รายงานวิจัยฉบับสมบูรณ์

โครงการ: การจำแนกและศึกษาคุณสมบัติของ carbohydrate-binding module จากเอนไซม์เชิงซ้อนไซลาโนไลติกและเซลลูโลไลติกที่ผลิตจาก *Paenibacillus xylaniclasticus*

โดย ดร.สมพิศ สอนโยธา และคณะ

กรกฎาคม พ.ศ. 2561

สัญญาเลขที่ TRG5680025

รายงานวิจัยฉบับสมบูรณ์

โครงการ: การจำแนกและศึกษาคุณสมบัติของ carbohydrate-binding module จากเอนไซม์เชิงซ้อนไซลาโนไลติกและเซลลูโลไลติกที่ผลิตจาก *Paenibacillus xylaniclasticus*

คณะผู้วิจัย	สังกัด
ดร.สมพิศ สอนโยธา	สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
รศ.ดร.กนก รัตนะกนกชัย	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี

สนับสนุนโดยสำนักงานกองทุนสนับสนุนการวิจัย

และ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

Abstract

Project Code: TRG5680025

Project Title: Identification and characterization of carbohydrate-binding module from xylanolytic-cellulolytic multienzyme complexs-producing bacterium, *Paenibacillus xylaniclasticus*

Investigator: Dr.Somphit Sornyotha King Mongkut's Institute of Technology Ladkrabang

E-mail Address: somphit.so@kmitl.ac.th, kssomphi@staff.kmitl.ac.th, somphitsorn@yahoo.com

Project Period: 2 years

A facultative anaerobic bacterium *Paenibacillus xylaniclasticus* strain TW1, isolated from an anaerobic digester fed with pineapple waste, produces an extracellular xylanolytic-cellulolytic multienzymes system when grown on corn hull under aerobic conditions. The crude enzyme was found to be capable of binding to insoluble xylan and cellulose and efficiently hydrolyzed raw agricultural residues. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) of a crude enzyme preparation exhibited at least 12 proteins could be bound to cellulose. Moreover, the results from matrix-assisted laser desorption/ionization time-of-flight/time-of-flight mass spectrometry (MALDI-TOF/TOF-MS) found that the peptide mass fingerprints of the unknown components from SDS-PAGE were almost identical as unknown proteins. This result suggested that the other proteins possibly present as new glycoside hydrolase family. The only two bands from SDS-PAGE could be identified as the glycoside hydrolase family 31 (CBP5) and family 11 (CBP12). Although CBP5 showed similarity to published glycoside hydrolase family 31 sequences from *P. curdlanolyticus* YK9 but no information on the carbohydrate-binding module (CBM) appearing in the glycoside hydrolase family 31 from *P. curdlanolyticus* YK9 was presented. In case of CBP12, the peptide sequence of CBP12 showed similarity to *P. curdlanolyticus* B6 xylanase family 11 with a calcium-dependent-carbohydrate-binding module family 36 (CBM36). However, the only CBM36 currently was characterized, that from *P. polymyxa* ATCC 842 xylanase 43A (*Pp*CBM36). In addition, CBM36 from *P. xylaniclasticus* TW1 showed only 50% amino acid sequence identity with CBM36 of the strain ATCC 842. Thus, the CBM36 from *P. xylaniclasticus* TW1 (*Px*TW1CBM36) was still attracted and, henceforth, was produced and characterized of this CBM.

*Px*TW1CBM36 was cloned, expressed, purified and studied for binding characteristics. The results found that *Px*TW1CBM36 displayed broad binding ability to carbohydrates which had a high affinity for xylans and insoluble cellulose. Interestingly, this is the first study that indicated that CBM36 had an affinity for insoluble cellulose. Although the ligand-binding apparatus of *Px*TW1CBM36 and *Pp*CBM36 display similarity, the binding capacity of *Px*TW1CBM36 do not perturb by the addition of the chelating agent such as ethylenediaminetetraacetic acid (EDTA). It is possible that calcium-dependency might not be the only mode of substrate binding by *Px*TW1CBM36. In addition, the binding characteristic of *Px*TW1CBM36 on carbohydrates embedded within plant cell walls was also elucidated. It displayed the strong recognition for ligands located in the epidermal tissue of sweet potato roots. This result shows that the cell wall components and the polysaccharides configuration in the plant cell wall might differ between tissues. Thus, this CBM can represent valuable tools with which to probe the intricate architecture of plant cell walls, and also possible as a new tool for immobilization enzyme for targeting enzymes to surface of plant tissue.

Furthermore, in this report not only carbohydrate-binding module but carbohydrate-active enzymes in *Paenibacillus xylaniclasticus* TW1 genome were also investigated. From nucleotide sequence, the data exhibited many carbohydrate-active enzymes in *P. xylaniclasticus* TW1 genome. It contained 28 xylanolytic enzymes, 9 cellulolytic enzymes and 3 other plant cell wall-degrading enzymes. In addition, not only CBM36 but *P. xylaniclasticus* TW1 contained many CBM which different family. As many reports, the complete and rapid hydrolysis of polysaccharides in agricultural residues requires not only cellulolytic enzymes but also the cooperation of many enzymes such as xylanolytic enzymes and side chain-cleaving enzymes such as deacetylase and α-L-arabinofuranosidase. Thus, this result explained that why crude enzyme produced from *P. xylaniclasticus* TW1 displayed the capable of binding to insoluble xylan and cellulose and efficiently hydrolyzed raw agricultural residues.

Keywords: Carbohydrate-binding module family 36, MALDI-TOF/TOF MS, *Paenibacillus xylaniclasticus*, Xylanolytic-cellulolytic multienzyme complexes

บทคัดย่อ

รหัสโครงการ: TRG5680025

ชื่อโครงการ: การจำแนกและศึกษาคุณสมบัติของ carbohydrate-binding module จากเอนไซม์เชิงซ้อนไซลาโนไลติกและเซลลูโลไลติกที่ผลิตจาก *Paenibacillus xylaniclasticus*

ชื่อนักวิจัย: ดร.สมพิศ สอนโยธา สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

E-mail Address: somphit.so@kmitl.ac.th, kssomphi@staff.kmitl.ac.th, somphitsorn@yahoo.com

