously for Os3BGlu6 [26]. The crude proteins were first purified by immobilized metal (Co²⁺) affinity chromatography (IMAC). The N-terminal thioredoxin, His₆ and S tags were then excised with TEV protease, and removed with a second IMAC column purification. Protein concentrations were estimated by the Bradford protein assay (Bio-Rad) with bovine serum albumin as the standard. The values from this assay were in the good agreement with those obtained from the 280 nm absorbance with the calculated extinction coefficient.

Determination of pH optimum for Os3BGlu6 and its mutants

The optimum pH of Os3BGlu6 and Os3BGlu6 M251N hydrolysis of pNPGlc were determined by incubating 1 μg of enzyme with 2 mM pNPGlc in 80 μl of 100 mM universal buffer (citric acid-disodium hydrogen phosphate), pH 2.0–11.0 in 0.5-pH-unit increments, at 30 °C for 10 min. The reactions were stopped by adding 100 μl of 2 M sodium carbonate. The released *p*-nitrophenol (pNP) was quantified by measuring the absorbance at 405 nm (A₄₀₅) with a microplate reader (Thermo Labsystems, Finland) and comparing it to that of a pNP standard curve.

The optimum pH of Os3BGlu6 and its mutants M251N, E178Q and E178A for hydrolysis of GA₄-GE were determined by incubating 1 μg of Os3BGlu6 or Os3BGlu6 M251N, or 5.0 μg of E178Q or E178A with 0.86 mM GA₄-GE in 80 μl of 100 mM universal buffer, pH 2.0–11.0 in 0.5-pH-unit increments, at 30 °C for 20 min. The reactions were stopped by boiling 1 min and cooled on ice immediately. The released glucose was quantified with a glucose oxidase assay [27].

Measurement of wild type and mutant Os3BGluc activities for hydrolysis of GA₄-GE and pNPGlc

The activities of Os3BGluc wild type and mutants to hydrolyze pNPGlc were measured as described for the pH optimum determination, but in 50 mM sodium acetate (NaOAc) buffer, pH 5.0. To measure hydrolysis activities toward GA₄-GE, Os3BGluc wild type and M251N mutant enzymes were incubated with GA₄-GE in 50 mM NaOAc, pH 5.0, but 50 mM 2-(N-morpholino)ethanesulfonic acid (MES) buffer was selected for the E178Q and E178A mutants to avoid transglucosylation. For determination of kinetic parameters, variable reaction times, enzyme amounts and substrate concentrations were tested to obtain the initial velocities. The K_m and k_{cat} were calculated from nonlinear regression of Michaelis–Menten plots with Graft 5.0 software (Erithacus Software, Horley, UK).

Identification of transglucosylation product with TLC, LC-MS and NMR

Transglucosylation reactions were studied for Os3BGluc wild type, M251N, E178Q and E178A mutants. 2 mM pNPGlc, 5 mM GA₄-GE or 2,4-dinitrophenyl-2-deoxy-2-fluoro- β -D-glucopyranoside (DNP2FG) donor was reacted with 1.0 μ g of enzymes in 50 mM MES buffer, pH 5; and 40 mM sodium azide or 0.96 mM free GA₄ was used as acceptor. At different times, 10 μ l aliquots of each reaction mix were removed, boiled 1 min, and kept on ice for TLC analysis. Reactions without enzymes were used as controls. Sample aliquots were spotted on TLC plates, which were developed with EtOAc–MeOH–H₂O (7.5:2.5:1.0) or CHCl₃–MeOH (7:3). The products were detected by staining with 10% sulfuric acid in ethanol followed by charring.

The new transglucosylation products of E178Q and E178A were analyzed by LC-MS. Samples were separated over a ZORBAX Eclipse XDB-C18 column (4.6 $\times$ 150 mm, 5 μ m, Agilent, USA) on an Agilent 1100 HPLC. A gradient of 0–80% MeOH in 0.05% formic acid was run over 20 min at a flow rate of 0.8 ml/min. The ion peaks and mass spectra were detected with an Agilent single quadrupole MSD mass spectrometer with the atmospheric pressure ionization-electro spray (API-ES) source in negative and positive ion modes. The scan range was 100–1000 m/z , and the fragmentor voltage was 70 V. The VCap was 3000 V for both positive and negative modes.

The transglucosylation product was collected from HPLC separations and dried with a vacuum centrifuge. The product was dissolved in acetone- d_6 , and ¹H NMR and gCOSY spectra were collected with a 300 MHz NMR spectrometer (Unity INOVA, Varian, USA).

Transglucosylation kinetics of the mutants of Os3BGluc

The pH optima of E178Q and E178A for transglucosylation were determined in 50 mM MES, pH 5.0–7.5, and 50 mM sodium acetate, pH 4.0–6.0 in 0.5-pH-unit increments. Two micrograms of E178Q or E178A were incubated with 2 mM GA₄-GE and 100 mM sodium azide, in the different pH buffers at 30 °C for 30 min. The reactions without sodium azide were used to measure hydrolysis activities. The reactions were stopped by boiling 1 min and cooled on ice immediately. The reaction mixes were centrifuged at 6000 rpm for 15 min and separated on LC-MS, as described in the preceding section. The GA₄ released was detected with a diode array detector (DAD) at 210 nm, and the amounts were calculated by comparison of the peak areas to a GA₄ standard curve.

The concentrations of donor GA₄-GE and acceptor sodium azide were varied to test their effects on transglucosylation kinetics. The concentration of sodium azide was varied from 0 to 400 mM, while the GA₄-GE donor was fixed at 2 mM. The reactions were per-

formed in 50 mM MES, pH 5, at 30 °C for 20 min. The turnover rates of GA₄ release per minute per unit enzyme (V_0/E_0) were calculated based on the GA₄ peak area. For the reactions varying the concentrations of GA₄-GE donor, sodium azide was fixed at 50, 100, 200 and 400 mM, and compared to reactions without sodium azide.

Crystallization of Os3BGluc E178Q

Purified Os3BGluc E178Q and Os3BGluc E394D were concentrated to 5 mg/ml and crystallized by hanging drop vapor diffusion with a precipitant of 13% polyethylene glycol 5000 monomethyl ester (PEG5000MME), 0.1 M Bis/Tris, pH 6.5, for Os3BGluc E178Q with pNPGlc and Os3BGluc E178Q with GA₄-GE, and 17% PEG5000MME, 0.1 M Bis/Tris, pH 6.5 for Os3BGluc E394D with GA₄-GE. The crystals were approximately 0.5 $\times$ 0.1 $\times$ 0.1 mm in size after growing 2 days. For cryo-protection, the crystals were transferred to a solution consisting of 22% PEG5000MME, 0.2 M Bis/Tris, pH 6.5, 18% (v/v) glycerol, which contained 10 mM GA₄-GE or pNPGlc.

X-ray diffraction data collection and processing

The data sets were collected at 100 K in a nitrogen stream (Cryojet; Oxford Instrument) on the PX 13B beamline at the National Synchrotron Radiation Research Center, Hsinchu, Taiwan, with a 1.0 Å wavelength X-ray beam and an ADSC Quantum 315 CCD detector. All the diffraction images were indexed, integrated and scaled with the HKL-2000 program [30]. The crystals of Os3BGluc E178Q soaked with GA₄-GE and pNPGlc belonged to the orthorhombic space group P2₁2₁2₁ and diffracted to 1.97 and 1.90 Å resolutions, respectively. Unit cell parameters for Os3BGluc E178Q with GA₄-GE were $a = 57.1$, $b = 91.2$, $c = 111.4$ Å, $\alpha = \beta = \gamma = 90^\circ$, while those with pNPGlc were $a = 57.0$, $b = 91.1$, $c = 111.2$ Å, $\alpha = \beta = \gamma = 90^\circ$ and for Os3BGluc E394D_GA₄-GE were $a = 56.5$, $b = 91.2$, $c = 111.2$ Å, $\alpha = \beta = \gamma = 90^\circ$. Data-collection and processing statistics for the Os3BGluc E178Q complex structures are presented in Table 1.

Structure refinement

The structures of Os3BGluc E178Q with GA₄-GE or pNPGlc and Os3BGluc E394D with GA₄-GE were solved via rigid body refinement with the wild type Os3BGluc structure (PDB: 3GNO) as a template model. A subset of 5% of the structure factor amplitudes was reserved for R_{free} determination. Manual rebuilding of the model was performed in Coot [31] and refinement with REFMAC5 [32]. The occupancy of the glucose and glycerol in the active site of Os3BGluc E178Q with GA₄-GE was refined by setting the occupancy for the two moieties at values that added to 1.00 or less and minimizing the positive and negative densities around the ligands in the weighted Fo–Fc map, R_{free} and B-values. The glucose bound conformation of E394 (conformation A) was set at the same occupancy as the glucose, while the sum of the occupancies of the two E394 conformations was held at 1.00. The overall quality of each model was evaluated with the program PROCHECK [33]. Additional structure determination and refinement statistics are presented in Table 1. All the structure figures were drawn with PyMOL (Schrödinger, USA).

Results

GA₄-GE hydrolysis by recombinantly expressed rice GH1 enzymes

GA₄-GE, which has previously been reported in rice [24], was synthesized and used to screen GH1 enzymes for GA₄-GE hydroly-

Table 1
Data-collection and structure-refinement statistics.

Dataset	Os3BGlu6 E178Q pNPGlc	Os3BGlu6 E178Q GA ₄ -GE
PDB code	3WBA	3WBE
Beamline	BL13B1	BL13B1
Wavelength (Å)	1.00	1.00
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Unit-cell parameters (Å)	<i>a</i> = 57.0 <i>b</i> = 91.1 <i>c</i> = 111.2	<i>a</i> = 57.1 <i>b</i> = 91.2 <i>c</i> = 111.4
Resolution range (Å)	30–1.90	30–1.97
Resolution outer shell(Å)	1.97–1.90	2.04–1.97
No. Unique reflections	46370	41777
No. Observed reflections	297256	297549
Completeness (%)	99.3 (100)	99.9 (100)
Average redundancy per shell	6.5 (6.3)	7.1 (7.2)
I/σ (I)	28.9 (4.5)	23.3 (6.6)
R _(merge) (%)	6.0 (41.0)	8.4 (27.6)
R _{factor} (%)	14.8	14.5
R _{free} (%)	17.9	18.0
No. of residues in protein	478	478
No. Protein atoms	3917	3952
No. Ligand atoms	11 (Glc)	11 (Glc)
No. Other hetero atoms	36 (GOL)	36 (GOL)
No. waters	534	491
Mean B-factor	18.9	22.6
Protein	16.6	20.6
Ligand	12.1	11.1
Other hetero atoms	35.3 (GOL)	44.2 (GOL)
Waters	34.3	36.8
r.m.s. bond deviations (length)	0.013	0.013
r.m.s. angle deviations (degrees)	1.290	1.264
Ramachandran plot		
Residues in most favorable regions (%)	88.9	88.6
Residues in additional allowed regions (%)	10.7	10.9
Residues in generously allowed regions (%)	0.5	0.5
Residues in disallowed regions (%)	0.0	0.0

sis activity. Five rice GH1 enzymes that have been expressed in *E. coli* were tested for the hydrolysis of pNPGlc and GA₄-GE. As shown in Table 2, Os3BGlu6 was found to have the highest hydrolysis activity to GA₄-GE among these enzymes. Although Os9B-Glu31 had a higher ratio of activity toward GA₄-GE compared to pNPGlc (0.27 vs. 0.07 for Os3BGlu6), it is primarily a transglycosidase [29] and has low activity toward both substrates.

Hydrolysis activities of Os3BGlu6 and its mutants

Mutation of the putative catalytic acid/base and nucleophile residues and aglycone-binding residue M251 had differential effects on pNPGlc and GA₄-GE hydrolysis. The mutation of the Os3B-Glu6 glutamate residue 178 to alanine (E178A) and glutamine (E178Q) and glutamate residue 394 to aspartate (E394D) and glutamine (E394Q) reduced the relative hydrolytic activity toward pNPGlc to <1%, while the M251N mutant reduced this activity to 52% (Table 3). The relative activity of M251N was 88.9% for GA₄-GE, and the catalytic efficiency (k_{cat}/K_m) followed the same pattern with decreases from 6.2 to 2.6 mM⁻¹ s⁻¹ for pNPGlc, and from 0.13 to 0.08 mM⁻¹ s⁻¹ for GA₄-GE compared to its wild type. For GA₄-GE, E178A had 22.2% and E178Q 12.5% of the relative hydrolytic activity of wild type Os3BGlu6 with both K_m and k_{cat} values lower than wild type. In contrast, the E394Q and E394D mutants had <2% wild type relative hydrolytic activity with GA₄-GE, similar to that with pNPGlc.

The Os3BGlu6 wild type and its M251N mutant were found to have high pNPGlc hydrolysis activity between pH 4.0 to 5.0, with

the highest value at pH 4.5, while the activity dropped quickly above pH 5.5 (Supplementary data Fig. S1A). With GA₄-GE as substrate, Os3BGlu6 wild type and M251N, E178Q and E178A mutants all had highest hydrolytic activities at pH 4.5 (Supplementary data Fig. S1B). The activities of Os3BGlu6 E178Q and E178A dropped somewhat more slowly at pH above 6.0 compared to Os3BGlu6 wild type and M251N when hydrolyzing GA₄-GE, with 50% maximal activities at approximately 6.5 and 6.0, respectively.

Identification of transglucosylation product with azide

Previously, we found that Os3BGlu6 had negligible transglycosylation activity in reactions with alcohols, such as *n*-octyl alcohol, despite its hydrolytic activity toward the corresponding glycosides (S. Seshadri and J.R. Ketudat Cairns, unpublished). When we tested Os3BGlu6 for the ability to transfer glucose to azide and GA₄ acceptors in MES buffer, it hydrolyzed pNPGlc to pNP and glucose, and no transglucosylation products were observed. In contrast, the acid/base mutants Os3BGlu6 E178Q and E178A produced transglucosylation products with sodium azide and no significant hydrolysis products were detected in overnight reactions (Supplementary data Fig. S2). However, when the E178Q mutant was used in an attempt to generate GA₄-GE by transglycosylation of GA₄ acceptor with pNPGlc donor, no transglycosylation product was detected (data not shown). When GA₄-GE was used as donor and sodium azide as acceptor for Os3BGlu6 E178Q and E178A, a prominent transglucosylation product and a smaller amount of glucose were observed (Supplementary data Fig. S3).

The transglucosylation product of E178A with GA₄-GE was identified with LC-MS (Fig. 1). The peak of newly formed product at retention time 3.69 min had the molecular mass consistent with the 205.17 a.m.u mass of β-D-glucopyranosyl azide (1-azido-β-D-glucoside) expected from the retaining mechanism of the β-glucosidase [34]. Since formic acid (HCO₂H) was used in the mobile phase of LC-MS, its adduct ion [M+HCO₂]⁻ was detected as the base peak along with [M-H]⁻ and [M+Cl]⁻ peaks (Supplementary data Fig. S4). The structure of the β-D-glucopyranosyl azide was also confirmed from its ¹H and gCOSY NMR spectra (data not shown). The ¹H spectrum peaks were assigned as: δ4.53, H1, d, $J_{2,1}$ = 9.0 Hz; δ3.85, H6, dd, $J_{6,6'}$ = 11.1 Hz; δ3.678, H6', dd, $J_{5,6'}$ = 4.8 Hz, $J_{6,6'}$ = 11.1 Hz; δ3.364–3.458, H3, H4, H5, m; δ3.192, H2, t, $J_{1,2}$ = $J_{3,2}$ = 9.0 Hz, which matched the published ¹H NMR data for β-D-glucopyranosyl azide [35]. The coupling constant between H1 and H2 of 9.0 Hz confirmed the β-D-glucopyranosyl azide had the “β” configuration.

Transglucosylation kinetics of the mutants of Os3BGlu6

The activity versus pH profiles for transglucosylation of azide with GA₄-GE donor were determined in sodium acetate and MES buffers (Supplementary data Fig. S5). For the Os3BGlu6 E178Q mutant, its optimum pH range was 5.0 to 6.0 in both buffers, slightly higher than its optimum pH for hydrolysis of GA₄-GE. The E178A mutant also had highest transglucosylation activity at pH 5.0 in the MES buffer, but was high at pH 4.0 in the sodium acetate. In the sodium acetate buffer, the release of pNP from pNPGlc by E178A was much higher than that in MES buffer, even without azide in the system.

Fig. 2 shows that when the concentration of donor GA₄-GE was fixed at 2 mM, the turnover rates, V_0/E_0 , of GA₄ for Os3BGlu6 E178A slowly increased with increasing concentrations of sodium azide and reached its maximum at 400 mM sodium azide. However, for Os3BGlu6 E178Q, the turnover rates of GA₄ decreased at higher azide concentrations after reaching its maximum at 100 mM sodium azide. Since transglucosylation products were the main products and little glucose was evident when the azide acceptor was

Table 2
GA₄-GE hydrolysis by recombinantly expressed rice GH1 enzymes.

Enzyme	Phylogenetic cluster ^a	Activity toward GA ₄ -GE (μmol Glc released/min/mg)	Activity toward pNPGlc (μmol pNP released/min/mg)	Ratio of activity toward GA ₄ -GE/pNPGlc
Os3BGlu6	At/Os1	0.185	2.6	0.07
Os3BGlu7 (BGlu1)	At/Os4	0.02	4.0	0.005
Os4BGlu12	At/Os7	0.035	130	0.003
Os4BGlu18	At/Os5	N.D.	0.94	–
Os9BGlu31	At/Os6	0.02 ^b	0.075 ^b	0.27

N.D. means not detectable.

^a Phylogenetic clusters from the eight sequence-based GH1 phylogenetic clusters shared by rice and *Arabidopsis*, as identified by Opassiri et al. [28].

^b Activity is primarily transglycosylation, rather than hydrolysis [29].

present compared to when it was not, the parameters were calculated from the released GA₄ without subtracting hydrolysis. The Os3BGlu6 E179Q and E179A mutants had similar kinetic parameters for sodium azide in MES buffer, in the presence of 2 mM GA₄-GE with apparent K_m of 3.71 ± 0.47 mM, k_{cat} of 0.48 ± 0.01 s⁻¹ and k_{cat}/K_m of 0.13 mM⁻¹ s⁻¹ for E178Q, and K_m of 3.1 ± 0.5 mM, k_{cat} of 0.204 ± 0.006 s⁻¹, and k_{cat}/K_m of 0.066 mM⁻¹ s⁻¹ for E178A.

The effects of the concentration of sodium azide on the transglycosylation kinetics were further studied by measuring k_{cat} and K_m for GA₄-GE at 50, 100, 200 and 400 mM sodium azide (Table 4). For Os3BGlu6 E178Q, the K_m for GA₄-GE was >18-fold higher and k_{cat} at least 27-fold higher than that without sodium azide at all concentrations tested and both the K_m and k_{cat} increased with azide concentration, while the k_{cat}/K_m of GA₄-GE was highest at 0.51 mM⁻¹ s⁻¹ in 100 mM sodium azide and decreased to 0.41 and 0.29 when the concentration of sodium azide was increased to 200 and 400 mM, respectively. For Os3BGlu6 E178A, the K_m was 3-fold higher and k_{cat} 15-fold higher with 200 mM sodium azide compared to without sodium azide, while the k_{cat}/K_m of GA₄-GE was the same at 200 mM and 400 mM sodium azide.

