รายงานวิจัยฉบับสมบูรณ์

Final Report

โครงการ การวิเคราะห์หา การผลิต และการทำวิศวกรรมเอนไซม์สำหรับการใช้ชีวมวล และการผลิตผลิตภัณฑ์ไกลโคไซด์

Identification, Production and Engineering of Enzymes for Biomass Utilization and Production of Glycoside Products

โดย ศ.ดร.เจมส์ เกตุทัต-คาร์นส์ และคณะ

Prof. Dr. James Ketudat-Cairns and colleagues

เมษายน 2565

April, 2022

สัญญาเลขที่ RSA6280073

รายงานวิจัยฉบับสมบูรณ์

Final Report

โครงการ การวิเคราะห์หา การผลิต และการทำวิศวกรรมเอนไซม์สำหรับการใช้ชีวมวล และการผลิตผลิตภัณฑ์ไกลโคไซด์

Identification, Production and Engineering of Enzymes for Biomass Utilization and Production of Glycoside Products

ศ.ดร.เจมส์ เกตุทัต-คาร์นส์ มหาวิทยาลัยเทคโนโลยีสุรนารี

Prof. Dr. James Ketudat-Cairns, Suranaree University of Technology

ดร.สลิลา เพ็งไธสง มหาวิทยาลัยเทคโนโลยีสุรนารี

Dr. Salila Pengthaisong, Suranaree University of Technology

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัย สถาบันวิจัยแสงซินโครตรอน (องค์การมหาชน) และ มหาวิทยาลัยเทคโนโลยีสุรนารี

Abstract (บทคัดย่อ)

This project investigated the utilization of carbohydrate, including polysaccharide residues from plant cell walls, which are the largest source of biomass in the world. We aimed to utilize mutants of *Tx*GH116 ß-glucosidase or other enzymes for direct generation of precursors for chemical synthesis of useful products and develop an efficient recycling system for cost-effective production of these compounds; to express and characterize ß-1,4-endoglucanases, cellobiosidases and ß-glucosidase from thermophilic and halophilic organisms and test their use in biomass degradation; to identify specific ß-xylanases and ß-xylosidase/α-L-arabinosidase enzymes for breakdown of grass-derived hemicellulose (arabinoxylan), such as rice straw; and test chemical modification, immobilization and engineering enzymes as a means of improving their utilization. For this work, all the enzymes were expressed in recombinant Escherichia coli or Pichia pastoris systems and purified by immobilized metal affinity chromatography (IMAC) and other protein column chromatography techniques. The *Tx*GH116 variants with nucleophile and acid/base mutated to various amino acids were tested for structure determination, and the nucleophile mutants were evaluated for glycosynthase activity and transglycosylation to produce α-azidoglucose. The upscale reaction was used to synthesize 280 mg of α-azidoglucose, which was used to synthesize effective human lysosomal α-glucosidase inhibitors. Moreover, mutations of all substrate-binding residues in the *Tx*GH116 active site were evaluated for catalytic characteristics and the structures of some of these enzymes determined to evaluate the importance of these residues and their mutability for enzyme engineering. Other family GH116 enzymes, along with GH1 and GH3 plant enzymes that had previously been expressed were evaluated for their activities and structures in the aim of engineering to generate new products from agricultural waste biomass. Furthermore, endoglucanases from 3 families, one family GH7 cellobiosidase, and a bacterial ß-glucosidase were evaluated for action on plant polysaccharides and pretreated biomass. The activity of a rice ß-xylosidase was evaluated and its action in generation of novel glycosides determined. To complement this activity for the production of products from biomass xylans, two bacterial ß-xylanases were produced and evaluated. Evaluation of immobilization for reuse of TxGH116 and its mutants demonstrated that it could allow reuse, although some mutants gradually lost activity. On the other hand, chemical modification to allow stability in ionic and organic solvents did not improve the activity or stability of the *Tx*GH116 ß-glucosidase, since its unmodified form is already resistant to these solvents. This

work generated 7 international journal publications, with others in process, while contributing to the training of 4 PhD students, a postdoc and a research assistant. The enzymes and knowledge generated in this work can be applied to generation of new products in the future.

Keywords : β-glucosidase, cellulase, xylanase, biomass conversion, agricultural waste
(คำหลัก)

บทคัดย่อ

โครงการวิจัยนี้ ทำการศึกษาเกี่ยวกับการใช้คาร์โบไฮเดรต ได้แก่ โพลิแซ็กคาไรด์จากผนังเซลล์พืชซึ่งเป็นแหล่งชีวมวลที่ใหญ่ที่สุดในโลก เรามีจุดประสงค์เพื่อใช้การกลายพันธุ์ของเอนไซม์ TxGH116 β-glucosidase หรือเอนไซม์อื่น ๆ สำหรับสร้างสารตั้งต้นในการสังเคราะห์ทางเคมีเพื่อให้ได้ผลิตภัณฑ์ที่มีประโยชน์และพัฒนาระบบรีไซเคิลที่มีประสิทธิภาพสำหรับการผลิตสารประกอบเหล่านี้ที่คุ้มค่า เพื่อทำการแสดงออกของยีนและทดสอบลักษณะของเอนไซม์ β-1,4-endoglucanases, cellobiosidases และ β-glucosidases จากจุลินทรีย์ที่เป็นเทอร์โมฟิลิกและฮาโลฟิลิก และทดสอบการย่อยสลายสารชีวมวล เพื่อวิเคราะห์หาเอนไซม์ β-xylanases และ β-xylosidase/α-L-arabinosidase ที่จำเพาะสำหรับการย่อยสลายของเฮมิเซลลูโลส (arabinoxylan) เช่น ฟางข้าว และทำการดัดแปลงทางเคมี การตรึง และเอนไซม์ทางวิศวกรรมเพื่อเพิ่มประสิทธิภาพการใช้ประโยชน์ของเอนไซม์เหล่านี้ สำหรับงานวิจัยนี้ เอนไซม์ทั้งหมดถูกแสดงออกในระบบรีคอมบิแนนท์ของเชื้อแบคทีเรีย Escherichia coli หรือยีสต์ Pichia pastoris และถูกทำให้บริสุทธิ์โดย immobilized metal affinity chromatography (IMAC) และเทคนิคโปรตีนคอลัมน์โครมาโตกราฟีอื่น ๆ เอนไซม์ TxGH116 กลายพันธุ์ที่ตำแหน่งกรดอะมิโนนิวคลีโอไฟล์และกรด/เบสไปเป็นกรดอะมิโนชนิดต่าง ๆ ได้รับการตรวจหาโครงสร้าง และตัวกลายพันธุ์ของนิวคลีโอไฟล์ถูกทดสอบการทำงานของไกลโคซินเทสและทรานส์ไกลโคซิเลชันเพื่อผลิต α-azidoglucose ขนาดของปฏิกิริยาการสังเคราะห์ถูกเพิ่มขนาดขึ้นเพื่อให้ได้ α-azidoglucose เป็น 280 มิลลิกรัม เพื่อใช้เป็นสารตั้งต้นในการสังเคราะห์ตัวยับยั้งการทำงานของ lysosomal α-glucosidase ของมนุษย์ที่มีประสิทธิผล นอกจากนี้ เอนไซม์กลายพันธุ์ของ TxGH116 ที่ตำแหน่งกรดอะมิโนทั้งหมดที่จับกับซับสเตรตในบริเวณแอกทีฟไซต์ได้รับการทดสอบการเร่งปฏิกิริยาและศึกษาโครงสร้างของเอนไซม์กลายพันธุ์บางตัวเพื่อประเมินความสำคัญของกรดอะมิโนเหล่านี้และการ กลายพันธุ์สำหรับการทำวิศวกรรมเอนไซม์ เอนไซม์ในตระกูล GH116 ชนิดอื่น ๆ รวมทั้งเอนไซม์ตระกูล GH1 และ GH3 จากพืชที่เคยทำการแสดงออกมาก่อนหน้านี้ได้รับการทดสอบการทำงานและหาโครงสร้างโดยมีวัตถุประสงค์ทางวิศวกรรมเอนไซม์เพื่อสร้างผลิตภัณฑ์ใหม่จากชีวมวลของเสียทางการเกษตร นอกจากนี้ endoglucanases จาก 3 ตระกูล โดยเฉพาะในตระกูล GH7 ได้แก่ cellobiosidase และ β-glucosidase จากแบคทีเรีย ได้รับการตรวจสอบการทำงานต่อโพลิแซ็กคาไรด์ของพืชและสารชีวมวลที่ผ่านการบำบัดแล้ว การทำงานของ β-xylosidase จากข้าว ได้รับการวิเคราะห์และตรวจสอบการทำงานในการสร้างไกลโคไซด์ชนิดใหม่ เพื่อเพิ่มประสิทธิภาพในการผลิตผลิตภัณฑ์จากไซแลนจากชีวมวล ได้มีการผลิตและตรวจสอบการทำงานของเอนไซม์ β-xylanase 2 ชนิดจากแบคทีเรีย การตรวจสอบการตรึงเอนไซม์และตัวกลายพันธุ์ของ TxGH116 เพื่อนำกลับมาใช้ใหม่ แสดงให้เห็นว่าสามารถนำมาใช้ซ้ำได้แต่การกลายพันธุ์บางตัวจะค่อยๆ สูญเสียความสามารถในการทำงานไป ในทางกลับกัน การดัดแปลงทางเคมีเพื่อให้เกิดความเสถียรในตัวทำละลายไอออนิกและอินทรีย์ไม่ได้เพิ่มความสามารถในการทำงานหรือความเสถียรของ TxGH116 β-glucosidase เนื่องจากเอนไซม์ในรูปแบบที่ยังไม่ได้ดัดแปลงมีความทนทานต่อตัวทำละลายเหล่านี้อยู่แล้ว จากงานวิจัยนี้ สามารถสร้างผลงานตีพิมพ์ในวารสารนานาชาติ 7 ฉบับ และยังมีผลงานที่อยู่ในระหว่างดำเนินการ

รวมทั้ง สามารถพัฒนาทักษะงานวิจัยของนักศึกษาปริญญาเอก 4 คน นักวิจัยหลังปริญญาเอก และผู้ช่วยวิจัย เอนไซม์และองค์ความรู้ที่สร้างขึ้นในงานวิจัยนี้สามารถนำไปประยุกต์ใช้กับผลิตภัณฑ์ใหม่ ๆ ได้ในอนาคต

Project Code (รหัสโครงการ): RSA6280073

Project Title: Identification, Production and Engineering of Enzymes for Biomass Utilization and Production of Glycoside Products

ชื่อโครงการ : โครงการ การวิเคราะห์หา การผลิต และการทำวิศวกรรมเอนไซม์สำหรับการใช้ชีวมวล และการผลิตผลิตภัณฑ์ไกลโคไซด์

Investigators: Professor Dr. James R. Ketudat-Cairns, Dr. Salila Pengthaisong

ชื่อนักวิจัย : ศ.ดร.เจมส์ เกตุทัต-คาร์นส์ ดร.สลิลา เพ็งไธสง

E-mail Address : cairns@sut.ac.th

Project Period (ระยะเวลาโครงการ) : 30 April, 2019 – 29 April, 2022

Objectives (วัตถุประสงค์)

The purpose of this study is to develop effective enzymes for the breakdown and utilization of lignocellulosic biomass, such as that obtained from agricultural waste. Specific goals are:

1. To utilize mutants of *Tx*GH116 β-glucosidase or other enzymes for direct generation of precursors for chemical synthesis of useful products and develop an efficient recycling system for cost-effective production of these compounds. This includes production of more efficient transglycosidases for product development.

2. To identify, express and characterize β-1,4-endoglucanase, cellobiosidase and β-glucosidase from thermophilic and halophilic organisms for testing in biomass degradation.

3. To identify specific β-xylanases and β-xylosidase/α-L-arabinosidase enzymes for breakdown of grass-derived hemicellulose (arabinoxylan), such as rice straw.

4. To modify these enzymes or the ones previously produced in our laboratory for use in minimally diluted ionic liquids after pretreatment of cellulose or biomass with these ionic liquids to improve cellulose degradation. Other approaches in enzyme engineering will also be tried to improve cellulosic biomass degradation.

Methods of Research (วิธีทดลอง)

1. *Tx*GH116 ß-glucosidase mutants and their utilization.

TxGH116 ß-glucosidase and mutants were produced and purified as previously published (Charoenwattanasatien et al., 2016). New mutations were made in the cDNA in the pET30a expression vector by QuikChange mutagenesis (Strategene, Agilent Technologies, USA). The protein was purified from crude soluble extracts of cells after heating at 50 or 60 °C, depending on the stability of the variant. For structural studies, after the first immobilized metal affinity chromatography step, the enzyme was digested with enterokinase to remove the tag and purified over S200 gel filtration size exclusion chromatography. The mutant proteins included those with mutations in catalytic acid/base and nucleophile mutants for testing activity and glycosyltransfer activity, glycone sugar-binding residues to assess their importance to catalysis and inhibition, residues interacting with the catalytic residues and active site gateway residues to assess effects on activity and inhibition. The x-ray crystal structures of mutants of interest were solved alone and with glucose and other ligands by the previously described methods (Charoenwattanasatien et al., 2016).

The catalytic nucleophile mutants TxGH116 E441A (previously characterized), E441D, E441G, E441Q and E441S were assessed for relative hydrolysis rates and then for transglycosylation with various alpha-fluoroglycoside donors and pNP-glycoside and oligosaccharide acceptors. The structures of E441A and E441G were solved in complexes with alpha-fluoroglucoside, cellobiose, and cellotriose by soaking the crystals obtained as noted above with the corresponding ligands in cryosolution (in the case of E441G with alpha-fluoroglucoside, this resulted in crystals with oligosaccharides attached, instead of the substrate that was soaked into the crystals) (Pengthaisong et al., 2021).

Previously, we had solved the structure of the TxGH116 D593A and D593N mutants in complexes with oligosaccharides. In this project, the structures were further refined and finalized and the final structures sent to Carme Rovira in U. Barcelona for QMMM calculations to determine the reaction mechanism, including the trajectory of the catalytic acid/base and sugar residue during the reaction. The D593N mutant structure with cellobiose was mutated back to D593 and energy minimized to serve as a starting point for these computations.

To utilize the transglycosylation activity of the TxGH116 nucleophile mutants, the transglycosylation reaction between *p*-nitrophenyl (pNP) ß-D-glucoside (pNPGlc) and sodium azide was optimized for the E441G mutant (Gorantla et al., 2021). Cellobiose was also demonstrated to act as a glucosyl donor in this reaction, but less efficiently. The scaled-up reaction was done in 100 mL 100 mM sodium azide in 100 mM citrate-phosphate buffer (final pH 4.6) with 600 mg pNPGlc and 100 mg TxGH116 E441G for 24 h at 45 °C. Reactions were carefully set-up and opened in a fume hood to avoid hydroazidic acid release and the reaction solution pH was adjusted to 8.0 with 2 M sodium carbonate before purification to remove hydroazidic acid, which might otherwise volatilize to a toxic gas. The alpha-azido-glucose product was purified by LH20 column chromatography in water to yield 280 mg pure alpha-azido-glucose. Various alkynes with different functional groups were synthesized by coupling propargylamine or propargyl alcohol with carboxylic acids through the well-known standard EDC/DMAC coupling reactions or coupling hydroxyls or amines with propargyl bromide with base catalyst (Gorantla et al., 2021). The alpha-azido-glucose and alkynes were coupled by copper-catalyzed Click reactions, under conditions that were appropriate to each compound and purified by standard chromatographic methods (Gorantla et al., 2021). The structures were verified by nuclear magnetic resonance (NMR) and high resolution mass spectrometry (HRMS). The products were then tested for inhibition of human lysosomal alpha-glucosidase in extracts of COS-7 cells that were transiently infected with the vector to overexpress the **α**-glucosidase (Ngiwsara et al., 2019) with initial assessment at 1.26 mM alpha-glucosyl triazole, followed by IC_{50} analysis for the compounds showing greatest inhibition.

2. Characterization of novel glycolytic enzymes

Seven cellulolytic enzymes, one beta-glucanase and one putative cellobiosidase were expressed in *E. coli* and initial characterization conducted. A putative cellulose 1,4-beta-cellobiosidase from *Isoptericola halotolerans* was identified as a potential salt-tolerant enzyme and cloned into pET32a and pET30a plasmids, then expressed in Origami(DE3), Rosetta-gami(DE3) and BL21(DE3) cells, but was no active protein was detected. A putative endoglucanase from *Thermobispora bispora* was identified and its expression vector in pET32a produced. The two enzymes were tested for activity on p-nitrophenyl (pNP) beta-D-cellobioside, avicel cellulose, alpha-cellulose,

carboxymethyl cellulose (CMC) and barley beta-glucan. The enzyme was characterized for its activity alone and in mixtures with commercial cellulase. Additional potential cellulolytic enzymes that were cloned and expressed included the *Thermobifida halotolerans* enzymes ThGH6A (NCBI: AHN09982.1), ThGH6B (NCBI: WP_119267886.1), and ThGH1B (NCBI: WP_068693385.1), along with the *Thermothelomyces thermophilus* enzyme TtGH7A. We later expressed the gene for the putative *Streptomyces diastaticus* GH6 endoglucanase WP_102931505.1 (SdGH6) that had been cloned in the pET30 vector. Of these enzymes, ThGH6A, ThGH1B and TtGH7A could be expressed and purified by IMAC, but ThGH6B and SdGH6 gave very low if any protein expression and were not characterized further.

Finally, we cloned the codon optimized genes for the relatively well-characterized GH5 enzyme *Trichoderma reesei* endoglucanase EG-II (TrCel5A, NCBI: P07982.1) and a related endoglucanase from *Thermothielavioides terrestris* (TtCel5A, NCBI: XP_003656246.1) into pET30a and pET32a for expression in *E. coli* and pPICZaBNH8 for expression in *Pichia pastoris*. The plasmids were transformed into *E. coli* BL21(DE3) and Origami(DE3) or the SMD1168H strain of *P. pastoris* and expressed by standard procedures (Opassiri et al., 2003; Luang et al., 2010). Expression screening showed high activity for *E. coli* expression, so the proteins were purified from *E. coli* and characterized for hydrolysis of natural and synthetic substrate.

In general, the enzymes were tested for hydrolysis of *p*-nitrophenyl (pNP) ß-D-glucoside, pNP-ß-cellobioside, barley ß-glucan, cellulose, and acid and base-treated agricultural residues, including rice straw, rice husk and sugarcane leaves. The pH optimum was generally determined with 0.1% barley ß-glucan as substrate in citrate-phosphate buffers ranging from pH 3 to 9. Temperature optima were determined by testing activity at various temperatures at or near the optimal pH. Hydrolysis rates were evaluated by dinitrosalicylic acid (DNS) reducing sugar assay and glucose oxidase/peroxidase assay. Tolerance to alcohols, detergents and metal salts were also evaluated in the standard assay. Reaction products were evaluated by silica gel TLC.

Further characterization and structural studies were also conducted on *Thermosynechococcus elongatus* GH116 (TeGH116) β-glucosidase. Additionally, a plant GH116 enzyme AtGCD3 was cloned from an *Arabidopsis thaliana* cDNA clone from the RIKEN plant resource into a modified pET30a expression vector (from Rung-Yi Lai). The protein was expressed, purified and submitted to crystallization trials. The

activity was tested in standard assays with methylumbelliferyl beta-D-glucoside and 4NPGLc. Preliminary crystals were sent to the NSRRC synchrotron in Taiwan and data collected remotely from SLRI. The TeGH116 structure was refined to completeness alone and in complexes with glucose and 2-fluoroglucoside, with Refmac5 and coot, followed by validation on the PDB homepage, but the AtGCD3 data was not sufficient for structure determination. A structure of Os4BGlu18 monolignol ß-glucosidase was completed and used for docking of substrates as described by Baiya et al. (2021).