ระยะเวลาโครงการ: 2 ปี

แบคทีเรีย *Paenibacillus xylaniclasticus* สายพันธุ์ TW1 ซึ่งเป็นแบคทีเรียที่สามารถเจริญได้ทั้งในสภาวะที่มีและไม่มีอากาศ (facultative anaerobic bacterium) ถูกคัดแยกจากถังหมักก๊าซชีวภาพที่มีเปลือกสับปะรดเป็นวัตถุดิบ สามารถผลิตเอนไซม์เชิงซ้อนในกลุ่มไซลาโนไลติกและเซลลูโลไลติก (xylanolytic-cellulolytic multienzyme complexes) เมื่อเพาะเลี้ยงในอาหารที่มีเปลือกข้าวโพดเป็นแหล่งคาร์บอนภายใต้สภาวะที่มีอากาศ โดยเอนไซม์เชิงซ้อนดังกล่าวสามารถยึดเกาะกับไซแลนและเซลลูโลสที่ไม่ละลายน้ำได้ และยังสามารถย่อยสลายวัสดุเหลือใช้ทางการเกษตรได้โดยไม่ต้องผ่านกระบวนการปรับสภาพอีกด้วย เมื่อตรวจสอบด้วยเทคนิค sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) พบว่าเอนไซม์เชิงซ้อนประกอบด้วยโปรตีนหน่วยย่อยอย่างน้อย 12 ชนิดที่สามารถยึดเกาะกับเซลลูโลสที่ไม่ละลายน้ำได้ และเมื่อนำโปรตีนทั้ง 12 ชนิดไปตรวจสอบด้วยเทคนิค matrix-assisted laser desorption/ionization time-of-flight/time-of-flight mass spectrometry (MALDI-TOF/TOF-MS) เพื่อระบุชนิดของโปรตีนดังกล่าว พบว่าโปรตีนที่ยึดเกาะกับเซลลูโลสที่ไม่ละลายน้ำได้ส่วนใหญ่ไม่สามารถระบุชนิดของโปรตีนได้ เป็นไปได้ว่าโปรตีนเหล่านี้อาจจัดอยู่ใน glycoside hydrolase แฟมิลี่ใหม่ที่ยังไม่มีรายงานในปัจจุบัน จากการศึกษาดังกล่าวพบโปรตีนเพียง 2 ชนิดที่สามารถระบุชนิดได้ โดยโปรตีน CBP5 มีความคล้ายคลึงกับเอนไซม์ glycoside hydrolase แฟมิลี่ 31 ที่ผลิตโดย *P. curdlanolyticus* YK9 แต่เอนไซม์ดังกล่าวไม่มีรายงานว่าพบ carbohydrate-binding module เป็นองค์ประกอบ ในขณะที่โปรตีน CBP12 มีความคล้ายคลึงกับเอนไซม์ไซลาเนสแฟมิลี่ 11 ที่ผลิตโดย *P. curdlanolyticus* B6 และมี carbohydrate-binding module แฟมิลี่ 36 (CBM36) เป็นองค์ประกอบ ซึ่งในปัจจุบันมีการศึกษาคุณสมบัติของ CBM36 จากเอนไซม์ไซลาเนส 43A ที่ผลิตโดย *P. polymyxa* ATCC 842 เพียงรายงานเดียว และ CBM36 จาก *P. xylaniclasticus* TW1 มีลำดับของกรดอะมิโนคล้ายคลึงกับ CBM36 ของ *P. polymyxa* ATCC 842 (*Pp*CBM36) เพียงร้อยละ 50 ดังนั้น CBM36 จาก *P. xylaniclasticus* TW1 จึงยังคงมีความน่าสนใจและได้ถูกนำไปศึกษาคุณสมบัติต่าง ๆ ต่อไป

CBM36 ของ *P. xylaniclasticus* TW1 (*Px*TW1CBM36) ได้ถูกโคลนและนำมาแสดงออก จากนั้นถูกทำให้บริสุทธิ์เพื่อนำไปศึกษาคุณสมบัติต่าง ๆ ต่อไป จากการศึกษาพบว่า *Px*TW1CBM36 สามารถยึดเกาะกับคาร์โบไฮเดรตได้หลายชนิด โดยสามารถยึดเกาะกับไซแลนและเซลลูโลสที่ไม่ละลายน้ำได้ดีที่สุด ซึ่งการศึกษานี้เป็นรายงานฉบับแรกที่พบว่า CBM36 สามารถยึดเกาะกับเซลลูโลสได้ และจากรายงานก่อนหน้านี้พบว่า CBM36 จัดเป็น calcium-dependent-carbohydrate-binding module ซึ่งต้องอาศัยแคลเซียมในการยึดเกาะกับสับสเตรท โดยจากการเปรียบเทียบลำดับกรดอะมิโนระหว่าง *Px*TW1CBM36 และ *Pp*CBM36 พบว่ากรดอะมิโนที่เกี่ยวข้องกับการยึดเกาะกับสับสเตรทของ CBM36 จากทั้ง 2 สายพันธุ์มีความคล้ายคลึงกัน แต่จากการศึกษานี้กลับพบว่าแม้จะมีการเติม EDTA ซึ่งเป็น chelating agent ลงไป แต่ *Px*TW1CBM36 ยังคงสามารถยึดเกาะกับคาร์โบไฮเดรตได้ ดังนั้นอาจเป็นไปได้ว่า *Px*TW1CBM36 นอกจากจะอาศัยแคลเซียมในการยึดเกาะกับสับสเตรทแล้ว ยังอาจมีกรดอะมิโนบริเวณอื่นที่มีส่วนเกี่ยวข้องกับการยึดเกาะกับสับสเตรทได้โดยไม่ต้องอาศัยแคลเซียม นอกจากนี้ยังได้ศึกษาความสามารถในการยึดเกาะของ *Px*TW1CBM36 กับคาร์โบไฮเดรตที่เป็นองค์ประกอบของผนังเซลล์ของรากมันเทศ พบว่า *Px*TW1CBM36 มีความจำเพาะในการยึดเกาะกับคาร์โบไฮเดรตที่เป็นองค์ประกอบของผนังเซลล์ที่บริเวณเนื้อเยื่อ epidermal tissue จากการศึกษาครั้งนี้ทำให้ทราบว่าผนังเซลล์ของเนื้อเยื่อพืชแต่ละบริเวณอาจมีชนิดและองค์ประกอบของคาร์โบไฮเดรตที่แตกต่างกัน ดังนั้น CBM36 ที่ถูกศึกษาในครั้งนี้อาจสามารถนำไปใช้เป็นเครื่องมือในการศึกษาชนิดและองค์ประกอบของคาร์โบไฮเดรตในผนังเซลล์พืชต่างๆ รวมทั้งสามารถนำไปเชื่อมต่อกับเอนไซม์เพื่อให้มีเป้าหมายในการย่อยสลายที่บริเวณพื้นผิวด้านนอกของเนื้อเยื่อพืชได้

นอกจากนี้ยังได้มีการศึกษาลำดับนิวคลีโอไทด์ของจีโนมที่สกัดจาก *P. xylaniclasticus* TW1 พบว่าจีโนมของแบคทีเรียสายพันธุ์ TW1 มียีนที่ผลิตเอนไซม์ในกลุ่มไซลาโนไลติกอย่างน้อย 28 ชนิด เอนไซม์ในกลุ่มเซลลูโลไลติกอย่างน้อย 9 ชนิด และเอนไซม์อื่น ๆ ที่ย่อยสลายผนังเซลล์พืชอีกอย่างน้อย 3 ชนิด และพบว่านอกจาก CBM36 แล้ว ยังสามารถพบ CBM แฟมิลี่อื่น ๆ อีกหลายแฟมิลี่ที่ผลิตได้โดยแบคทีเรียสายพันธุ์ TW1 ซึ่งจากรายงานหลายฉบับได้กล่าวไว้ว่าการที่จุลินทรีย์จะสามารถย่อยสลายวัสดุเหลือใช้ทางเกษตรได้อย่างมีประสิทธิภาพนั้น จำเป็นต้องอาศัยการทำงานร่วมกันของเอนไซม์ทั้งในกลุ่มเซลลูโลไลติก กลุ่มไซลาโนไลติก รวมทั้งเอนไซม์ที่ย่อยสลายกิ่งก้านต่าง ๆ อาทิเช่น เอนไซม์ deacetylase และ α-L-arabinofuranosidase เป็นต้น ด้วยเหตุนี้จึงแสดงให้เห็นว่าเพราะเหตุใดเอนไซม์ที่ผลิตจาก *P. xylaniclasticus* TW1 จึงมีความสามารถในการยึดเกาะได้ทั้งกับไซแลนและเซลลูโลสที่ไม่ละลายน้ำ และมีความสามารถในการย่อยสลายวัสดุเหลือใช้ทางการเกษตรชนิดต่าง ๆ ที่ไม่ผ่านกระบวนการปรับสภาพได้อย่างมีประสิทธิภาพ

คำหลัก: Carbohydrate-binding module แฟมิลี่ 36, MALDI-TOF/TOF MS, *Paenibacillus xylaniclasticus*, เอนไซม์เชิงซ้อนในกลุ่มไซลาโนไลติกและเซลลูโลไลติก

Acknowledgement

Project research over many years is not accomplished without the support of myriad of people. It is impossible to thank everyone who contributed to my happiness and success. To those unnamed here, your gestures and assistance have not gone unnoticed by me.