Structure of Os3BGlu6E178Q soaked with either GA₄-GE or pNPGlc

When the Os3BGlu6 E178Q (acid/base) mutant was soaked with GA₄-GE, electron density for a glucosyl residue covalently bound to the catalytic nucleophile was seen, rather than a GA₄-GE molecule. However, when the glucosyl residue was built into the density, a patch of positive density protruded in the axial position from glucose carbon 4 and another positive density patch corresponded to the position of the E394 sidechain in the apo Os3BGlu6 structure (PDB: 3GNO). When the two sidechain positions and glucosyl moiety were set at 50% occupancy, a positive density defining a glycerol molecule was observed at positions overlapping the glucosyl C2, O2, C3, O3 and C4, but pointing away from O4. Thus, the glucose was present with partial occupancy, with a glycerol molecule from the solvent sharing the same positions as glucose carbons 2, 3 and 4 (PDB: 3WBE; Figs. 3B and 4B). The catalytic nucleophile was found in both the position previously seen for the covalent complex of Os3BGlu6 with 2-deoxy-2-fluoroglucoside (G2F) and that found in the apo enzyme (Figs. 3B, 4B, and 5C) [26]. Another crystal soaked in 2 mM pNPGlc also showed a covalently bound α-D-glucosyl residue, but with high occupancy (PDB: 3WBA; Figs. 3A and 4A). Since the structure appeared the same, but the ligand density was clearer in the pNPGlc structure (Fig. 4A), this structure best represents the covalent complex of the Os3BGlu6 E178Q mutant with an unmodified glucosyl residue. In both structures, the protein structure was well defined from G11 to T488, except for the

loop of residues 331 to 337, which was poorly defined by the electron density, as was the N-terminus (similar to the previously reported wild type structure). Soaking of Os3BGlu6 E394D crystals with GA₄-GE resulted in electron density in both the glycone and aglycone subsites, but the occupancy was too low and the ligand density too poorly defined to build a reliable GA₄-GE complex structure.

In the covalent intermediate structure, the anomeric carbon of glucose is covalently bonded with Oε1 of E394 at a distance of 1.3 Å (Figs. 3A and 4A). The noncovalent interactions between the surrounding amino acids and the hydroxyls of the glucose ring, which took a ⁴C₁ chair conformation in the –1 subsite, included ten hydrogen bonds. The O2 atom of glucose forms hydrogen bonds with H132Nε2 (3.30 Å), N177 Oε1 (3.01 Å) and Q178Nε2 (2.99 Å), while O3 of glucose hydrogen bonds with Q31 Oε1 (2.66 Å) and W452Nε1 (2.85 Å). The glucosyl O4 is hydrogen bonded to Q31Nε2 (2.91 Å) and E451 Oε1 (2.61 Å), and O5 of glucose forms a hydrogen bond with Y321 (2.70 Å), while O6 forms a hydrogen bond with Oε2 of the same residue (2.60 Å). A water molecule (Water52) is also hydrogen bonded to O2 of glucose (2.80 Å) and E394 Oε1 (2.98 Å) (Fig. 4A). A glycerol molecule was observed in the active site of this Os3BGlu6 E178Q structure in two conformations in the +1 subsite.

Discussion

Hydrolysis of GA₄-GE by Os3BGlu6 and its mutants

Thirty-four GH1 genes encoding potentially active rice β-glucosidases have been reported, and these were broken into eight sequence-based phylogenetic clusters that were shared between rice and *Arabidopsis thaliana* [28]. The five rice GH1 enzymes tested each represented a different phylogenetic cluster, from which we have been able to express active enzymes. Four of the tested enzymes had the ability to free GA₄ from GA₄-GE, and could potentially contribute to this release *in planta*. Although it also hydrolyzes β-1,2- and β-1,3-linked gluco-disaccharides and octyl-β-D-glucoside with high activity and may play other roles *in planta* [26], Os3BGlu6 had the highest activity toward GA₄-GE and was chosen to characterize the hydrolysis of the glucosyl ester.

Although much characterization has been done of the kinetics of hydrolysis of alkyl and aryl glycosides and oligosaccharides by β-D-glucosidases and their catalytic acid/base and nucleophile mutants [34–37], little description of the kinetics of glucosyl ester hydrolysis is available in the literature. Kiso et al. [38] compared the activities of 3 bacterial, 2 fungal and almond β-glucosidases for hydrolysis of 1-O-(p-hydroxybenzoyl) β-D-glucose (pHBGlc), and found that *Caldocellum saccharolyticum* β-glucosidase had the highest activity of the commercial enzymes tested with 21% activity for pHBGlc relative to pNPGlc. Almond was reported to have approx. 5% relative activity for pHBGlc, somewhat less than the 7% of Os3BGlu6 for GA₄-GE compared to pNPGlc. In contrast to *Caldocellum saccharolyticum* and almond β-glucosidase, which had similar V_{max} for pNPGlc and pHBA but nearly 10-fold and 15-fold higher K_m for pHBA, respectively [39], Os3BGlu6 had similar K_m values for pNPGlc and GA₄-GE, but a 26-fold lower apparent k_{cat} for GA₄-GE relative to pNPGlc. Although the large aglycone of the GA₄-GE is not strictly comparable to that of pNPGlc, it shows that, in the case of Os3BGlu6, the rate of the hydrolysis and not the binding step is the distinguishing feature of the ester compare to the glucoside.

It might be expected that the relatively low pK_a of the leaving group would make the glucosyl ester relatively insensitive to the presence of a catalytic acid for the initial glycosylation step (the GA₄ pK_a is 4.3, compared to 7.2 for pNPGlc), similar to 2,4-dinitro-

Table 3
Hydrolysis activities of Os3BGlu6 and its mutants.

Enzyme	pNPGlc				GA ₄ -GE			
	Relative activity (%)	<i>K_m</i> (mM)	<i>k_{cat}</i> (s ^{−1})	<i>k_{cat}/K_m</i> (mM ^{−1} s ^{−1})	Relative activity (%)	<i>K_m</i> (mM)	<i>k_{cat}</i> (s ^{−1})	<i>k_{cat}/K_m</i> (mM ^{−1} s ^{−1})
Os3BGlu6	100	6.3 ± 0.4 ^a	38.9 ± 0.9 ^a	6.2 ^a	100	5.8 ± 0.6	0.75 ± 0.04	0.13
Os3BGlu6-M251N	51.6	6.5 ± 0.7	15.4 ± 0.9	2.6	88.9	14.6 ± 2.0	1.2 ± 0.1	0.08
Os3BGlu6-E178Q	0.3	n.m.	n.m.	n.m.	12.5	0.09 ± 0.01	0.03 ± 0.001	0.33
Os3BGlu6-E178A	0.7	n.m.	n.m.	n.m.	22.2	3.7 ± 0.4	0.09 ± 0.005	0.03
Os3BGlu6-E394D	0.9	n.m.	n.m.	n.m.	1.4	n.m.	n.m.	n.m.
Os3BGlu6-E394Q	0.8	n.m.	n.m.	n.m.	0.6	n.m.	n.m.	n.m.

n.m.: not measured.

For the relative activity with pNPGlc, 0.5 μg of Os3BGlu6 wild type or M251 N, or 3 μg of E178A, E178Q, E394D or E394Q was incubating with 4 mM pNPGlc in NaOAc, pH 4.5 (final volume 50 μl) at 30 °C for 10 min. For the relative activity with GA₄-GE, 1.0 μg of Os3BGlu6 wild type or M251 N, or 3 μg of E178A, E178Q, E394D or E394Q was incubated with 1.7 mM GA₄-GE in NaOAc, pH 4.5 (final volume was 50 μl), at 30 °C for 20 min. The relative activities were calculated based on A₄₀₅ and the amount of the enzymes. For determination of the kinetic parameters, 50 mM NaOAc buffer was used for Os3BGlu6 wild type and M251 N, for both substrates, but 50 mM MES buffer was selected for the E178Q and E178A mutants.

^a Data from Seshadri et al. (2009).

phenyl glycosides [34,41], but other influences of having the ester bond in apposition to the glycosidic bond are not so clear. The Os3BGlu6 wild type had an optimum pH for hydrolysis of both pNPGlc and GA₄-GE at pH 4.5, indicating that the low p*K_a* of the leaving group carboxylate did not affect the pH optima. Although the Os3BGlu6 E178Q and E178A mutants had low activities toward pNPGlc, their GA₄-GE hydrolysis activities were significant and dropped slowly above pH 6.0, compared to wild type Os3BGlu6, consistent with their designations as catalytic acid/base mutants. The pH dependence of enzymatic reactions is generally considered to reflect ionizations of acid/base groups involved in catalysis [39,40]. The Os3BGlu6 E178Q and E178A mutants lack the ionizable group to donate a proton in the glycosylation step or extract a proton in the deglycosylation step, resulting in lower pH dependence at high pH.

In contrast to the *Agrobacterium* and rice Os3BGlu7 β-glucosidases [37,41], the acid/base glutamate to alanine (E178A) mutant was more active than the acid/base to glutamine mutant (E178Q). This was evidently due to the buildup of an intermediate on the enzyme, since both the *K_m* and *k_{cat}* dropped dramatically in the E178Q mutant. The X-ray crystallographic structure (Fig. 3) indicated this intermediate was the covalent intermediate formed in the glycosylation half reaction (Fig. 5). The glutamine sidechain can provide a hydrogen bond to stabilize the leaving group in glycosylation and acceptor in deglycosylation of the enzyme, giving it

higher activity [42], but it evidently could not facilitate water as an acceptor in Os3BGlu6. Without E178, the low p*K_a* of GA₄ allows it to be released in the glycosylation step without general acid-assistance more effectively than pNPGlc, for which glycosylation appeared to be the rate-limiting step. The lack of deprotonation of the water molecule in the deglycosylation step makes Os3BGlu6 E178A and Os3BGlu6 E178Q relatively effective at transglycosylation with GA₄-GE donor compared to their wild type.

When the structure of wild type Os3BGlu6 was resolved by X-ray crystallography, the M251 residue was suggested to block binding of straight cellobiosaccharides in the active site cleft [26]. The mutation of M251 to asparagine resulted in 9 to 24-fold increases hydrolysis efficiency for oligosaccharides in Os3BGlu6 M251N compared to wild type [43]. So, it was of interest to see whether this mutation would affect the hydrolysis of less polar substrates like pNPGlc and GA₄-GE as well. The catalytic efficiencies (*k_{cat}/K_m*) for M251N were reduced 2.4-fold for pNPGlc, and 1.6-fold for GA₄-GE compared to wild type. The *K_m* to pNPGlc for both Os3BGlu6 and Os3BGlu6 M251N is almost the same, and the smaller *k_{cat}* of Os3BGlu6 M251N is the main cause for its lower *k_{cat}/K_m* compared to the wild type. In contrast, for GA₄-GE, the higher *K_m* for Os3BGlu6 M251N is the main reason for its lower *k_{cat}/K_m* compared to the wild type, suggesting that the larger aglycone of GA₄-GE may interact with residue M251, which may stabilize binding of large hydrophobic aglycones.

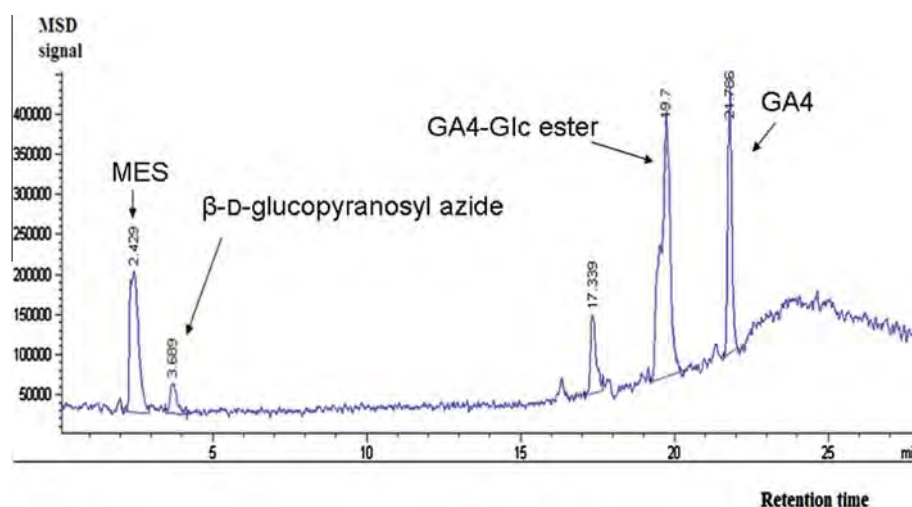


Fig. 1. The LC-MS chromatogram of the transglucosylation reaction of Os3BGlu6 E178A with GA₄-GE as donor and sodium azide as acceptor. The reaction components were separated by C18 reverse phase HPLC, as described in the methods. The detector was an Agilent MSD single quadrupole mass spectrometer with API-ES source in negative ion mode, with a scan range of 50–800 *m/z*.

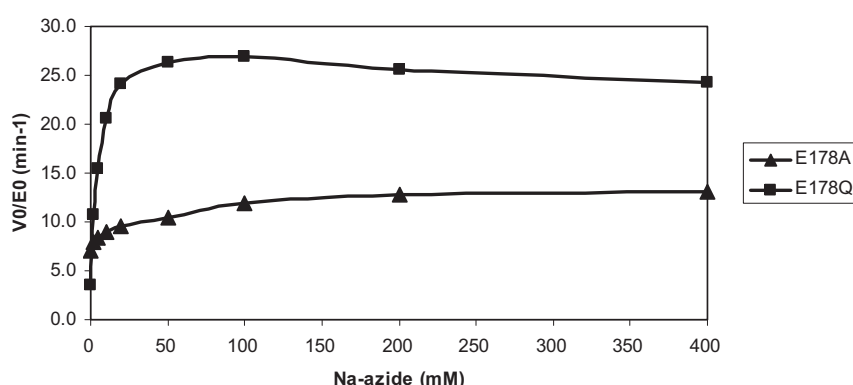


Fig. 2. Rates of transglucosylation of azide acceptor with GA₄-GE donor by the Os3BGlu6 E178A and E178Q acid/base mutants. The concentrations of sodium azide acceptor were varied from 0 to 400 mM, while donor GA₄-GE was fixed at 2 mM. The reactions were performed in 50 mM MES, pH 5, at 30 °C for 20 min. The turnover rates V_0/E_0 of GA₄ (μmole of GA₄ release per minute per μmole of enzyme) were calculated based on the area of GA₄ HPLC peaks. The total area of GA₄ was used without subtracting the hydrolysis reaction.

Table 4

Transglucosylation kinetics of Os3BGlu6 E178Q and Os3BGlu6 E178A for GA₄-GE donor at various fixed concentrations of sodium azide acceptor.

Enzyme	Sodium azide (mM)	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (mM ⁻¹ s ⁻¹)
Os3BGlu6 E178Q	0	0.09 ± 0.01	0.03 ± 0.001	0.33
	50	1.64 ± 0.15	0.81 ± 0.04	0.49
	100	1.69 ± 0.13	0.86 ± 0.04	0.51
	200	2.07 ± 0.31	0.84 ± 0.08	0.41
	400	3.97 ± 0.65	1.16 ± 0.14	0.29
Os3BGlu6 E178A	0	3.7 ± 0.4	0.09 ± 0.005	0.03
	200	11.6	1.39	0.12
	400	13.36	1.56	0.12

The transglucosylation rates were calculated from the released GA₄ without subtracting hydrolysis product.

Transglucosylation activities of Os3BGlu6E178Q and Os3BGlu6E178A

When azide was included in the reaction, both Os3BGlu6 E178A and E178Q could catalyze transglucosylation reactions from GA₄-

GE to produce β-D-glucosyl azide, as shown schematically in Fig. 5. In this case, Os3BGlu6 E178Q had higher activity than E178A. For the Os3BGlu6 E178A, the GA₄ releasing activity in sodium acetate was much higher than that in MES buffer, even without azide in the system, indicating that acetate may also act as a nucleophile and/or substitute acid/base to rescue the activity, as was seen in the work of Wang et al. [35].

The transglucosylation turnover rate, V_0/E_0 , of GA₄ increased very significantly as the concentration of azide was increased. The rate increase with increasing concentration of azide was also observed by Wang et al. [35]. They concluded that addition of azide as a competitive nucleophile increased k_{cat} values 100–300 fold for substrates for which the rate-limiting step is deglycosylation, such as GA₄-GE with Os3BGlu6 E178Q and Os3BGlu6 E178A. As the concentration of azide is increased, the rate of deglycosylation increases, until glycosylation becomes rate determining. The build-up at the E178Q covalent intermediate was released by reaction with the azide, which was apparently facilitated by interaction of Q178 with the azide. Although the K_m values for azide for the

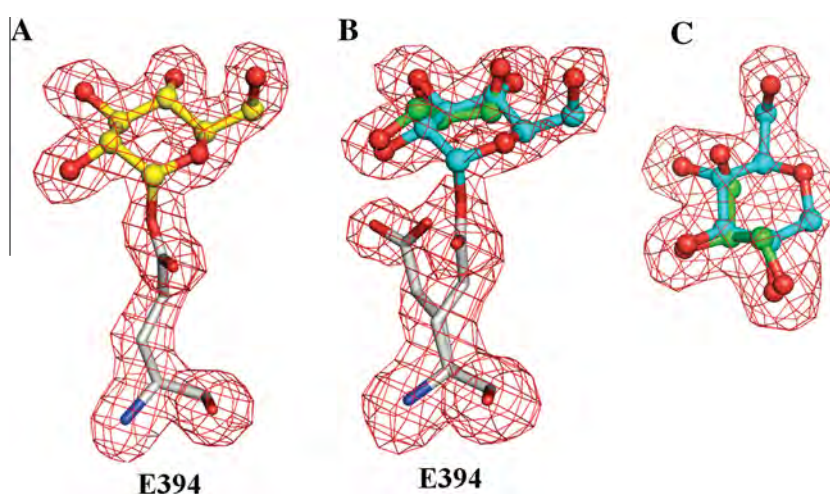


Fig. 3. The electron density map (Omit $F_{obs} - F_{calc}$ contoured at 3σ) of glycosyl-enzyme intermediate of the Os3BGlu6 E178Q with pNPGlc (A) and GA₄-GE (B). (A) The glucose of Os3BGlu6 E178Q soaked with pNPGlc are represented by balls and sticks with carbon in yellow for glucose. The nucleophile residue (E394) is represented by sticks with carbon in gray, oxygen in red and nitrogen in dark blue. (B) The mixed structures of apo Os3BGlu6 E178Q and Os3BGlu6 E178Q covalently bound to Glc from GA₄-GE was modeled with two conformations of E394 and glucose and glycerol occupying the same position, by refining the occupancy to 0.5 for glucose and glycerol and each of the E394 conformations. The glucose and glycerol are shown in ball and stick representation with carbon in cyan for glucose and in green for glycerol. The nucleophile residue (E394) is represented as sticks with carbon in gray, oxygen in red and nitrogen in dark blue. (C) A top view of the glucosyl ring and the overlapping glycerol electron density for Os3BGlu6 E178Q with GA₄-GE shown in B to clarify the different positions of the glucosyl 4-OH and glycerol 3-OH. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

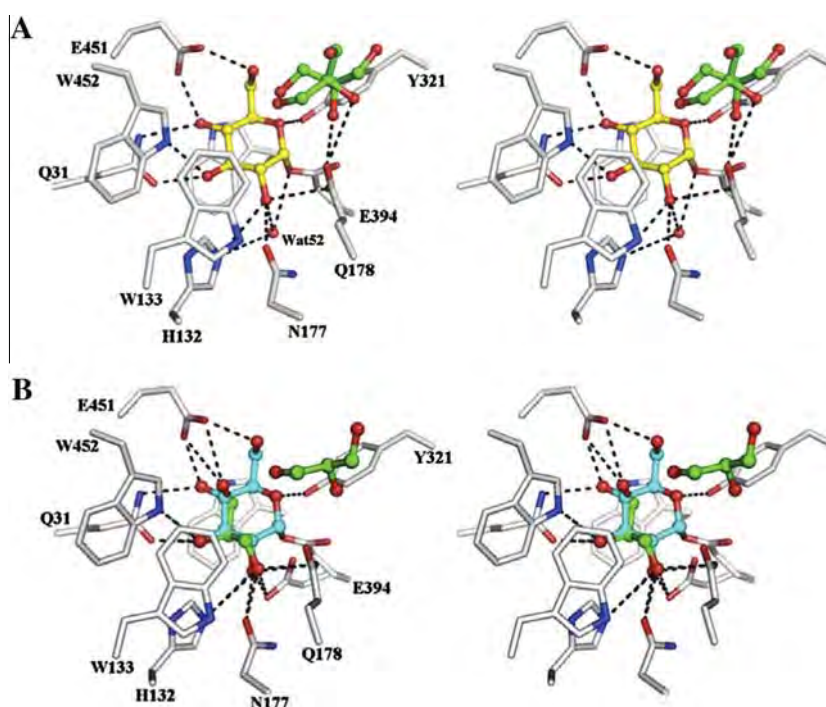


Fig. 4. Binding of glucose in the active site of Os3BGlu6 E178Q. A and B, Stereo view of protein–ligand interaction in the active site of the Os3BGlu6 E178Q soaked with pNPGlc and with GA₄-GE, respectively. The amino acids surrounding the –1 subsite are represented as sticks with carbon in gray, nitrogen in dark blue and oxygen in red. The covalent glucosyl moiety is represented as balls and sticks with carbon in yellow for Os3BGlu6 E178Q soaked with pNPGlc and in cyan for Os3BGlu6 E178Q soaked with GA₄-GE, and oxygen in red. The two glycerols in both complex structures are represented as balls and sticks with carbon in green and oxygen in red. Hydrogen bonding interactions of glucose and glycerol with amino acid residues are shown as black dotted lines.

two mutants were similar, the k_{cat}/K_m values for sodium azide as acceptor with 2 mM GA₄-GE donor of $0.13 \text{ mM}^{-1} \text{ s}^{-1}$ for Os3BGlu6 E178Q and $0.066 \text{ mM}^{-1} \text{ s}^{-1}$ for Os3BGlu6 E178A indicate that the polar glutamine at residue 178 in Os3BGlu6 E178Q supports transglycosylation by azide better than the nonpolar alanine residue in Os3BGlu6 E178A. A stronger interaction with the azide in the Os3BGlu6 E178Q mutant is also supported by the fact that it is inhibited by high concentrations of azide, while the E178A mutant is not.