Further work was done to define the activity of the rice enzyme Os12BGlu31, for which mutations of its gene disrupt pollen development (Shim et al., 2022). The enzyme was expressed as a maltose-binding protein fusion protein in E. coli and in secreted form from pPICBαBNH8 in *P. pastoris* and tested for activity toward pNPGlc, cello- and laminari- oligosaccharides at various pH and temperatures.

We continued to characterize the rice family GH3 exoglucanase enzyme OsExoI and OsExoII. OsExoI was expressed from a codon-optimized gene from pPCZaBNH8 in *P. pastoris* SMD1168H, while OsExoII was expressed in *E. coli* Origami(DE3) from the pET32a vector. Due to low expression and proteolytic susceptibility of OsExoII, it was purified by IMAC, S200 gel filtration size exclusion and phenyl Sepharose chromatography without removal of the tag. The enzymes were tested for hydrolysis of soluble ß-glucans and for transglycosylation of various primary, secondary and tertiary alcohols. Products were evaluated by TLC and purified for structural verification by NMR on a Bruker Avance 500 spectrometer.

3. Characterization of xylanolytic enzymes

We continued to optimize the purification of the previously cloned OsXyl1 family GH3 ß-xylosidase expressed in *P. pastoris* and secreted into the medium by ion exchange, hydrophobic interaction and size exclusion chromatography. The enzyme was then tested for hydrolysis of pNP-ß-D-xylopyranoside (pNPXyl), pNP-α-L-arabinopyranoside, pNP-α-L-arabionofuranoside, ß-1,4-xylooligosaccharides and oat and beechwood xylan. The pNP glycoside hydrolysis was evaluated by alkalinization of the reaction with 50% volume 2 M sodium carbonate and measuring the absorbance at 405 nm (A405). Hydrolysis of xylooligosaccharides was evaluated with a xylose oxidase kit from Megazyme Corp., while ß-xylan hydrolysis was evaluated by DNS assay.

The OsXyl1 was further tested for transglycosylation of various alcohols with pNPXyl as a xylose donor. Acetone was used as a cosolvent for longer chain alcohols.

Genes encoding two family GH5 ß-xylanases from *Thermothelomyces heterothallicus* (ThXyln5A, NCBI: QBC65279.1) and the previously characterized *Thermomyces lanuginosus* (TlXyln5A, NCBI: O43097.1, PDB: 1YNA) without the putative signal peptides were synthesized and inserted in pET32a by Gene Universal. The vectors were transformed into BL21(DE3) and Origami(DE3) and tested for expression upon induction with various amounts of IPTG. After large scale culture in the appropriate host, the proteins were purified from the soluble fractions by IMAC and tested for hydrolysis of oat and beechwood xylans by detection of products by DNS reducing sugar assay and TLC. They were then tested for hydrolysis of pretreated rice straw and rice husk in the same manner.

4. Immobilization and chemical modification to improve enzyme usage

In order to economize on *Tx*GH116 enzyme and it mutants, the immobilized enzymes were made for increased recovery, reuse and stability. The *Tx*GH116 E441G and D593A with His-tag were immobilized with IMAC resin for transglucosylation to produce α- and β-glucosylazides for synthesis of α- and β-D-glucosyl-1,2.3-triazoles. The transglucosylation activities of the immobilized enzymes were optimized by vary protein, pNPGlc and NaN_3 concentrations, and characterized the stability and reuse of the immobilized enzymes, as well as, the transglucosylation reaction of the immobilized E441G was tested in large scale. Furthermore, the *Tx*GH116 and E441G with His-tag were immobilized on CNBr-activated Sepharose™ 4B. The immobilized enzymes were characterized pH and temperature optima, temperature stability and recycle.

In order to develop cellulose-degrading enzymes that tolerate ionic liquids, solvents and/or high temperatures, the *Tx*GH116 was modified by coupling aspartic and glutamic acid residues on external surface with N,N′-bis(2-aminoethyl)-1,3-propanediamine and N-(3-dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride, resulting [C-TxGH116]. Surfactant ethylene glycol ethoxylate lauryl ether was added to the C-*Tx*GH116 to yield the [C-TxGH116][S] nanoconstruct. The TxGH116 and [C-*Tx*GH116][S] were assayed hydrolysis of 4NPGlc in 100 mM sodium acetate, pH 5.5 and 0, 10, 30, 50 and 70% ionic liquid (1-butyl-1-methylpyrrolidinium

bis(trifluoromethylsulfonyl)imide) at 60, 70, 80 and 90°C for 15 min. The TxGH116 was assayed activity with β-cellulose and avicel (pretreated in 100% ionic liquid at 120°C for 3 h) in 100 mM NaOAc, pH 5.5 and 70% ionic liquid (1-butyl-1-methylpyrrolidinium bis(trifluoromethylsulfonyl)imide) at 60°C for 18 hr. Moreover, the hydrolysis activity of *Tx*GH116 was tested with 4NPGlc in 0.1 M sodium acetate, pH 5.5 and 0, 10, 20, 30 and 50% acetone at 60°C for 15 min. The hydrolysis activity of TxGH116 was also assayed in 10% acetone and 0, 10, 20, 30 and 40% octanol. Transglucosylation activity of 0.025 mg/ml *Tx*GH116 with 10 mM pNPGlc, 10% acetone and 20% octanol was assayed in 50 mM sodium phosphate buffer, pH 5.5. at 37°C for 18 hr and at 60°C for 15 min, 1, 2, 3, 4 and 5 hr. The products were analyzed by TLC.

Results (ผลการทดลอง)

1. Structure-function, engineering and application of GH116 ß-glucosidases

In the past, we expressed, characterized and solved the structure of TxGH116 b-glucosidase, the first for a glycoside hydrolase family 116 enzyme, which allowed reasonable modeling of other GH116 enzymes like human GBA2 (Sansenya et al., 2014; Charoenwattanasatien et al., 2016). Furthermore, we had previously demonstrated that its acid/base mutant, TxGH116 D593A, could be used to generate beta-azidoglucose, which could be used to generate beta-glucosidase inhibitors by click chemistry (Gorantla et al., 2019). During this work, we continued to study TxGH116 to study more about the unusual position of the catalytic acid/base residue, determine the roles and importance of the residues surrounding the glycon sugar in the active site (Huang et al., 2022), and characterize and utilize catalytic nucleophile mutants for glycosynthase and transglycosylation reactions (Pengthaisong et al., 2021; Gorantla et al., 2022).

For the TxGH116 acid/base mutants, we finalized the crystal structures of TxGH116 D593A alone and in complexes with cellobiose and laminaribiose and TxGH116 D593N alone and in complexes with 2-fluoroglucoside (covalent intermediate), cellobiose and laminaribiose. These structures were provided to Carme Rovira of the University of Barcelona, whose group computationally mutated the D593N cellobiose complex to the wildtype, energy minimized and calculated the free energy landscape of the substrate in the active site, along with the reaction trajectory (Figure 1). The reaction mechanism confirmed that the protonation of the glycosidic bond occurs via a perpendicular mechanism that is slightly to the cis side of the glycosidic bond oxygen, relative to the endocyclic oxygen of the glycon ring, and suggested the transfer of the proton occurs relatively late in the reaction course.

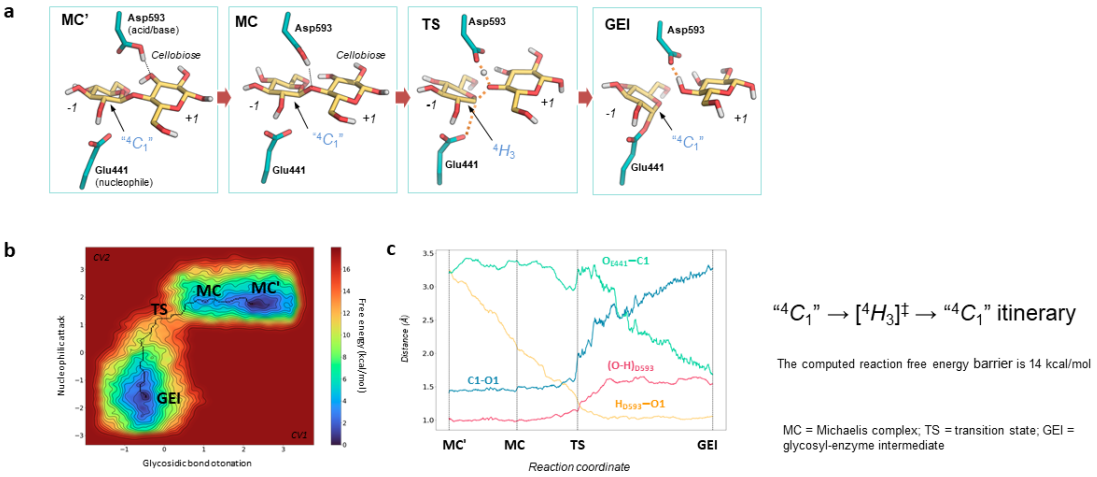
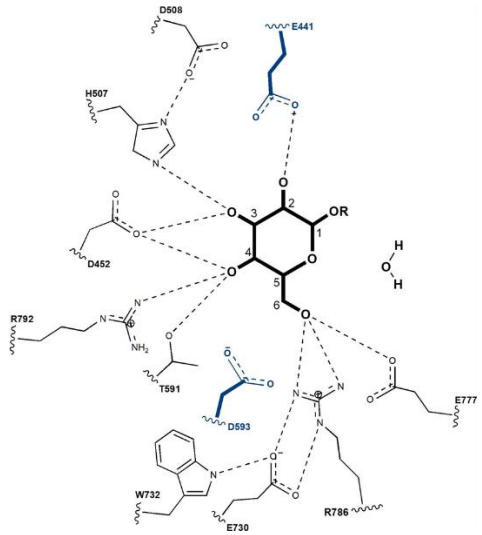


Figure 1. TxGH116 reaction mechanism for cellobiose hydrolysis. (a) Snapshots of the sugar and catalytic residues at critical points during the glycosylation half reaction of the hydrolysis of cellulose. MC' is preliminary Michaelis complex, MC is activated Michaelis complex, TS is transition state, and GEI is glycosyl enzyme intermediate formed at the end of the half reaction. (b) Free energy landscape for the reaction showing the relative energies of the different sugar shapes in the active site. (c) Lengths of bonds broken, formed or stretched during the reaction.

To get a better idea for the roles and importance of the residues making polar contacts to the sugar in the active site of TxGH116 (Figure 2) and test whether any could be modified to engineer the enzyme for lower glucose inhibition or other desired properties, each was mutated and the effects on enzyme activity characterized (Huang et al., 2022). The mutations that affected charges and charge environments were often found to shift the activity versus pH profile and sometimes the protein stability, as reflected in the optimum temperature (Figure 3). Table 1 shows that most of the mutations in residues making polar contacts with the glycon sugar caused drastic decreases in activity, as reflected by decreases in the specificity constants for pNPGlc (k_{cat}/K_m). This was sometimes a result of increased K_m, but often caused by decreased k_{cat}, except for residues interacting with the C6H_2OH group, for which mutations tended to increase k_{cat}.



Residue	Interactions
E777	H-bond to C6OH, along with R786, interacts with R792
R786	H-bonds to C6OH, Glc2 C2OH, Y445
W732	CH-pi interactions with C6, H-bond to E730
E730	2 H-bonds to R786 guanidinyl group
R792	H-bonds to C4OH, E777
T591	H-bonds to C4OH
D452	H-bonds to C3OH, C4OH
H507	H-bonds to C3OH

Figure 2. Amino acids that interact with the glucose glycon in the active site of *Tx*GH116. The amino acid sidechains are shown in lines and the glucose heavy atoms in bold with "OR" indicating the aglycone position. Putative hydrogen bonds are shown as dashed lines and the description of the roles of the amino acids based on the structure are given to the right.

To understand some of these results better, we solved the structures of TxGH116 solved the structures of the R786A, R786K, E730A and E730Q mutants in complexes with glucose. The structures with R786 mutations seemed to show that the entrance to the active site was larger, thereby allowing faster entrance of reactants and exit of products (Figure 4). For the E730A mutant, the active site sidechains looked similar to the wildtype (Figure 5), but its Km for pNPGlc increased 20-fold, increased k_{cat} 2-fold and the K_i for the mutant increased to 19 mM from 4.5 mM in wildtype. In contrast, the E730Q mutation displaced the Arg786 sidechain to disrupt glucose C6H_2OH binding (Figure 5) and increased its K_m for pNPGlc 40-fold (Table 1), but had no significant effect on glucose inhibition. As such, the static structures in the crystals were not immediately explain the kinetic behavior of the mutant enzymes, although the loss of interactions is likely to generate a less stable and more dynamic interaction with the substrate.

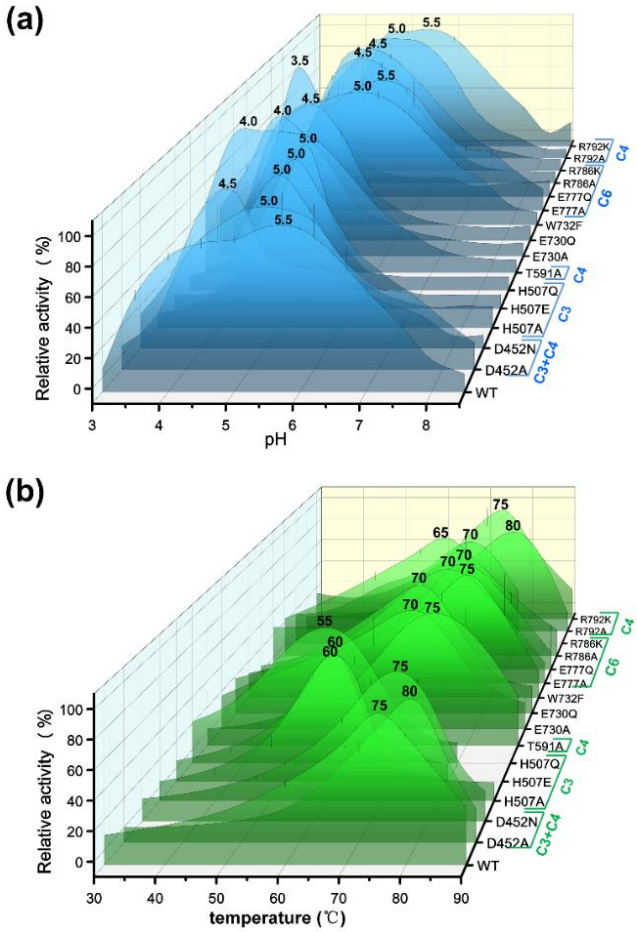
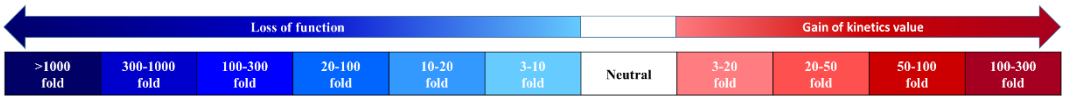


Figure 3. Optimal pH (a) and temperature (b) analysis of mutants and wild type (WT) *Tx*GH116. All assays were carried out in triplicate. The 3-D stacked graph was made by Origin 8.0 with Bézier curves. C3, C4, C6, and C3 + C4 represent the glucose carbons carrying hydroxyls interacting with the side chains of the amino acid residues that were mutated.

Table 1. Kinetics of TxGH116 glycon-binding mutants with p-nitrophenyl β-D-glucopyranoside (pNPGlc) substrate

Protein	K_M(mM)	k_{cat}(s^{-1})	k_{cat}/K_M ($mM^{-1}s^{-1}$)
WT	0.17±0.014	41.34±0.63	243.18
D452A	28.33±1.29	0.12±0.0021	0.0043
D452N	11.66±0.70	0.0081±0.00010	0.00069
H507A	23.01±0.71	4.72±0.051	0.21
H507E	9.93±0.62	0.010±0.00020	0.001
H507Q	33.30±4.30	12.20±0.63	0.37
T591A	12.61±0.45	3.63±0.037	0.29
E730A	3.79±0.25	80.81±1.38	21.32
E730Q	7.21±0.49	55.64±1.12	7.72
W732F	0.47±0.010	62.33±0.39	132.61
E777A	29.95±0.63	0.12±0.0010	0.0039
E777Q	6.97±0.22	0.02±0.00020	0.0028
R786A	7.89±0.57	171.08±3.96	21.68
R786K	7.30±0.71	88.85±2.70	12.17
R792A	4.53±0.38	0.0082±0.00010	0.0018
R792K	5.32±0.44	0.053±0.0013	0.0099



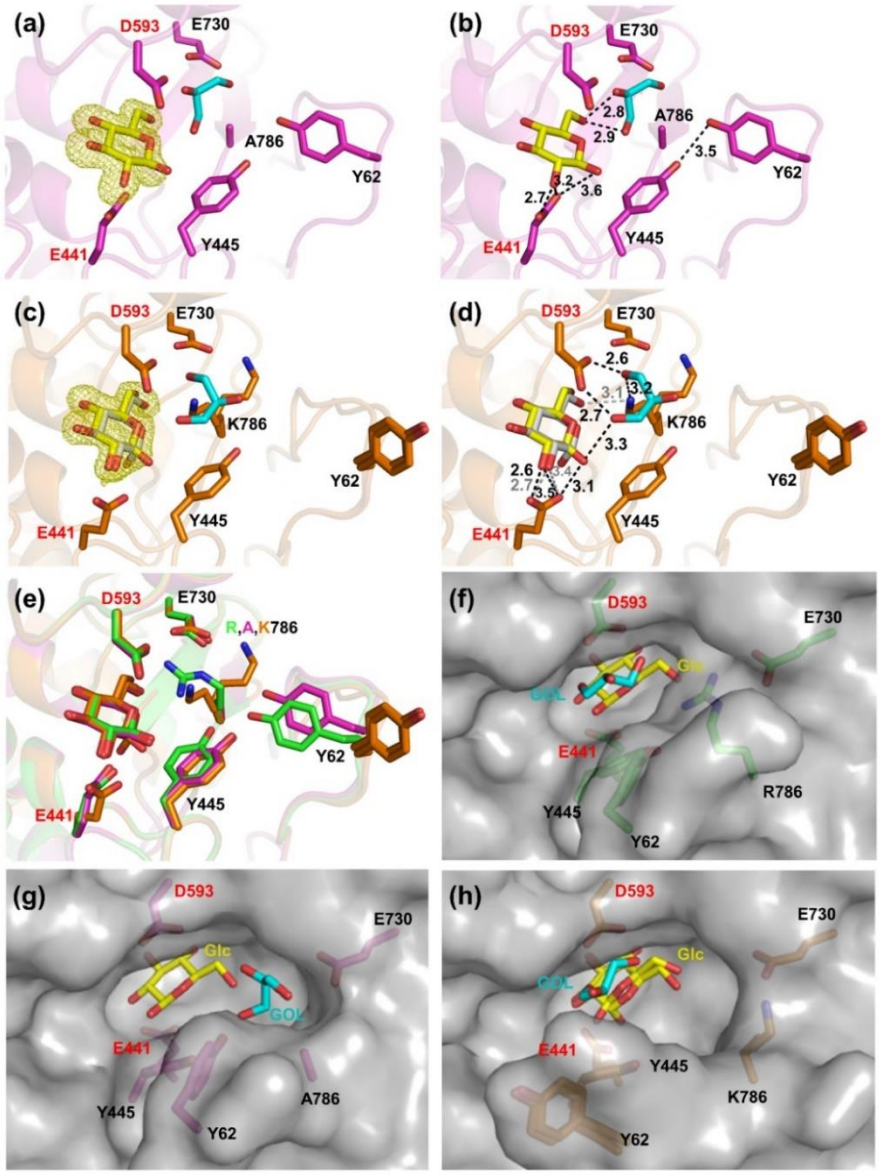


Figure 4. Active site comparison between *Tx*GH116 R786A and R786K and the wild type (WT) with glucose ligand. The glucose ligand weighted $|F_O\text{-}F_C|$ electron density maps calculated with the ligands omitted and contoured at 3**σ** are shown in the active sites of the glucose complexes of R786A (**a**) and R786K (**c**). Active site interactions with glucose and glycerol molecules are shown for R786A (**b**) and R786K (**d**), in which, with two anomers, glucose ligands were found (alpha 80%, in grey and beta 20%, in yellow). (**e**) Superposition of active site residues of WT (PDB: 5BX5, green), R786A (magenta), and R786K (orange). Slot-like active site entrance comparison between WT (**f**), R786A (**g**) and R786K (**h**). R786A and R786K have an enlarged entrance space, with an extra glycerol molecule in the place of the original arginine side chain. The potential hydrogen bonds between the glucose (Glc) ligands (yellow sticks) and surrounding residues are displayed as black dashed lines. Glycerol (GOL) molecules are shown as cyan sticks, and catalytic residues are labeled in red. From Huang et al. (2022).