First and foremost, I would like to thank Assoc. Prof. Dr. Khanok Ratanakhanokchai for years of constant guidance and encouragement. His example will continue to guide me through the future years while their extraordinary ability to handle many the tasks is an inspiration to me.

I would like to thank Prof. Dr. Shuichi Karita and Assoc. Prof. Dr. Tohru Tsuchiya of Mie University Japan, for being a wonderful mentor about gene manipulations and for critically evaluating my research through the years.

This work would not have been complete without the financial support from the Thailand Research Fund (TRF Grant) and King Mongkut's Institute of Technology Ladkrabang.

Finally, I am very grateful to my family and friends, who have been of great support and help during my research. I wish to express my sincere and heartfelt gratitude to them for will power.

Contents

Contents (Cont.)

List of tables

List of figures

List of figures (Cont.)

Chapter 1 Introduction

1.1 Introduction to the research problem and its significance

Efficient conversion of lignocellulosic biomass such as agricultural wastes into fuel ethanol has become a world priority for producing environmentally friendly renewable energy at reasonable price for the transportation sector. Fuel ethanol can be utilized as oxygenate of gasoline elevating its oxygen content, allowing a better oxidation of hydrocarbons and reducing the amounts of greenhouse gas emissions into the atmosphere. Moreover, lignocellulosic biomass, especially agricultural wastes are attractive feedstock for bioethanol production, due to Thailand is an agricultural country, a large amount of agricultural residues are produced. Thus, these materials will provided sufficient supply for sustainable development.

The use of agricultural wastes as a renewable source of energy via breakdown to sugars that can then be converted to liquid fuel is of great interest. Agricultural wastes contain a complex mixture of polysaccharides, such as cellulose, hemicellulose (xylan and galactomannan), pectic substances (galacturonan and arabinogalactan), and other polysaccharides (e.g., type II arabinogalactan and fucoxyloglucan). The hemicellulose and pectin polysaccharides, as well as the aromatic polymer lignin, interact with the cellulose fibrils, creating a rigid structure strengthening the plant cell wall. Therefore, complete and rapid hydrolysis of these polysaccharides requires not only β-1,4-glucosidic chain-cleaving enzymes such as endo-β-1,4-glucanase, cellobiohydrolase, and β-glucosidase but also the cooperation of many enzymes such as xylanolytic enzymes and side chain-cleaving enzymes, such as β-1,4-xylanase and α-L-arabinofuranosidase, respectively (Murashima et al., 2002).

Various anaerobic bacteria have been reported to produce cellulases and xylanases containing carbohydrate-binding modules (CBMs) like cellulose-binding domains (CBDs) and/or xylan-binding domains (XBDs) which are associated with a discrete, high-molecular-weight cellulolytic or xylanolytic enzyme complex known as the cellulosome (Bayer et al., 2004). These complexes are dedicated to hydrolyzing agricultural substances because of their ability to bind to insoluble cellulose and/or xylan via CBDs and/or XBDs, respectively. The CBMs effect binding to the substrate surface, presumably to facilitate hydrolysis by bringing the catalytic domain in close proximity to the insoluble substrate. Thus, a tight interaction between the enzymes and their

substrates via CBMs and the cooperation of multiple plant cell wall-degrading enzymes into a multienzyme complex has advantages over single enzyme systems.

Previous work in our laboratory showed that a facultatively anaerobic bacterium *P. xylaniclasticus* strain TW1 could be produced xylanolytic-cellulolytic multienzyme complex under aerobic cultivation (Tachaapaikoon et al., 2012a, 2012b) which easy-to-handle than other multienzyme-producing anaerobacteria. Furthermore, xylanolytic-cellulolytic multienzyme complex from *P. xylaniclasticus* strain TW1 had a more efficient non-pretreated-agricultural residues-degradation than those from *P. curdlanolyticus* B-6 and *Clostridium thermocellum* ATCC 27405 and commercial cellulases (data not shown). Therefore, xylanolytic-cellulolytic multienzyme complex produced from *P. xylaniclasticus* strain TW1 is a potential candidate for the production of fermentable sugars from agricultural wastes. Nevertheless, although it could degrade the agricultural waste effectively, the production of enzyme from this bacterium is still limited by the low yield concentration. Then, the large cultivation was handled for the saccharification, it was one of the main factors for determination of the cost of enzyme. For this solution, reuse or recycling enzyme will be needed. The CBM immobilization is one of the ideas for recycling enzymes. But some CBMs could be bound to carbohydrates tightly, and could not release from the carbohydrates under moderate condition. Anyway, understanding properties of CBM from *P. xylaniclasticus* strain TW1 is a key for the next steps. Hence, characterization of CBM from xylanolytic-cellulolytic multienzyme complex-producing bacterium, *P. xylaniclasticus* strain TW1 will be required for solving the problem and improving the yield of sugars and enzyme costs.

1.2 Objectives

1.2.1 To identify the secreted enzymes appear to contain CBM by using affinity purification and applying MALDI-TOF/TOF MS.

1.2.2 To study the binding characteristics and specificities of CBM against isolated polysaccharides.

1.2.3 To indicate the substrate available to the parent enzymes during the initial rate limiting phase of the degradative process by seeking the recognition site of CBM on carbohydrates embedded within plant cell walls.

1.3 Scope of research

The research were consist of a set of tasks grouped into two Macro-Tasks:

MT 1: Isolation and identification of carbohydrate-binding proteins

MT 1.1: Isolation of carbohydrate-binding proteins by using affinity purification on insoluble substrate

MT 1.2: Identification of CBM-containing protein by using MALDI-TOF/TOF MS

MT 2: Production and characterization CBM-Cyan Fluorescent Protein (CFP)

MT 2.1: Cloning, expression and purification of CBM-CFP

MT 2.2: Substrate-binding assays

MT 2.3: CBM-labeling

1.4 Expected benefits

The binding ability of enzyme to substrate is very important role for degradation and hydrolysis. Recently, xylanolytic-cellulolytic multienzyme complex of *P. xylaniclasticus* strain TW1 was characterized and reported. It showed the high efficient to degraded plant cell wall. Therefore, the unique binding ability of multienzyme complex from *P. xylaniclasticus* strain TW1 was monitored in this study. It was able to get the new knowledge and understanding the binding properties of enzyme by this approach and also possible to get a new tool for immobilization enzyme for targeting enzymes to surface of plant tissue.

Chapter 2 Literature review

2.1 lignocellulosic biomass and compositions of lignocellulosic biomass

Lignocellulosic biomass is the non-starch, fibrous part of plant material useful for its renewability and abundance (Perlack et al., 2005). Lignocellulosic biomass like wood, grass and agricultural wastes, such as rice straw, rice husk, sugarcane bagasse, corn stover, corn hull and etc. is a potential raw material for bioconversion to simple sugars and subsequently to desired chemicals like ethanol.