Structures of Os3BGlu6 covalent intermediate complexes

Although we intended to obtain a noncovalent GA₄-GE complex with an Os3BGlu6 mutant, the high reactivity of this substrate resulted in poor density in the E394D nucleophile mutant active site and the covalent glucosyl-enzyme intermediate in the E178Q acid/base mutant. Fig. 6 shows that the positions and interactions made by amino acid residues surrounding the native glucosyl moiety in the active site of Os3BGlu6 E178Q with glucose from GA₄-GE

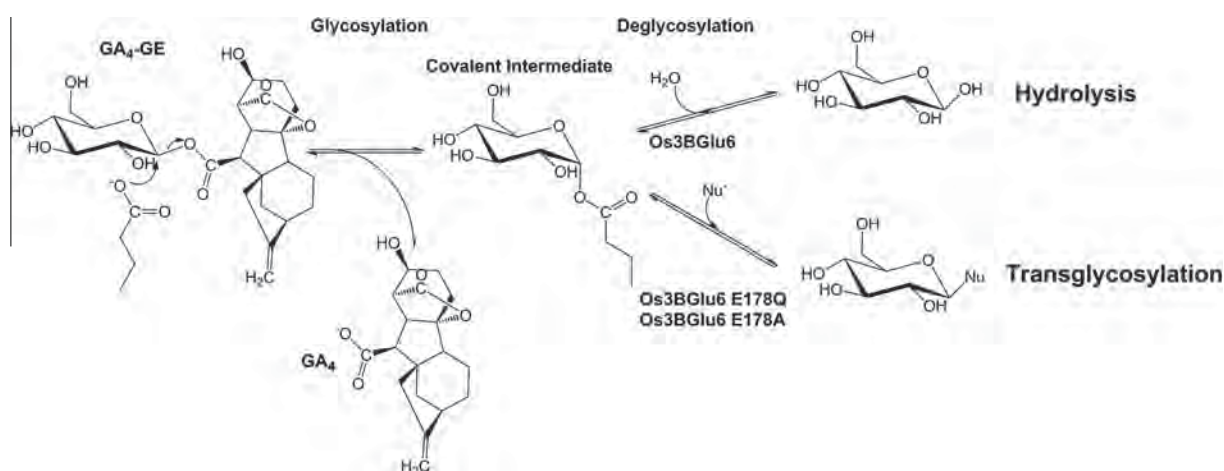


Fig. 5. Schematic of the hydrolysis and transglycosylation reactions catalyzed by Os3BGlu6 and its acid/base mutants with GA₄-GE. The first step, glycosylation, is allowed to proceed without acidic assistance to release the GA₄, due to its low pK_a . The covalent intermediate, seen in the crystal structures, builds up in the acid/base mutants E178A and E178Q, but the presence of a small nucleophile (Nu[–]) such as azide promotes deglycosylation of the mutant enzymes by transglycosylation. In contrast, the wild type Os3BGlu6 only appears to catalyze hydrolysis.

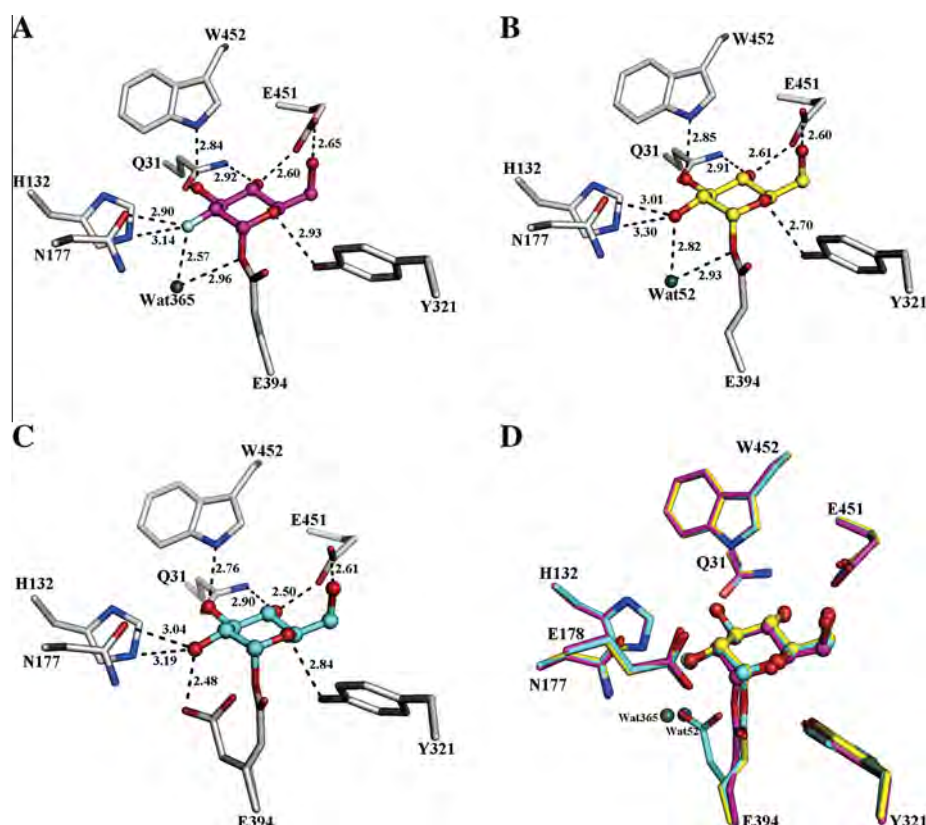


Fig. 6. Comparison of the active site region of Os3BGlu6 bound with G2F (A) and those of the structures of Os3BGlu6 E178Q soaked with pNPGlc (B) and Os3BGlu6 E178Q soaked with GA₄-GE (C). The G2F and glucosyl moieties are represented as balls and sticks with carbon in pink for G2F, in yellow for Os3BGlu6 E178Q soaked with pNPGlc and cyan for Os3BGlu6 E178Q soaked with GA₄-GE, oxygen in red and fluorine in pale cyan. The amino acid residues surrounding the ligands are represented by sticks with carbons in gray, nitrogens in dark blue and oxygens in red for all three structures. (D) Superposition of the three structures in (A–C), showing the nucleophile residues covalently bound to the glucosyl moiety at the –1 subsite of the Os3BGlu6E178Q structures soaked with pNPGlc (yellow carbons) and GA₄-GE (cyan carbons) and to the G2F moiety in Os3BGlu6 (3GNP, pink carbons). The unbound position of the nucleophile is represented by the alternative conformation of Os3BGlu6 E178Q soaked with GA₄-GE. The molecules are represented as in (A–C), except for the protein carbon colors.

(PDB: 3WBE) or pNPGlc (PDB: WBA) were similar to those interacting with the 2-deoxyl-2-fluoroglucopyranoside residue (G2F) in the structure of wild type Os3BGlu6 with G2F (3GNR). In the structure of Os3BGlu6/G2F complex [26], the density of the 2-deoxy-2-fluoroglucosyl residue matched a relaxed ⁴C₁ chair conformation, as did the glucosyl moiety in the Os3BGlu6 E178Q covalent intermediate structure (Fig 6D). The distance between the anomeric carbon of glucose and the Glu394 nucleophile residue of Os3BGlu6 E178Q was also similar to that in the G2F covalent complex (1.3 Å). A water molecule was observed between the 2-OH of glucose or fluorine atom of G2F and N177 and E394 in the Os3BGlu6 E178Q/Glc and Os3BGlu6/G2F structures. A water molecule in this position has been shown to be critical for the hydrolysis and glycosynthase reactions of Os3BGlu7 [43]. The hydrogen bonding pattern between glucosyl residues at the –1 subsite of the Os3BGlu6 E178Q/Glc covalent complex structure is similar when compared to that of the Os3BGlu6/G2F complex (Fig. 6A, B and C). The average distances between the 2-OH of the glucose and 2-fluorine of G2F moieties and polar atoms of the surrounding amino acid residues are similar (Fig 6A–D).

Many structures of GH1 glycosyl-enzyme intermediates with a 2-deoxy-2-fluoroglucosyl (G2F) moiety bound to the nucleophilic residue have been reported, including one for Os3BGlu6 (PDB code 3GNR) [26,44–49]. These GH1 glycosyl-enzyme intermediates used 2,4-dinitrophenyl 2-deoxy-2-fluoroglucoside, with which the substitution of an electronegative fluorine atom for a hydroxyl group adjacent to the reaction center at C-2 destabilizes the transition states and decreases the rates of both glycosylation and deglycosy-

lation [50]. The human cytosolic β -glucosidase acid/base mutant glucose complex (2ZOX) is the only previously reported GH1 structure with the nucleophile covalently bound to native glucose [51]. The hydrogen bonding pattern between G2F molecule and the enzyme at the –1 subsite is conserved in other reported GH1 covalent intermediates, but this pattern lacks the hydrogen bonds for which the 2-OH is the proton donor seen in human cytosolic β -glucosidase and Os3BGlu6 E178Q covalent glucoside complex structures. In general, the reports of G2F complexes do not report hydrogen bonds to the 2-F group, but the conservation of the water in proximity to the fluorine in these complexes supports the idea that the fluorine may accept a hydrogen bond from the water and other nearby hydrogen donors. On the other hand, the loss of possible hydrogen bonds from the 2-OH to O δ of N177 and O ϵ 2 of E394 in Os3BGlu6, which are seen in the native glucosyl-enzyme intermediates (Fig. 6), may result in a further loss of stabilization of the transition state in the 2-F-glucoside reaction [51].

Conclusion

This work identified Os3BGlu6 as a β -glucosidase with relatively high activity to hydrolyze GA₄-GE. Comparison of the activity of Os3BGlu6 and its acid/base mutants, E178Q and E178A, for hydrolysis of pNPGlc and GA₄-GE suggested that hydrolysis of esters and glycosides is similar and the low pK_a of the carboxyl leaving group is what differentiates it, although the rate of the wild type enzyme is much slower with the ester. Despite the observation of little or no transglucosylation activity for wild type

Os3BGlu6, its acid/base mutants were still rescued by azide transglycosylation to form β -D-glucosyl azide. The covalent intermediate of Os3BGlu6 E178Q with unmodified α -D-glucoside was observed with either pNPGlc or GA₄-GE donor substrate, supporting the covalent double displacement mechanism for this enzyme with either glycoside or glycosyl ester substrates.

Acknowledgments

The authors are grateful to Assist. Prof. Dr. Thanaporn Manyum for providing her laboratory for initial synthesis work and Ms. Supaporn Baiya for providing Os4BGlu18 β -glucosidase. Suranaree University of Technology (SUT) and the Center for Scientific and Technological Equipment (CSTE) are thanked for supporting YH in her Ph.D. studies. The research funding was provided by SUT and the National Research University Project of the Commission on Higher Education of Thailand, along with the Thailand Research Fund (Grant BRG5380017).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.abb.2013.06.005>.

References

- [1] T. Fowler, in: A. Esen (Ed.), *Biochemistry and Molecular Biology*, American Chemical Society, Washington, DC, 1993, pp. 56–65. ACS Symposium Series 533.
- [2] J.E. Poulton, *Plant Physiol.* 94 (1990) 401–405.
- [3] L. Duroux, F.M. Delmotte, J.M. Lancelin, G. Keravis, C. Jay-Alleand, *Biochem. J.* 333 (1998) 275–283.
- [4] B. Brzobohatý, I. Moore, P. Kristoffersen, L. Bako, N. Campos, J. Schell, K. Palme, *Science* 262 (1993) 1051–1054.
- [5] A. Falk, L. Rask, *Plant Physiol.* 108 (1995) 1369–1377.
- [6] D.P. Dharmawardhana, B.E. Ellis, J.E. Carlson, *Plant Physiol.* 107 (1995) 331–339.
- [7] M. Reuveni, Z. Sagi, D. Eynor, A. Hetzroni, *Plant Sci.* 147 (1999) 19–24.
- [8] L. Barleben, X. Ma, J. Koepke, G. Peng, H. Michel, J. Stöckigt, *Biochim. Biophys. Acta.* 1747 (2005) 89–92.
- [9] B. Henrissat, *Biochem. J.* 280 (1991) 309–316.
- [10] B.L. Cantarel, P.M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, *Nucleic. Acids Res.* 37 (2009) D233–D238.
- [11] J.R. Ketudat Cairns, A. Esen, *Cell. Mol. Life Sci.* 67 (2010) 3389–3405.
- [12] B. Cobucci-Ponzano, V. Aurilia, G. Riccio, B. Henrissat, P.M. Coutinho, A. Strazzulli, A. Padula, M.M. Corsaro, G. Pieretti, G. Pocsfalvi, I. Fiume, R. Cannio, M. Rossi, M. Moracci, *J. Biol. Chem.* 285 (27) (2010) 20691–20703.
- [13] W. Schliemann, *J. Plant Physiol.* 116 (1984) 123–132.
- [14] K.H. Lee, H.L. Piao, H.Y. Kim, S.M. Choi, F. Jiang, W. Hartung, I. Hwang, J.M. Kwak, I.J. Lee, I. Hwang, *Cell* 126 (2006) 1109–1120.
- [15] K. Kai, K. Wakasa, H. Miyagawa, *Phytochemistry* 68 (2007) 2512–2522.
- [16] S. Wakuta, S. Hamada, H. Ito, H. Matsuura, K. Nabeta, H. Matsui, *Phytochemistry* 71 (2010) 1280–1288.
- [17] Z.Y. Xu, K.H. Lee, T. Dong, J.C. Jeong, J.B. Jin, Y. Kanno, D.H. Kim, S.Y. Kim, M. Seo, R.A. Bressan, D.J. Yun, I. Hwang, *Plant Cell* 24 (2012) 2184–2199.
- [18] Y. Koda, E.A. Omer, T. Yoshihara, H. Shibata, S. Sakamura, Y. Okazawa, *Plant Cell Physiol.* 29 (1988) 1047–1051.
- [19] T. Yoshihara, M. Amanuma, T. Tsutsumi, Y. Okumura, H. Matsuura, A. Ichihara, *Plant Cell Physiol.* 37 (1996) 586–590.
- [20] S. Wakuta, S. Hamada, H. Ito, R. Imai, H. Mori, H. Matsuura, K. Nabeta, H. Matsui, *J. Appl. Glycosci.* 58 (2011) 67–70.
- [21] P.J. Davies, *Plant Hormones: Physiology, Biochemistry and Molecular Biology*, Kluwer Academic Publishers, 1995.
- [22] G. Schneider, in: A. Crazier (Ed.), *The Biochemistry and Physiology of Gibberellins*, vol. 1, Praeger Publishers, New York, 1983, pp. 389–456.
- [23] G. Schneider, E. Jensen, C.R. Spray, B.O. Phinney, *Proc. Natl. Acad. Sci. USA* 89 (1992) 8045–8048.
- [24] M. Koshioka, E. Minami, H. Saka, R.P. Pharis, L.N. Mander, *Gibberellins*, Springer, New York, 1991. 264–272.
- [25] K. Hiraga, T. Yokota, N. Takahashi, *Phytochemistry* 13 (1974) 2371–2376.
- [26] S. Seshadri, T. Akiyama, R. Opasiri, B. Kuaprasert, J.R. Ketudat Cairns, *Plant Physiol.* 151 (1) (2009) 47–58.
- [27] R. Opasiri, J.R. Ketudat Cairns, T. Akiyama, O. Wara-aswapati, J. Svasti, A. Esen, *Plant Sci.* 165 (2003) 627–638.
- [28] R. Opasiri, B. Pomthong, T. Onkoksoong, T. Akiyama, A. Esen, J.R. Ketudat Cairns, *BMC Plant Biol.* 6 (2006) 33.
- [29] S. Luang, J.-I. Cho, B. Mahong, R. Opasiri, T. Akiyama, K. Phasai, J. Komvongsa, N. Sasaki, Y. Hua, Y. Matsuba, Y. Ozeki, J.-S. Jeon, J.R. Ketudat Cairns, *J. Biol. Chem.* 288 (2013) 10111–10123.
- [30] Z. Otwinowski, W. Minor, *Meth. Enzymol.* 276, Academic Press, New York, (1997) 307–326.
- [31] P. Emsley, K. Cowtan, *Acta Crystallogr. D* 60 (2004) 2126–2132.
- [32] G.N. Murshudov, A.A. Vagin, E.J. Dodson, *Acta Crystallogr. D* 53 (1997) 240–255.
- [33] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, *Appl. Cryst.* 26 (1993) 283–291.
- [34] Q. Wang, D. Trimburt, R. Graham, R.A.J. Warren, S.G. Withers, *Biochemistry* 34 (1995) 14554–14562.
- [35] Q. Wang, R.W. Graham, D. Trimburt, R.A.J. Warren, S.G. Withers, *J. Am. Chem. Soc.* 116 (1994) 11594–11595.
- [36] L. Mackenzie, G. Sulzenbacher, C. Divne, T.A. Jones, H.F. Woldike, M. Schulein, S.G. Withers, G. Davies, *Biochem. J.* 335 (1998) 409–416.
- [37] W. Chuenchor, S. Pengthaisong, R.C. Robinson, J. Yuvaniyama, J. Svasti, J.R. Ketudat Cairns, *J. Struct. Biol.* 173 (2011) 169–179.
- [38] T. Kiso, S. Kitahata, K. Okamoto, S. Miyoshi, H. Nakano, *J. Biosci. Bioeng.* 90 (2000) 614–618.
- [39] J.B. Kempton, S.G. Withers, *Biochemistry* 31 (1992) 9961–9969.
- [40] L.P. McIntosh, G. Hand, P.E. Johnson, M. Joshi, M. Korner, L.A. Plesniak, L. Ziser, W.W. Wakarchuk, S.G. Withers, *Biochemistry* 35 (1996) 9958–9966.
- [41] J. Müllegger, M. Jahn, H.M. Chen, R.A. Warren, S.G. Withers, *Protein Eng. Des. Sel.* 18 (2005) 33–40.
- [42] S. Sansenya, J. Maneesan, *Carbohydr. Res.* 351 (2012) 130–133.
- [43] J. Wang, S. Pengthaisong, J.R. Ketudat Cairns, Y. Liu, *Biochim. Biophys. Acta.* 1834 (2013) 536–545.
- [44] W. Burmeister, S. Cottaz, H. Driguez, R. Iori, S. Palmieri, B. Henrissat, *Structure* 5 (1997) 663–675.
- [45] P. Isorna, J. Polaina, L. Latorre-García, F.J. Cañada, B. González, J. Sanz-Aparicio, *J. Mol. Biol.* 371 (2007) 1204–1218.
- [46] T.M. Gloster, S. Roberts, V.M.A. Ducros, G. Perugino, M. Rossi, R. Hoos, M. Moracci, A. Vasella, G.J. Davies, *Biochemistry* 43 (2004) 6101–6109.
- [47] D.L. Zechel, S.P. Reid, D. Stoll, O. Nashiru, R.A. Warren, S.G. Withers, *Biochemistry* 42 (2003) 7195–7204.
- [48] W. Chuenchor, S. Pengthaisong, R.C. Robinson, J. Yuvaniyama, W. Oonant, D.R. Bevan, A. Esen, C.J. Chen, R. Opasiri, J. Svasti, J.R. Ketudat Cairns, *J. Mol. Biol.* 377 (2008) 1200–1215.
- [49] S. Sansenya, R. Opasiri, B. Kuaprasert, C.J. Chen, J.R. Ketudat Cairns, *Arch. Biochem. Biophys.* 510 (2011) 62–72.
- [50] J. Noguchi, Y. Hayashi, Y. Baba, N. Okino, M. Kimura, M. Ito, Y. Kakuta, *Biochem. Biophys. Res. Commun.* 374 (2008) 549–552.
- [51] B. Rempel, S. Withers, *Glycobiology* 18 (2008) 570–586.