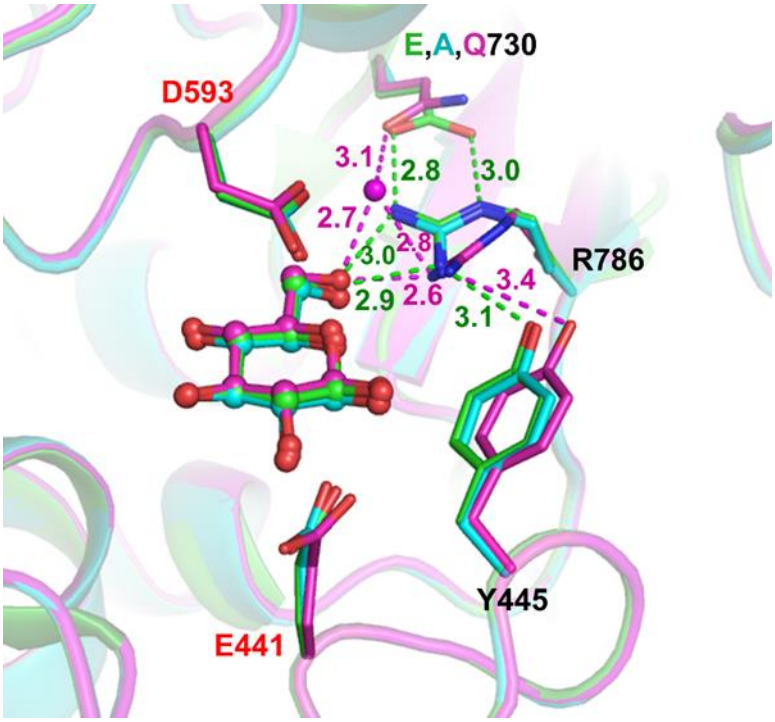


Figure 5. Structures of the active sites of TxGH116 E730A (cyan) and E730Q (magenta) with glucose ligand. The glucose ligand molecules are positioned similar to that in the wild type (green, PDB: 5BX5 [17]). One water molecule that takes the position near that vacated by the R786 residue guanidinyl group in E730Q mutant is shown as a magenta ball. Hydrogen bonds and close polar contacts are shown as dashed lines with distances marked (green for wild type and magenta for E730Q). From Huang et al. (2022).

One interesting aspect of the mutations of TxGH116 glucose C6-binding mutants is to look at the effect on sugar specificity, given that xylopyranoside lacks this functional arm and some GH115 enzymes act as ß-xylosidases. Another aspect is that ß-glucosidases and ß-galactosidase tend to bind the $C6H_2OH$ in a different position relative to the rest of the sugar, although *Tx*GH116 sugar binding looks more like that of ß-galactosidases, likely due to the position of the catalytic acid/base. When the TxGH116 R786A and R786K mutants were tested on pNPGlc, cellobiose, pNP-cellobiose, pNP-ß-D-galactopyranoside and pNP-ß-D-xylopyranoside, both mutants showed higher activity on xyloside compared to their parent enzyme, but relatively poorer activity on ß-galactoside compared to pNPGlc (Figure 6). Unexpectedly, the R786A mutant also performed better on pNP-cellobioside, probably due to wider entrance to the active site, since the first glucose must exit before the second glucose can be cut off to release p-nitrophenol, which is detected.

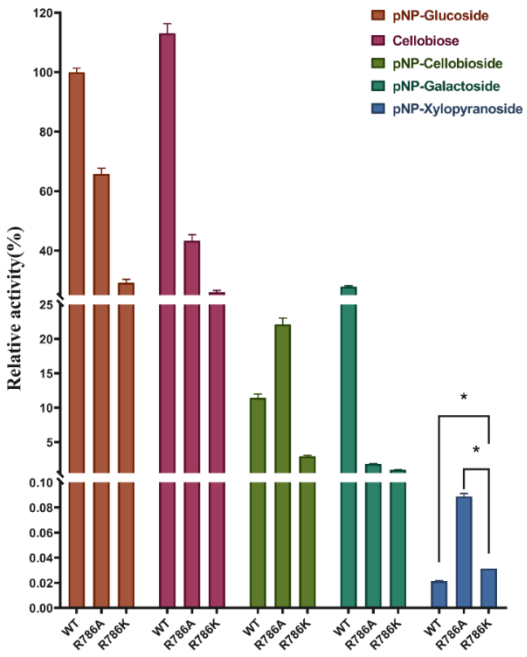


Figure 6. Relative activity on pNP-ß-D-glucopyranoside, cellobiose, pNP-ß-cellobioside, pNP-ß-D-galactopyranoside, and pNP-ß-D-xylopyranoside. The assay con-tained 1 mM substrate in 50 mM sodium acetate buffer, pH 5.5, at 60 °C. Relative activity is based on pNP released from pNP-ß-glycosides or glucose release from cellobiose. Data are expressed as means of three independent reactions ± SD. Stars represent one-way ANOVA significance at $p < 0.05$. From Huang et al. (2022).

Another aspect of work on *Tx*GH116 was to characterize nucleophile mutations effects on transglycosylation and glycosynthase activities, where the transglycosylation starts from the regular beta-configured substrate and results in an alpha-glucoside product, but glycosynthase reactions start from alpha-glucoside with a small aglycone (α-fluoroglucoside) to produce beta-configured products. In priciple, either can produce useful products, although we had a particular interest in α-azidoglucose from the transglycosylation reaction to use for production of α-glucosyl triazoles as potential α-glucosidase inhibitors by click chemistry.

In the absence of acceptor, the E441D, E441G, E441Q and E441S mutants exhibited 28,000-, 200,000-, 170,000- and 39,000- fold lower pNPGlc cleavage activity compared to wild type enzyme, respectively. The E441G and E441S mutants could be rescued with the small nucleophiles azide, formate and acetate, similar to E441A (Charoenwattanasatien et al., 2016), but E441D and E441E could not, likely due to steric hindrance. The E441A, E441D, E441G and E441S mutants also exhibited glycosynthase activity to transfer glucose from alpha-fluoroglucoside to cellobiose (Figure 7A). Since E441G and E441S had highest activity (Figure

8B), they were tested with various monosaccharide fluorides and pNP-glycosides as donors and acceptors and found to have broad specificity for which sugars they could use, as shown for E441G in Figure 8 C and D. Analysis of products and activity in the presence of cyclophellitol, which inhibits wildtype transferase activity indicated that the E441D mutant had a small amount of wild type transglycosylation contamination, but most of its activity came from glycosynthase activity, which was not previously reported for glycosidase nucleophile E to D mutants.

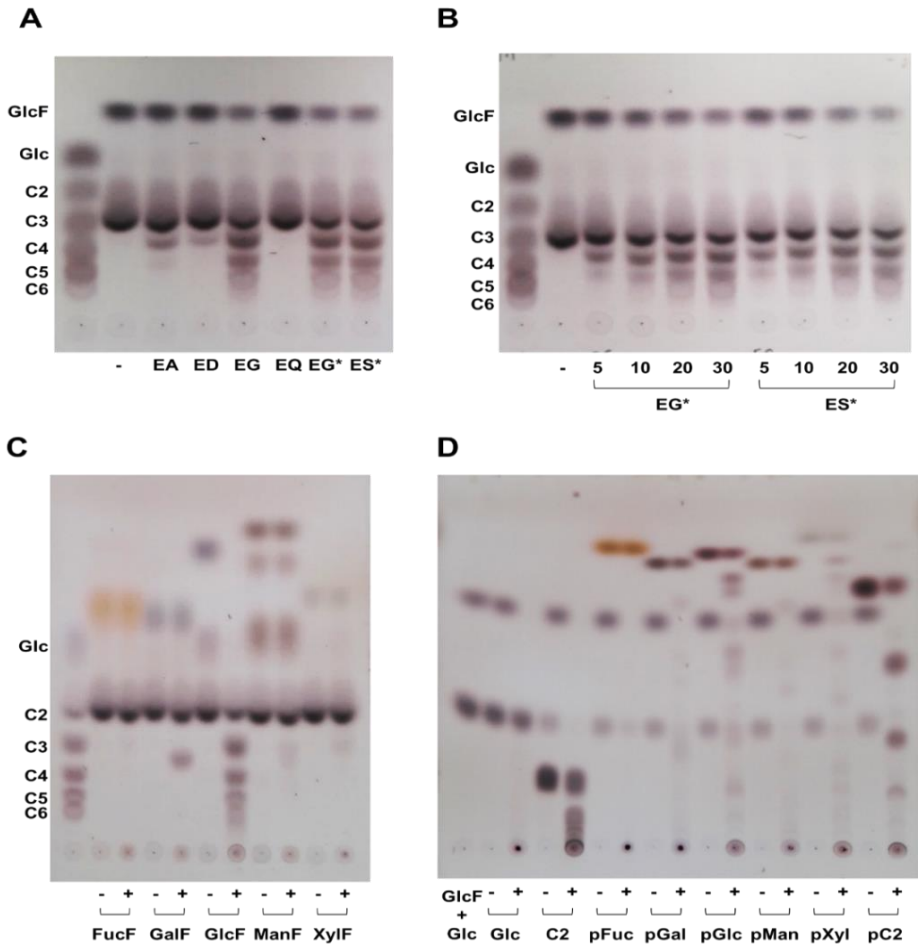


Figure 7. TLC detection of transglucosylation catalyzed by TxGH116 nucleophile mutants. (A) The 1 μM TxGH116 E441A, E441D, E441G, E441Q and E441S were incubated with 10 mM **α**-GlcF donor and 10 mM cellobiose acceptor in 150 mM sodium phosphate buffer pH 5.5 at 37°C for 30 min., and (B) vary incubation times: 5, 10, 20 and 30 min. Then, 40 μM E441G was incubated with 10 mM cellobiose and 10 mM donors (C) or 10 mM **α**-GlcF and 10 mM acceptors (D) at 37°C for 18 hr. The products of the reaction mixtures were monitored by silica gel TLC using ethyl acetate:acetic acid:water (2:1:1) as solvent for (A, B and C) or

EtOAc:MeOH:H2O (7:2.5:1) for (D). Plates were visualized by exposure to 10% sulfuric acid in ethanol followed by charring. EG* and ES* are E441G and E441S with N-terminal tag. Glc, C2, C3, C4, C5 and C6 are glucose, cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose standards. FucF, α-FucF; GalF, α-GalF; GlcF, α-GlcF; ManF, α-ManF and XylF, α-XylF are donors. Glc, glucose; C2, cellobiose; pFuc, 4NPFuc; pGal, 4NPGal; pGlc, 4NPGlc; pMan, 4NPMan; pXyl, 4NPXyl and pC2, 4NP-cellobiose are acceptors.

To investigate the structural basis for the TxGH116 glycosynthase activities, the structure of TxGH116 E441A and E441G with α-glucosyl fluoride (α-GlcF), cellobiose and cellotriose and E441A alone were determined by X- ray crystallography. The structure of α-GlcF in the E441A mutant showed the -1 subsite (glycon-binding site), but the structure of the E441G mutants soaked with α-GlcF showed cellotriose in one protein and cellotetraose in the other protein in the asymmetric unit, which evidently represented the autocondensation product, which could be seen as long oligosaccharides that precipitated in reactions in solution and sat at the application spot of the TLC. Hydrolysis of these compounds with endoglucanase generated cello-oligosaccharides, although a few ß-1,3-linked oligosaccharides also seemed to be present. These structures were similar to the structures of E441A and E441G soaked with cellobiose and cellotriose and in each case glycerol was found in the -1 subsite with the oligosaccharides starting from the +1 subsite (Figure 8).

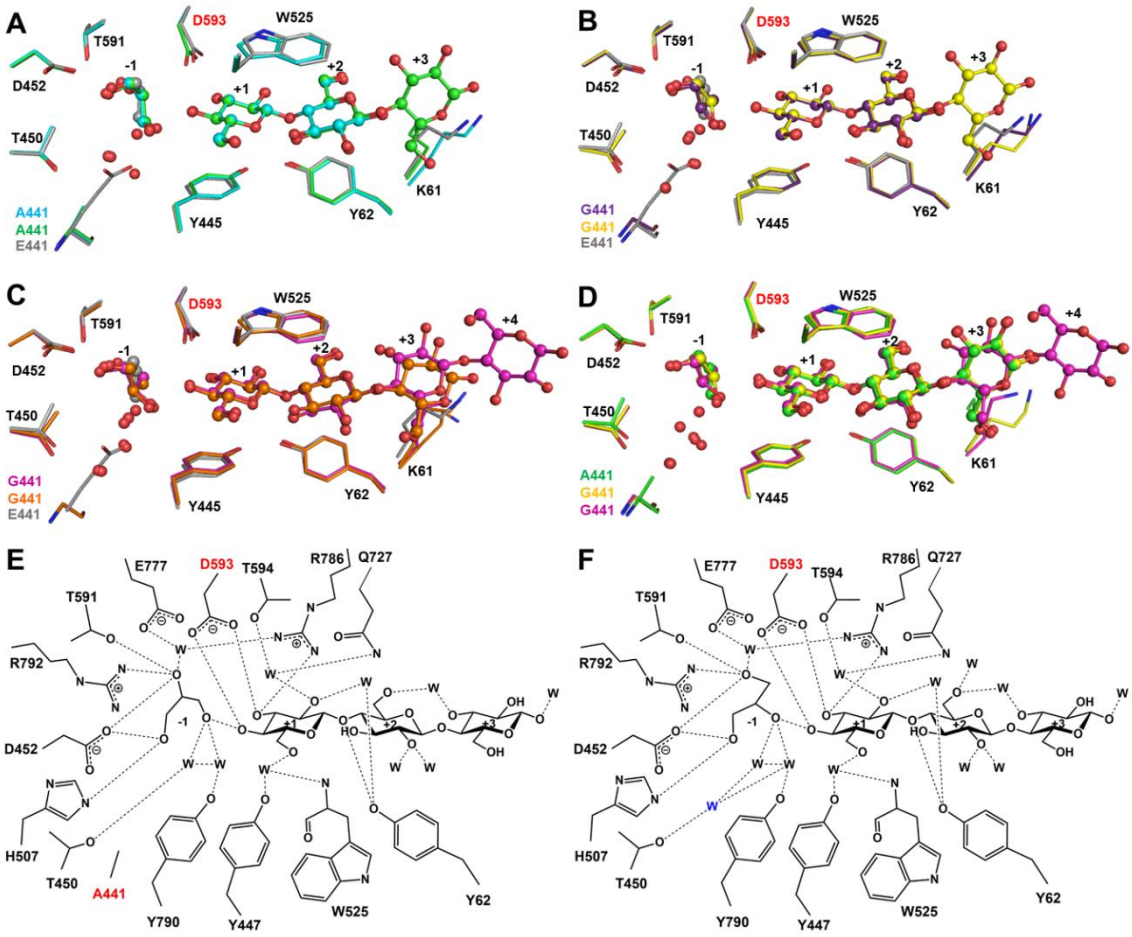


Figure 8. Comparison of the active sites of the TxGH116 E441A and E441G glycosynthases with cellooligosaccharides and free TxGH116. (A) The superimposition of E441A with cellobiose (PDB: 7DKU, cyan carbons) and cellotriose (PDB: 7DKV, green carbons) and free wild type enzyme (PDB: 5BVU, grey carbons). (B) The superimposition of E441G with cellobiose (PDB: 7DKX, violet carbons) and cellotriose (PDB: 7DKY, yellow carbons) and free wild type enzyme (PDB: 5BVU, grey carbons). (C) The superimposition of E441G with **α**-GlcF (PDB: 7DKW) molecule A exhibiting 4 glucosyl residues (pink carbons) and molecule B exhibiting 3 glucosyl residues (orange carbons) from autocondensation product and free wild type enzyme (PDB: 5BVU, grey carbons). An active site water molecule is bound at the same location as the catalytic nucleophile in the free wild type enzyme in each glycosynthase structure. (D) The superimposition of E441A and E441G with cellotriose (PDB: 7DKV green and PDB: 7DKY yellow carbons, respectively) and E441G with 4 residues of autocondensation product (PDB: 7DKW Mol A, pink carbons). The glycerol and cellooligosaccharides are represented as balls and sticks and the water molecules are represented as balls. Oxygen is shown in red, nitrogen in blue. (E and F) Hydrogen bonds between the glucose residues of cellotriose, surrounding amino acid residues and the network of water molecules (W) in the active sites of TxGH116

E441A (PDB: 7DKV, E) and E441G (PDB: 7DKY, F). Hydrogen bonds with measured distances of 3.2 Å or less are shown as dashed lines to indicate the hydrophilic interactions. From Pengthaisong et al. (2021).

Taking together the structure of TxGH116 E441A complexes with **α**-GlcF and with cellobiose, we were able to construct a plausible structure of the active site for the glycosynthase reaction (Figure 9). In this complex, the 4-hydroxylgroup of the nonreducing sugar residue is within 2.5 Å of the anomeric carbon opposite the fluoride leaving group and in close proximity for the basic assistance from the catalytic acid/base. This gives a clear explanation why the TxGH116 glycosynthase mutants prefer production of ß-1,4-glycosidic linkages. This would allow production of nanocellulose on a fine scale, which could be used in various fiber applications.

For further application of the E441G mutant, which showed highest activity for transglycosylation reactions that rescue the activity, we optimized the reaction for production of **α**-azidoglucose with pNPGlc donor and demonstrated cellobiose could also act as a glucose donor, although with lower efficiency (Gorantla et al., 2021). We then increased the volume of the reaction and developed the purification procedure to produce 280 mg pure **α**-azidoglucose. A set of alkyne with various functional groups was generated to demonstrate the usefulness of using this unprotected glycosylazide for synthesis, since no deprotection reactions that might destroy other groups were used. By 3 different click reaction strategies that depended on the alkyne, we produced the glucosyltriazoles, as shown in Figure 10. With the **α**-glucosyltriazoles in hand, we went onto test their activity on human lysosomal **α**-glucosidase overexpressed in Cos7 cells, in collaboration with the Laboratory of Biochemistry, Chulabhorn Research Institute. We found that 6 glucosyltriazoles gave stronger inhibition of human lysosomal **α**-glucosidase than the commercial inhibitor acarbose, with the best, AGT14, having an IC50 65-times lower than acarbose in our assay system (Table 2). We further went on to test the inhibitors on pig blood **α**-glucosidase, which is similar to human intestinal **α**-glucosidase (amylomaltase/maltobiase). As seen in Figure 11, the inhibitors were less effective than acarbose on the pig blood enzyme.