Lignocellulosic biomass has three basic components consisting of cellulose, hemicelluloses and the lignin fraction (Carvalheiro et al., 2008). In addition, there are low amounts of acids, salts, and minerals while in some, there may be minute concentrations of proteins, as shown in Figure 2.1 (Higuchi, 1997). The percentage compositions of these constituents vary but cellulose is commonly the largest fraction, comprising of 30 - 50%, hemicellulose at 20 - 30% and lignin at 15 - 20% of total the dry matter. Cellulose and hemicellulose are macromolecules from different sugars, whereas lignin is an aromatic polymer synthesized from phenylpropanoid precursors. The composition and percentages of these polymers vary from one plant species to other. Moreover, the composition within a single plant varies with age, stage of growth, and other conditions.

2.1.1 Cellulose

Cellulose is a long chain polymer with repeating units of D-glucose. In the cellulose chain, the glucose units are in 6-membered rings, called pyranoses. In nature, the glucose units will form acetal linkages between the C-1 of glucose monomer and the C-4 of the next monomer called β-1,4-glycosidic bond (Figure 2.2). The basic molecular structure is a linear polymer consisting of up to 104 D-glucose molecules that are arranged in fibrils. The fibrils consist of several parallel cellulose molecules stabilized by hydrogen bonds. Although individually the hydrogen bonds are relatively weak, collectively they become a strong associative force as the degree of polymerization increases. The fibrils are organized into a “paracrystalline” state, thus adding to the structural rigidity of cellulose. Over three-quarters of the cellulose structure is considered to exist in these enzymatically resistant, crystalline regions and the remainder comprises the relatively easily hydrolyzable “amorphous” areas (Bayer et al., 1998).

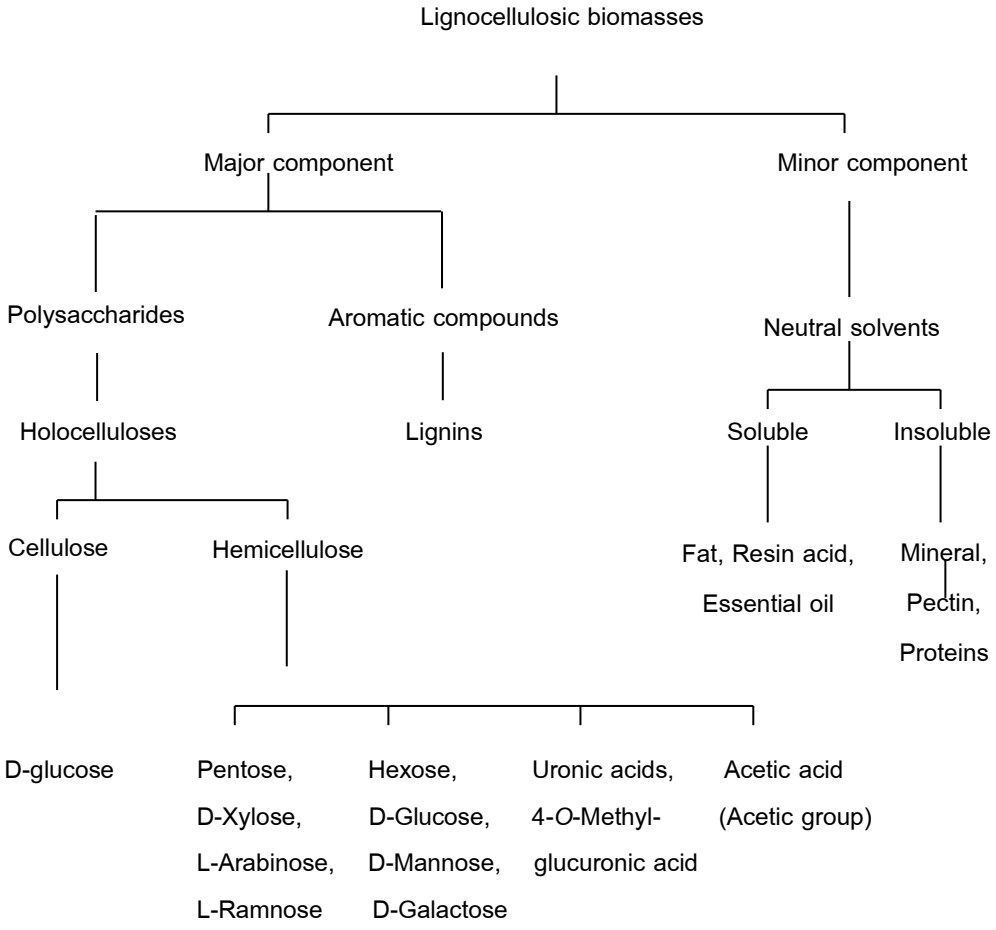


Figure 2.1 Chemical composition of lignocellulosic biomass (Higuchi, 1997).

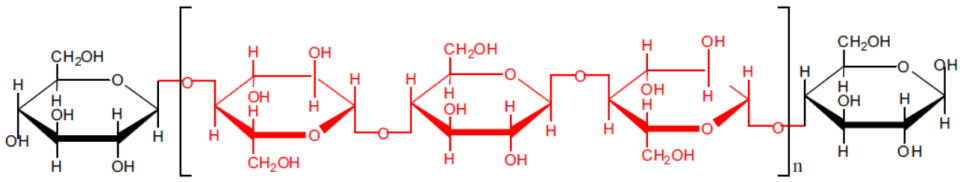
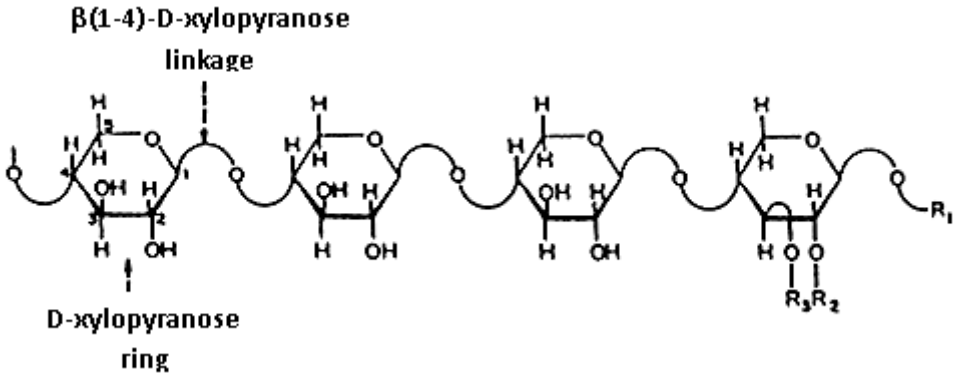


Figure 2.2 Conformation of cellulose (Juturu and Wu, 2014).

2.1.2 Hemicellulose

Hemicelluloses are low-molecular-weight polysaccharides and usually considered to be structural polysaccharides. Hemicelluloses are heteroglycans and one of the three major naturally plant biomass. Together with cellulose and lignin, hemicellulose built up the supporting material in plant cell wall (Puls and Schuseil, 1993). Xylan is a major component of hemicellulose. It

occurs as heteropolysaccharides, containing different groups in the backbone chain and in the side chain (Biely, 1985). The common substituents found in the xylan backbone are acetyl, arabinosyl, and glucuronosyl residues (Figure 2.3).