Structural analysis and insights into glycon specificity of the rice GH1 Os7BGlu26 β -D-mannosidase

Anupong Tankrathok^a, Javier Iglesias-Fernández^b, Sukanya Luang^{a,c}, Robert Robinson^{d,e,f},
Atsuo Kimura^g, Carme Rovira^{b,h}, Maria Hrmova^c, and James Ketudat Cairns^{a,i,*}

^aSchool of Biochemistry, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima, 30000, Thailand, ^bComputer Simulation and Modeling Laboratory and Institut de Química Teòrica i Computacional (IQTUB), Parc Científic de Barcelona, Baldori Reixac 10-12, 08028, Barcelona, Spain, ^cAustralian Centre for Plant Functional Genomics, School of Agriculture, Food and Wine, University of Adelaide, Glen Osmond, Australia, ^dInstitute of Molecular and Cell Biology, 61 Biopolis Drive, 138673, Singapore, ^eDepartment of Biochemistry, National University of Singapore, 8 Medical Drive, Singapore 117597, ^fSchool of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore, 637551, ^gResearch Faculty of Agriculture, Hokkaido University, Sapporo, Hokkaido, 060-8589, Japan, ^hInstitució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, 08018 Barcelona, Spain and ⁱLaboratory of Biochemistry, Chulabhorn Research Institute, Bangkok 10210, Thailand

Correspondence email: cairns@sut.ac.th

Keywords: X-ray crystallography; structural analysis; glycoside hydrolase; β -mannosidase; rice

Synopsis

The structures of rice Os7BGlu26 β -D-mannosidase, its product complex with β -D-mannose, and Os7BGlu26 computationally docked with 4-nitrophenyl β -D-mannoside suggest that the D-mannose pyranose ring makes a conformational transition from ⁰S₂ skew boat to B_{2,5} boat to ¹S₅ skew boat during deglycosylation of the enzyme in hydrolysis. These structures led to the hypothesis that residues interacting with the catalytic acid/base play a role in determining β -D-mannosidase activity, which was supported by site-directed mutagenesis and kinetic analysis.

Abstract

Rice Os7BGlu26 is a family GH1 glycoside hydrolase with a 3-fold higher $k_{\text{cat}}/K_{\text{m}}$ value for 4-nitrophenyl β -D-mannoside (4NPMan) than for 4-nitrophenyl β -D-glucoside (4NPGlc). To investigate selectivity for β -D-mannoside and β -D-glucoside substrates, we elucidated the structure of apo Os7BGlu26 at a resolution of 2.20 Å and one of Os7BGlu26 with mannose at a resolution of 2.45 Å from isomorphous crystals in the P2₁2₁2₁ space group. The $(\beta/\alpha)_8$ barrel structure is similar to other GH1 family structures, but with a narrower active site cleft. The Os7BGlu26 structure with D-mannose corresponds to a product complex, with β -D-mannose in the ¹S₅ skew boat conformation. Docking of the ¹S₃, ¹S₅, ²S_O, and ³S₁ pyranose ring conformations of 4NPMan and 4NPGlc substrates into the active site of Os7BGlu26 indicated lowest energies in ¹S₅ and ¹S₃ skew boat conformations. Comparison of these docked conformers with other rice GH1 structures revealed differences in the residues interacting with the catalytic acid/base between enzymes with and without β -D-mannosidase activities. Mutation of Tyr134 to Trp in Os7BGlu26 resulted in a lower ratio of $k_{\text{cat}}/K_{\text{m}}$ for 4NPMan to that for 4NPGlc, while Tyr134 mutation to Phe increased this ratio 13-fold. Mutation of Cys182 to Thr decreased both activity and selectivity for β -D-mannoside. We conclude that interactions with the catalytic acid/base play a significant role in glycon selection.

1. Introduction

β -D-Mannosidases (β -D-mannopyranoside mannohydrolase, E.C. 3.2.1.25) hydrolyze β -glycosidic linkages between non-reducing β -D-mannosyl residues and the neighboring aglycons or oligosaccharides. β -D-Mannosidases are found in a variety of organisms, including archaea, bacteria, animals, fungi, and plants. In plants, these enzymes are present during and following seed germination in legumes (McCleary & Matheson, 1975), lettuce (*Lactuca sativa* L.) (Ouellette & Bewley, 1986) and tomato (*Lycopersicon esculentum* Mill.) (Mo & Bewley, 2002). Hrmova and colleagues (2006) showed that barley (*Hordeum vulgare* L.) beta-glucosidase isoenzyme II (also called BGQ60 and β II, and designated here as HvBII), appears to work as a β -D-mannosidase in concert with β -mannanase to hydrolyze barley seed β -D-mannans, so they gave it the new name HvMannos.

β -D-Mannosidases belong to glycoside hydrolase (GH) families GH1, GH2 and GH5 in CAZy (Carbohydrate-Active enZYme Database, <http://www.cazy.org>; Cantarel *et al.*, 2009). These families fall within the GH-A clan, members of which adopt a $(\beta/\alpha)_8$ barrel structure with two catalytic glutamic acid residues, the acid/base and the nucleophile, located at the C-terminal ends of β -strands 4 and 7, respectively. Currently, only two β -D-mannosidase structures have been reported, both from bacteria, GH2 Man2A from *Bacteroides thetaiotaomicron* VPI-5482 (Tailford *et al.*, 2008)

and GH5 Man5A from *Cellvibrio mixtus* NCIMB 8633 (Dias *et al.*, 2004). While no GH1 β -D-mannosidase structure has yet been reported, structures for 39 other GH1 hydrolases are available, including 5 from archaea, 17 from bacteria, and 17 from eukaryotes, most of which represent β -D-glucosidases (E.C. 3.2.1.21). A few other GH1 structures represent bacterial β -D-glycosidases with broad substrate specificity, two are thioglucosidases (E.C. 3.2.1.147), five are 6-phospho- β -glucosidases (E.C. 3.2.1.86), and one is a 6-phospho- β -galactosidase (E.C. 3.2.1.85).

It is believed that β -D-mannosides and β -D-glucosides, which differ only in that the 2-hydroxyl is equatorial in D-mannose and axial in D-glucose, are hydrolyzed via different conformational trajectories, based on X-ray crystallographic (Vocadlo & Davies, 2008) and conformational free energy landscape analysis (Biarnés *et al.*, 2007; Adrèvol *et al.*, 2010) data, as shown in Fig. 1. During hydrolysis, β -D-glucopyranosyl rings are thought to undergo a conformational change via a 4H_3 half-chair. The ring appears to be primed to form this half-chair by its distortion to a 1S_3 skew boat upon binding the enzyme and upon completing the glycosylation step of hydrolysis, and the glucose in the covalent glycosyl-enzyme intermediate is found in a relaxed 4C_1 conformation (Davies *et al.*, 1998; 2012) (Fig. 1A). In contrast, crystal structures of the β -D-mannosidase complexes with substrate and transition-state-based inhibitors reveal a 1S_5 skew boat in the Michaelis complex, which proceeds through a $B_{2,5}$ boat near the transition state to a 0S_2 skew boat in the covalent complex (Ducros *et al.*, 2002; Tailford *et al.*, 2008) (Fig. 1B). The β -D-mannoside has been reported in the 1S_5 skew boat conformation in Michaelis complexes with a GH2 β -mannosidase (PDB code 2WBK; Offen *et al.*, 2009) and a GH26 β -mannanase (1GVY; Ducros *et al.*, 2002). The 1S_3 skew boat conformation of β -D-glucoside substrates has been published in several enzymes, including the GH1 sorghum dhurrinase (PDB entry 1V03, Verdoucq *et al.*, 2004), rice (*Oryza sativa*) Os3BGlu6 (3GNP, Seshadri *et al.*, 2009), termite (*Neotermes koshunensis*) β -D-glucosidase NKBgl (3AI0, Jeng *et al.*, 2011), and rice Os4BGlu12 (3PTQ, Sansenya *et al.*, 2011). In the case of rice BGlu1 (designated here with its systematic name of Os3BGlu7), the nonreducing β -D-glucopyranosyl ring in the oligosaccharide complexes with catalytic mutants was reported to be between the 1S_3 and 4H_3 or closely related 4E conformations, again supporting the trajectory of 1S_3 to 4H_3 (Chuenchor *et al.*, 2011). Calculations of the energetics of the conformational transitions in solution show the same conformational preferences of 1S_3 to 4H_3 to 4C_1 for β -D-glucosides and 1S_5 to $B_{2,5}$ to 0S_2 for β -D-mannosides, suggesting that enzymes may tend to bind the lowest energy forms as they catalyze hydrolysis (Biarnés *et al.*, 2007; Adrèvol *et al.*, 2010).

Plant β -D-mannosidases fall in a single amino acid sequence-based phylogenetic cluster of GH1, which also contains β -D-glucosidases with β -D-mannosidase activity (Opassiri *et al.*, 2006; Kuntothom *et al.*, 2009). The amino acid sequences of rice Os7BGlu26 and three closely related rice

β -D-glycosidase isoenzymes (Os3BGlu7, Os1BGlu1 and Os3BGlu8) are grouped in this phylogenetic cluster with barley HvBII (Hrmova *et al.*, 1996; 1998), Arabidopsis BGLU44 (Xu *et al.*, 2004) and tomato LeMside (Mo & Bewley, 2002) β -D-mannosidases, with which Os7BGlu26 shares 82%, 66% and 66% amino acid sequence identity, respectively (Opassiri *et al.*, 2006; Kuntothom *et al.*, 2009). Within this group, only Os3BGlu7 has an elucidated structure, but it hydrolyzes 4NPGlc with a 34-fold higher $k_{\text{cat}}/K_{\text{m}}$ value than 4NPMan (Opassiri *et al.*, 2004). In contrast, rice Os7BGlu26 and barley HvBII, which have closely related amino acid sequences, as shown in Fig. 2, hydrolyze 4NPMan with 3-fold and 12-fold higher $k_{\text{cat}}/K_{\text{m}}$ values than 4NPGlc, respectively (Kuntothom *et al.*, 2009). The Os7BGlu26 β -D-mannosidase also hydrolyzes manno oligosaccharides and celooligosaccharides, and the natural glycosides dhurrin, D-amygdalin and *p*-coumaryl alcohol β -D-glucoside.

Given the difference seen in hydrolysis of the glucoside and mannoside substrates, it is of interest to understand the basis for β -D-mannosidase catalysis in GH1 enzymes that have both β -D-mannosidase and β -D-glucosidase activities, and their preference for β -D-mannoside *versus* β -D-glucoside substrates. Saturation transfer difference nuclear magnetic resonance (STD-NMR) showed HvBII β -D-mannosidase bound 4NP β -D-thioglucoside in either a 1S_3 or 3S_5 conformation and 4NP β -D-thiomannoside in a relaxed 4C_1 chair (Kuntothom *et al.*, 2010). Quantum mechanics/molecular mechanics (QM/MM) simulations for Os3BGlu7 β -D-glucosidase and HvBII β -D-mannosidase binding indicated their preference to bind 1S_3 skew boat conformations of 4NPGlc, 4NPMan and 4-nitrophenyl β -D-thiomannoside, and the 4C_1 chair conformation of β -D-thioglucoside in the Michaelis complex. Notably, Kuntothom and colleagues (2010) used a homology model of the HvBII β -D-mannosidase for the QM/MM simulations, due to lack of a GH1 β -D-mannosidase structure. To investigate the molecular mechanism for β -D-mannosidase specificity in plant GH1 β -D-mannosidases, we determined the Os7BGlu26 β -D-mannosidase crystal structure and probed the residues involved by computational docking and mutagenesis.

2. Materials and methods

2.1. Protein expression and purification

The pET32a/Os7BGlu26 plasmid, which includes the *Os7BGlu26* cDNA in frame to produce an N-terminally thioredoxin and His-tagged Os7BGlu26 fusion protein (Kuntothom *et al.*, 2009), was transformed into *E. coli* strain Rosetta-gami(DE3) cells. The cells were cultured in low salt LB (Lennox) media containing 50 $\mu\text{g/ml}$ ampicillin, 15 $\mu\text{g/ml}$ kanamycin, 12.5 $\mu\text{g/ml}$ tetracyclin and 34 $\mu\text{g/ml}$ chloramphenicol. When the culture optical density at 600 nm reached 0.4–0.5, protein expression was induced with 0.3 mM IPTG for 24 h, at 20°C. Cell pellets were collected by centrifugation and suspended in the extraction buffer (50 mM Tris-HCl, pH 8.0, 150 mM sodium

chloride, 200 µg/ml lysozyme, 1% (v/v) Triton-X 100 mM, 1 mM PMSF, 4 µg/ml DNase I) at approximately 25 °C, 30 min. Centrifugation removed insoluble debris, and the protein was purified from the soluble extract by Immobilized Metal-Affinity Chromatography (IMAC) on cobalt-equilibrated IMAC resin (GE Healthcare). The resin was washed with the equilibration buffer (150 mM NaCl, 50 mM Tris-HCl, pH 8.0), followed by 20 mM imidazole in the equilibration buffer, and eluted with 250 mM imidazole in the equilibration buffer. The fractions of Os7BGlu26 containing β-D-glucosidase activity, as judged by 4NPGlc hydrolysis, were pooled and imidazole removed by dialysis in 150 mM NaCl, 20 mM Tris-HCl buffer, pH 8.0. The dialysed preparation was concentrated in a 30 kDa molecular mass cut off (MWCO) Centricon centrifugal filter (Millipore). The N-terminal fusion tag was removed from the Os7BGlu26 fusion protein by cleavage with 2 ng enterokinase (New England Biolabs) per 1 mg of a fusion protein at 23°C for 18 h, followed by a second round of IMAC. The flow-through fractions containing β-glucosidase activity were pooled and the protein purity was analysed by SDS-PAGE. The Os7BGlu26 was dialysed and concentrated with a 30 kDa MWCO Centricon filter to obtain approximately 6 mg/ml of Os7BGlu26 protein in the 150 mM NaCl, 20 mM Tris-HCl buffer, pH 8.0.

2.2. Protein crystallization

Before crystallization, purified Os7BGlu26 was filtered through an Ultrafree-M 0.22 µm filter (Millipore) (4,000 g, 4°C, 5 min). Crystallization conditions were screened by the microbatch under oil method at 288 K with precipitants from the Crystals Screen High Throughput HR2 kits (Hampton Research). After optimization of the crystallization conditions, the crystals were grown in 0.8 M K₂Na tartrate, 0.1 M Na HEPES, pH 7.5. For the complex with ligand, the crystals were soaked in 400 mM D-mannose in the precipitant solution. Prior to data collection, the crystals were soaked in cryoprotectant containing the precipitant, 400 mM D-mannose and 20% (v/v) glycerol. The crystals were flash vitrified and stored in liquid nitrogen.

2.3. Data collection, processing and structure refinement

The X-ray data for Os7BGlu26 were collected at the BL13B1 beamline at the National Synchrotron Radiation Research Center (NSRRC), Hsinchu, Taiwan, with a 1.0-Å wavelength X-ray beam and an ADSC Quantum 315 CCD detector. The crystals were maintained at 110 K in a cold stream of nitrogen throughout data collection. All data sets were indexed, integrated, and scaled with the HKL-2000 package (Otwinowski & Minor, 1997). The native structure was solved by molecular replacement with the rice Os3BGlu7 β-glucosidase structure (PDB code 2RGL) used as a search model and the *MOLREP* program (Vagin & Teplyakov, 1997) in the *CCP4* suite of programs (Collaborative Computing Project Number 4, 1994), followed by refinement with *REFMAC5*

(Murshudov *et al.*, 1999). The structure of the complex with D-mannose was solved by rigid body refinement of the native structure in *REFMAC5*. Model building was performed with the *COOT* program (Emsley *et al.*, 2004). For the Os7BGlu26 mannose complex structure, initial refinement and model building was performed with a resolution cutoff of 2.25 Å. However, due to the low completeness of the refined data in the outer shells, the model was refined to the 2.45 Å cutoff to achieve acceptable completeness parameters (Table 1). The quality of the final model was assessed with *PROCHECK* (Laskowsky *et al.*, 1993) and *MOLPROBITY* (Chen *et al.*, 2010). Graphic representations of structures were generated in *PyMol* (Schrödinger LLC).

2.4. Docking calculations

The 4NPMan and 4NPGlc ligands were docked in to the active site of Os7Glu26 via a Lamarckian genetic search algorithm, as implemented in the *Autodock* 4.2 program (Morris *et al.*, 2009). The ligands were docked into four different conformations (1S_3 , 1S_5 , 2S_0 , 3S_1), which are representative of the different catalytic itineraries followed by glycoside hydrolases (Vocadlo & Davies, 2008). Ligand conformations were constructed manually by adding the 4NP portion to the corresponding sugar conformations obtained from previous studies (Biarnés *et al.*, 2007; Adrèvol *et al.*, 2010). All ligands were geometry optimized with Density Functional Theory and the *CPMD* program (*CPMD* program v. 3.15). A restraint on the sugar ring was added to maintain the desired conformation. All calculations were performed on the Os7BGlu26 enzyme excluding all crystallographic water molecules. Protonation states for histidine residues were assigned based on the hydrogen bond environment and E179 was modeled as protonated because of its role as an acid/base residue. Gasteiger charges were assigned to the protein and ligand atoms using *AutoDock tools*. One hundred *AutoDock* runs were performed for each one of the substrates tested to calculate the binding energy, holding the enzyme but not the ligands fixed. A grid with dimensions 40 x 40 x 40 Å centered on the catalytic residues acid/base (E179) and nucleophile (E389) was used.

2.5. Site-directed mutagenesis

Os7BGlu26 mutants were constructed with the QuikChange site-directed mutagenesis kit (Stratagene) with the pET32a/Os7BGlu26 plasmid as the template. The following oligonucleotide primers were used for mutagenesis: for the E179Q mutation, 5'-GAC TGG TTT ACC TTC AAT CAG CCG AGA TGC GTT GCT G-3' and its reverse complement; for Y134W, 5'-GCA AAC CTC TAC CAC TGG GAC CTA CCA TTA GCA C-3' and its reverse complement; for C182T, 5'-CTT CAA TGA GCC GAG AAC CGT TGC TGC TCT AGG-3' and its reverse complement; for Y134F, 5'-CGC AAA CCT CTA CCA CTT TGA CCT ACC ATT AGC AC-3' and its reverse complement and for C182A, 5'-CTT CAA TGA GCC GAG AGC CGT TGC TGC TCT AGG-3' and its reverse

complement (the mutated bases are underlined in each case). All mutant plasmids were sequenced through the entire Os7BGlu26 coding region in both directions.