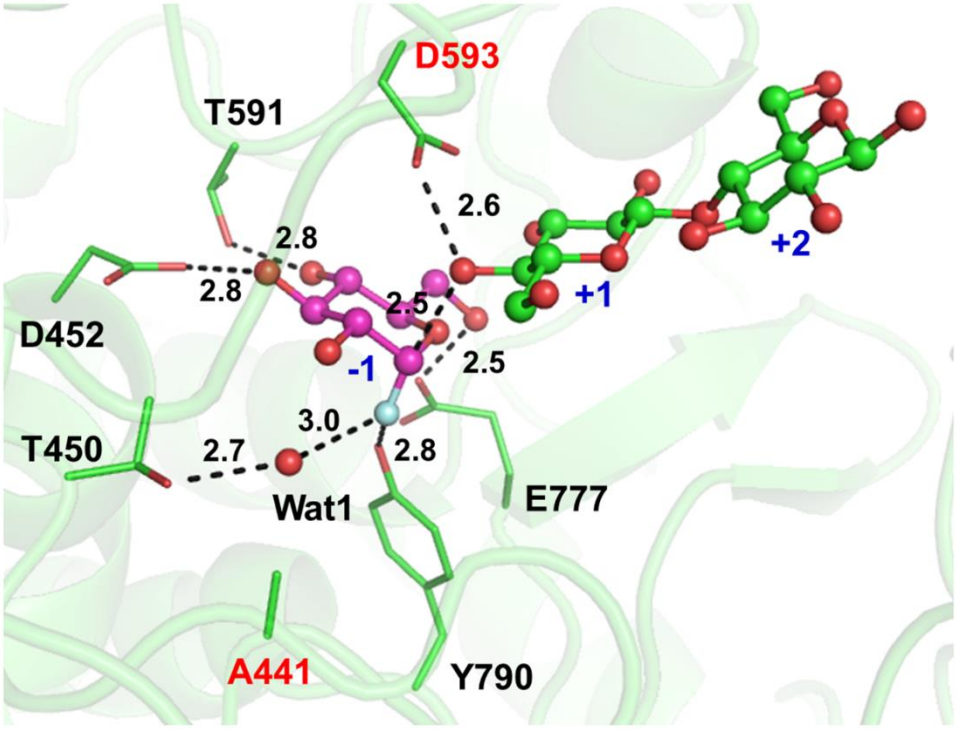


Figure 9. Plausible Michaelis complex active site for TxGH116 E441A glycosynthase based on combining the α-GlcF and cellobiose complex structures. Black dashed lines represent distances between the atoms that can indicate the distance for nucleophilic attack on the anomeric carbon, hydrogen bonds or hydrogen bonds as a prelude to proton extraction.

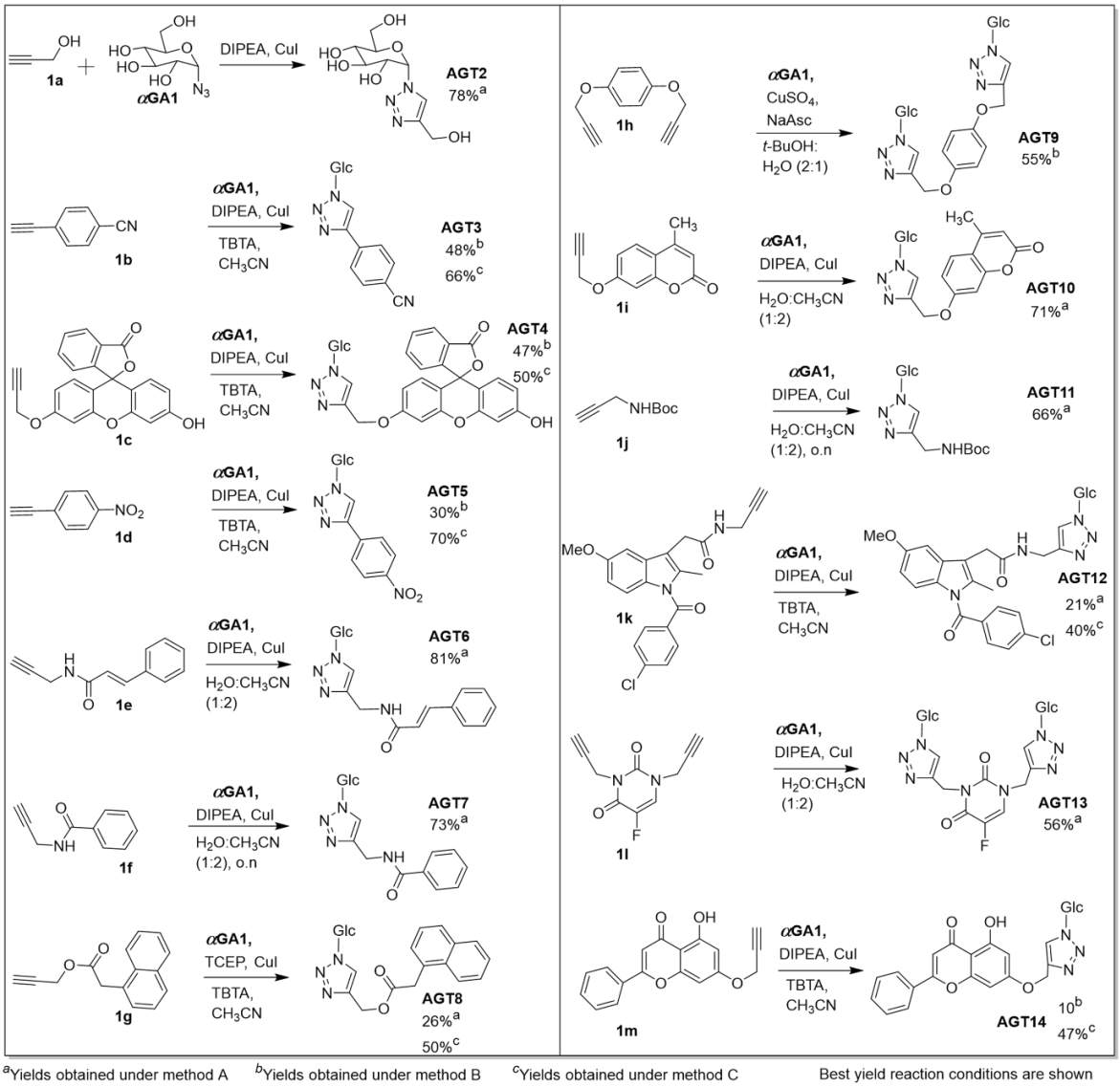


Figure 10. Glucosyltriazoles synthesized from the alpha-azidoglucose and the conditions used for click reactions to synthesize them.

Table 2. Inhibition of human lysosomal alpha-glucosidase by glucosyltriazoles.

Entry	Compound	α-Glucosidase inhibition [%][a]	IC_{50} [μM][b]
1	Acarbose	66	1100 ± 0.10
2	αGA1	09	ND
3	AGT2	12	ND
4	AGT3	46	ND
5	AGT4	100[c]	18 ± 0.002
6	AGT5	57	ND
7	AGT6	56	ND
8	AGT7	47	ND
9	AGT8	72	447 ± 0.035
10	AGT9	87	227 ± 0.025
11	AGT10	80	137 ± 0.015
12	AGT11	52	ND
13	AGT12	77	483 ± 0.021
14	AGT13	18	ND
15	AGT14	96	17 ± 0.001

[a]Percentage inhibition of lysosomal α-glucosidase in comparison to extracts incubated with 1.43 mM acarbose, αGA1, and AGT2 to AGT14. [b]IC_{50} values were determined by measuring human lysosomal α-glucosidase activity at different concentrations of the best inhibitors, AGT4, AGT8, AGT9, AGT10, AGT12 and AGT14. [c]The absorbance and fluorescence of AGT4 interfered with activity detection at the initial concentration, but AGT4 did show high inhibition at lower concentrations, which did not interfere with the assay.

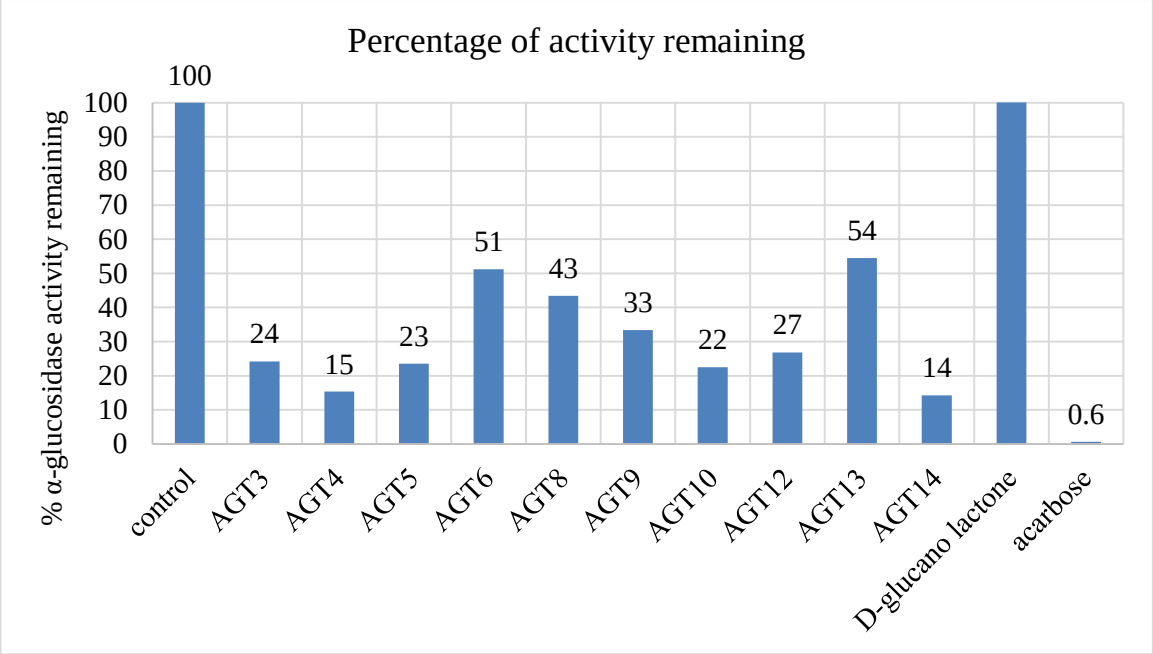


Figure 11. Percentage enzyme activity remaining after inhibition by acarbose, δ-glucano lactone and alpha-D-glucosyl triazoles derivative compounds.

2. Investigation of new GH116 enzymes.

In the previous grant, we had expressed, purified, characterized and started to solve the x-ray crystal structure of TeGH116, a novel family GH116 ß-glucosidase from the cyanobacterium *Thermosynechococcus elongatus*. During this project, we completed the structure and validated it, as well as solving structures with some ß-glucosyltriazoles.

In this time, we further characterized a plant GH116 protein, AtGCD3, the product of the At4G10090 gene locus, which was also expressed in *E. coli* and characterized as a glucosylceramidase by Dai et al. (2020), which we gave the more systematic name At4GH116 (GH116 protein encoded on *Arabidopsis thaliana* chromosome 4). Since mutations in the gene for this enzyme had little effect on glucosylceramide concentrations in the plant, but those in gene for the family GH1 enzyme Os3BGlu6 had significant effects on glucosylceramidase activity, the role of this enzyme as a glucosylceramidase is doubtful. We therefor characterized the enzyme purified from *E. coli* expression and tested its activity on a number of substrates. Preliminary screening by TLC showed that it had high activity on flavonoid glucosides, so we characterized its activity on these substrates by kinetic reactions. This showed that it had high activity on flavonoid 7-O-glucosides, as shown in Table 3. The relatively high activity on the synthetic substrate pNP-ß-D-galactoside suggests that it might also act on flavonoid 7-O-galactosides, although we did not test these substrates. The

efficiency on apigenin 7-O-glucoside as indicated by the specificity constant (k_{cat}/K_m) of 1100 $mM^{-1}s^{-1}$ is over 200,000-fold better than reported by Dai et al. (2020) for the synthetic fluorogenic glucosyl ceramide C_6-NBD-GlcCer (0.0053 $mM^{-1}s^{-1}$), suggesting that flavonoid glycosides are more likely to be natural substrates.

Table 3. Kinetics of At4GH116 on synthetic and natural substrates.

Substrate	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1}s^{-1}$)
*p*NP-β-D-glucopyranoside	0.077±0.005	1.35±0.008	17.4
*p*NP-β-D-galactopyranoside	0.264±0.016	2.43±0.027	9.2
*p*NP-α-L-arabinofuranoside	0.795±0.058	0.37±0.010	0.48
*p*NP-β-D-xylopyranoside	0.516±0.045	0.043±0.001	0.084
Apigenin-7-O-β-D-glucoside	0.049+0.0.003	56.5+0.76	1133
Quercetin-7-O-β-D-glucoside	0.048+0.0.003	34.9+0.53	735
Kaempherol-7-O-β-D-glucoside	0.077+0.0.008	17+0.39	220
Luteolin-7-O-β-D-glucoside	0.125+0.0.009	13.3+0.18	106
Kaempherol-3-O-β-D-glucoside	0.037+0.0.002	1.98+0.02	54.3

3. Characterization of novel glucoside/glucan degrading enzymes

In this work, we characterized novel ß-glucosidases, endoglucanases and cellobiosidases from plants, bacteria and fungi. Since the polymers cellulose and mixed-like beta-glucans make-up a large fraction of the grass cell wall, such as in rice straw, rice husk and sugarcane bagasse and leaves, these enzymes were expected to help in biomass degradation. In addition, determination of the function of rice enzymes might help in understanding the construction and breakdown of the cell wall in plants, allowing its manipulation for biomass release.

One enzyme that we characterized from rice was Os12BGlu38. When the gene for this enzyme was knocked out, pollen development was aborted during the late stages of development, similar to the effect of knockout of several cell-wall-proteins expressed at this stage (Shim et al., 2022). Os12BGlu was expressed in *E. coli* as maltose binding protein (MBP) and thioredoxin fusions, with the MBP fusion giving the best expression and purification. It was also expressed in secreted form in P. pastoris and was also purified in active form. However, all the active forms had low activity. Nonetheless, they could be shown to have activity to hydrolyzed pNPGlc, cellobiose, cellotriose, and cellotetraose, as well as a few glycosides, including sitosterol glucoside. It was thereby demonstrated that is had ß-glucosidase activity and postulated it may have a role in cell wall synthesis at the late stage of pollen development when other ß-glucosidases that might play similar roles are not expressed. However, the natural substrate and possible interaction with other proteins remains to be clarified.

Another rice GH1 ß-glucosidase that was previously characterized is the monolignol ß-glucosidase Os4BGlu18. In this project, we finalized the x-ray crystal structure of Os4BGlu18, which allowed docking of monolignols in the active site (Baiya et al., 2021). This showed what the residues are that interact with the monolignol and provide specificity to the enzyme (Figure 12). Modeling of other monolignol ß-glucosidases based on the Os4BGlu18 crystal structure indicated that the residues interacting with the monolignol aglycon part are not conserved, but have similar characteristics that will allow binding by aromatic stacking and a few hydrogen bonds. Since monolignols that are released from their glucosides can be rapidly polymerized by oxidation in the cell wall, these enzymes can increase lignin in cell walls under stress conditions, which will make the biomass more difficult to break down.

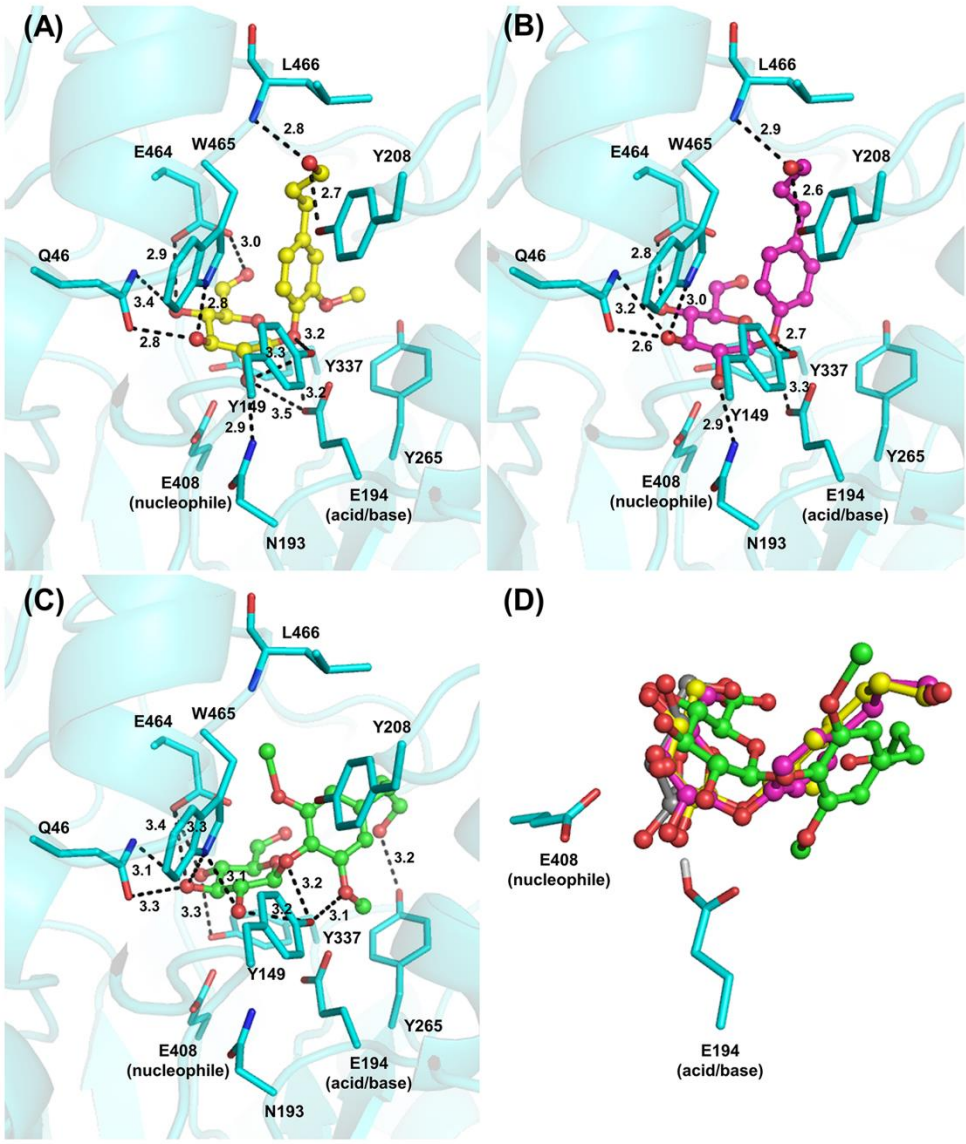


Figure 12. Molecular docking analysis of rice Os4BGlu18 with its natural substrates ligands. The interactions between coniferin (A), p-coumaryl alcohol glucoside (B), and syringin (C) and the amino acid residues in the active site cleft are shown. Hydrogen bonds (less than 3.5 Å between donor and acceptor atoms) are represented as black dashed lines and the amino acids which make these interactions are shown as sticks. (D) Superimposition of the active sites containing coniferin (yellow carbons), p-coumaryl alcohol glucoside (magenta carbons), syringin (green carbons), and glucono δ-lactone (grey carbons) showing that they are in nearly the same position and orientation for the glycone part. Oxygen is shown in red, nitrogen in blue and hydrogen in white.