R_1 : ρ Coumaric acid;

α-D-Glucuronic acid

α-D-Arabinofuranose

R_2 : Acetyl group;

α-D-Glucuranopyranose [Cα Lignin]

4-O-Methyl-α-D-Glucuranopyranose [Cα Lignin]

α-L-Arabinofuranose [Cα and Cδ Lignin, Ferulic acid, Acetyl group, ρ Coumaric acid]

Cα Lignin

R_3 : Acetyl group;

α-L-Arabinofuranose [Cα and Cδ Lignin, Ferulic acid, Acetyl group, ρ Coumaric acid]

β-D-Galactopyranosyl (1,5) α-L-Arabinofuranose

β-D-Xylopyranosyl (1,2) α-L-Arabinofuranose

α-L-Arabinofuranosyl (1,2; 1,3; 1,2,3 Arabinofuranose) n

β-D- Galactopyranosyl (1,4) D-Xylopyranosyl (1,2) α-L-Arabinofuranose

Figure 2.3 Chemical structure of xylans (Kulkarni et al., 1999).

Homoxylans, on the other hand, consist exclusively of xylosyl-residues. This type of xylan is not widespread in nature and has been isolated from esparto grass and tobacco stalks. Wood xylans exist as *O*-acetyl-4-*O*-methylglucuronoxylans in hardwoods (Figure 2.4) or arabino-4-*O*-methylglucuronoxylans in softwood (Figure 2.5). The degree of polymerization of hardwood xylans (150-200) is higher than that of softwood (70-130). Hardwood xylans are highly acetylated (e.g.,

birch xylan) presence of these acetyl groups is responsible for the partial solubility of xylan in water. Acetyl groups are readily removed when xylan is subjected to alkali extraction.

The cereal xylans are made up of D-glucuronic acid and/or its 4-*O*-methyl ether and arabinose (Figure 2.6). Endospermic arabinoxylans of annual plants, also called pentosan, are more soluble in water and dilute alkali than xylans of lignocellulosic materials because of their branched structure.

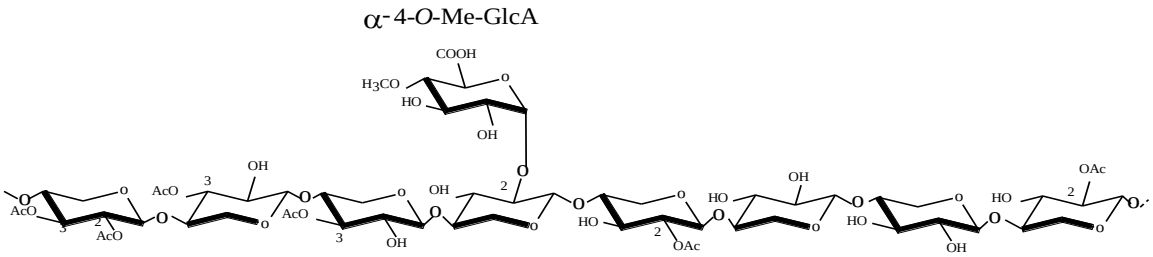


Figure 2.4 Composition of *O*-acetyl-4-*O*-methylglucuronoxylans (hardwoods xylan). Numbers indicate the carbon atom at which substitutions take place. Ac: Acetyl group; α-4-*O*-Me-GlcA: α-4-*O*-methylglucuronic acid (Sunna and Antranikian, 1997).

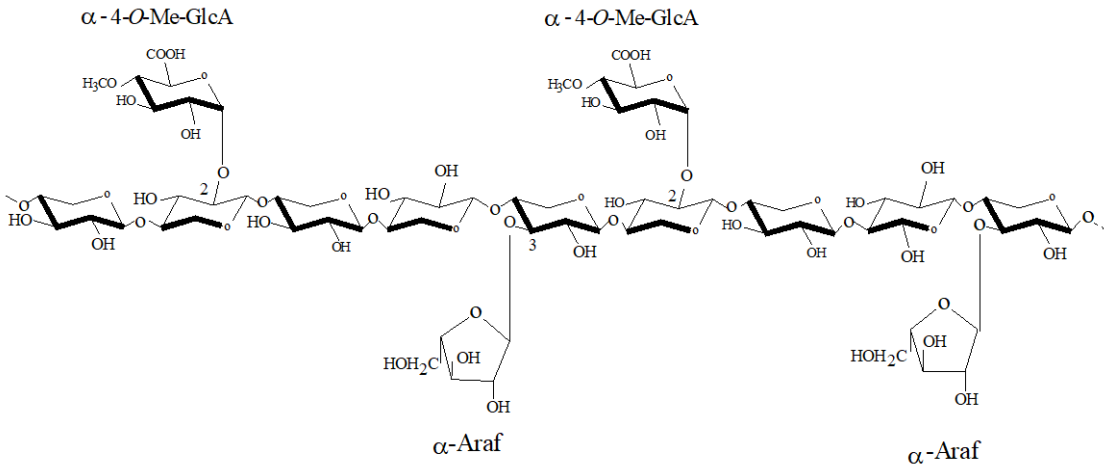


Figure 2.5 Composition of arabino-4-*O*-methylglucuronoxylans (soft wood xylan). Numbers indicate the carbon atom at which substitutions take place. α-Araf: α-Arabinofuronose; α-4-*O*-Me-GlcA: α-4-*O*-methylglucuronic acid (Sunna and Antranikian, 1997).

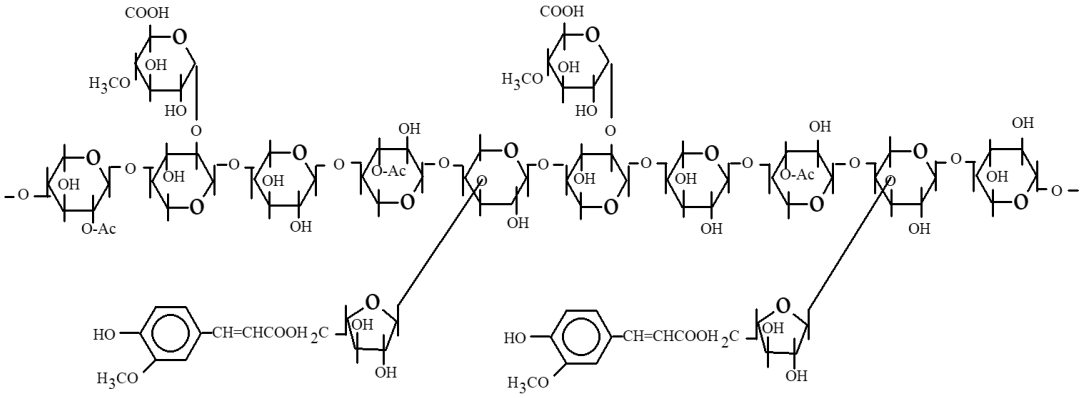


Figure 2.6 Composition of 4-*O*-methylglucuronoxylans (cereal xylan) (Coughlan and Hazlewood, 1993).

2.1.3 Lignin

Lignin is a complex aromatic polymer and has a role in cementing the polysaccharide components in cell walls both chemically and physically thus increasing the mechanical strength of wood as a composite material and its decay resistance toward microorganisms. Lignins are generally classified into three major groups based on their structural monomer units shown in Figure 2.7. Gymnosperm lignin is a dehydrogenation polymer of coniferyl alcohol. Angiosperm lignin is a mixed dehydrogenation polymer of coniferyl and sinapyl alcohols and grass lignin is composed of a mixed dehydrogenation polymer of coniferyl, sinapyl and *p*-coumaryl alcohols (Higuchi, 1978).