2.6. Kinetic studies

The kinetic parameters of the enzymes with the 4NPGlc and 4NPMan substrates were determined from triplicate assays containing 0.05 – 18.5 µg enzyme, substrates at concentrations from 0.01 to 30 mM, and 1 µg/µl BSA in 50 mM sodium acetate buffer, pH 5.0, in a total volume of 140 µl, at 30 °C for reaction intervals that had linear initial velocities (pseudo-first order rate or V_o). Reactions were stopped by alkalization with 70 µl of 0.4 M sodium carbonate and the absorbance at 405 nm was read and compared to a 4-nitrophenolate standard curve in the same buffer. The kinetic parameters were calculated by nonlinear regression of the Michaelis-Menten plots with *GraFit 5.0* (Erithacus Software, Horley, Surrey, UK). The Gibbs free energy change for transition state binding was calculated as $\Delta\Delta G_{S^*mut} = -RT[\ln(k_{cat}/K_m)_{mutant} - \ln(k_{cat}/K_m)_{wildtype}]$ (Fersht *et al.*, 1987). The inhibition constant (K_i) of Os7BGlu26 for inhibition by HEPES was determined at 30°C by incubating 2 µg enzyme in 50 mM sodium acetate buffer, pH 5.0, containing 1 µg/µl BSA, and with 0–300 mM HEPES for 10 minutes. The residual enzyme activities were monitored by assaying activity against 0.1, 0.5, 1.0 and 1.5 mM 4NPMan. Inhibition constants K_i were calculated by linear regression of a plot of the apparent K_m/V_{max} values (slopes of Lineweaver-Burk plots) *versus* inhibitor concentrations.

3. Results and Discussion

3.1. Os7BGlu26 protein and crystal production

Expression of the soluble 66 kDa Os7BGlu26 fusion protein with N-terminal thioredoxin and His-tags was improved in *E. coli* strain Rosetta-gami(DE3), compared to the previously reported expression of the same construct in the Origami(DE3) strain (Kuntothom *et al.*, 2009). Purification by IMAC, followed by removal of the N-terminal fusion tag by enterokinase cleavage and adsorption of the tag to IMAC resin, yielded the 50 kDa Os7BGlu26 protein with approximately 90% purity on SDS-PAGE (Appendix A, Fig. S1). This protein was screened for crystallization, and Os7BGlu26 crystals with dimensions of 90 x 20 x 20 µm were observed within one week in microbatch screening with a precipitant of 0.8 M K,Na tartrate, 0.1 M Na HEPES, pH 7.5. When the pH and salt concentrations of the precipitant were optimized in hanging-drop vapor diffusion, a single crystal with dimensions of 160 x 25 x 25 µm was obtained within 5 days in 0.58 M K,Na tartrate, 0.1 M Na HEPES, pH 7.25 (Appendix A, Fig. S2).

3.2. Structure and model quality

The apo Os7BGlu26 crystal diffracted X-rays to 2.20 Å resolution, and belonged to the orthorhombic $P2_12_12_1$ space group. Its unit-cell parameters were $a = 68.1$ Å, $b = 71.7$ Å, $c = 136.7$ Å. Diffraction of the crystal soaked in D-mannose gave an isomorphous dataset processed to 2.45 Å resolution and unit cell parameters of $a = 68.0$ Å, $b = 73.6$ Å, $c = 134.0$ Å. The data collection statistics for both crystals are summarized in Table 1. The asymmetric units of both crystals were estimated to contain one molecule, with a Matthews coefficient (V_M) of 3.03 Å³ Da⁻¹ (Matthews, 1968) and a solvent content of 59.5% for the apo Os7BGlu26 crystal, and a V_M of 3.05 Å³ Da⁻¹ and a solvent content of 59.7% for the D-mannose-soaked crystal.

The fold of the structure of Os7BGlu26 is a classic TIM (β/α)₈-barrel, similar to other GH1 enzymes (Fig. 3A). The structure placed the highly conserved E179 and E389 at the C-terminal ends of β -strands 4 and 7, respectively. These residues are positioned at the bottom of the active site cleft, as observed for the catalytic acid/base and nucleophile residues in other members of GH clan A (Jenkins *et al.*, 1995; Henrissat *et al.*, 1995). Further, nucleophilic rescue of mutants of the corresponding residues confirmed they are the catalytic residues in the closely related Os3BGlu7 β -D-glucosidase (Hommalai *et al.*, 2007; Chuenchor *et al.*, 2011). The four variable loops that have been reported to account for much of GH1 structural and functional diversity (Sanz-Aparicio *et al.*, 1998) connect the β -strands and α -helices at the carboxy-terminal side of the core barrel structure. These loops are: loop A (A28-D68), loop B (E179-T209), loop C (H317-P366) and loop D (N390-D406) (Figs 2 and 3A). Although, no electron density was observed for the 14-residues from the N-terminal fusion tag (A-M-A-D-I-T-S-L-Y-K-K-A-G-S-A) and the five C-terminal residues (S-K-K-R-N), five amino acid residues of the fusion protein linker region (A-A-P-F-T) and residues 1-478 of the mature Os7BGlu26 gave clear electron density for the structure. Two cis-peptide bonds were found between A194 and P195, and between W436 and S437, as seen in other plant GH1 enzymes (Barrett *et al.*, 1995). The conserved disulfide bond found in nearly all plant GH1 enzymes was present on loop B between C198 and C201. The conserved active site tryptophan, W444, fell in the outlier region of the Ramachandran plot, while it is found in a similar outlier or borderline region in other GH1 enzymes (Czjzek *et al.*, 2000; Chuenchor *et al.*, 2008). Other Ramachandran statistics were similar to other plant GH1 structures.

Fig. 3B shows that a glycerol molecule, originating from the cryoprotectant, was bound in the active site. This molecule hydrogen bonded to Q32, Y318, E389, E443 and two water molecules, which also interacted with H133 and W444, and the catalytic residues E179 and E389. The distance between C δ carbons of the catalytic acid/base, E179, and nucleophile, E389, was 4.9 Å, consistent with the distance expected for the retaining mechanism of glycoside hydrolases (Rye & Withers, 2000). The carboxyl oxygen of E179, the acid/base residue, had a close contact of 2.7 Å with the Y134 hydroxyl. In addition, the conserved W436 and F452 residues at the -1 subsite provided an

aromatic platform for sugar binding and the additional hydrophobic interactions with the substrate, respectively (Fig. 3B and 3C), as noted for other GH1 enzymes (Czjzek *et al.*, 2000). A HEPES molecule from the crystallization buffer was found in the substrate binding cleft at the +1, +2 and +3 subsites, which were defined for binding of β -(1,4)-linked D-glucosyl residues in Os3BGlu7 (Chuenchor *et al.*, 2011). The HEPES molecule was most favourably modeled in two alternate conformations, both of which hydrogen bonded to E179 and Tyr360 (Fig. 3C). HEPES was found to bind to the Os7BGlu26 with a non-competitive inhibition constant K_i of 81.4 mM and ΔG -6.3 kJ mol⁻¹ (Appendix A, Fig. S3). The combined occupancy of the two alternate HEPES molecules was constrained to be 1.0 in the refinement. This artificial full occupancy is reflected by higher temperature factors ($B = 28.9$ to 46.4 for the individual HEPES atoms) (Table 1).

3.3. Structural comparison of Os7BGlu26 with other rice GH1 structures

In comparison to other rice GH1 structures, the active site of Os7BGlu26 has the narrowest shape, with a gate of 9.9 Å x 12.5 Å (from the atom centers of Q337 O ϵ 1 to Y360 O η and from N190 N ϵ 2 to Y346 O η , respectively), due to the presence of F192 on loop B and Y360 on loop C, which constrict the substrate route into the active site. In comparison, these parameters for the Os3BGlu7 (10.3 Å x 18.1 Å from Q187 C δ to W358 C β and from Y341 C ϵ 2 to L442 C β , respectively) Os3BGlu6 (12.1 Å x 20.8 Å from A189 C β to W366 C β and from L 342 C δ 2 to A 454C α , respectively) and Os4BGlu12 (10.1 Å x 19.8 Å from K203 C ϵ to W365 C β and from L348 C δ 1 to N452 C α) β -D-glucosidases indicate that these enzymes have broader substrate-binding clefts (Fig. 4.). This suggests that a preferred substrate for Os7BGlu26 should be the one with a small aglycon or a straight and narrow chain oligosaccharide, consistent with its hydrolysis of β -(1,4)-linked manno- and gluco-oligosaccharides, dhurrin, D-amygdaalin, and *p*-coumaryl alcohol glucoside (Appendix B) (Kuntothom *et al.*, 2009).

Although the four rice GH1 enzymes have different substrate specificities, their overall structures are very similar. Upon superposition of these structures, Os7BGlu26 had the RMSD value of 0.576 Å over 437 C α atoms with rice Os3BGlu7 β -glucosidase (2RGL), with which it shares 63% sequence identity. The other RMSD values are 0.481 Å over 379 C α residues with Os3BGlu6 (3GNO, 52% sequence identity), and 0.522 Å over 391 C α residues with Os4BGlu12 (3PTK, 51% sequence identity).

3.4. Os7BGlu26 in complex with the D-mannose hydrolysis product

To provide evidence for the interactions between Os7BGlu26 and D-mannosyl glycon, D-mannose was soaked in the Os7BGlu26 crystal. The resolution of this complex structure dataset was limited at 2.45 Å, which resulted in a calculated $F_o - F_c$ OMIT map in which the mannose residue

could be placed unambiguously (Fig. 5A). The density clearly showed the presence of D-mannose in the β -anomeric configuration and the 1S_5 conformation, indicating that Os7BGlu26 binds specifically the β -anomer of D-mannose. The β -D-mannose hydrogen bonded to Q32, H133, Y134, N178, E179, E389, Y318, E443 and W444 (Fig. 5B).

Previous structures of β -D-mannosidase from GH2 and β -D-mannanase from GH26 include the Michaelis complex, transition state analogue and covalent intermediate complexes, however this is the first report of a product complex. Free D-mannose in the active site of Os7BGlu26 was distorted to the 1S_5 skew boat, which is the same conformation as that reported for the Michaelis complexes of the enzymes hydrolyzing β -D-mannosides (Ducros *et al.*, 2002; Offen *et al.*, 2009). Superposition of the β -D-mannose complex and native structure showed that the position of the β -D-mannose was in essentially the same position as the glycerol and water molecules in the native Os7BGlu26 structure (Fig. 5C). Nearly all of the residues in the -1 subsite were in the same positions in the two structures, with C δ of the catalytic nucleophile, E389, displaced by 0.4 Å. The superimposition of the Os7BGlu26/ β -D-mannose complex structure with that of the Os3BGlu7 complex with cellopentaose (PDB entry 3F5K) showed that the 1S_5 configured 1- β -D-mannose orientation was similar to that of the 1S_3 configured non-reducing terminal β -D-glucosyl ring of the cellopentaose in the -1 subsite. A HEPES molecule occupied the +1, +2, and +3 subsites defined for β -(1,4)-linked glucosyl residues in the Os3BGlu7 cellopentaose complex, apparently in multiple conformational states (Fig. 5D). This structure indicates that the β -D-mannose hydrolysis product may be retained in the -1 subsite in a 1S_5 conformation that suggests a conformational pathway of 0S_2 to $B_{2,5}$ to 1S_5 for the deglycosylation step of Os7BGlu26 β -D-mannoside hydrolysis, since the deglycosylation step in the retaining mechanism is essentially the reverse of the glycosylation step shown in Fig. 1.

3.5. Docking studies of the Michaelis complex

To provide further evidence for the conformational itinerary in the glycosylation step of Os7BGlu26, 4NPGlc and 4NPMAN were computationally docked in the active site of the Os7BGlu26 structure in the four starting conformations observed for pyranoside rings in glycoside hydrolase mechanisms (Volcadlo & Davies, 2008). The predicted binding energies for all the conformations tested are shown in Figure 6C. The binding energies for the 4NPGlc conformations are higher (more negative) than those calculated for 4NPMAN (Fig. 6). Therefore, the glucose derivative molecule binds more tightly to the enzymatic cavity than the mannose derivative. This is consistent with the lower K_m values observed for a β -D-glucoside (K_m 0.124 mM) compared to a β -D-mannoside (K_m 0.48 mM) (Table 2).

Both 4NPMAN and 4NPGlc show similar trends of binding affinity with respect to the sugar conformation (Fig. 6). However, whereas 4NPGlc binds preferentially to the enzyme in a 1S_3

conformation (consistently with the experimental and theoretical evidence for a 1S_3 - 4H_3 - 4C_1 catalytic conformational itinerary for β -glucosyl hydrolases), both 1S_3 and 1S_5 have a similar stability for 4NPMMan. Together with the extensive literature supporting that the hydrolysis of β -D-mannosides follow a 1S_5 - $B_{2,5}$ - 0S_2 itinerary (Ducros *et al.*, 2002; Tailford *et al.*, 2008; Offen *et al.*, 2009; Ardèvol *et al.*, 2010), the results obtained suggests that the two substrates might follow a different conformational itinerary for catalysis (1S_3 - 4H_3 - 4C_1 for 4NPGlc and 1S_5 - $B_{2,5}$ - 0S_2 for 4NPMMan). In fact, the 1S_5 mannoside substrate was found more frequently than 1S_3 glucoside substrate in the docking calculations, which may explain the 10-fold higher k_{cat} observed for the former.

Fig. 6 shows the enzyme complexes with the 1S_5 β -D-mannoside (4NPMMan) and the 1S_3 β -D-glucoside (4NPGlc). The substrate is nicely accommodated in the binding cavity in each case. The mannose molecule forms hydrogen bonds with Q32, Y134, N178, E179, R181, Y318, E443 and W444 (Fig 6B) and the glucose molecule is hydrogen bonded to Q32, H133, N178, E179, E443 and W444 (Fig 6A). The catalytic residues (E179 and E389) are well oriented for catalysis in both complexes (the carboxylic acid hydrogen of E179 points towards the glycosidic oxygen and the nucleophile is within 3–3.5 Å distance from the anomeric carbon), in agreement with quantum chemical studies of glycosidic bond hydrolysis (Petersen *et al.*, 2010; Biarnés *et al.*, 2011).

3.6. Comparison of Os7BGlu26 with covalent intermediate complexes of other GH1 enzymes

To learn more about the residues effecting glycon specificity of Os7BGlu26, we endeavored to produce a covalent intermediate complex, as it was previously achieved for Os3BGlu7 (Chuenchor *et al.*, 2008). Attempts to soak the mechanism based inhibitors 2,4-dinitrophenyl 2-deoxy-2-fluoroglucoside (dNPG2F) and 2,4-dinitrophenyl 2-deoxy-2-fluoromannoside (dNPM2F) into the Os7BGlu26 crystals resulted in a release of 2,4-dinitrophenolate ion, as judged by a yellow chromophore. However, no glycon density was seen in the active site with either ligand. Pretreatment of Os7BGlu26 with dNPG2F or dNPM2F in sodium acetate buffer, pH 5.0, resulted in rapid hydrolysis with negligible inhibition of the enzyme at concentrations that strongly inhibited other rice GH1 enzymes (data not shown). Although an 0S_2 skew boat mannosyl covalent complex with a GH26 β -D-mannanase has been reported (Ducros *et al.*, 2002), no such structure has been reported in the GH1 family. Therefore, the free Os7BGlu26 structure was superposed with the rice GH1 structures in complex with bound 2-deoxy-2-fluoroglucoside (G2F) moiety. This moiety occupies a low-energy 4C_1 chair conformation and is covalently bound to the catalytic nucleophile residue (Fig. 7). We conjectured that the glucoside and mannoside substrates are likely to bind in the same positions, and in the same orientations in enzymes that belong to the same family, although the conformations of these sugars may differ, as reported by Kuntothom and colleagues (2010), and supported by the computational docking studies in the previous section.

The identities and placement of nearly all amino acid residues in direct contact with glycons in subsite -1 are conserved in the structures of rice GH1 isoenzymes, including the Os7BGlu26 β -D-mannose complex structure and the complexes with 4NPGlc and 4NPMAN generated by molecular docking. The one exception is a tyrosine residue found in all characterized plant GH1 enzymes with β -D-mannosidase activity, *i.e.* Y134 of Os7BGlu26 that corresponds to Y131 of Os3BGlu7 (Fig. 7A), to Y136 of barley HvBII, to Y136 of tomato LeMside and to Y137 of Arabidopsis AtBGLU44. This tyrosine residue is substituted by tryptophan in Os3BGlu6 (W133) and Os4BGlu12 (W134), which lack β -D-mannosidase activity (Fig. 2 and 7B) (Seshadri *et al.*, 2009; Opassiri *et al.*, 2010), and most other plant β -D-glucosidases, including all others with known crystallographic structures, although a few others have tyrosine, phenylalanine or smaller residues in this position (data not shown). It is noteworthy that this tyrosine residue makes a very close contact through a hydrogen bond with the catalytic acid/base residue. Inspection of the superimposed structures revealed that the acid/base residues of Os3BGlu6 and Os4BGlu12 are displaced slightly from the positions of the acid/base residues in Os3BGlu7 and Os7BGlu26, to form contact through hydrogen bonds with T181 of Os3BGlu6 and T182 of Os4BGlu12 in the +1 subsite, whereas Os7BGlu26 and Os3BGlu7 have C182 and I179, respectively, in corresponding positions. The equivalent residues from barley HvBII, tomato LeMside, and Arabidopsis AtBGLU44, V184, V184 and V185, respectively, do not have polar groups to form contact through hydrogen bonds to the catalytic acid/base residue. Although this position is not unique in plant GH1 enzymes, most of those with known structures have threonine in this position, including maize Glu1, sorghum Dhr1, wheat and rye benzoxazinone glucoside β -D-glucosidases, and *R. serpentina* strictosidine and raucafricine β -D-glucosidases. White clover cyanogenic β -D-glucosidase (1CBG) has glycine in this position, but its catalytic acid base maintains a position similar to those of Os3BGlu6, Os4BGlu12, and the other β -D-glucosidase structures with threonine in this position. Aside from the two differences noted above, the -1 subsite architecture is similar in the four rice GH1 structures. Therefore, amino acid residues in other subsites or in the surrounding layers of residues outside catalytic sites might be important for the glycon specificity. Mutations of residues in the layers surrounding the active site residues have been shown to modulate β -D-fucosidase *versus* β -D-glucosidase activities in an insect GH1 β -D-glucosidase (Mendoça & Marana, 2011). A recent attempt to increase the β -D-mannosidase activity of a plant β -D-glucosidase by mutagenesis of residues that were apparently close to the glycon in a homology model showed only marginal increases in the ratios of $k_{\text{cat}}/K_{\text{m}}$ for 4NPMAN versus 4NPGlc, leading the authors to similarly speculate that the shape of the active site may be more critical than the residues that directly interact with the glycon (Ratananikom *et al.*, 2013).

3.7. Kinetic studies of glycon specificity mutants of Os7BGlu26

To investigate function of the catalytic acid/base Glu179 and other key residues in the structure of Os7BGlu26 that were adjudged to play possible roles in Os7BGlu26 glycon-specificity (cf. previous section), we mutated the E179 to Q (E179Q), Y134 to W (Y134W) and F (Y134F), C182 to T (C182T) and C182 to A (C182A) residues. We have also evaluated effects of the Y134W, Y134F, E179Q, C182T and Y134W/C182T mutations on hydrolysis of the β -D-glucoside and β -D-mannoside substrates in precise kinetic terms. The kinetic parameters and the Gibbs free energy changes for the Os7BGlu26 mutations are presented in Table 2.

The Os7BGlu26 catalytic acid/base mutant E179Q showed lower activity than wild-type, as indicated by the low k_{cat}/K_m for both 4NPMAN and 4NPGlc. However, this mutant affected β -D-mannosidase more than β -D-glucosidase activity, since its $\Delta\Delta G$ with 4NPMAN was $+9.0 \text{ kJ mol}^{-1}$, whereas the $\Delta\Delta G$ for 4NPGlc was $+8.1 \text{ kJ mol}^{-1}$. This change was driven by the 44-fold reduction in the k_{cat} , with a little change in K_m for 4NPMAN, although the K_m value increased 8-fold and k_{cat} decreased nearly 3-fold for 4NPGlc. It should be noted that the mutation of the catalytic acid/base to glutamine has been shown to have relatively high activity for substrates with good leaving groups, such as 4NP or 2,4-dinitrophenolate, which have low pK_a and less need for protonation by the catalytic acid base (Müllegger *et al.*, 2005; Chuenchor *et al.*, 2011). The use of acetate buffer, which can act as a substitute base or nucleophile, facilitates the deglycosylation step, so that this mutation has relatively mild effects in this assay. On the other hand mutation of E179 to alanine resulted in poor yields of soluble protein in this expression system, so this mutant could not be characterized (data not shown).