Although we finished functional and structural characterization of rice enzymes in this work, the main focus was on microbial biomass degrading enzymes and the first to the studied

was *Thermobispora bispora Tb*Cel12A endoglucanase (Kuntothom and Ketudat Cairns, 2020). *Tb*Cel12A endoglucanase was identified by searching a rice compost metagenomic database with a GH12 cellulase and then comparing the sequence of one putative endoglucanase identified with the NCBI Genbank nr database. Upon expression in *E. coli*, this enzyme was found to have high activity on pNP-ß-cellobioside, which made it easy to check various parameters, such as temperature and pH optima and inhibition or activation by various compounds. *Tb*Cel12A had a temperature optimum of 65 °C and retained 86% of its activity after 16 h at 60 °C. Its pH optimum was 5.5. It also retained 93% of its activity after freeze drying and rehydrating. *Tb*Cel12A hydrolyzed barley (1,3),(1,4)-linked β-glucan faster than carboxymethyl cellulose (CMC). The main product of celloligosaccharide hydrolysis was cellobiose (Figure 13A). Products from barley β-glucan hydrolysis migrated at positions similar to cellobiose, cellotriose and cellotetraose on TLC, but β-1,3-linked glucose residues are likely to be found in some of these, based on the structure of the glucan. *Tb*Cel12A did not hydrolyze the β-1,3-linked glucose polymers laminarihexose, lichenin, and linamarin, nor did it hydrolyze beechwood xylan (plate assay). *Tb*Cel12A also hydrolyzed avicel and α-cellulose to release glucose, cellobiose, and cellotriose, but with a low hydrolysis rate compared to the soluble glucans (Figure 13B). The enzyme was tolerant of 10% ethanol and methanol and most metal ions tested, other than Cu^{2+} and Fe^{2+}. Although it had relatively low activity compared to commercial cellulase, a combination of *Tb*Cel12A with commercial cellulase released more reducing sugar from rice straw than either enzyme alone.

We attempted to express and characterize 4 more endoglucanases/cellobiosidases from GH family GH6 and successfully characterized one from *Thermobifida halotolerans* (NCBI accession: KJ101553.1), which we designated *Th*GH6 endocellulase. This enzyme had a pH optimum of 5.5 and temperature optimum of 55 °C. It retained 80% of its activity when incubated at 50 °C for 24 h. The enzyme was tolerant to some metal ions, but lost activity in 1 mM $HgCl_2$, $ZnCl_2$, $FeSO_4$, and $CuSO_4$. Interestingly, *Th*GH6 maintained over 80% activity in 20% methanol, ethanol or isopropanol, suggesting it may be relatively tolerant for co-saccharification/fermentation applications. When products of purified and natural substrates were analyzed by TLC, *Th*Gh6 produced mainly cellobiose and cellotriose, although some glucose was released in the case of phosphoric acid swollen cellulose and mixed linkage oligosaccharides were released in substrates with mixed linkage glucans (Figure 14).

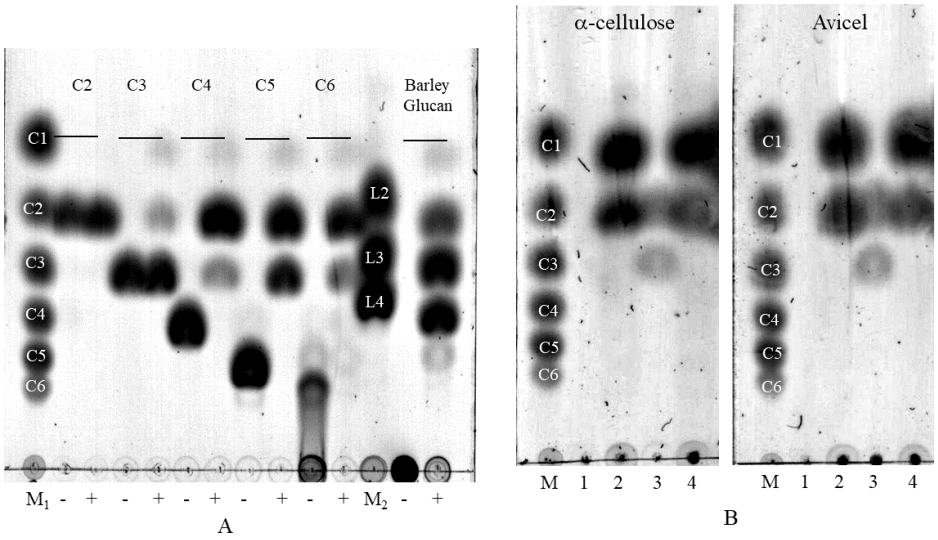


Figure 13. Products from *Tb*Cel12A hydrolysis of cellooligosaccharides and polysaccharides. (A) Products of hydrolysis of 1 mM cellooligosaccharides DP 2-6 and 1% barley mixed linkage β-glucan by *Tb*Cel12A; (B) Products from hydrolysis of 1% α-cellulose and avicel by commercial cellulase or *Tb*Cel12A. Reactions containing 20 pmol *Tb*Cel12A were incubated at optimal conditions and stopped by boiling. The silica-gel TLC was developed and products were detected by wetting the TLC plate with 10% v/v sulfuric acid in ethanol and heating at 120°C. Standards: C1: Glucose, C2: cellobiose, C3: cellotriose. C4: cellotetraose, C5: cellopentose, C6: cellohexaose, L2: laminaribiose, L3: laminaritriose, L4: laminaritetraose, Glucan: barley mixed linkage β-glucan, Minus signs show negative controls, and plus signs are TbCel12A reactions. In (B), lane 1: negative control, lane 2: products from commercial cellulase, lane 3: products from TbCel12A hydrolysis, lane 4: products from the combination of commercial cellulase and *Tb*Cel12A.

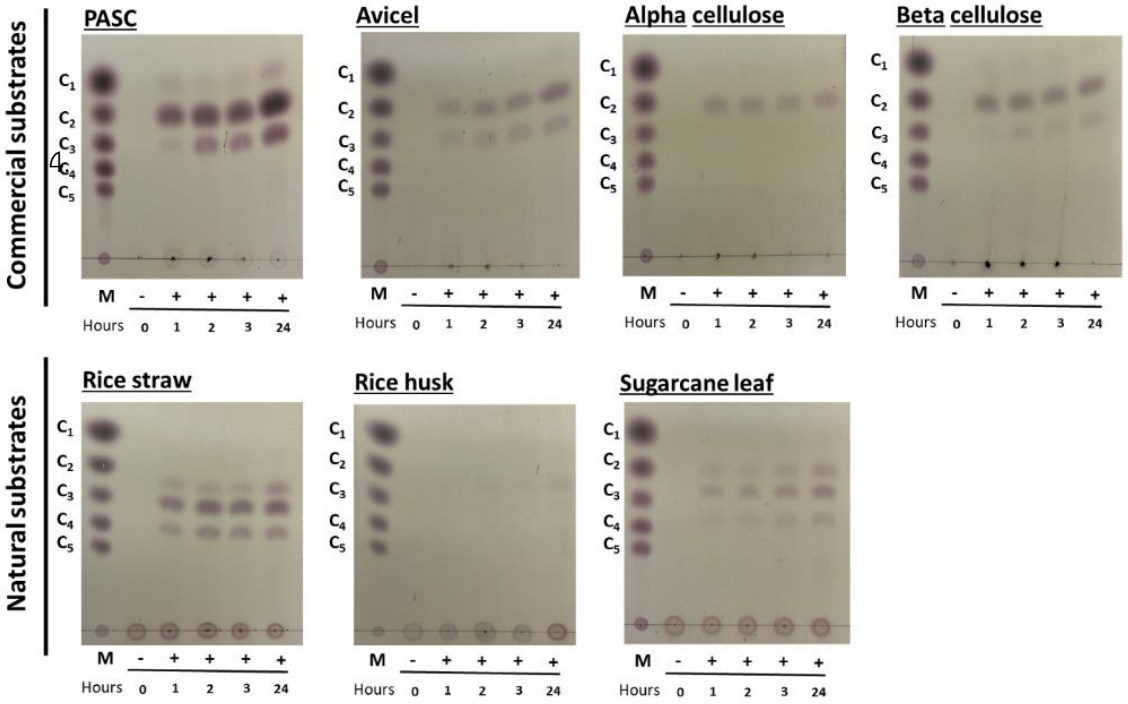


Figure 14. TLC analysis of products released from various substrates by *Th*GH6. PASC is phosphoric acid swollen cellulose, Avicel is cellulose, and Rice straw, Rice husk and Sugarcane leaf were powdered and pretreated with base. C1-6 represent cellooligosaccharides as in Figure 15. The + and – represent digestion with enzyme or not.

While GH6 cellobiosidases typically cut cellobiose off from the nonreducing end of beta-glucans and most exoglucanases also act on this end, GH7 enzymes have been reported to cut cellobiose off from the reducing end, suggesting there could be complementarity between these enzymes. So, we expressed and characterized a putative GH7 cellobiosidase, TtCel7A from *Thermothelomyces thermophilus* as a thioredoxin-fusion protein from pET32a vector in *E. coli*. TtCel7A had optimal activity at pH 5.5 and at 40-50 °C. It was stable at 40 and 50 °C for 24 h, but lost activity within 1 h at 60 °C. As shown in Figure 15, although TtCel7A primarily released cellobiose from cellooligosaccharides, it could also release a small amount of glucose from oligosaccharides with degree of polymerization of 3 (cellotriose) and higher. The release of cellotriose and glucose, in addition to cellobiose from cellotetraose, cellotetraose from cellopentaose and cellopentaose from cellohexaose demonstrated that it was not an obligatory cellobiosidase and can cut at other positions as well. When tested on polysaccharides, the primary product was cellobiose, but cellotriose and glucose were also detected with phosphoric acid swollen cellulose (PASC), rice straw and sugarcane leaf (Figure

18). Although the enzyme activity was interesting, TtCel7A did not seem to improve the action of enzyme mixtures, including when mixed with commercial cellulase.

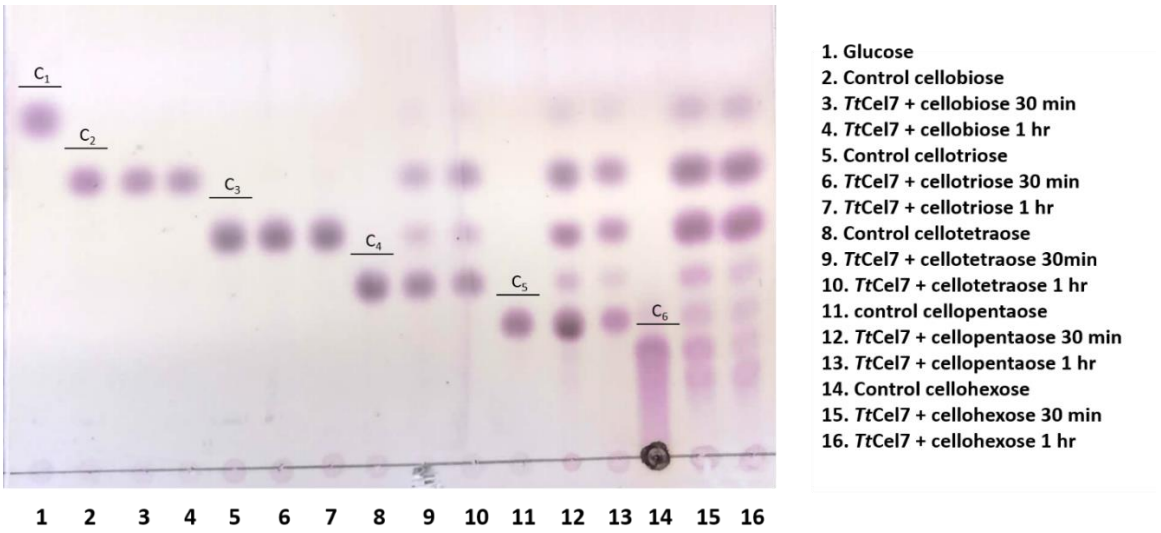


Figure 15. Thin-layer chromatography of the hydrolysis products of TtCel7A on cellooligosaccharides. Reaction products were detected after 30 min and 60 min of incubation at the optimal reaction condition.

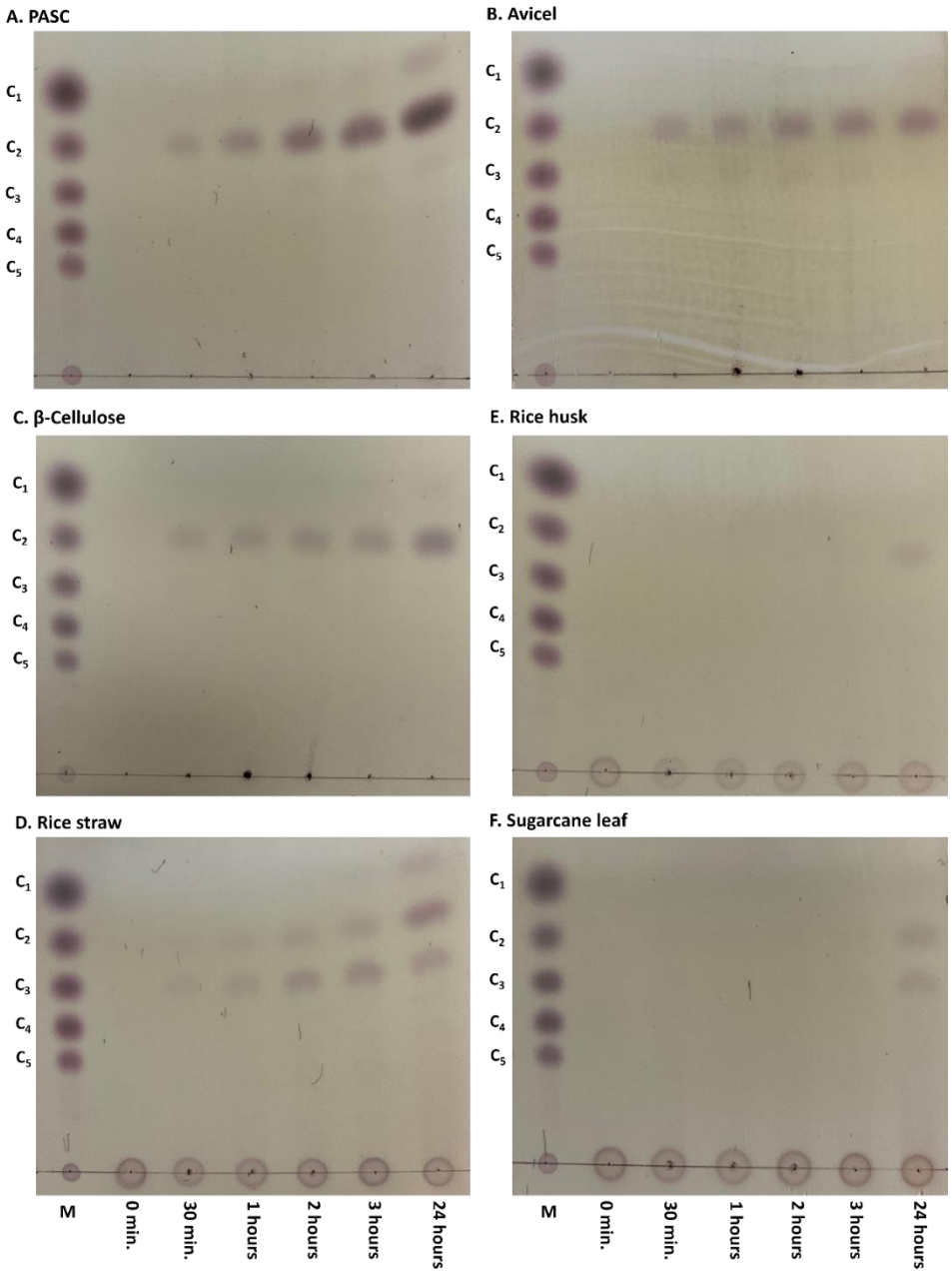


Figure 16. Thin layer chromatography showing the reaction products of TtCel7A with polysaccharide and agricultural residue substrates. Products at various time points were evaluated on silica gel TLC with carbohydrate staining using 10% sulfuric acid in ethanol.

Since the endoglucanases and cellobiosidases that we expressed did not seem to be as active as commercial cellulases, we decided to try to express relatively well characterized and active GH family 5 (GH5) endoglucanases from the fungi *Trichoderma reesei* and *Thermothielavioides terrestris*. Previously, *T. reesei Tr*Cel5A (previously designated EG II) had primarily been expressed in yeast without the N-terminal carbohydrate binding module, while

the *Tt*Cel5A was purified from its native fungi. Crystal structures are available for the catalytic domains of both proteins.

We expressed the full-length mature TrCel5A and TtCel5A proteins in *E. coli* and P. *pastoris*, but the expression was not obvious from *P. pastoris*, despite it being a fellow fungus, so we characterized the full-length proteins as thioredoxin fusion proteins expressed in *E. coli* strain Origami(DE3). Although much of each protein was found in the insoluble fractions, a significant amount was found in the soluble fractions and could be purified by IMAC, as shown in Figure 17. When tested on polysaccharide substrates, the enzymes could hydrolyze a range of ß-glucans, with the soluble polysaccharide barley ß-glucan being the most rapidly hydrolyzed (Figure 18). Using this barley ß-glucan, *Tr*Cel5A showed a pH optimum of 7.0 and temperature optimum of 55 °C, while *Tt*Cel5A showed a pH optimum of 6.0 and temperature optimum of 60 °C.

When *Tt*Cel5A, *Tr*Cel5A, their mixture and mixtures with *Tx*GH116 and *Tt*Cel7A were tested on pretreated agricultural residues, the two enzymes showed similar reaction profiles (Figure 19). When mixed with *Tx*GH116, the soluble products were all converted to glucose, but adding *Tt*Cel7A had little effect on the hydrolysis products.

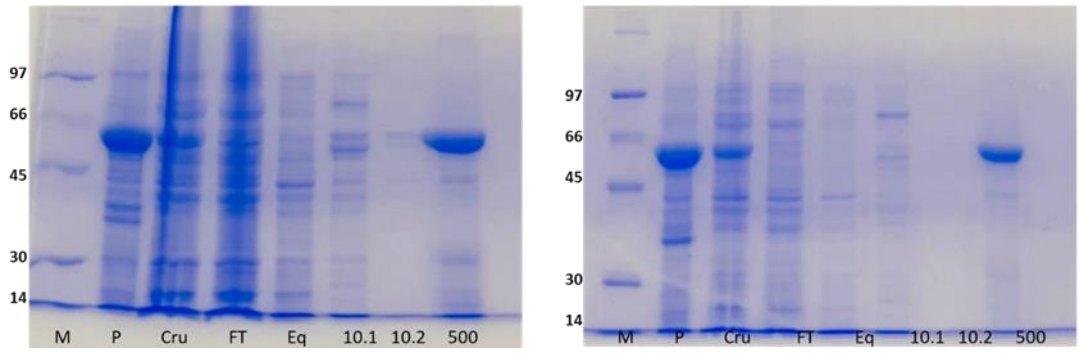


Figure 17. SDS-PAGE analysis of *Tt*Cel5A (left) and *Tr*Cel5A (right) expression as thioredoxin fusion proteins in Origami(DE3) *E. coli*. Lanes are: M, low MW protein markers; P, insoluble pellet of cell extract; Cru, crude soluble extract, FT, flow through fraction from IMAC column, Eq, wash of IMAC column with equilibration buffer; 10.1 and 10.2, washes with 10 mM imidazole in equilibration buffer; 500, elution of IMAC column with 500 mM imidazole.