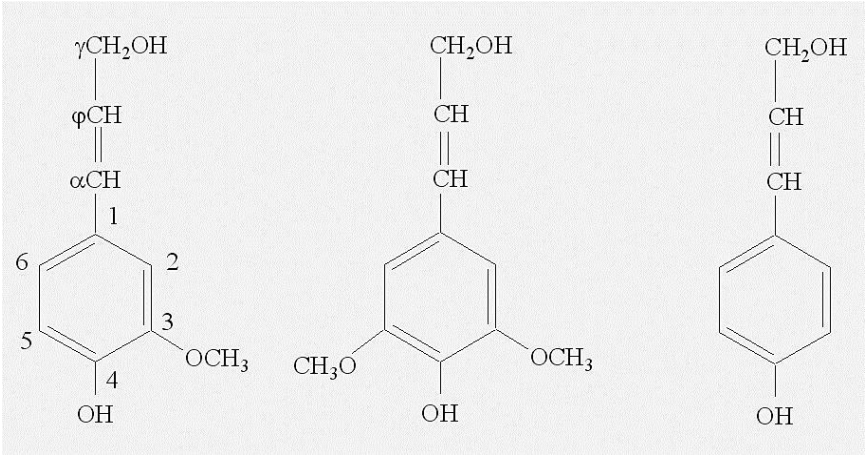


Figure 2.7 Monomer unit of three major groups are coniferyl (A), sinapyl (B) and *p*-coumaryl (C) alcohols (Higuchi, 1978).

2.2 Enzymes required to degrade lignocellulosic biomass

Even though lignocellulosic biomass especially agricultural residues could be converted by chemical methods, it is still controversial since the energy input for pretreatment and chemicals used in the process cost extensively. Therefore, attention is paid to enzymatic method which is less expensive and reduces the environmental impacts.

A number of processes that use lignocellulolytic enzymes or are based on microorganisms could lead to interesting, environmentally friendly technologies. The relationship between cellulose and hemicellulose in the cell walls of higher plants is much more intimate than was previously thought. It is possible that molecules at the cellulose-hemicellulose boundaries and those within the crystalline cellulose domains require different enzymes for efficient hydrolysis. Each polymer is degraded by a variety of microorganisms which produce a battery of enzymes that work synergically. If so, this may help to explain why cellulolytic microorganisms typically synthesize a range of different cellulases with apparently overlapping specificities and why some xylanases carry substrate-binding domains with affinity for cellulose. In any event, a cellulolytic and xylanolytic microbe is not in pure culture, but exists in a mixture with other bacterial and/or fungal species. One or more stains serve as the central polymer degraders, which give their respective simple sugar and other degradation product.

A large variety of enzymes with different specificities are required to degrade all components of lignocellulosic biomass. Table 2.1 gives a brief overview of the types of enzymes that are required to degrade complex lignocellulosic biomasses. However, there are indications that many other proteins may contribute to lignocellulosic degradation in ways that are not yet clearly understood, such as glycoside hydrolases from family 61, expansins and swollenin (Banerjee et al., 2010, Harris et al., 2010, Merino and Cherry, 2007). Han and Chen (2007) also demonstrated that a protein from *Zea mays* had a synergistic interaction with cellulases and increased degradation of biomass even though it did not have any detectable cellulase activity by itself.

Table 2.1 Example of the main enzymes required to degrade lignocellulosic biomass (Dyk and Pletschke, 2012).

Lignocellulosic component	Enzyme hydrolysis
Cellulose	Endoglucanase, cellobiohydrase, β-glucosidase
Hemicellulose	Endo-xylanase, acetyl xylan esterase, β-xylosidase, endo-mannanase, β-mannosidase, α-L-arabinofuranosidase, α-glucuronidase, ferulic acid esterase, α-galactosidase, *p*-coumaric acid esterase
Lignin	Laccase, manganese peroxidase, lignin peroxidase
Pectin	Protopectinase, pectin methyl esterase, pectate lyase, polygalacturonase, rhamnogalacturonan lyase

2.2.1 Cellulolytic enzymes

According to intensive research on *Trichoderma reesei* cellulase system, the concept of cellulolysis is that cellulose is hydrolyzed by the synergistic action of three type of enzymatic activities: (1) endoglucanases, (2) exoglucanases or cellobiohydrolases and (3) β-glucosidase (Table 2.2). Along with these three type of cellulases some accessory enzymes such as lytic oxidases act along with the endoglucanases releasing oligomers with reducing ends and aids in depolymerization of recalcitrant cellulose (Juturu and Wu, 2014).

In addition, based on sequence comparison, the Carbohydrate-Active enZYmes Database (CAZy) provides classification of all cellulases involved in glycoside hydrolase (GH) families that utilize the same catalytic mechanism of acid-base catalysis, with inversion or retention of glucose anomeric configuration. According to the CAZy Database, endoglucanases are grouped in the GH families 5-9, 12, 44-45, 48, 51, 74, and 124. Exoglucanases or cellobiohydrolases are grouped in GH families 1, 3, 5-7, 9 and 48 and β-glucosidases in GH families 1, 3, 5, 9, 30 and 116 (www.cazy.org; Juturu and Wu, 2014).

Table 2.2 Components of aerobic fungal cellulases and their mode of action on the cellulose chain (Bhat and Bhat, 1997).

Enzyme Type	EC Code	Synonym	Mode of Action
Endo-(1,4)-β-D-glucanase	EC 3.2.1.4	Endoglucanase or Endocellulase	--G--G--G--G-- Cleaves linkages at random
Exo-(1,4)-β-D-glucanase	E.C. 3.2.1.176 EC 3.2.1.91	Cellobiohydrolase	G--G--G--G--G- Release cellobiose either from reducing or non-reducing end
Exo-(1,4)-β-D-glucanase	EC 3.2.1.74	Exoglucanase or glucohydrolase	G--G--G--G--G- Release glucose from non-reducing end
β-Glucosidase	EC 3.2.1.21	Cellobiase	G--G, G--G--G--G Release glucose from cellobiose or short chain cello-oligosaccharides

2.2.2 Hemicellulolytic enzymes

As the hemicelluloses to be hydrolyzed are heteropolymers, several different enzymes are needed for their complete hydrolysis. Its hydrolysis requires the action of a complex enzyme system as shown in Figure 2.8. This is usually composed of β-1,4-xylanase, β-xylosidase, α-L-arabinofuranosidase, β-glucuronidase, acetylxylan esterase and phenolic acid esterases. The enzymes involved in the degradation of xylan and their modes of action were summarized in Table 2.3. The xylanase attacks the main chains of xylans and it plays an important role in hydrolysis of xylan backbones while β-xylosidase hydrolyzes xylooligosaccharides to xylose. Arabinofuranosidase and β-glucuronidase remove the arabinose and 4-*O*-methyl glucuronic acid

substituents, respectively, from the xylan backbone. Esterases hydrolyze the ester linkages between xylose units of the xylan and acetic acid (acetylxylan esterase) or between arabinose side chain residues and phenolic acids, such as ferulic acid (ferulic acid esterase) and *p*-coumaric acid (*p*-coumaric acid esterase). All these enzymes act cooperatively to convert xylan to its constituent sugar. Heteroxylans contain different substituent groups in the backbone and side chain. Therefore, the degradation of such a complex polysaccharide may involve synergistic action between the different components of the xylanolytic enzyme system (Sunna and Antranikian, 1997).