The Os7BGlu26 Y134W mutant had a lesser effect on the β -mannosidase activity (Table 2), as characterized by a $\Delta\Delta G$ of -0.9 kJ mol^{-1} . However, it showed a nearly 6-fold increase in K_m to 2.37 mM that was compensated by a nearly 7-fold increase in k_{cat} to 2.4 s^{-1} . Since this mutant caused a nearly 4-fold decrease in K_m and a little change in k_{cat} for 4NPGlc, it appeared to improve binding of 4NPGlc and the transition state of the first covalent step of its hydrolysis (as judged by its $\Delta\Delta G$ of -3.5 kJ mol^{-1}). It is of note that the k_{cat}/K_m values of the Os7BGlu26 Y134W mutant for 4NPGlc and 4NPMAN are nearly equal, although it will still hydrolyze 4NPMAN much faster than 4NPGlc at substrate concentrations above 0.01 mM. For Os7BGlu26 C182T, the K_m of this mutant was increased by approximately 25 to 30-fold for both substrates, but the k_{cat} increased 2.6-fold for 4NPGlc *versus* only 1.45-fold for 4NPMAN. For this reason, the $\Delta\Delta G$ values were $+7.1$ and $+6.2 \text{ kJ mol}^{-1}$ for 4NPMAN and 4NPGlc, respectively. To verify whether these changes were due to the introduction of the hydrogen-bonding threonine or simply to removing the cysteine sulfhydryl group, we also mutated C182 to alanine. The C182A mutation had minor effects on both 4NPMAN ($\Delta\Delta G$ of $+1.4 \text{ kJ}$) and 4NPGlc ($\Delta\Delta G$ of $+0.8 \text{ kJ mol}^{-1}$), suggesting the major effect of the C182T mutation was from introduction of the T182 sidechain, which could hydrogen bond with the catalytic acid/base. The

double mutant Os7BGlu26 Y134W/C182T had kinetic parameters and $\Delta\Delta G$ similar to the single mutant C186T for 4NPGlc, and the K_m value for 4NPMMan was also similar to Os7BGlu26 C186T, but the k_{cat} was 40% lower in the double mutant. Hence, Y134W and C182T combined to give non-additive effects with 4NPMMan ($\Delta\Delta G$ of $+8.3 \text{ kJ mol}^{-1}$), which could be attributed to the fact that both mutated residues affect the position of the catalytic acid/base (Mildvan *et al.*, 1992; 2004). The results of our mutagenesis imply that the hydrogen-bonding of neighboring residues to the catalytic acid/base have a significant effect on the relative hydrolytic rates of β -D-mannoside *versus* β -D-glucoside substrates.

To differentiate whether the effect of the Y134W mutation was due to increase in size of the aromatic group or loss of the hydrogen bond from the hydroxyl group to the catalytic acid/base, Y134 was mutated to F instead of W in Os7BGlu26 Y134F. The Y134F mutant showed a similar K_m to the wild type enzyme for 4NPMMan and a 1.4-fold increase in the K_m with 4NPGlc compared to the wild-type enzyme. Since the k_{cat} value was 4.4-fold that of wild-type for 4NPGlc, the k_{cat}/K_m of this mutant increased 3-fold to give a $\Delta\Delta G$ of -2.9 kJ mol^{-1} , which is a slightly smaller change than that observed in the Y134W mutant. However, the k_{cat} value for 4NPMMan increased 36-fold and k_{cat}/K_m 39-fold compared to the wild-type enzyme to give a $\Delta\Delta G$ value of -9.2 kJ mol^{-1} for 4NPMMan. Thus, the selectivity of this mutant for 4NPMMan over 4NPGlc is improved more than 37-fold, in terms of the relative k_{cat}/K_m values, suggesting the smaller steric bulk of the aromatic residue is a critical factor in the preference for β -mannosidase over β -glucosidase substrates.

In summary, the determination of the structure of a plant GH1 β -D-mannosidase Os7BGlu26 and structural investigations of the residues interacting with the glucoside or mannoside substrates indicated that the shape of the active site and interactions with surrounding residues are critical for glycon specificity. The mutations described here had differential effects on the k_{cat}/K_m values of 4NPMMan *versus* 4NPGlc, indicating the residues interacting with the catalytic acid/base play a role in determining which of these is hydrolyzed more rapidly. Future structural and enzyme kinetics studies of binding of substrate and transition state analogues will be useful to further illuminate the interactions that differentiate β -D-mannosidase and β -D-glucosidase activities in GH1.

Acknowledgements We thank Professor Stephen Withers from the University of British Columbia for providing 2,4-dinitrophenyl 2-deoxy-2-fluoromannoside. We are also grateful to the Synchrotron Light Research Institute (SLRI) for providing reagents for crystal screening. This work was supported by the Thailand Research Fund Grant BRG5380017, the Strategic Scholarships for Frontier Research Network Ph.D. Program and the National Research University Project to Suranaree University of Technology from the Commission on Higher Education of Thailand. RCR thanks the Agency for Science, Technology and Research (A*STAR), Singapore for support. Data collection was carried out

at the National Synchrotron Radiation Research Center (NSRRC), a national user facility supported by the National Science Council of Taiwan, ROC. The Synchrotron Radiation Protein Crystallography Facility is supported by the National Research Program for Genomic Medicine.

References

- Ardèvol, A., Biarnés X., Planas, A. & Rovira, C. (2010). *J. Am. Chem. Soc.* **132**, 16058–16065.
- Barrett, T., Suresh, C. G., Tolley, S. P., Dodson, E. J. & Hughes, M. A. (1995). *Structure* **3**, 951–960.
- Biarnés, X., Ardèvol, A., Planas, A., Rovira, C., Laio, A. & Parrinello, M. (2007). *J. Am. Chem. Soc.* **129**, 10686–10693.
- Biarnés, X., Ardèvol, A., Iglesias-Fernández, J., Planas, A. & Rovira, C. (2011). *J. Am. Chem. Soc.* **133**, 20301–20309.
- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B. (2009). *Nucleic Acids Res.* **37**, D233–D238.
- Chen V. B., Arendall W. B., Headd J. J., Keedy, D. A., Immormino, R. M., Kapral, G. P., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010) *Acta Cryst.* **D66**, 12–21.
- Chuenchor, W., Pengthaisong, S., Robinson, R. C., Yuvaniyama, J., Oonasant, W., Bevan, D. R., Esen, A., Chen, C. J., Opasiri, R., Svasti, J. & Ketudat Cairns, J. R. (2008). *J. Mol. Biol.* **377**, 1200–1215.
- Chuenchor, W., Pengthaisong, S., Robinson, R. C., Yuvaniyama, J., Svasti, J. & Ketudat Cairns, J. R. (2011). *J. Struct. Biol.* **173**, 169–179.
- CPMD program, Copyright IBM Corp. 1990–2003, Copyright MPI für Festkörperforschung, Stuttgart 1997–2001. URL: <http://www.cpmc.org>.
- Czjzek, M., Cicek, M., Zamboni, V., Bevan, D. R., Henrissat, B. & Esen, A. (2000). *Proc. Natl Acad. Sci.* **97**, 13555–13560.
- Davies, G. J., Mackenzie, L., Varrot, A., Dauter, M., Brzozowski, A. M., Schülein, M. & Withers, S. G. (1998). *Biochemistry* **37**, 11707–11713.
- Davies, G. J., Planas, A. & Rovira, C. (2012). *Acc. Chem. Res.* **45**, 308–316.
- Dias, F. M., Vincent, F., Pell, G., Prates, J. A., Centeno, M. S., Tailford, L. E., Ferreira, L. M., Fontes, C. M., Davies, G. J. & Gilbert, H. J. (2004). *J. Biol. Chem.* **279**, 25517–25526.
- Ducros, V. M., Zechel, D. L., Murshudov, G. N., Gilbert, H. J., Szabó, L., Stoll, D., Withers S. G. & Davies, G. J. (2002). *Angew. Chem. Int. Ed. Engl.* **4**, 2824–2827.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Fersht, A. R., Leatherbarrow, R. J. & Wells, T. N. (1987). *Biochemistry* **26**, 6030–6038.
- Gouet, P., Robert, X. & Courcelle, E. (2003). *Nucleic Acids Res.* **31**, 3320–3323.
- [Henrissat, B.](#), [Callebaut, I.](#), [Fabrega, S.](#), [Lehn, P.](#), [Mornon, J. P.](#) & [Davies, G.](#) (1995). *Proc. Natl Acad. Sci.* **92**, 7090–7094.

- Hommalai, G., Withers, S. G., Chuenchor, W., Ketudat Cairns, J. R. & Svasti, J. (2007). *Glycobiology* **17**, 744–753.
- Hrmova, M., Harvey, A. J., Wang, J., Shirley, N. J., Jones, G. P., Stone, B. A., Høj, P. B. & Fincher, G. B. (1996). *J. Biol. Chem.* **271**, 5277–5286.
- Hrmova, M., MacGregor, E. A., Biely, P., Stewart, R. J. & Fincher, G. B. (1998). *J. Biol. Chem.* **273**, 11134–11143.
- Hrmova, M., Burton, R. A., Biely, P., Lahnstein, J. & Fincher, G. B. (2006). *Biochem. J.* **399**, 77–90.
- Jeng, W. Y., Wang, N. C., Lin, M. H., Lin, C. T., Liaw, Y. C., Chang, W. J., Liu, C. I., Liang, P. H. & Wang, A. H. (2011). *J. Struct. Biol.* **173**, 46–56.
- Jenkins, J., Lo Leggio, L., Harris, G. & Pickersgill, R. (1995). *FEBS Lett.* **362**, 281–285.
- Kuntothom, T., Luang, S., Harvey, A. J., Fincher, G. B., Opassiri, R., Hrmova, M. & Ketudat Cairns, J. R. (2009). *Arch. Biochem. Biophys.* **491**, 85–95.
- Kuntothom, T., Raab, M., Tvaroska, I., Fort, S., Pengthaisong, S., Cañada, J., Calle, L., Jiménez-Barbero, J., Ketudat Cairns, J. R. & Hrmova, M. (2010). *Biochemistry* **49**, 8779–8793.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
- Leah, R., Kigel, J., Svendsen, I. & Mundy, J. (1995). *J. Biol. Chem.* **270**, 15789–15797.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McCleary, B. V. & Matheson, N. K. (1975). *Phytochemistry* **14**, 1187–1194.
- Mendoça, L. M. F. & Marana, S. R. (2011). *Biochim. Biophys. Acta.* **1814**, 1616–1623.
- Mildvan, A. S., Weber, D. J. & Kuliopulos, A. (1992). *Arch. Biochem. Biophys.* **294**, 327–340.
- Mildvan, A. S. (2004). *Biochemistry* **43**, 14517–14520.
- Mo, B. & Bewley, J. D. (2002). *Planta* **215**, 141–152.
- Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. & Olson, A. J. (2009). *J. Comput. Chem.* **16**, 2785–2791.
- Murshudov, G. N., Lebedev, A., Vagin, A. A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst.* **D55**, 247–255.
- Müllegger, J., Jahn, M., Chen, H. M., Warren, R. A. & Withers, S. G. (2005). *Protein Eng. Des. Sel.* **18**, 33–40.
- Offen, W. A., Zechel, D. L., Withers, S. G., Gilbert, H. J. & Davies, G. J. (2009). *Chem. Commun.* **18**, 2484–2486.
- Opassiri, R., Hua, Y., Wara-Aswapati, O., Akiyama, T., Svasti, J., Esen, A. & Ketudat Cairns, J. R. (2004). *Biochem. J.* **379**, 125–131.
- Opassiri, R., Pomthong, B., Onkoksoong, T., Akiyama, T., Esen, A. & Ketudat Cairns, J. R. (2006). *BMC Plant Biol.* **6**, 33–52.

- Opassiri, R., Maneesan, J., Akiyama, T., Pomthong, B., Jin, S., Kimura, A. & Ketudat Cairns, J. R. (2010). *Plant Sci.* **179**, 273-280.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307-326.
- Ouellette, B. F. F. & Bewley, J. D. (1986). *Planta* **169**, 333-338.
- Petersen, L., Ardèvol, A., Rovira, C. & Reilly, P. J. (2010) *J. Am. Chem. Soc.* **132**, 8291-8300.
- Ratanakikom, K., Choengpanya, K., Tongtubtim, N., Charoenrat, T., Withers, S. G. & Konsaeree, P. T. (2013). *Carbohydr. Res.* **373**, 35-41.
- Rye, C. S. & Withers, S. G. (2000). *Curr. Opin. Chem. Biol.* **5**, 573-580.
- Sansanya, S., Opassiri, R., Kuaprasert, B., Chen, C. J. & Ketudat Cairns J. R. (2011). *Arch. Biochem. Biophys.* **510**, 62-72.
- Sanz-Aparicio, J., Hermoso, J. A., Martinez-Ripoll, M., Lequerica, J. L. & Polaina, J. (1998). *J. Mol. Biol.* **275**, 491-502.
- Seshadri, S., Akiyama, T., Opassiri, R., Kuaprasert, B. & Ketudat Cairns J. R. (2009). *Plant Physiol.* **151**, 47-58.
- Tailford, L. E., Offen, W. A., Smith, N. L., Dumon, C., Morland, C., Gratien, J., Heck, M. P., Stick, R. V., Blériot, Y., Vasella, A., Gilbert, H. J. & Davies, G. J. (2008). *Nature Chem. Biol.* **4**, 306-312.
- Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022-1025.
- Verdoucq, L., Morinière, J., Bevan, D. R., Esen, A., Vasella, A., Henrissat, B. & Czjzek, M. (2004). *J. Biol. Chem.* **279**, 31796-31803.
- Vocadlo, D. J. & Davies, G. J. (2008). *Curr. Opin. Chem. Biol.* **12**, 539-555.
- Xu, Z., Escamilla-Treviño, L., Zeng, L., Lalgondar, M., Bevan, D., Winkel, B., Mohamed, A., Cheng, C.L., Shih, M.C., Poulton, J. & Esen, A. (2004). *Plant Mol. Biol.* **55**, 343-367.

Table 1 Data-collection and processing statistics.

Values in parentheses are for the outer shell.

	Native Os7BGlu26	Os7BGlu26/mannose complex
PDB code	4JHO	4JIE
Wavelength (Å)	1.00	1.00
Resolution range (Å)	30 – 2.20 (2.28 - 2.20)	30 – 2.45 (2.54 - 2.45)
Completeness (%)	99.9 (99.9)	95.8 (79.9)
Average redundancy per shell	5.5 (5.5)	6.1 (4.4)
$R_{\text{merge}}^{\dagger}$ (%)	8.8 (49.4)	14.9 (45.6)
$\{I/\sigma(I)\}$	18.0 (3.2)	11.9 (3.2)
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters (Å)	$a = 68.1, b = 71.7, c = 136.7$	$a = 68.0, b = 73.6, c = 134.0$
No. of unique reflections	34370	24116
No. of observed reflections	188040	147585
No. of molecules per ASU	1	1
R_{factor} (%)	17.6	15.3
$R_{\text{free}}^{\dagger\dagger}$ (%)	21.9	19.5
No. of protein atoms	3955	3960
No. of water molecules	305	234
No. of ligand atoms	None	12
No. of non solvent hetero atoms	51	51
r.m.s.d. bonds (Å)	0.010	0.010
r.m.s.d. angles (°)	1.155	1.212
Mean B -factor		
- Protein	22.9	23.7
- non solvent hetero atom	45.0	46.9
- solvent	32.8	30.8

- D-mannose	None	31.4
Ramachandran plot (%)		
Most favored	88.5	88.0
Allowed region	11.3	11.7
Outlier region	0.2	0.2

$$^{\dagger} R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}.$$

$^{\dagger\dagger} R_{\text{free}}$ represents the residual factor calculated from approximately 5% of the data that was not used in the refinement.

Table 2 Kinetic parameters of wild type Os7BGlu26 and its mutants using the 4NPMAN and 4NPGlc substrates.

	substrate	K_m (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_m (s ⁻¹ mM ⁻¹)	Mutant/wild type k_{cat}/K_m ratio	$\Delta\Delta G_{\text{s}^*\text{mut}}$ (kJmol ⁻¹)
Wild type	4NPMAN	0.48 ± 0.003	0.35 ± 0.004	0.714		
	4NPGlc	0.124 ± 0.002	0.029 ± 0.0001	0.237		
E179Q	4NPMAN	0.40 ± 0.003	0.008 ± 0.0004	0.02	0.028	+9.0
	4NPGlc	1.07 ± 0.02	0.0102 ± 0.0004	0.0095	0.040	+8.1
Y134W	4NPMAN	2.37 ± 0.04	2.4 ± 0.09	1.01	1.4	-0.9
	4NPGlc	0.032 ± 0.002	0.0306 ± 0.0006	0.959	4.0	-3.5
C182T	4NPMAN	12.4 ± 0.78	0.51 ± 0.03	0.043	0.060	+7.1
	4NPGlc	3.7 ± 0.14	0.078 ± 0.005	0.021	0.089	+6.1
Y134W/C182T	4NPMAN	11.56 ± 1.70	0.30 ± 0.03	0.026	0.036	+8.3
	4NPGlc	2.95 ± 0.17	0.058 ± 0.005	0.0198	0.084	+6.3
Y134F	4NPMAN	0.45 ± 0.02	12.5 ± 0.6	27.7	39	-9.2
	4NPGlc	0.167 ± 0.009	0.127 ± 0.003	0.759	3.2	-2.9
C182A	4NPMAN	2.46 ± 0.23	1.02 ± 0.07	0.414	0.58	+1.4
	4NPGlc	1.39 ± 0.11	0.242 ± 0.016	0.174	0.73	+0.8

$$\Delta\Delta G_{\text{s}^*\text{mut}} = -RT[\ln(k_{\text{cat}}/K_m)_{\text{mutant}} - \ln(k_{\text{cat}}/K_m)_{\text{wildtype}}] \text{ (Fersht } et al., 1987).$$

Figure 1 A proposed pyranose ring itinerary of β -D-glucosides (A) and β -D-mannosides (B) during the glycosylation step of hydrolytic pathway (Davies *et al.*, 2003; Vocadlo and Davies, 2008).

Figure 2 Multiple sequence alignment of rice Os7BGlu26, HvBII, LeMside, AtBGlu44, Os3BGlu7, Os3BGlu6 and Os4BGlu12. The amino acid sequences were aligned with *ClustalW*, and the secondary structure of Os7BGlu26 was aligned on the top of the alignment with the *ESPrpt* program (Gouet *et al.*, 2003). Stars indicate the catalytic acid/base and nucleophilic residues, while black arrowheads mark the amino acid residues mutated in this study. The Genbank accession codes for the sequences are: Os7BGlu26, ACF35791; HvBII, AAA87339; LeMside, AAL37714; AtBGLU44, Q9LV33; Os3BGlu7, AC091670; Os3BGlu6, AC146619; and Os4BGlu12, AAAA02014151.

Figure 3 Structure of rice Os7BGlu26 β -D-mannosidase. (A) Cartoon representation of the overall structure of rice Os7BGlu26. The catalytic nucleophile and acid/base residues are represented as red sticks (E389 and E179), glycerol and HEPES are represented as green and cyan ball-and-sticks, respectively. Loops A, B, C and D are indicated in red, green, orange and magenta, respectively. The N-terminus is indicated in black. (B) Amino acid residues around the -1 subsite with contacts mediated through hydrogen bonds cover distances between 2.4 and 3.1 Å, and are indicated as black dashed lines. (C) Amino acid residues interacting with the HEPES molecule in two alternate positions in the active site. The contacts through hydrogen bonds between the protein and HEPES are indicated as black dashed lines.