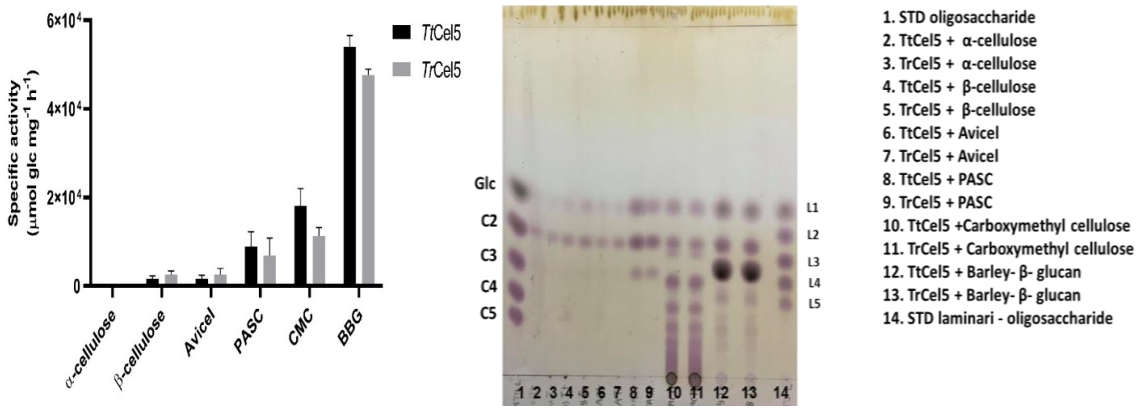


Figure 18. The hydrolysis of polysaccharide, including α-cellulose, β-cellulose, Avicel cellulose, PASC, carboxymethyl cellulose (CMC) and barley β-glucan, by *Tt*Cel5A and *Tr*Cel5A. Each reaction initially contained 1 mg of polysaccharide and 10 μg of *Tt*Cel5 and *Tr*Cel5 The reaction product amounts were determined by DNS assay. Thin-layer chromatography of the hydrolysis products of TtCel5A and *Tr*Cel5A are shown on the right. The reaction products were detected after 16 hours of hydrolysis.

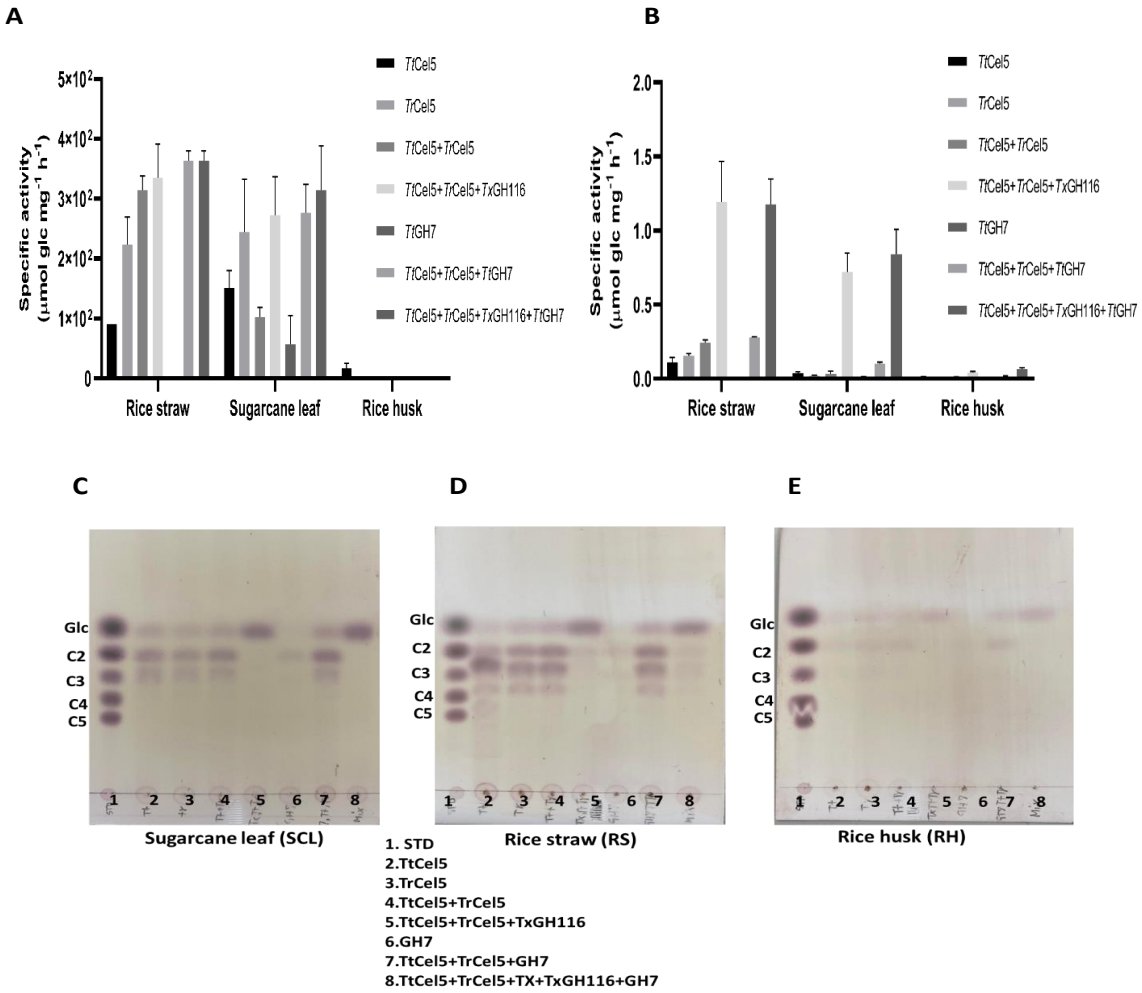


Figure 19. Hydrolysis of pretreated agricultural residues with *Tt*Cel5A, *Tr*Cel5A and their mixtures with *Tx*GH116 ß-glucosidase and *Tt*Cel7A. A. Assay of reducing sugars with DNS assay. B. Assay of glucose released with glucose oxidase assay. TLC of reaction products from hydrolysis of sugarcane leaf (C), rice straw (D) and rice husk (E) by *Tt*Cel5A, *Tr*Cel5A, *Tt*Cel7A, *Tx*GH116 ß-glucosidase and their combinations are shown below.

Since *Tx*GH116 has relatively high glucose inhibition, we tried to produce a GH1 ß-glucosidase that we hoped would be less inhibited. To find a ß-glucosidase that also worked at higher pH, we mined the sequences from *Thermobifida halotolerans* to find one that was not yet characterized, and identified the NCBI accession WP_068693385.1, as an enzyme that had not yet been characterized and named it *Th*GH1B, since another *T. halotolerans* GH1 enzyme was previously characterized. *Th*GH1B could be expressed in soluble form in *E. coli*, similar to the previously described enzymes, and the recombinant protein had pH optimum of 5.5 and temperature optimum of 55 °C with pNPGlc substrate, but was stable at 40 °C, but not at higher temperatures. This was not the thermostable and alkalophilic enzyme that we

expected. It had activity on cellooligosaccharides with the highest activity on cellotriose. Its K_m and k_{cat} for pNPGlc were 0.61 mM and 117 s-1 (k_{cat}/K_m 192 $mM^{-1}s^{-1}$), but it was less efficient on cellobiose with a K_m of 23 mM and k_{cat} of 136 s^{-1} (k_{cat}/K_m 5.9 $mM^{-1}s^{-1}$). Although it had higher glucose tolerance than *Tx*GH116, the glucose competitive inhibition constant (K_i) of 36 mM still indicates significant inhibition. Therefore, in the future, we may investigate ß-glucosidases that have been reported to have high inhibition constants to see if they improve the production of glucose from agricultural waste.

Another set of enzymes that were further investigated in this study were the rice family GH3 enzymes OsExoI and OsExoII. OsExoI could only be produced in the yeast *P. pastoris*, similar to its barley counterpart HvExoI (Luang et al., 2010), but OsExoII was expressed in *E. coli*, though at low yield, as described in previous grants. While trying to improve OsExoII purification, we found that it had significantly higher activity in fractions containing ethylene glycol, and this was found to result from the transglycosylation of the ethylene glycol (Prawisut et al., 2020). Therefore, we tested its ability of OsExoI and OsExoII expressed in pichia and *E. coli*, respectively, to transglycosylate several alcohols, including methanol, ethanol, isopropanol, n-butanol, propargyl alcohol and n-octanol. At the same time, we expressed OsExoI in *Pichia pastoris* and purified and characterized it and also tested its ability at transglycosylation. As can be seen in Figures 20 and 21, OsExoI could transglycosylate p-nitrophenyl-β-D-glucoside (pNPGlc), glucose and various alcohols. The enzyme could also use oligosaccharides and soluble glucans as glucose donors for glycosylation. OsExoII expressed in *E. coil* tended to have similar but somewhat better activity, although n-octanol and isopropanol were not glycosylated efficiently. We were able to purify and verify the structures of the products by NMR. In the presence of p-nitrophenyl-β-D-glucoside (pNPGlc) donor, the enzyme produces pNP-β-gentiobiose and lower amounts of pNP-cellobiose and pNP-laminaribiose and some trisaccharides, as well as oligosaccharides (mainly gentiobiose) from the transglycosylation of the glucose hydrolysis product (Figure 20). The glycosides of the various alcohols in Figure 21 were also verified by NMR. With cellobiose donor, the enzyme can produce similar amounts of glycosides, which tend to be more separated from the substrate, making chromatographic purification easier. Since the enzyme from a single round of IMAC purification is sufficient for this work, we were able to avoid losses of enzyme during purification.

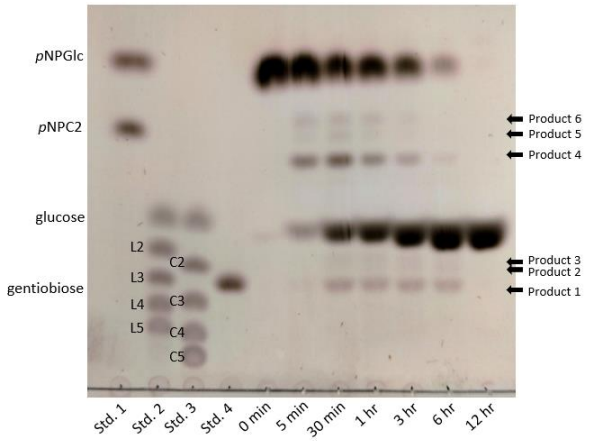


Figure 20. Transglycosylation of pNPGlc substrate by OsExoI from expression in *Pichia pastoris*. Products 4, 5 and 6 were verified to be pNP-ß-gentiobioside, pNP-ß-cellobioside, and pNP-ß-laminaribioside by NMR. The reaction products at various time points are shown.

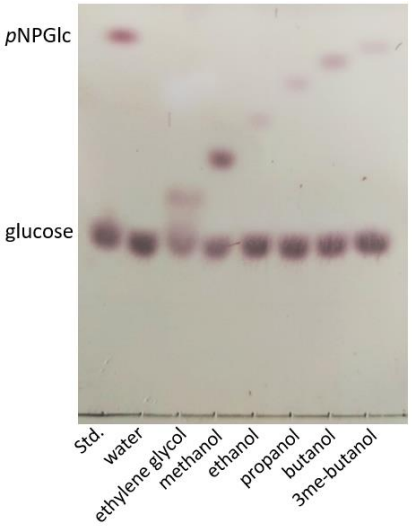


Figure 21. Glycosylation of various alcohols by OsExoI using pNPGlc donor analyzed by TLC. After complete breakdown of the pNPGlc, the simple glycosides appear to be relatively stable to digestion. The alcohols that were transglycosylated to form the products are given at the bottom.

In addition to these work on the rice enzymes, we continued our collaboration on barley ExoI characterization, in particular work on the structures of mutations related to the tryptophan clamp at the entrance to the active site. The structures were submitted to computational chemists to simulate the effect on the reaction, showing that this clamp plays an important role in progressive hydrolysis of oligosaccharides and polysaccharides (Luang et al., in revision for Nature Communications).

4. Studies on xylanolytic enzymes

We continued on a GH3 xylosidase, which was expressed in *Pichia pastoris* and purified from the medium, designated OsXyl1. The purification of this protein was improved to allow characterization studies. This enzyme had a pH optimum of 4.0 and temperature optimum of 60-65 °C in a half-hour assay, but was not stable at 60 °C, although it retained over 50% of its activity for 4 h at 50 °C. It was inhibited xylose with a competitive inhibition constant of 1.41 ± 0.06 mM, but other sugars showed little inhibition at this concentration, including L-

arabinose, D-glucose, D-galactose, D-mannose and D-fructose. It showed 27% and 24% relative activity on pNP-α-L-arabinopyranoside and arabinofuranoside, relative to pNP-ß-D-xylopyranoside, but less than 10% relative activity on other pNP-glycosides. It also had activity on ß-1,4-linked xylooligosacchardies, as shown in Figure 22. With pNP-ß-D-xylopyranoside, OsXyl1 had a specificity constant (k_{cat}/K_m) of 19 $mM^{-1}s^{-1}$, while no other pNP-glycosides had values higher than 0.01 $mM^{-1}s^{-1}$ and xylotriose specificity constant was 4.2 $mM^{-1}s^{-1}$. Thus, OsXyl1 can act on xylan-derived oligosaccharides, which are likely natural substrates for the enzyme, given that pNP-ß-D-xylopyranoside is not found in nature. It may have application to biomass xylan breakdown, although we will investigate bacterial and fungal enzymes that may be easier to produce in the future.

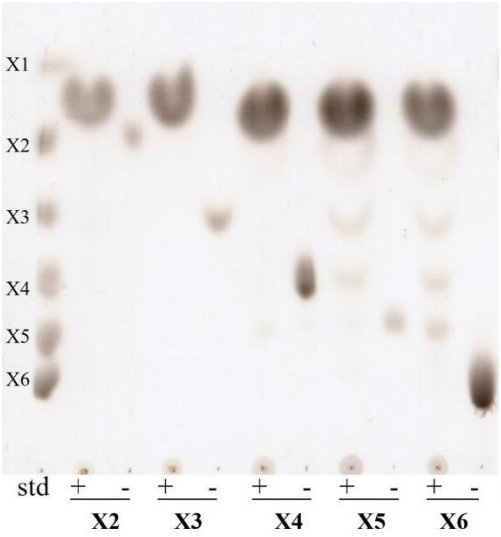


Figure 22. Hydrolysis of ß-1,4-linked xylooligosaccharides by OsXyl1. The TLC shows products of overnight reactions with xylobiose (X2), xylotriose (X3), xylotetraose (X4), xylopentaose (X5) and xylohexaose (X6).

To further explore biomass degradation, we expressed 2 putative archaeal beta-xylanases from glycoside hydrolase family GH11, a well-studied enzyme from *Thermomyces lanuginosus* and one from *Thermothelomyces heterothallicus*. Initial expression of ThXyln11A with a simple N-terminal his-tag from pET30 gave no significant soluble protein, but transferring to pET32a, to produce the proteins with an N-terminal thioredoxin tag in Origami(DE3) bacteria allowed us to produce soluble proteins. The ThXyln11A was produced at about 5 mg/liter culture and showed high activity on at pH 6 at 60-65 °C on oat and birchwood xylan, but no activity on barley ß-glucan, as expected. The ThXyln11A had similar activity, but its temperature optimum was only 40 °C with highest activity at pH 6-7.5, and it was not stable,

except at 30 °C. A preliminary test showed the TlXyln11A could be used to pretreat rice straw to allow release of more glucose when cellulose was applied. In the future, these enzymes will be used to release xylooligosaccharides for further degradation by β-xylosidases, such as OsXyl1.

5. Immobilization and modification of TxGH116 variants

Immobilized TxGH116 E441G and D593A on IMAC resin had transglucosylation activities to transfer glucose from 4NPGlc to azide to produce α- and β-glucosyl azides, respectively. The immobilized E441G and D593A enzymes were stable at 50°C, as shown the relative activity by measuring the 4NP release (Figure 23), while the immobilized enzymes showed slightly decrease transglucosylation activities after recycling more than 5 times (Figure 24). For the large scale, 5 ml of immobilized E441G (150 mg) was incubated with 10 mM 4NPGlc and 1 M NaN_3 in 30 ml reaction at 50°C with shaking. The precipitation of the immobilized E441G was observed in the reaction.

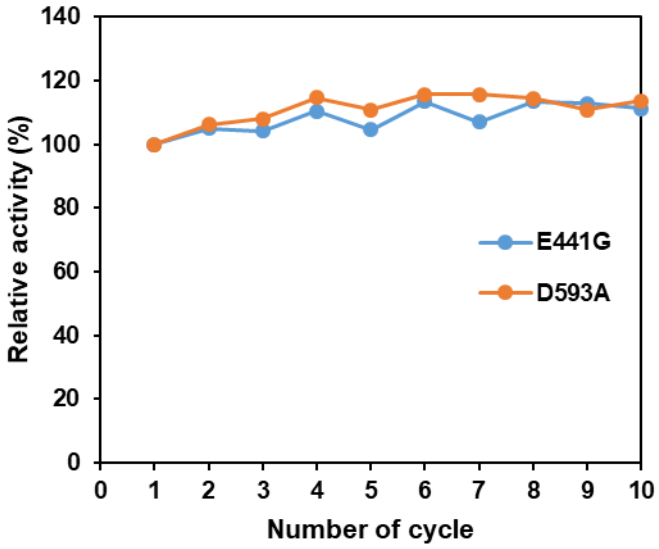


Figure 23. Stability of immobilized TxGH116 E441G and D593A on IMAC at 50°C during recycling for up to 10 times.

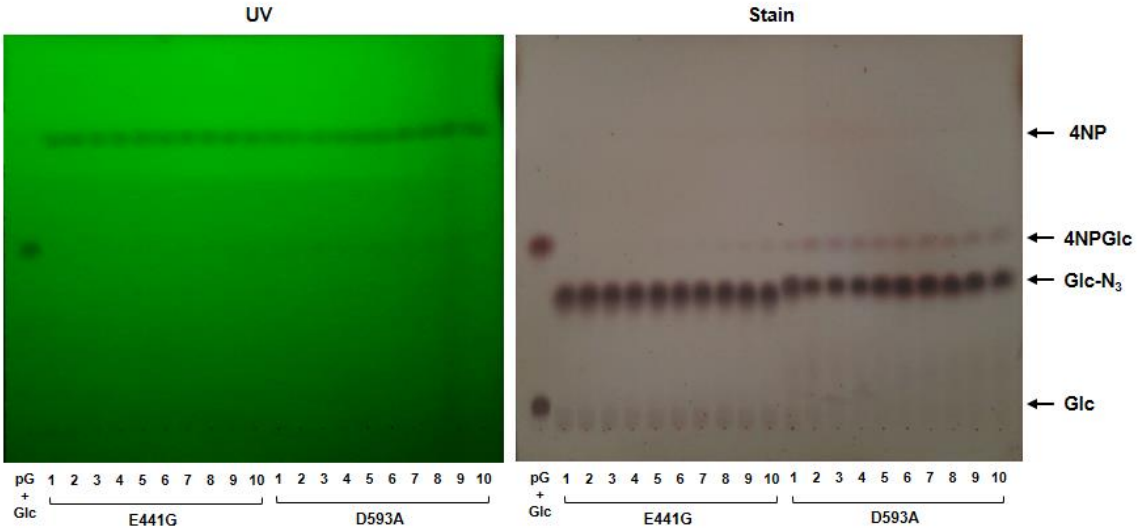


Figure 24. Azidoglucose products during transglucosylation reactions by recycled *Tx*GH116 E441G and D593A immobilized on IMAC resin. Reactions contained 10 mM 4NPGlc and 500 mM (for E441G) or 100 mM (for D593A) NaN_3 in 50 mM MES pH 5.5 at 50°C after recycling for up to 10 times.