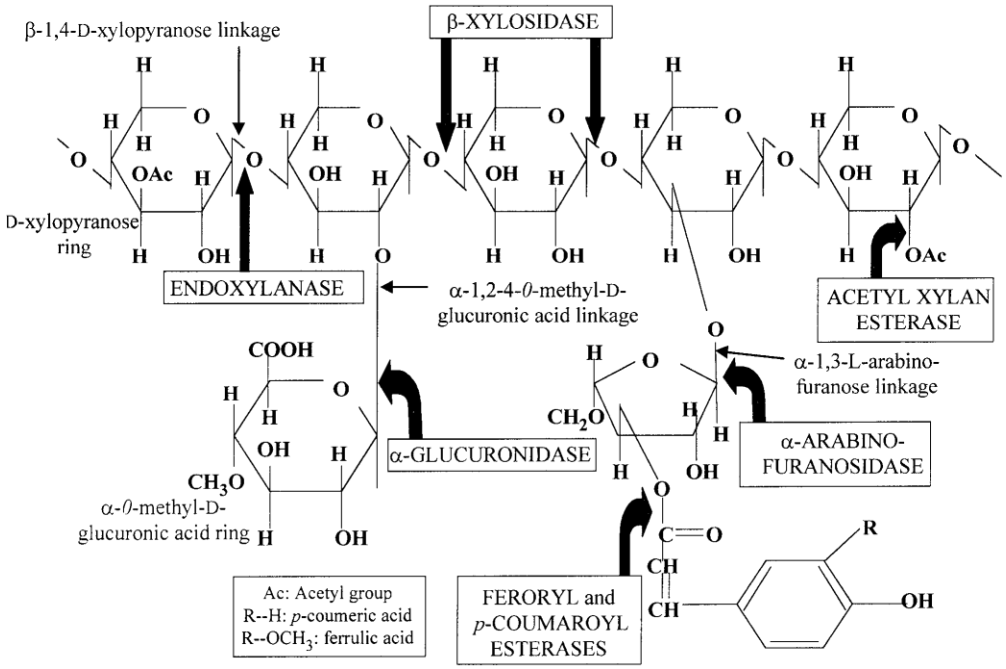


Figure 2.8 The xylanolytic enzymes involved in the degradation of xylan (Beg et al., 2001).

Table 2.3 Enzymes involved in the hydrolysis of hemicellulose (Modified from Dyk and Pletschke, 2012).

Enzyme	Mode of action
Xylanase	Hydrolyzed mainly internal β-1,4-xylose lingkage of the xylan backbone.
β-Xylosidase	Releases xylose from xylobiose and xylooligosaccharide.
β-Arabinofuranosidase	Hydrolyzes terminal nonreducing α-arabinofuranose from arabinoxylans.
β-Glucuronidase	Releases glucuronic acid from glucuronoxylans.
Acetylxylan esterase	Hydrolyzes acetylester bonds in acetyl xylans.
Ferulic acid esterase	Hydrolyzes feruloylester bonds in xylans.
p-Coumaric acid esterase	Hydrolyzes *p*-coumaryl ester bonds in xylans.

2.3 Multienzyme complex

For years, it was believed that microbial cellulase systems consisted solely of numerous types of free cellulases that act synergistically on the insoluble substrate. Many cellulase systems, particularly in aerobic microorganisms, seem to be characterized by free enzymes. The rumen is a highly anaerobic environment that represents one of the most active sites for breakdown of plant cell wall material in nature. In recent years, multienzyme complexes, have been identified and described in many anaerobic, thermophilic, cellulolytic bacterium. The bacterium adhesion to cellulose is accomplished by means of a discrete multifunctional, multicomponent cell surface protein complex, known as the cellulosome which is exquisitively designed for efficient binding and hydrolysis of the substrate (Lamed et al., 1983). The production of the multienzyme complex may have a number of advantages for the effective hydrolysis of cellulosic substances:

1. Synergism is optimized by the correct ratio between the components, which is determined by the composition of the complex.

2. Non-productive adsorption is avoided by the optimal spacing of components working together in synergistic fashion.

3. Competitiveness in binding to a limited number of binding sites is avoided by binding the whole complex to a single site through a strong binding domain with low specificity.

4. A halt in hydrolysis on depletion of one structural type of multienzyme complex at the site of adsorption is avoided by the presence of other enzymes with different specificity.

The multienzyme complexes attach both to cell envelope and to the substrate, mediating the proximity of the cell to the cellulose (Ohara et al., 2000). Thus multienzyme complex is efficient to hydrolyze the complex substrates (Figure 2.9).

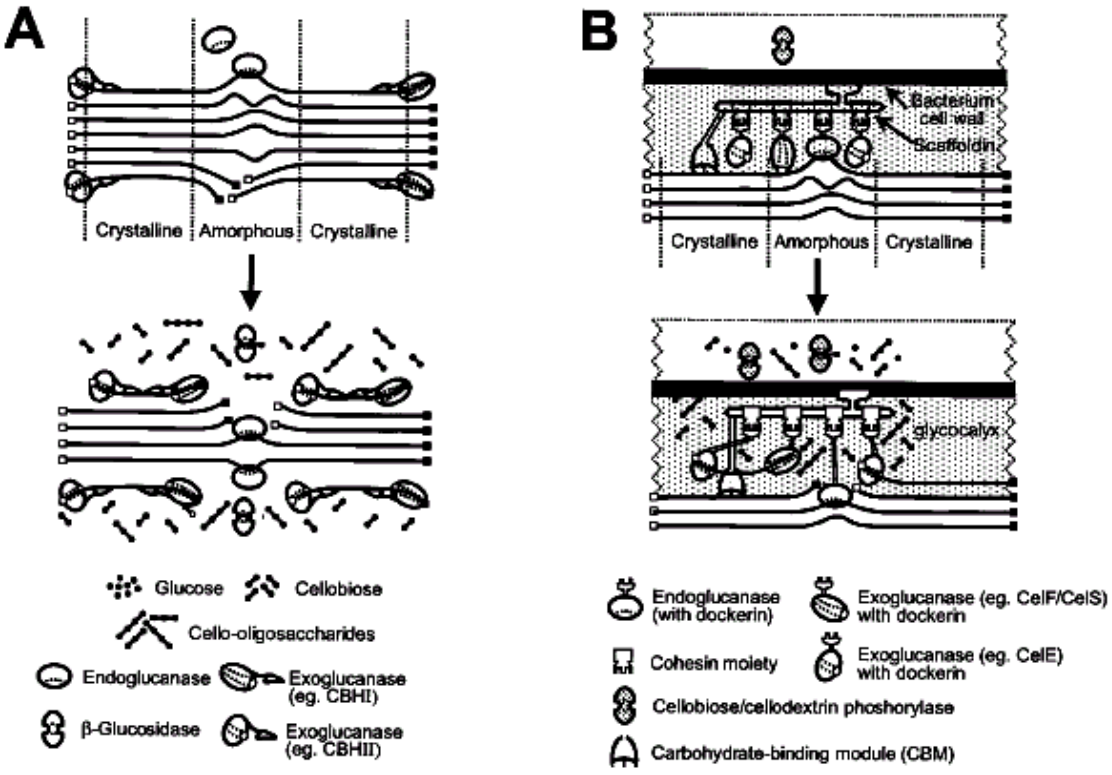


Figure 2.9 Schematic representation of the hydrolysis of amorphous and microcrystalline cellulose by noncomplexed (A) and complexed (B) cellulase systems. The solid squares represent reducing ends, and the open squares represent nonreducing ends. Amorphous and crystalline regions are indicated. Cellulose, enzymes, and hydrolytic products are not shown to scale (Lynd et al., 2002).