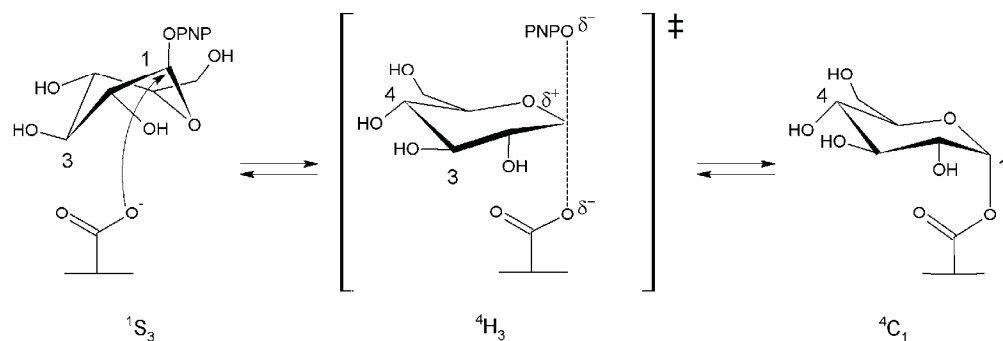
Figure 4 Comparison of the active site clefts of the rice GH1 structures, Os7BGlu26, Os3BGlu7 (PDB code: 2RGL), Os3BGlu6 (PDB code: 3GNO) and Os4BGlu12 (PDB code: 3PTK). The dashed lines indicate the width and breadth of the substrate binding cleft at the entrance to the active site, as described in the main text. *PyMol* was used to generate qualitative vacuum electrostatic charges to color the surfaces with red for negatively and blue for positively charged areas.

Figure 5 The structure of rice Os7BGlu26 in complex with β -D-mannose. (A) The An unbiased $F_o - F_c$ (OMIT) map of β -D-mannose is represented as a blue mesh contoured at 3σ . (B) Amino acid residues interacting with β -D-mannose in the active site. The contacts through hydrogen bonds between the protein and β -D-mannose are indicated by black dashed lines, while β -D-mannose is shown in cyan ball-and-stick representation. The superimposition of the Os7BGlu26/ β -D-mannose complex (orange) over the native Os7BGlu26 structure (blue) is shown in (C). The superimposition of the Os7BGlu26/ β -D-mannose complex (carbons in blue) with the Os3BGlu7 cellopentaose complex (carbons in violet) is shown in (D). β -D-mannoside, HEPES and cellopentaose are indicated in ball-and-stick representation with carbons in cyan, blue and green, respectively. The glucosyl residue-binding subsites observed in the Os3BGlu7 cellopentaose complex are marked by -1, +1, +2, +3, and +4.

Figure 6 Structures of Os7BGlu26 in complex with 4NPGlc (A.) and 4NPMAN (B.) obtained by docking. The distance between the carboxylic hydrogen atom of the catalytic acid/base residue (E179) and the glycosidic oxygen is 3.08 Å for the glucoside and 3.58 Å for the mannoside. The distance between the Oε1 of the nucleophile (E389) and the anomeric carbon is 2.98 Å for the glucoside and 3.0 Å for the mannoside. The binding energies of different conformers of 4NPMAN and 4NPGlc in complex with Os7BGlu26 are shown in C.

Figure 7 Comparison of the active sites of the rice Os7BGlu26 β-D-mannosidase and other rice GH1 β-D-glucosidases with bound 2-deoxy-2-fluoroglucoside (G2F) moieties. (A) Superimposition of Os7BGlu26 with Os3BGlu7 (PDB code: 2RGM), both of which have significant β-D-mannosidase activity. (B) Superimposition of Os7BGlu26 with Os3BGlu6 (PDB code: 3GNR) and Os4BGlu12 (PDB code: 3PTM), which do not have significant β-D-mannosidase activity, in complex with G2F. The structures are shown in stick representations with carbons colored blue in Os7BGlu26, orange in Os3BGlu7, green in Os3BGlu6 and magenta in Os4BGlu12. The inhibitors are presented in ball-and-stick representations in the same colors as the corresponding protein.

A.



B.

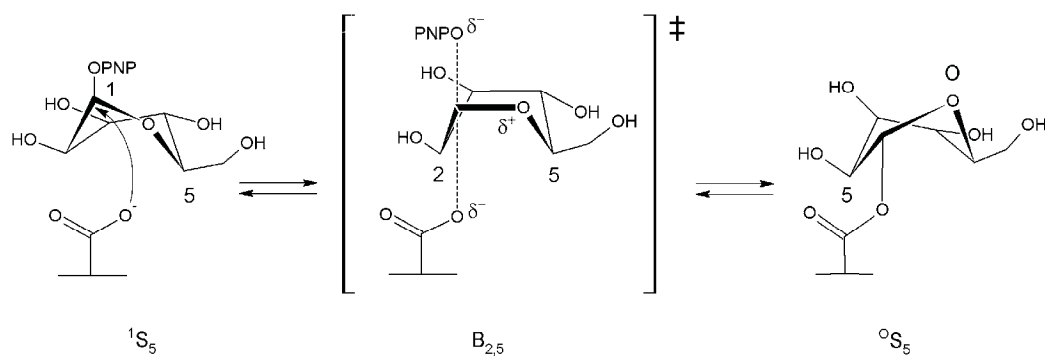


Figure 1 A proposed pyranose ring itinerary of β -D-glucosides (A) and β -D-mannosides (B) during the glycosylation step of hydrolytic pathway (Davies *et al.*, 2003; Vocadlo and Davies, 2008).

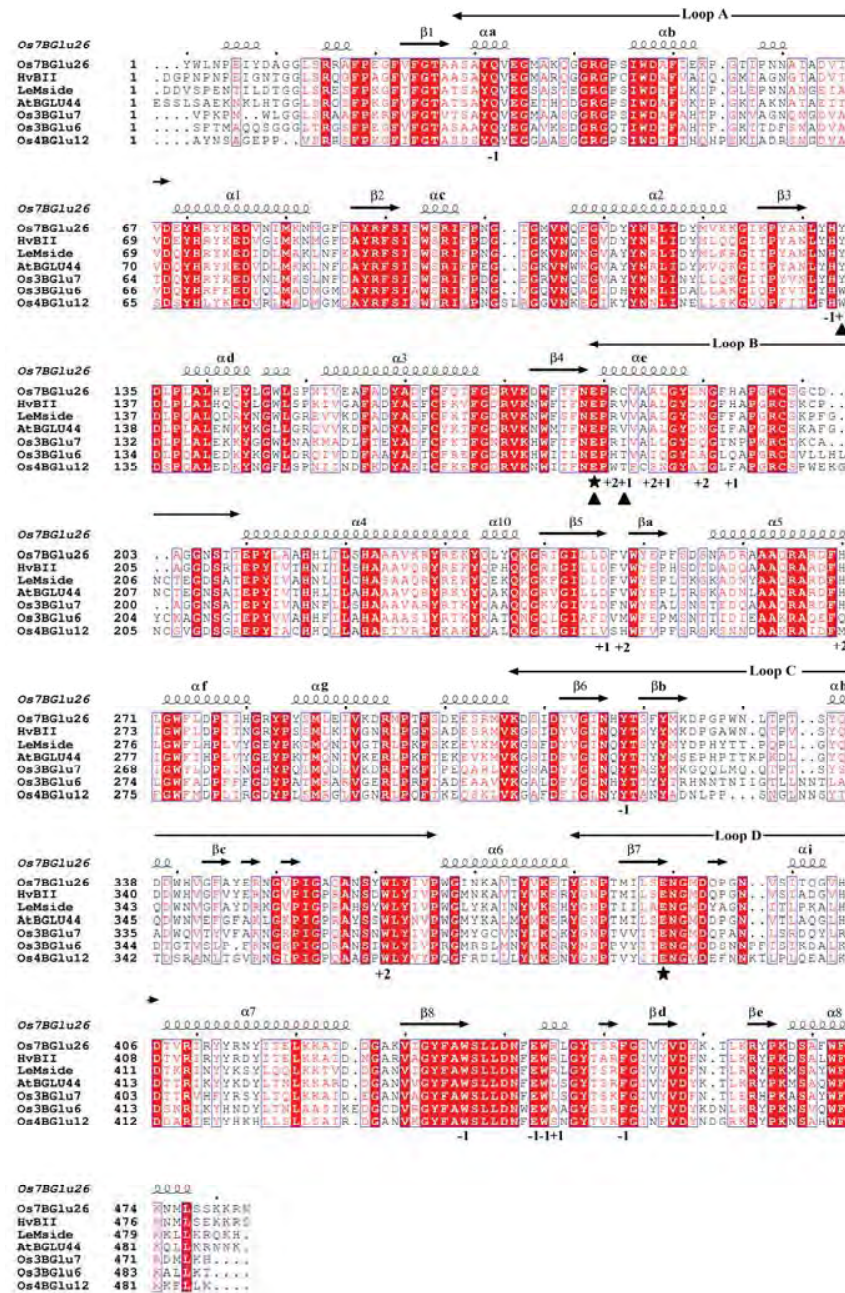


Figure 2 Multiple sequence alignment of rice Os7BGlut26, HvBII, LeMside, AtBGLU44, Os3BGLu7, Os3BGLu6 and Os4BGLu12. Amino acid sequences were aligned with *ClustalW*, and the secondary structure of Os7BGlut26 was represented on the top of the alignment with the *ESPrpt* program (Gouet et al., 2003). Stars indicate catalytic acid/base and nucleophilic residues, and black arrowheads mark the amino acid residues mutated in this study. The GenBank entry accessions are: Os7BGlut26, ACF35791; HvBII, AAA87339; LeMside, AAL37714; AtBGLU44, Q9LV33; Os3BGLu7, AC091670; Os3BGLu6, AC146619; and Os4BGLu12, AAAA02014151.

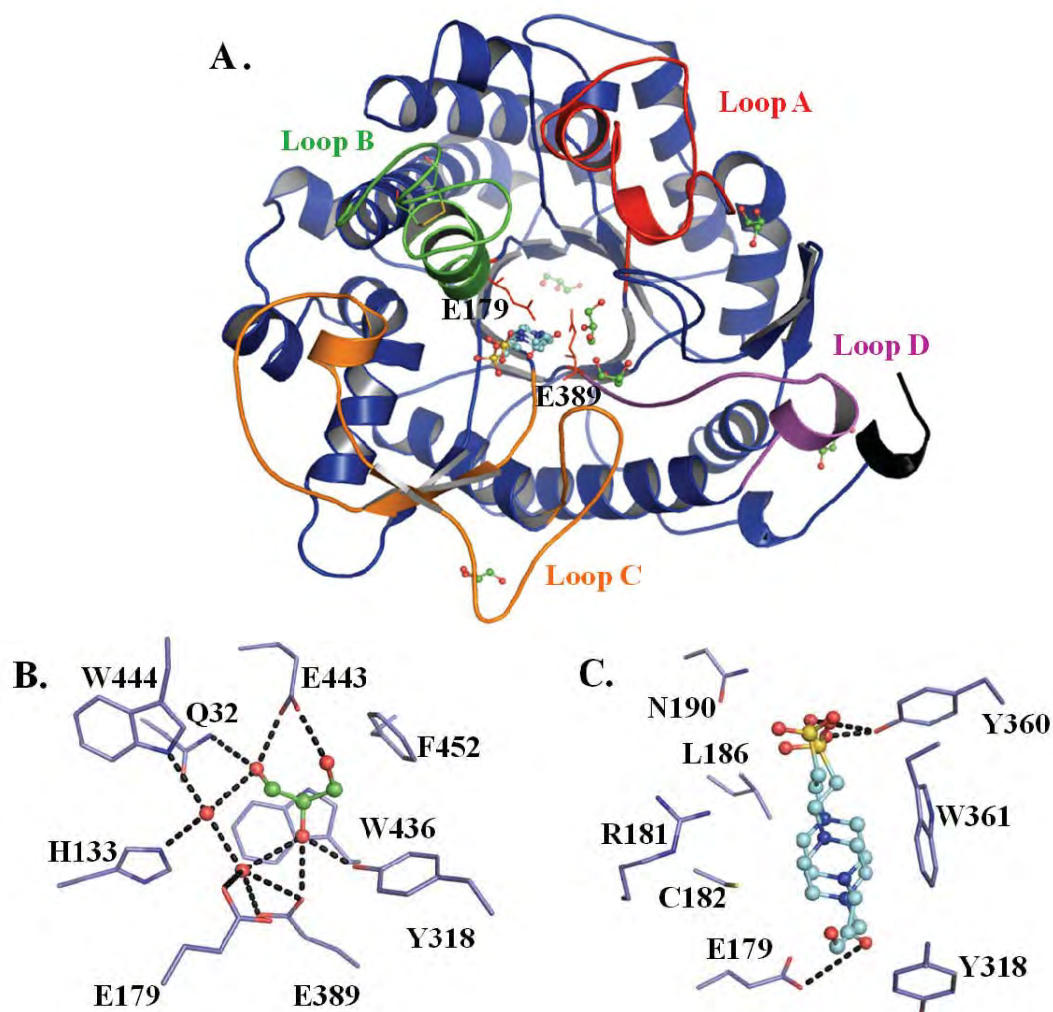


Figure 3 Structure of rice Os7BGlu26 β -D-mannosidase. (A) Cartoon representation of the overall structure of rice Os7BGlu26. The catalytic nucleophile and acid/base residues are represented as red sticks (E389 and E179), glycerol and HEPES are represented as green and cyan ball-and-sticks, respectively. Loops A, B, C and D are indicated in red, green, orange and magenta, respectively. The N-terminus is indicated in black. (B) Amino acid residues around the -1 subsite with contacts mediated through hydrogen bonds cover distances between 2.4 and 3.1 Å, and are indicated as black dashed lines. (C) Amino acid residues interacting with the HEPES molecule in two alternate positions in the active site. The contacts through hydrogen bonds between the protein and HEPES are indicated as black dashed lines.

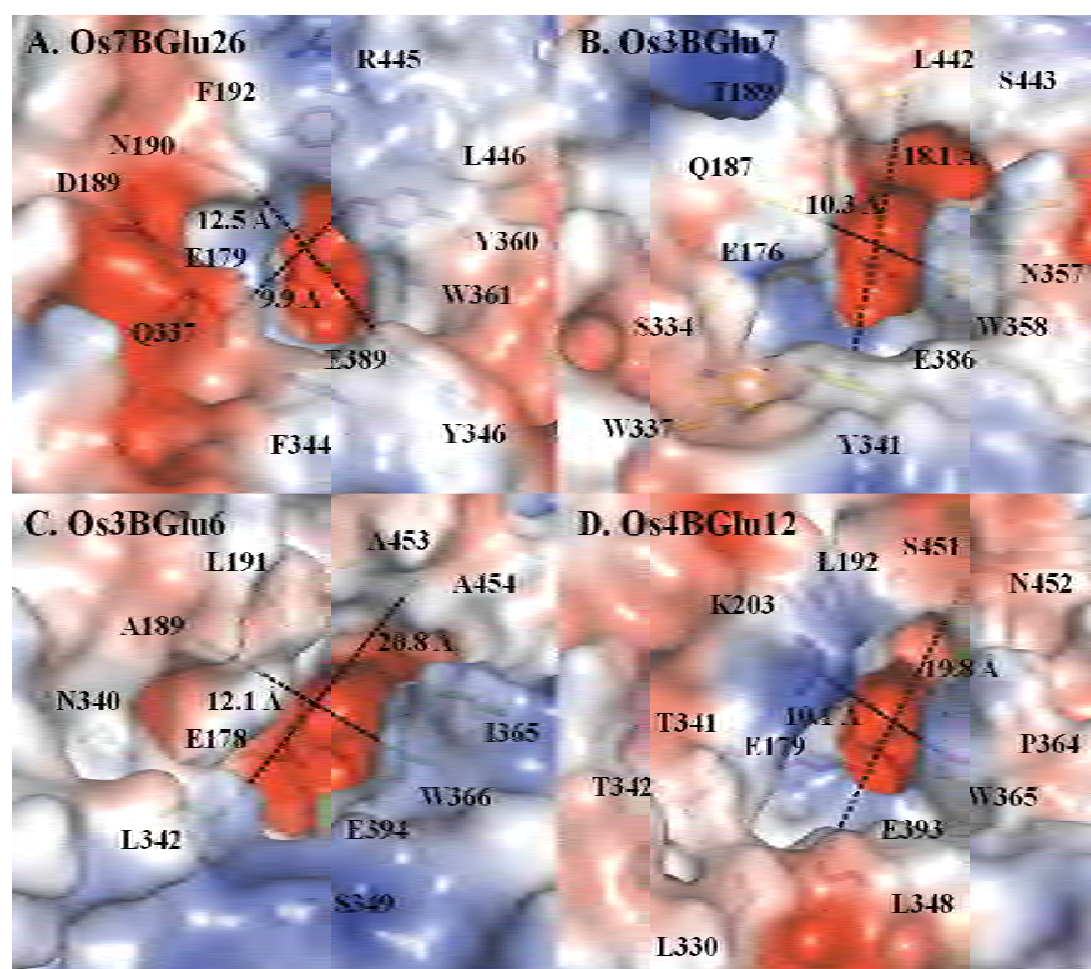


Figure 4 Comparison of the active site clefts of the rice GH1 structures, Os7BGlu26, OS3BGlu7 (PDB code: 2RGL), Os3BGlu6 (PDB code: 3GNO) and Os4BGlu12 (PDB code: 3PTK). The dashed arrows indicate the width and breadth of the substrate binding cleft at the entrance to the active site, as described in the main text. *PyMol* (Schrödinger LLC) was used to generate qualitative vacuum electrostatic charges to color the surfaces with red for negatively and blue for positively charged areas.

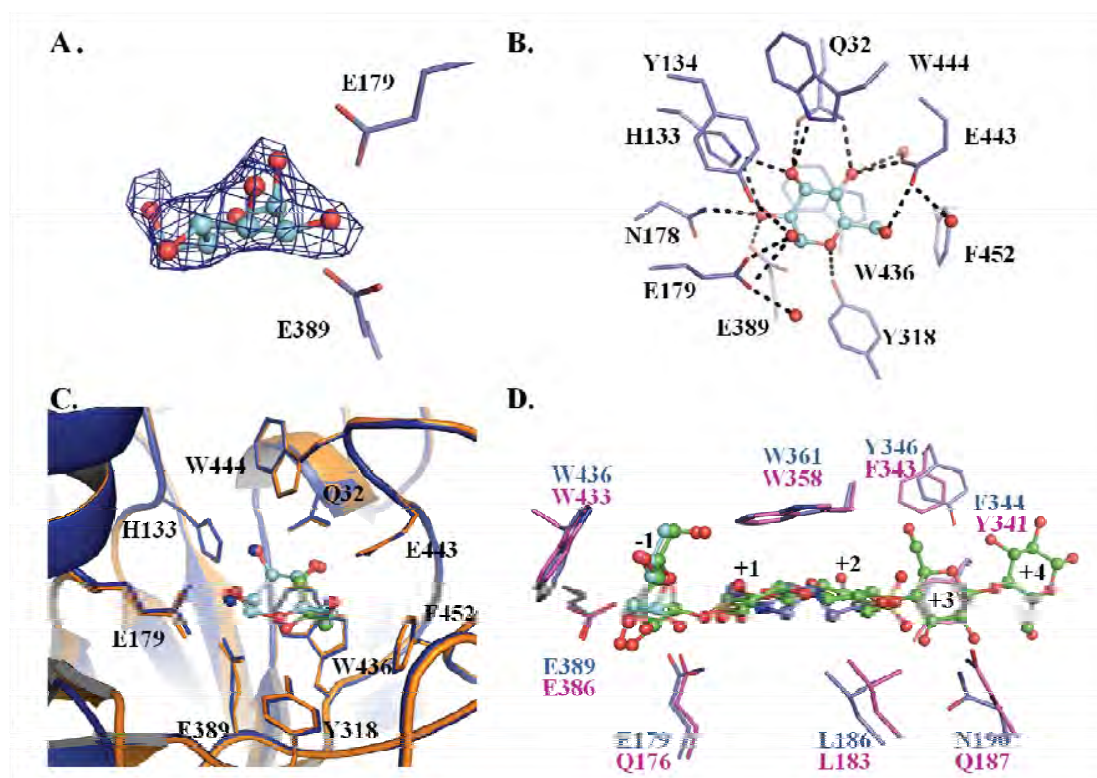


Figure 5 The structure of rice Os7BGlu26 in complex with β -D-mannose. (A) An unbiased $F_o - F_c$ (omit) map of β -D-mannose is represented as a blue mesh contoured at 3σ . (B) Amino acid residues interacting with β -D-mannose in the active site. The contacts through hydrogen bonds between the protein and β -D-mannose are indicated by black dashed lines, while β -D-mannose is shown in cyan ball-and-stick representation. The superimposition of the Os7BGlu26/ β -D-mannose complex (orange) over the native Os7BGlu26 structure (blue) is shown in (C). The superimposition of the Os7BGlu26/ β -D-mannose complex (carbons in blue) with the cellopentaose/Os3BGlu7 complex (carbons in violet) is shown in (D). β -D-mannoside, HEPES and cellopentaose are indicated in ball-and-stick representations with carbons in cyan, blue and green, respectively. The glucosyl residue-binding subsites observed in the cellopentaose complex are marked by -1, +1, +2, +3, and +4.