The immobilized TxGH116 and E441G on CNBr-activated Sepharose™ 4B were produced for hydrolysis and transglucosylation, respectively. The pH optimum of immobilized TxGH116 was decreased from pH 5.5 to pH 5.0 and the pH profile was narrow, while the temperature optimum and temperature stability of immobilized TxGH116 was similar to free enzyme. Both immobilized and free TxGH116 had no activity at 70 and 75°C after incubated for 24 and 6 h (Figure 25A), respectively, while the enzymes was stable at 60° and 65°C, retaining up to 90% and 65% activity after incubated for 3 days (Figure 25B), respectively. The activity of immobilized enzyme remained more than 90% when incubated the immobilized enzyme at 60°C for 15 min and recycled for 20 times (Figure 26).

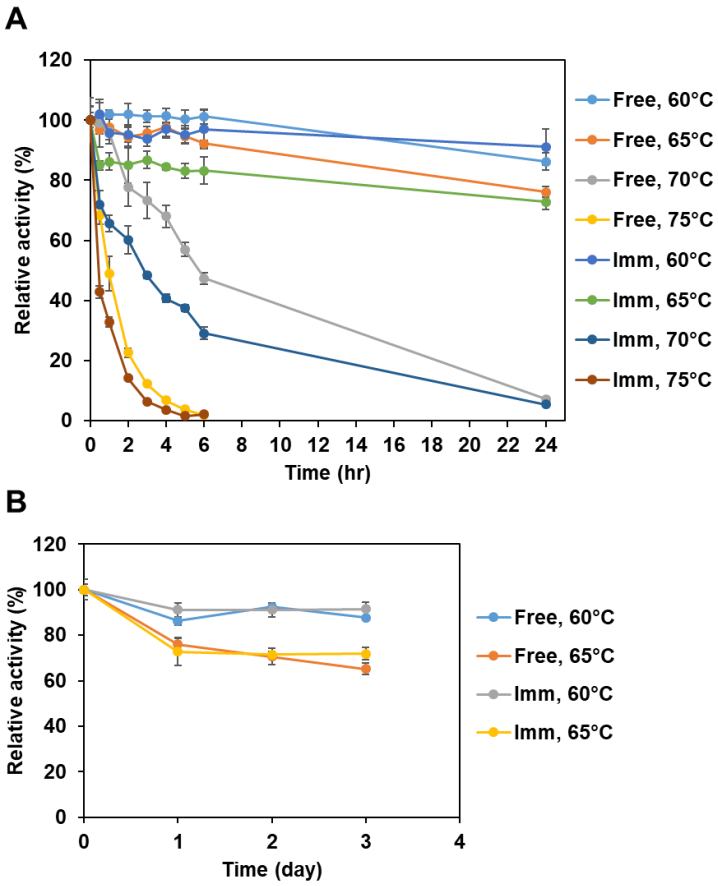


Figure 25. Stability of free TxGH116 and TxGH116 immobilized on CNBr-activated Sepharose 4B. (A) Activity after incubation from 1-24 hours at various temperatures. (B) Activity remaining after incubation at 60 and 65 ℃ for 1-3 days. Imm stands for immobilized enzyme, while Free is enzyme that is free in solution.

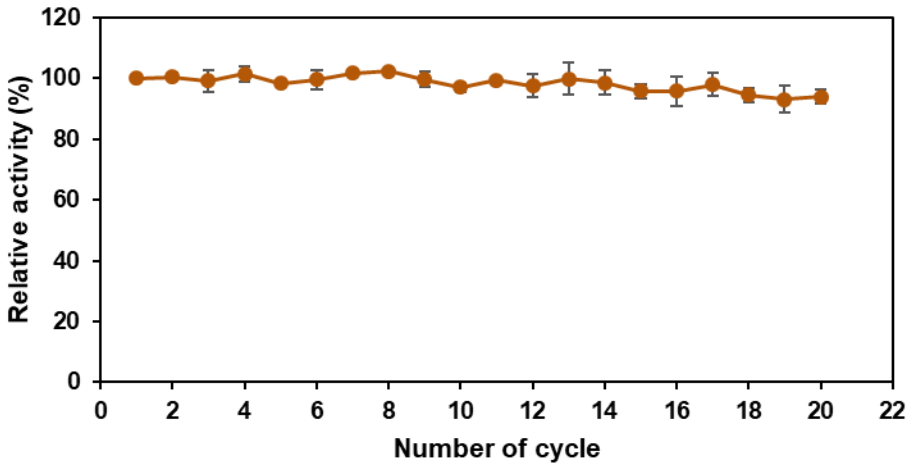


Figure 26. Activity of TxGH116 immobilized on CNBr-activated Separose 4B in each reaction during recycling up to 20 cycles.

The immobilized TxGH116 E441G on CNBr-activated Sepharose™ 4B showed transglucosylation activity to produce α-glucosylazide. The pH optimum of immobilized TxGH116 E441G was pH 4.5 and the pH profile was similar to free enzyme. The activity of immobilized E441G slightly decreased at 50°C and remained 70% after incubated for 10 days, while the free enzyme was stable at 50°C for at least 10 days. The immobilized E441G retained 50% of its activity on the 10^{th} cycle of reaction at 50°C for 20 hr and recycled for 10 times (Figure 27). So, unfortunately the immobilized mutant was not as robust as the wild type enzyme but could be reused, although with decreasing activity.

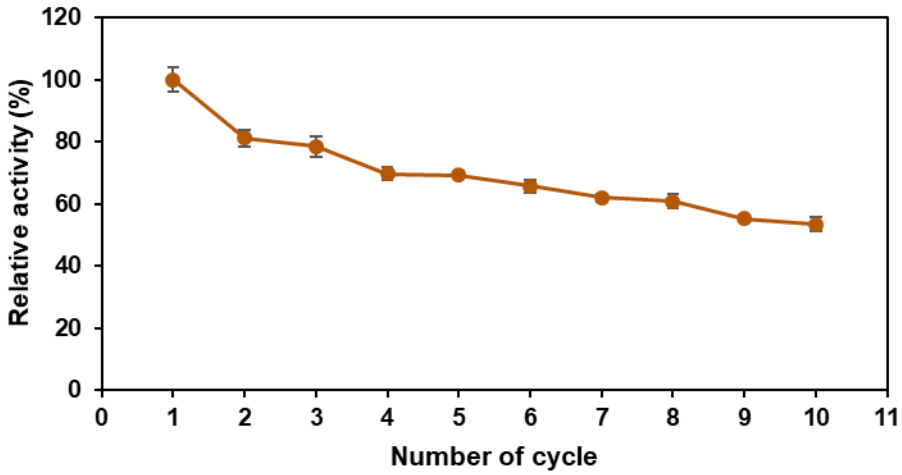


Figure 27. Relative activity of TxGH116 E441G immobilized on CNBr-activated Sepharose™ 4B during multiple reaction cycles. Each cycle was a reaction at 50°C for 20 h.

When tested at various temperatures and various concentrations of ionic liquid, unmodified TxGH116 and [C-TxGH116-S], which had been modified to enhance solubility and activity in ionic liquid, showed highest hydrolysis activity at 80°C, but the activity decreased when increasing the ionic liquid concentration (Figure 30). However, the [C-TxGH116-S] showed lower hydrolysis activity than free TxGH116 in ionic liquid. Both of their activities at 90°C remained about 60% in 70% ionic liquid (Figure 28). The TxGH116 activity with pretreated ß-cellulose and avicel in ionic liquid showed that the TxGH116 had no activity toward both ß-cellulose and avicel.

Since the unmodified TxGH116 had high activity in ionic liquid, we tested its activity in other solvent systems, especially for transglycosylation reactions where decreasing water activity is an advantage. The TxGH116 had highest hydrolytic activity toward 4NPGlc in 10% acetone (Figure 29A), and showed similar hydrolytic activity in 10% acetone and 0 to 20% octanol and the activity slightly decreased at 30 and 40% octanol (Figure 29B). The *Tx*GH116 could synthesize octyl glucoside at 37 °C and 60 °C, and the product increased when increased incubation time at 60 °C (Figure 30).

We concluded that unlike the previous publication with another enzyme (Brogan et al., 2018), the chemical modification of *Tx*GH116 and E441G mutant could not enhance the pH and temperature stabilities in aqueous and ionic liquid. However, the unmodified enzyme has relatively high activity in ionic liquids and organic solvents.

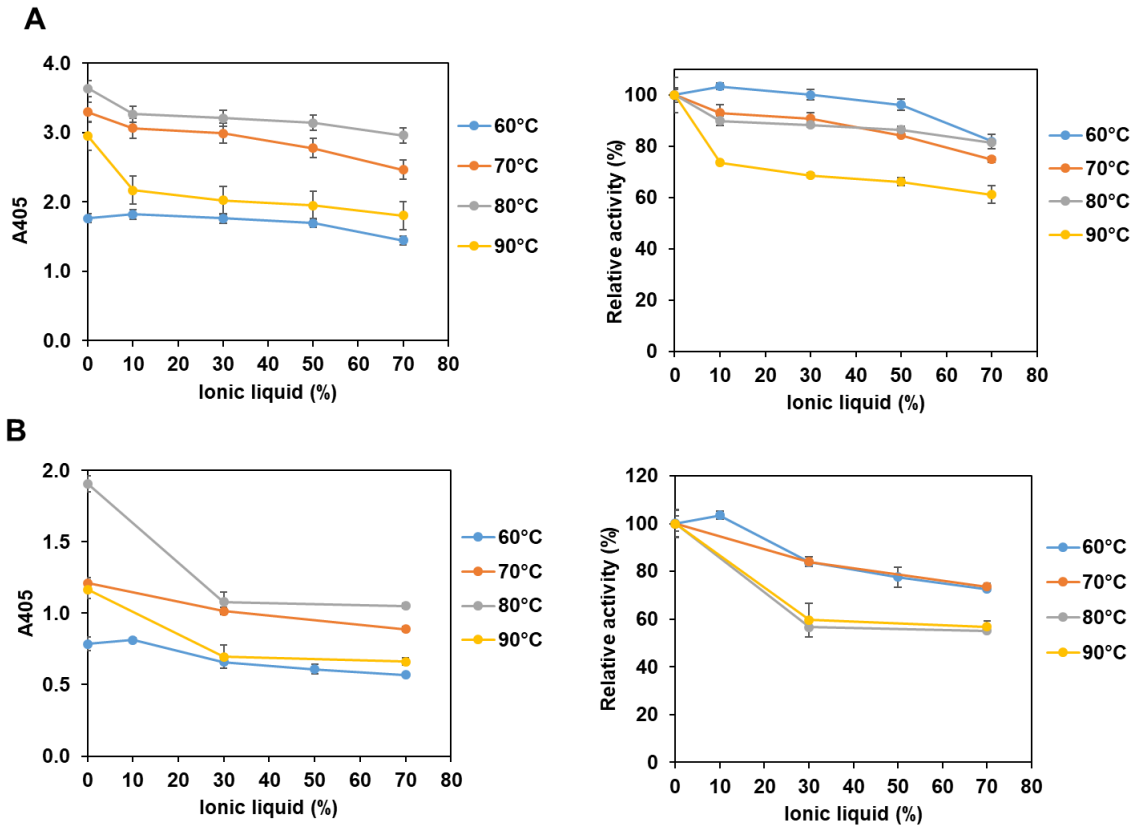


Figure 28. Hydrolysis activities of *Tx*GH116 (A) and [C-*Tx*GH116-S] (B) in ionic liquid shown in A405 absorbance of the released pNP product (left) and relative activity (right).

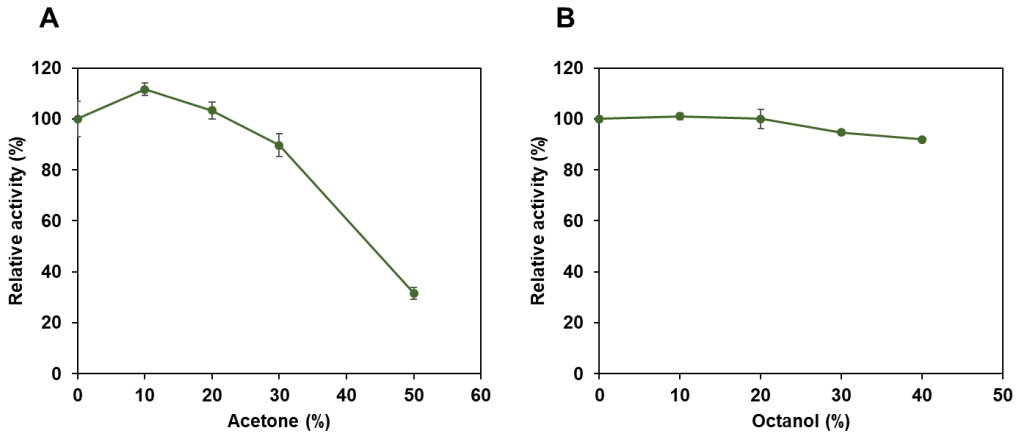


Figure 29. Hydrolysis activity of *Tx*GH116 in acetone (A), and 10% acetone and octanol (B).

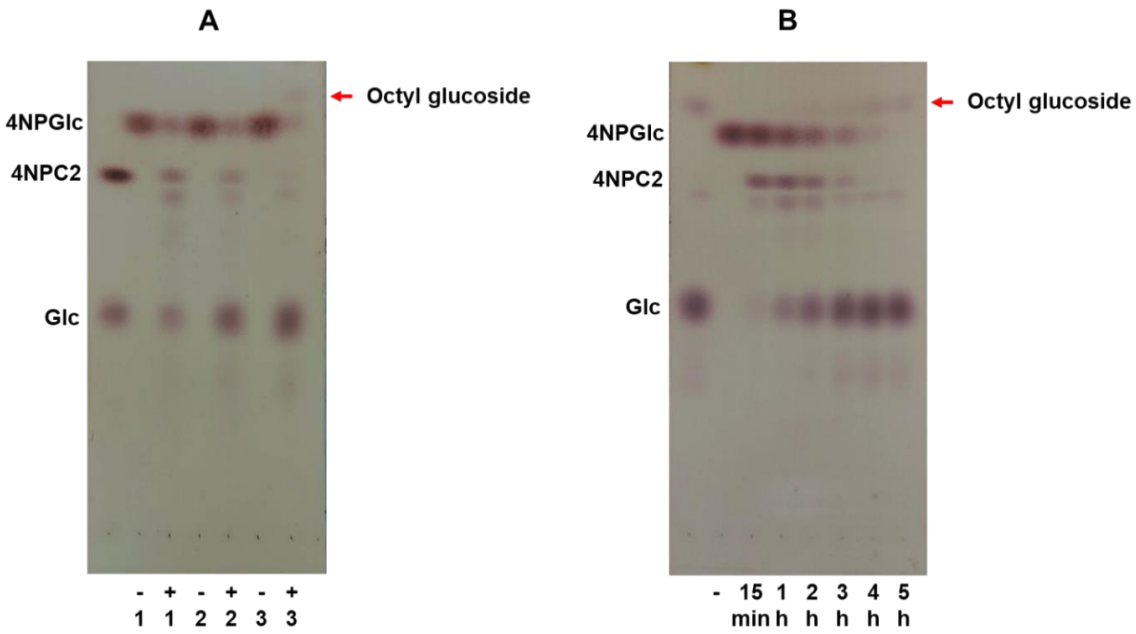


Figure 30. Transglycosylation activity of *Tx*GH116 with pNPGlc without and with octanol acceptor. (A) Transglucosylation activity of 0.025 mg/ml *Tx*GH116 with 10 mM pNPGlc in 50 mM sodium phosphate buffer, pH 5.5. The reactions of **1**: 0% acetone and 0% octanol, **2**: 10% acetone, 0% octanol and **3**: 10% acetone and 20% octanol without (-) and with (+) enzyme were incubated at 37°C for 18 h. (B) Transglucosylation activity of 0.025 mg/ml TxGH116 with 10 mM pNPGlc, and 10% acetone and 20% octanol in 50 mM sodium phosphate buffer, pH 5.5. The reactions were incubated at 60°C for 15 min, 1, 2, 3, 4 and 5 h.

Discussion and Conclusions (สรุปและวิจารณ์ผลการทดลอง)

During this project, we characterized and applied mutated variants of TxGH116, further evaluated and structurally analyzed family GH1, GH3 and GH116 β-glucosidases and transglucosidases, produced endoglucanases from families GH5, GH6, GH7 and GH12 and xylanases from GH11 to evaluate their action on agricultural biomass, and tested chemical modification and immobilization of TxGH116 as a means of improving its utilization. These studies generated 7 publications, along with several manuscripts in process, and provided us with enzymes to apply to production of glycosides and biomass degradation in future studies.

For the first part, we characterized mutations of TxGH116 in the catalytic nucleophile and acid/base, interacting residues, residues binding to the glycon sugar and residues in the active site mouth. The activity of the nucleophile mutants to produce oligosaccharides via glycosynthase reactions and glycosides, such as **α**-azidoglucose was characterized with novel observations, particularly for the glutamate to aspartate mutant, which unexpectedly displayed glycosynthase activity (Pengthaisong et al., 2021). Moreover, production of **α**-azidoglucose was optimized and scaled up to allow production of **α**-glucosyltriazoles, some of which were found to be strong inhibitors of human lysosomal **α**-glucosidase (Gorantla et al., 2021) and weaker inhibitors of pig blood **α**-glucosidase, a surrogate for human intestinal digestive **α**-glucosidase. This process was patented, but because it developed more quickly than expected, the patent was submitted on the previous grant, but the characterization for publishing and synthesis of additional **α**-glucosyltriazoles was completed during this grant period.

Among the β-glucosidases that were structurally characterized in this grant, the TeGH116 structure was completed based on new data at higher resolution, the GH1 Os4BGLu18 monolignol β-glucosidase structure was completed and used for docking and the Os9BGlu31 W243L structure was originally solved to allow more rigorous engineering. For functional characterization, the Os12BGlu38 enzyme was characterized, although all recombinant enzymes had low activity, at least it was enough to confirm its action as a β-glucosidase, possibly involved in pollen cell wall recycling. Most of these and other rice enzymes were not applicable to biomass degradation due to low activity on cellobiose. So, we generated and characterized one bacterial GH1 β-glucosidase, which had somewhat higher glucose tolerance than TxGH116 and reasonable action on cellobiose, but it was less tolerant to high temperature and other conditions than expected and appeared to give less advantage

than TxGH116 when combined with endoglucanases in biomass conversion. On the other hand, the GH3 enzymes OsExoI and OsExoII have good activity on cellobiose, but are a bit more difficult to produce. However, they show good promise for converting glucose in cellobiose into glucosides, some of which could be detergents or be applied to further synthesis. Characterization of the mechanism of these GH3 enzymes was achieved in collaboration with Maria Hrmova, with whom we expect to publish another paper soon. Overall, this work has helped us to understand more about how the enzymes work in the plant and find new applications for them *in vitro*.

In order to generate the substrates for the b-glucosidases and exoglucanases, we generated several endoglucanases and cellobiosidases that could help breakdown cellulose and other glucans in plant cell wall biomass. Initial expression of a GH12 endoglucanase showed that it was active on pretreated biomass, but could not compete effectively with commercial cellulases, which are optimized mixtures of endoglucanases, although it could improve the breakdown of pretreated rice straw, under certain conditions. To find more enzymes to increase effectiveness, we considered several bacterial GH6 enzymes, but could only express the *Th*GH6 protein effectively. We also produced the fungal GH7 enzyme TtGH7A, since it was expected to release cellobiose from the reducing end, while other enzymes work on the nonreducing end, so it might have synergistic action. However, it was not found to increase the hydrolysis of biomass much when mixed with other enzymes.

Finally, we decided to express the well characterized *Tr*Cel5A and a related enzyme from a thermophilic fungus, *Tt*Cel5A to see whether they had higher activity. In the past, the TrCel5A was expressed primarily without its N-terminal carbohydrate binding module (CBM) in yeast like *P. pastoris*, but we expressed the full-length mature proteins and found expression in *E. coli* was more successful than *P. pastoris* in this format. The two enzymes had similar activity, which was higher than that of our previously characterized endoglucanases, but much lower than commercial cellulase mixtures, such as Celex CTech-2 from Novozyme. We will need to compare the activity with and without the CBM in the future, to see whether the non-CBM format that was previously used is more effective and what are the relative benefits of each form.

Another aspect of the project was looking at xylan active enzymes. During this project, we continued to work on the plant family GH3 ß-xylosidase OsXyl1 from rice, and were finally able to purify enough for characterization. The enzyme seems to be fairly specific for ß-xyloside and xylo-oligosaccharides, although it can hydrolyze **α**-L-arabinosides. It could also

transglycosylate alcohols to make xyloside products. Since using pNP-xyloside for such reactions is relatively expensive and we wanted enzymes for xylan breakdown, we produced the well-characterized fungal xylanase TlXyln11A and the related enzyme ThXyln11A in our *E. coli* system. The TlXyln11A has a known crystal structure and has previously been produced in both *E. coli* and yeast and has been engineered, while the *Th*Xyln11A has only been produced from its native source, to our knowledge. Both enzymes could effectively be expressed in our *E. coli* system and were able to hydrolyze xylan, although the *Tl*Xyln11A was expressed more highly, had higher temperature tolerance and activity than *Th*Xyln11A. Although the fact that the well-characterized enzyme being a stronger candidate depletes the novelty of the initial experiments, the availability of the effective xylanase in the lab will allow us to combine it with our ß-xylosidase and other enzymes for biomass breakdown and xylan utilization.

Perhaps the least successful part of the project was the chemical modification of the TxGH116 enzyme, although the immobilization worked to some degree. The immobilization allowed the enzyme to be put in a solid phase that could readily be reused, thereby saving on cost of production of enzyme and aiding product purification. However, the greatest immediate interest was to convey this property to the E441G nucleophile mutant to allow large scale production of **α**-azidoglucose, which is difficult to synthesize chemically but is readily produced by this enzyme, although it requires a large amount of the mutant enzyme. Although the E441G enzyme could be immobilized and reused effectively on IMAC resin, this did not work at large scale under to low pH conditions optimal for this enzyme, since the metal and His-tag will have the same charge under these conditions and the repulsion from the resin resulted in precipitation of protein in the reaction. When it was immobilized on CNBr-activated Sepharose 4B, the enzyme could be reused, but gradually lost activity during cycling, unlike the wild type enzyme, so it was less effective than we would like.

Chemical modification was supposed to produce a sort of solvent shell around the protein, which would aid is solution in ionic solvents, and this could also allow use at high temperature, according to the previous publication (Brandon et al., 2018). However, we found that TxGH116 was already quite tolerant of ionic solvents and the modification did not seem to improve the tolerance to the solvent or to high temperature and gave a lower activity of protein. Although the lower activity may be partly due to an issue of quantifying the amount of enzyme after modification, the lack of improvement does not justify the cost of the

modification. Nonetheless, the project did allow us to determine that the TxGH116 beta-glucosidase is quite tolerant of ionic liquids and organic solvents.

Overall, this project allowed us to explore the mechanism of TxGH116 and application of its mutants, as well as characterize the activities of other b-glucosidases and glycoside hydrolases that are active on agricultural biomass, such as rice straw and husk and sugarcane leaves. These enzymes can be used in the future for isolation of glucose and xylose as feedstock for biorefinery applications and for production of new products, such as the α-glucosidase inhibitors and other glycosides produced in this project. It also contributed to the training of at least 4 PhD students and an RA and a postdoc for future work in bio-circular-green economy, and generated 7 publications to date, with more in progress.

Future Research Directions (ข้อเสนอแนะสำหรับงานวิจัยในอนาคต)

Several enzymes and processes generated in this project have strong implications for use in the future. The production of a-azido glucose, which is an expensive commodity when generated chemically can be scaled up and its ecological impact improved by finding ways to recycle the azide left over from the reaction, along with various side products. Scaling up will require both recycling and careful containment to achieve our goal of efficient and ecologically friendly production of large amounts of azidoglucose, which can be used for inhibitor synthesis or sold to scientists around the world to allow them to explore its applications. Finalization of the characterization of other TxGH116 variants will also be published to notify the world of its unique mechanistic characteristics that can open the minds of other investigators to new possibilities.

Now that we have a number of biomass active enzymes in hand, we can try to develop an efficient process for releasing sugars and other valuable components of agricultural biomass found in Thailand, including rice straw and sugarcane leaves. We can further compare to the components found in commercial enzyme mixtures to see whether we need to add other activities or enzymes to achieve an efficient mixture. Moreover, we can try to use the xylan-active enzymes to pretreat for cellulases and vice versa, in order to generate steps where monosaccharides are only released from glucans or xylans for efficient use in further processes. Moreover, we can obtain more enzymes, such as lytic polysaccharide monooxygenases (LPMO) and cellulases with activities different from those of the ones we expressed to help in the breakdown of biomass polysaccharides. The identification of enzymes from microorganisms found in Thai biomass may also help, in terms of commercial application

with fewer intellectual property issues, although Thai groups have worked in this area for many years.

There is still room for improving the production of glycosides, especially production using biomass or biomass-derived polysaccharides as the glycosyl donor, rather than synthetic substrates. In particular, the production and application of xylosides seems less thoroughly explored compared to glucosides. Nonetheless production of glucosides can also be of interest in many cases, since they can improve the solubility of potential drug and nutraceutical compounds. For antibiotics, our experience tells us that this approach is not so good for drugs that act inside bacteria, apparently due to difficulty getting the glycosides into the cell, but they can be used for targeting enzymes on the surface of the cell. Exploring the use of our glycosidases and transglucosidase, as well as glycosyltransferases with appropriate recycling of the donor substrate will be well worth exploring in the future.

References

Baiya S, Pengthaisong S, Kitjaruwankul S, Ketudat Cairns JR 2021. Structural analysis of rice Os4BGlu18 monolignol β-glucosidase. PLoS ONE 16(1): e0241325. https://doi.org/10.1371/journal.pone.0241325.

Brogan APS, Le LB, Hallet JP. (2018) Non-aqueous homogeneous biocatalytic conversion of polysaccharides in ionic liquids using chemically modified glucosidase. Nature Chemistry https://doi.org/10.1038/s41557-018-0088-6.

Charoenwattanasatien R, Pengthaisong S, Breen I, Mutoh R, Sansenya S, Hua Y, Tankrathok, A, Wu L, Songsiriritthigul C, Tanaka H, Williams SJ, Davies GJ, Kurisu G., Ketudat Cairns JR. (2016) Bacterial ß-glucosidase reveals the structural and functional basis of genetic defects in human glucocerebrosidase 2 (GBA2). ACS Chemical Biology 11: 1891–1900.

Dai GY, Yin J, Li KE, Chen DK, Liu Z, Bi FC, Rong C, Yao XN. (2020) The Arabidopsis AtGCD3 protein is a glucosylceramidase that preferentially hydrolyzes long-acyl-chain glucosylceramides. Journal of Biological Chemistry 295(3): 717–728.

Gorantla JN, Pengthaisong S, Choknud S, Kaewpuang T, Manyum T, Promarak V, Ketudat Cairns JR. (2019) Gram scale production of 1-azido-β-D-glucose via enzyme catalysis for the synthesis of 1,2,3-triazole-glucosides. RSC Advances 9: 6211-6220. DOI: 10.1039/c9ra00736a

Gorantla J, Santhi M, Pengthaisong S, Ngiwsara L, Sawangareetrakul P, Chokchaisiri S, Kittakoop P, Svasti J, Ketudat Cairns J. (2021) Chemoenzymatic and protecting-group-free synthesis of 1,4-substituted 1,2,3-triazole-α-D-glucosides with potent inhibitory activity towards lysosomal α-glucosidase. ACS Omega 6(39): 25710–25719. https://doi.org/10.1021/acsomega.1c03928

Huang M, Pengthaisong S, Charoenwattanasatien R, Thinkumrob N, Jitonnom J, Ketudat Cairns JR. (2022) Systematic functional and computational analysis of glucose-binding residues in glycoside hydrolase family GH116. Catalysts 12(3): 343. https://doi.org/10.3390/catal12030343

Koga J, Yazawa M, Miyamoto K, Yumoto E, Kubota T, Sakazawa T, Hashimoto S, Sato M, Yamane H. (2021) Sphingadienine-1-phosphate levels are regulated by a novel glycoside hydrolase family 1 glucocerebrosidase widely distributed in seed plants. Journal of Biological Chemistry 297(5): 101236. https://doi.org/10.1016/j.jbc.2021.101236

Luang S, Hrmova M, Ketudat Cairns JR. (2010) High-level expression of barley β-D-glucan exohydrolase HvExoI from a codon-optimized cDNA in *Pichia pastoris*. Protein Expression and Purification 73, 90-98. DOI: 10.1016/j.pep.2010.04.011

Ngiwsara L, Wattanasirichaigoon D, Tim-Aroon T, Rojnueangnit K, Noojaroen S, Khongkraparn A, Sawangareetrakul P, Ketudat-Cairns JR, Charoenwattanasatien R, Champattanachai V, Kuptanon C, Pangkanon S, Svasti J. (2019) Clinical course, mutations and its functional characteristics of infantile-onset Pompe disease in Thailand. BMC Medical Genetics 20(1): 156.

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 Science 165: 627-638.

Pengthaisong S, Hua Y, Ketudat Cairns JR. (2021) Structural basis for transglycosylation in glycoside hydrolase family GH116 glycosynthases. Archives of Biochemistry and Biophysics 706: 108924 https://doi.org/10.1016/j.abb.2021.108924

Prawisut A, Choknud S, Ketudat Cairns JR. Expression of Rice β-exohydrolase II (OsExoII) in *Escherichia coli*, purification and characterization. Protein Expression and Purification (2020) 175 105708.

Sansenya S, Mutoh R, Charoenwattanasatien R, Kurisu G, Ketudat Cairns JR. 2015. Expression and crystallization of a bacterial glycoside hydrolase family 116 beta-glucosidase from *Thermoanaerobacterium xylanolyticum*. Acta Crystallographica Section F--Structural Biology and Crystallization 71 (1), 41-44. (doi:10.1107/S2053230X14025461)

Shim SH, Mahong B, Lee SY, Kongdin M, Lee C, Kim YJ, Qu G, Zhang D, Ketudat Cairns JR, Jeon JS. 2022. Rice β-glucosidase Os12BGlu38 is required for synthesis of intine cell wall and pollen fertility. Journal of Experimental Botany 73(3): 784–800. https://doi.org/10.1093/jxb/erab439

Output จากโครงการวิจัยที่ได้รับทุนจาก สกว.

1. Publications in International Journals **ผลงานตีพิมพ์ในวารสารวิชาการนานาชาติ (ระบุชื่อผู้แต่ง ชื่อเรื่อง ชื่อวารสาร ปี เล่มที่ เลขที่ และหน้า) หรือผลงานตามที่คาดไว้ในสัญญาโครงการ**

1.1. Prawisut A, Choknud S, Ketudat Cairns JR. Expression of Rice β-exohydrolase II (OsExoII) in Escherichia coli, purification and characterization. *Protein Expression and Purification* (2020) 175 105708. (ISI JCR 2019 Impact factor: 1.513)

1.2. Kuntothom T, Ketudat Cairns JR. Expression and characterization of TbCel12A, a thermophilic endoglucanase with potential in biomass hydrolysis. *Biocatalysis and Agricultural Biotechnology* (2020) 30: 101835. (not currently ranked in ISI)

1.3. Baiya S, Pengthaisong S, Kitjaruwankul S, Ketudat Cairns JR Structural analysis of rice Os4BGlu18 monolignol β-glucosidase. *PLoS ONE* (2021) 16(1): e0241325. (ISI JSR 2019 impact factor: 2.870)

1.4. Pengthaisong S, Hua Y, Ketudat Cairns JR. Structural basis for transglycosylation in glycoside hydrolase family GH116 glycosynthases. *Archives of Biochemistry and Biophysics* (2021) 706: 108924. https://doi.org/10.1016/j.abb.2021.108924 (Accepted 10 May, 2021) (ISI JSR 2020 impact factor: 4.013).

1.5. Gorantla J, Santhi M, Pengthaisong S, Ngiwsara L, Sawangareetrakul P, Chokchaisiri S, Kittakoop P, Svasti J, Ketudat Cairns J. Chemoenzymatic and protecting-group-free synthesis of 1,4-substituted 1,2,3-triazole-α-D-glucosides with potent inhibitory activity towards lysosomal α-glucosidase. *ACS Omega* (2021) 6(39): 25710–25719. https://doi.org/10.1021/acsomega.1c03928 (Accepted 9 September, 2021) (ISI JSR 2020 impact factor: 3.512).

1.6. Shim SH, Mahong B, Lee SY, Kongdin M, Lee C, Kim YJ, Qu G, Zhang D, Ketudat Cairns JR, Jeon JS. Rice β-glucosidase Os12BGlu38 is required for synthesis of intine cell wall and pollen fertility. *Journal of Experimental Botany* (2022) 73(3): 784–800. https://doi.org/10.1093/jxb/erab439 (Accepted 22 September, 2021) (ISI JSR 2020 impact factor: 6.992)

1.7. Huang M, Pengthaisong S, Charoenwattanasatien R, Thinkumrob N, Jitonnom J, Ketudat Cairns JR. Systematic functional and computational analysis of glucose-binding residues in glycoside hydrolase family GH116. *Catalysts* (2022) 12(3): 343. https://doi.org/10.3390/catal12030343 (Accepted 14 March, 2022) (ISI JCR 2020 IF: 4.146)

2. Application of Research Results (การนำผลงานวิจัยไปใช้ประโยชน์)
   - Commercial applications (เชิงพาณิชย์ (มีการนำไปผลิต/ขาย/ก่อให้เกิดรายได้ หรือมีการนำไปประยุกต์ใช้โดยภาคธุรกิจ/บุคคลทั่วไป)) We were able to produce α-azidoglucose in an amount that is nearly adequate for commercialization, but we still need to improve the process to upscale it, as noted in the Discussion and Conclusions.
   -
   - Policy effects เชิงนโยบาย (มีการกำหนดนโยบายอิงงานวิจัย/เกิดมาตรการใหม่/เปลี่ยนแปลงระเบียบข้อบังคับหรือวิธีทำงาน) –not applicable.
   - Public Output เชิงสาธารณะ (มีเครือข่ายความร่วมมือ/สร้างกระแสความสนใจในวงกว้าง) We extended our collaboration with groups around the world, however, due to the covid-19 pandemic, we were less involved with the public.
   - Academic output เชิงวิชาการ (มีการพัฒนาการเรียนการสอน/สร้างนักวิจัยใหม่) This project helped to train 4 PhD students, 1 postdoc and 1 research assistant. The methods and thought processes that they developed can be used to build a BCG economy based on the utilization of biomass from agricultural waste, as well as development of new chemical products.
3. Other Output (อื่นๆ (เช่น ผลงานตีพิมพ์ในวารสารวิชาการในประเทศ การเสนอผลงานในที่ประชุมวิชาการ หนังสือ การจดสิทธิบัตร))

3. 1. Presentations at International and National Conferences

3.1.1. Ketudat Cairns J. (2020) Carbohydrate-active enzymes from rice and related compounds: Discovery & application. The 2nd International Seminar on Smart Molecules for Natural Resources. Universitas Brawijaya, Malang, Indonesia (online) 25-26 August, 2020. Keynote speech 5.

3.1.2. Ketudat Cairns JR. (2020) Chemical applications of beta-glucosidases and related enzymes. International Seminar on Chemistry. ITS Sukolilo Campus, Surabaya, Indonesia (online) 7-8 October, 2020. Keynote speech 3.

3.1.3. Pengthaisong S, Ketudast Cairns JR. (2020) Structure-function relationships in oligosaccharide synthesis by Thermoanaerobacterium xylanolyticum TxGH116 glycosynthase. The 15th International Symposium of the Protein Society of Thailand, Chulabhorn Research Institute Convention Center, 4-6 November, 2020. PP021.

3.1.4. Huang M, Pengthaisong S, Ketudat Cairns JR. The roles of the glycone sugar-binding residues of TxGH116 in glucose binding and substrate hydrolysis. The 15th International Symposium of the Protein Society of Thailand, Chulabhorn Research Institute Convention Center, 4-6 November, 2020. PP022.

3.1.5. Beagbandee C, Charoenwattanasatien R, Pengthaison S, Ketudat-Cairns JR. Characterization and structure determination of TeGH116, a cyanobacterial glycoside hydrolase family 116 beta-glucosidase. The 15th International Symposium of the Protein Society of Thailand, Chulabhorn Research Institute Convention Center, 4-6 November, 2020. PP023.

3.1.6. Choknud S, Prawisut A, Ketudat Cairns JR. (2020) Transglucosylation activity of rice - exoglucanse II (OsExoII). The 15th International Symposium of the Protein Society of Thailand, Chulabhorn Research Institute Convention Center, 4-6 November, 2020. PP037.

3.1.7. Ketudat-Cairns JR, Svasti, J. The Asia Pacific Protein Association and Collaboration with Asia-Pacific Partners. APPA/PSSJ Joint Workshop: Toward International Cooperation in Protein Science, Annual Symposium of the Protein Science Society of Japan, 17 June, 2021, Zoom, Japan.

3.1.8. Beagbandee C, Jatooratthawichot P, Ketudat Cairns JR. (2021) Characterization of a cellobiohydrolase from the thermophilic bacterium *Thermothelomyces thermophilus* (*Tt*Cel7) produced in recombinant *Escherichia coli*. 16th International Online Mini-Symposium of the Protein Society of Thailand. November 17-18, 2021. Faculty of Science, Mahidol University (online). Poster PP60. Poster PP059. Abstracts & Proceedings pg. 78.

3.1.9. Pengthaisong S, Ketudat Cairns JR. Immobilization of *Thermoanaerobacterium xylanolyticum Tx*GH116 and E441G nuclophile mutant. 16th International Online Mini-Symposium of the Protein Society of Thailand. November 17-18, 2021. Faculty of Science, Mahidol University (online). Poster PP61.