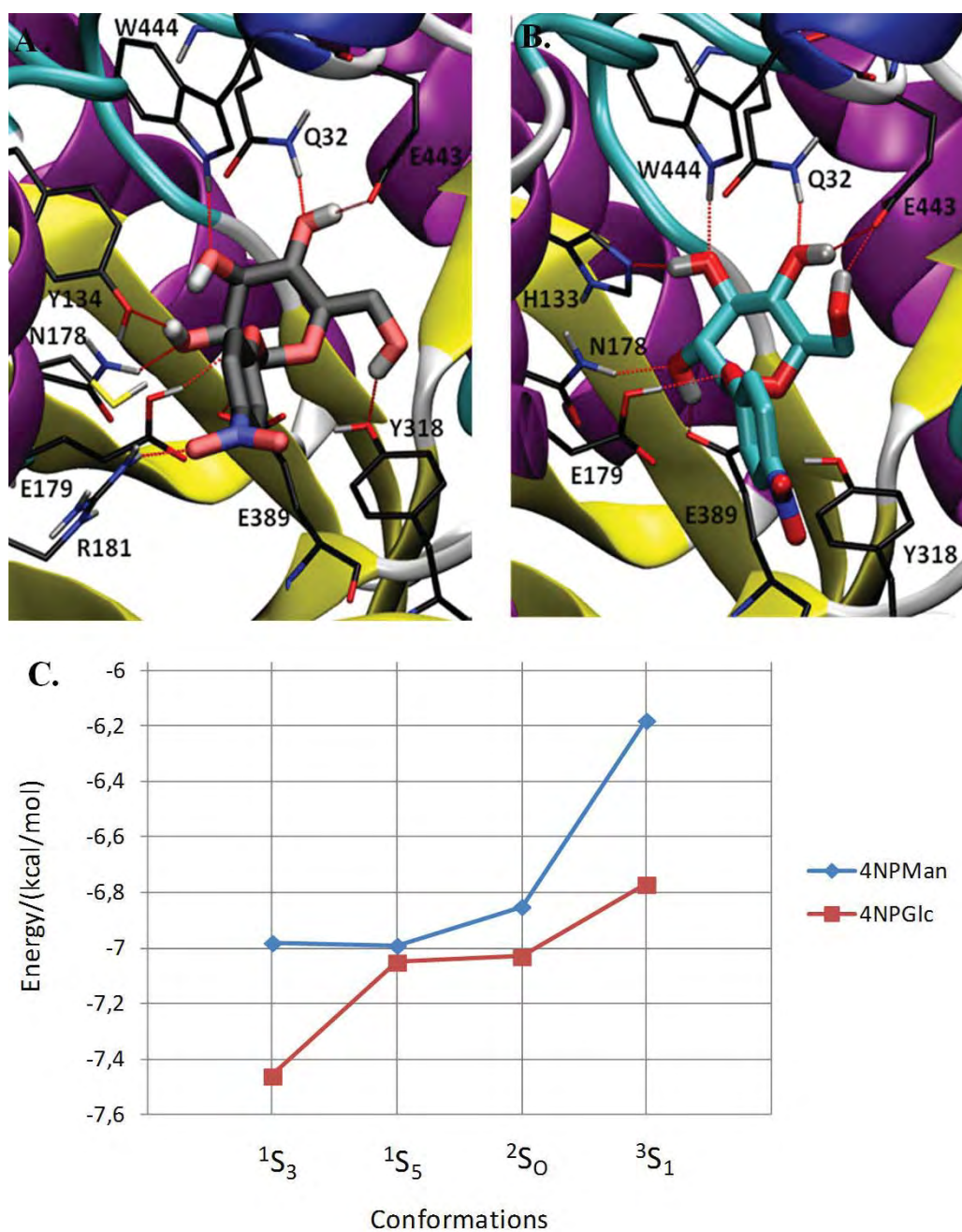


Figure 6 Structures of Os7BGlu26 in complex with 4NPGlc (A.) and 4NPMAN (B.) obtained by docking. The distance between the carboxylic hydrogen atom of the catalytic acid/base residue (E179) and the glycosidic oxygen is 3.08 Å the glucoside and 3.58 Å for the mannoside. The distance between the O ϵ 1 of the nucleophile (E389) and the anomeric carbon is 2.98 Å for the glucoside and 3.0 Å for the mannoside. The binding energies of different conformers of 4NPMAN and 4NPGlc in complex with Os7BGlu26 are shown in C.

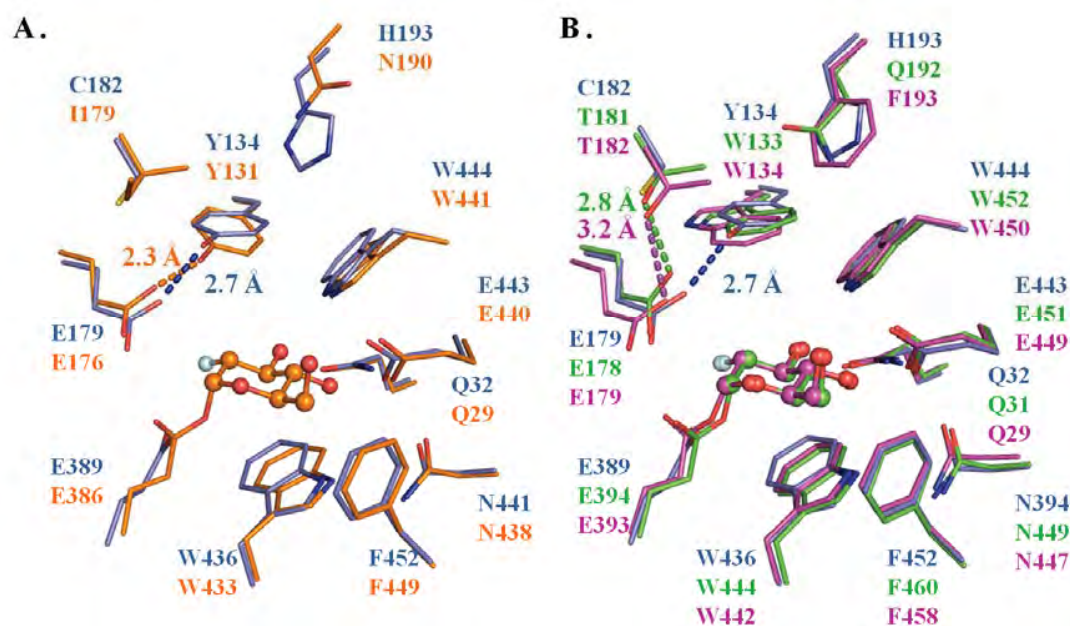


Figure 7 Comparison of the active sites of the rice Os7BGlu26 β -D-mannosidase and other rice GH1 β -D-glucosidases with bound 2-deoxy-2-fluoroglucoside (G2F) moieties. (A) Superimposition of Os7BGlu26 with Os3BGlu7 (PDB code: 2RGM), both of which have significant β -D-mannosidase activity. (B) Superimposition of Os7BGlu26 with Os3BGlu6 (PDB code: 3GNR) and Os4BGlu12 (PDB code: 3PTM), which do not have significant β -D-mannosidase activity, in complex with G2F. The structures are shown in stick representations with carbons colored blue in Os7BGlu26, orange in Os3BGlu7, green in Os3BGlu6 and magenta in Os4BGlu12. The inhibitors are presented in ball-and-stick representations in the same colors as the corresponding protein.

Appendix A.

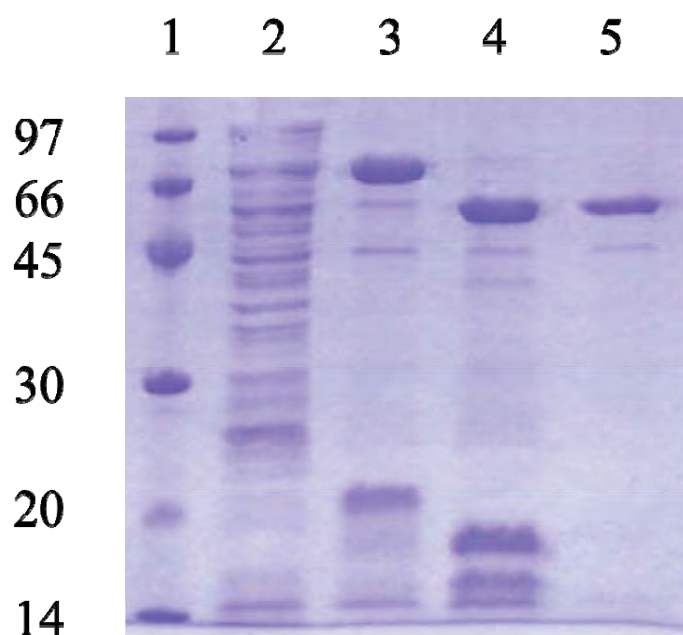


Figure S1. SDS-PAGE of Os7BGlu26 throughout purification. Lane 1, Bio-Rad low molecular-mass markers (masses shown at left in kDa); Lane 2, crude extract of soluble proteins from *E. coli* cells; Lane 3, the N-terminal thioredoxin/His₆-tagged Os7BGlu26 fusion protein after the first IMAC step; Lane 4, the products of digestion of the thioredoxin/His₆-tagged Os7BGlu26 protein by enterokinase; Lane 5, purified Os7BGlu26 after the second IMAC step.

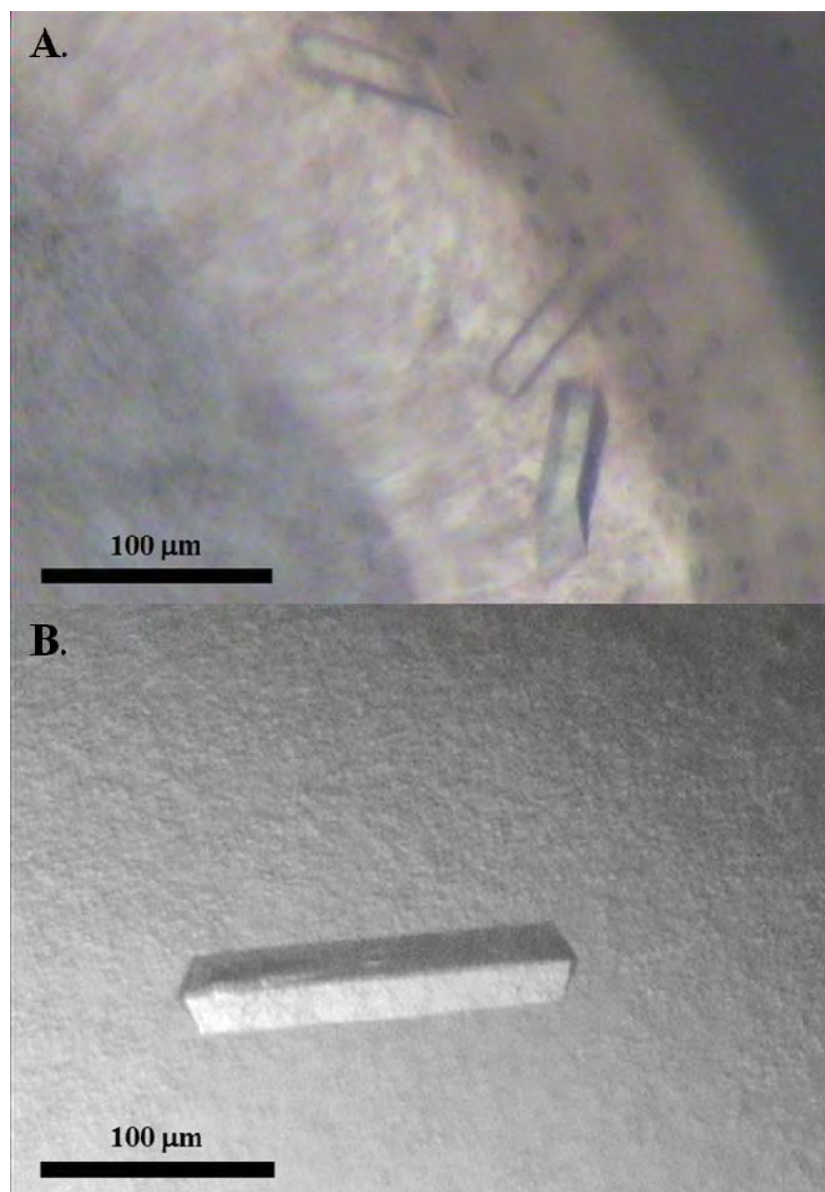


Figure S2. The Os7BGlu26 crystals. (A), Crystals in 0.1 M sodium HEPES, pH 7.5, containing 0.8 M Na,K tartrate grown in microbatch.under oil. (B), A crystal grown after optimization in a hanging-drop with 0.1 M sodium HEPES, pH 7.25 containing 0.58 M Na,K tartrate.

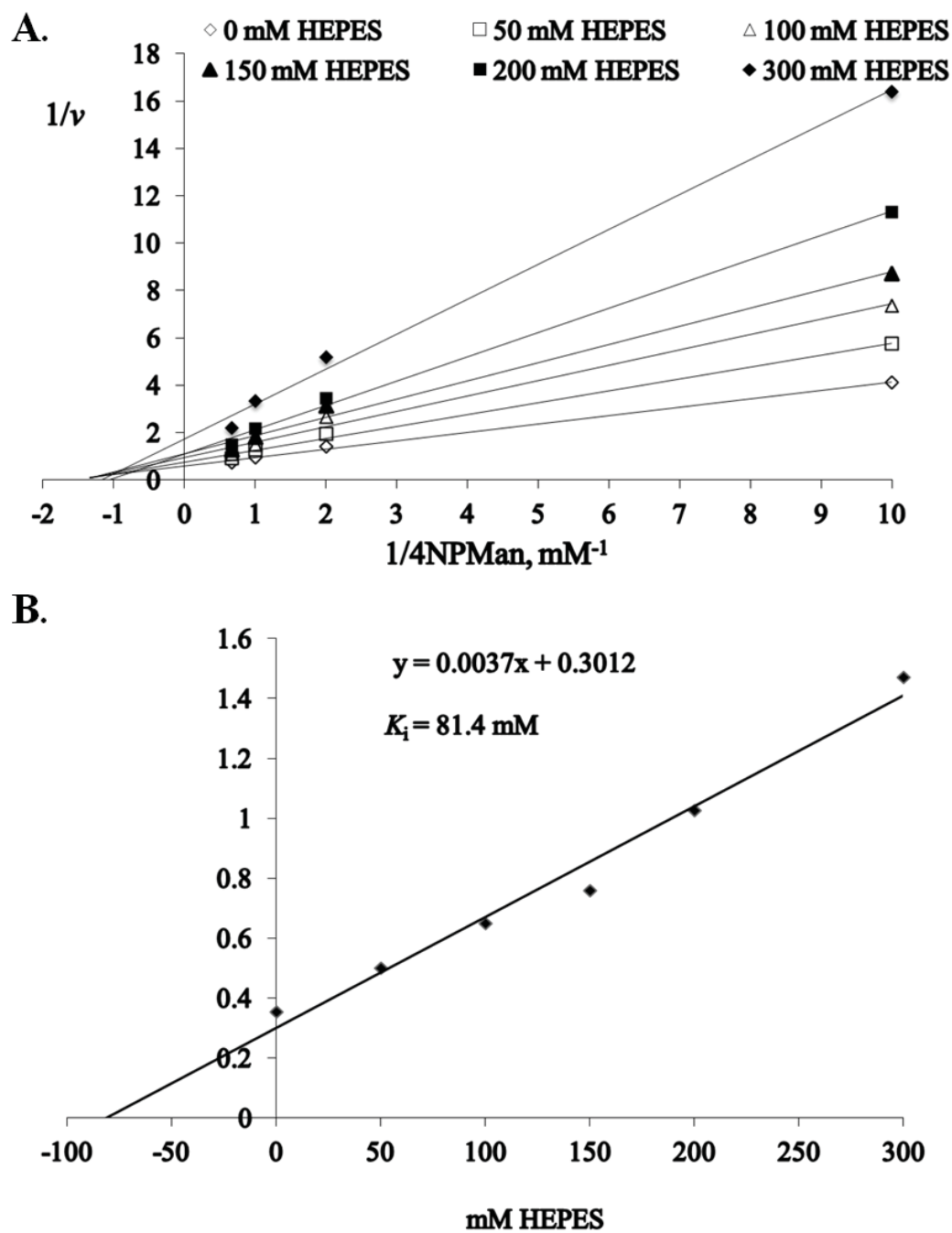
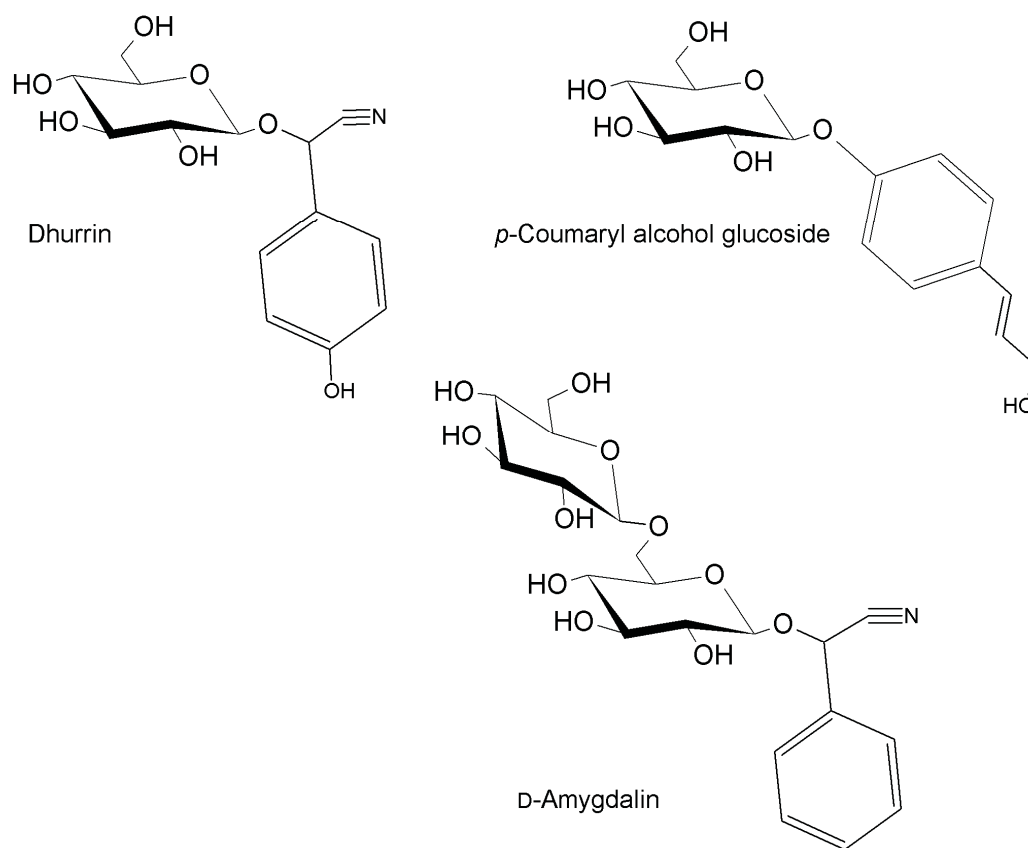


Figure S3. Noncompetitive inhibition of Os7BGlu26 by HEPES. (A) The $1/S$ vs $1/v$ plot (Lineweaver-Burk plots). (B) To determine K_i values, the slopes of inhibitions obtained from Lineweaver-Burk plots were plotted against the HEPES concentration and the data were subjected to linear regression ($r^2=0.979$).



Appendix B. Chemical structures of natural substrates hydrolyzed by Os7BGlu26.