2.3.1 Concept of multienzyme complex

The cellulosome consists of numerous different enzymatic subunits, which are held together into the complex via a separate, noncatalytic subunit, called scaffoldin (the core of the cellulosome structure) (Bayer et al., 1994). The scaffoldin subunit is structured into a series of

functional domains, one of which is a carbohydrate-binding module (CBM). This CBM is very similar to the CBMs derived from simple microbial cellulases and xylanases. The letter CBMs usually deliver a single type of catalytic domain to the substrate (Gilkes et al., 1991). In contrast, the CBM from scaffoldin subunits of some cellulosome anchored the cellulosome into the cell surface, and that it was responsible for organizing the enzymic components into the complex (Figure 2.10). Some bacterial cellulases also contain a relatively small domain that was initially described in *Clostridium thermocellum*. This domain exhibits a docking function, and compiles a 22-residue reiterated (or duplicated) segment, amino acids. This type of domain, known as dockerin, mediates the attachment of the catalytic subunit to the cellulosome in *Clostridium thermocellum* by binding with cohesin in scaffoldin.

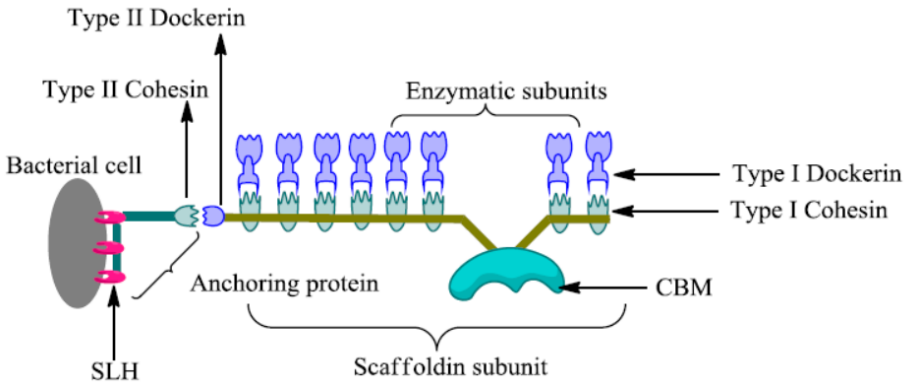


Figure 2.10 Schematic architecture of bacterial cellulosome (Juturu and Wu, 2014).

The original discovery of the cellulosome in *Clostridium thermocellum* was based on a combination of biochemical, immunological, ultrastructural and genetic technique (Bayer et al., 1983; Lamed et al., 1983). Ultrastructural evidence showed the multisubunit composition and dimensions of the cellulosome and their organization on the cell surface in the form of polycellulosome organelles. Later, traditional molecular biological approaches led to the cloning, sequencing and expression of numerous cellulases and some of which turned out to be cellulosome components (Beguin and Lemaire, 1996).

On the other hand, xylanosomes, having a structure analogous to the cellulosomes, are discrete, multifunctional, multienzyme complexes found in several microorganisms and play an important role in the degradation of hemicellulose (Sunna and Antranikian, 1997; Jiang et al.,

2005). The extracellular xylanosome complex B from *Butyvibrio fibrisolvens* H17c exists as a multisubunit protein aggregate with a molecular mass of more than 669 kDa (Lin and Thomson, 1991). *Clostridium papyrosolvens* C7 produces a multicomplex cellulase-xylanase system, which consists of seven protein complexes with molecular mass range from 500 to 660 kDa.

2.3.2 Components of multienzyme complex

2.3.2.1 Enzyme subunits (The modular nature of the cellulase and related enzyme)

The enzyme subunits of multienzyme complex are glycosyl hydrolase, which hydrolyze oligosaccharide and polysaccharides (Henrissat and Bairoch, 1996; Davies and Henrissat, 1995). In this sense, the cellulosomal enzyme subunits are not different from related free cellulases; both free and cellulosomal enzymes contain common types of catalytic domains from the same collection of glycosyl hydrolase families. The major difference between the two types of enzymes is that all cellulosomal enzymes known to date contain a dokerin domain, which mediates its integration into the cellulosome complex, whereas noncellulosomal enzymes apparently lack such a domain (Bayer et al., 1994).

The dockerin domains are usually connected to one another by linking segments, known as linkers. These segments are frequently (but not individual) rich in poline, and threonine and/or serine residues and are often glycosylated. The linkers bear some resemblance to cell wall extensions in plants, and they may play a similar structural role in the cellulase (Fierobe et al., 1999).

2.3.2.2 The scaffoldin subunit (cellulosome-intregating protein)

The scaffoldin subunit plays a critical rule in cellulosome assembly by its repetitive cohesin domains, each of which interacts with the dockerin domain of the individual cellulosomal enzymes. A carbohydrate-binding module (CBM) of scaffoldin protein is responsible for mediating the binding of the cellulosome to the substrate. In *Clostridium thermocellum*, the scaffoldin is believed to bind to another type of cellulosome-related protein (i.e., cell surface anchoring proteins) that carry C-terminal surface-layer homology domains. The binding is mediated.

The scaffoldin subunits are very large, modular polypeptides. Each of them contain a single family-IIIa CBM, one or more conserved hydrophilic modules of unknown function, and most importantly, multiple copies of cohesin domains. Several cellulosomal enzymes also bear

CBMs as part of their structure, although the presence of CBMs seems not to be a definitive cellulosomal characteristic, and its function is not necessarily that of targeting agent. For this purpose, the cellulosomal enzymes pedant collectively on a special CBM, borne by a separate scaffolding subunit. Not only a cellulose-binding domain (CBD) but also other carbohydrate-binding domain had investigated. The modular organization of the scaffoldins appears to represent two basic types, the *Clostridium thermocellum* scaffoldin carries as internal CBM and a C-terminal type-II dockerin domain, which presumably anchors it to the cell surface, whereas the other species have N-terminal CBMs and lack a type-II dokerin (Bayer et al., 2004).

2.3.2.3 Carbohydrate-binding modules

The carbohydrate-binding modules (CBMs) are defined as contiguous amino acid sequence within a carbohydrate-active enzyme with a discreet fold having carbohydrate-binding activity. A few exceptions are CBMs in cellulosomal scaffoldin proteins and rare instances of independent putative CBMs. The requirement of CBMs existing as modules within larger enzymes, sets this class of carbohydrate-binding protein apart from other non-catalytic sugar binding proteins such as lectins and sugar transport proteins.

CBMs were previously classified as cellulose-binding domains (CBDs) based on the initial discovery of several modules that bound cellulose. However, additional modules in carbohydrate-active enzymes are continually being found that bind carbohydrates other than cellulose yet otherwise meet the CBM criteria, hence the need to reclassify these polypeptides using more inclusive terminology. CBMs are classified into families based on amino acid sequence similarity and the CAZy database assigns CBM into 83 families (www.cazy.org; Cantarel et al. 2009). Up to date, some of the families (families 7 and 33) have been reclassified and hence contain no entries (www.cazy.org). Within families, CBMs are also subclassified into three types: type A (surface-binding CBMs), type B (glycan-chain-binding CBMs) and type C (small-sugar-binding CBMs) (Oliveira et al, 2015).

Type A CBMs or surface-binding CBMs (Figure 2.11-A, D, F, H) have a flat hydrophobic binding surface comprised of aromatic residues. The planar architecture of the binding sites is consistent with the flat surfaces of insoluble, crystalline polysaccharides such as cellulose or chitin (Boraston et al., 2004). The interaction of type A CBMs with crystalline cellulose is related with positive entropy, indicating that the thermodynamic forces that drive the binding of CBMs to

crystalline ligands are relatively unique among CBMs. Another feature of type A CBMs is their little or no affinity for soluble polysaccharides (Bolam et al., 1998; Shoseyov et al., 2006; Oliveira et al, 2015).

Type B CBMs or glycan-chain-binding CBMs (Figure 2.11-B) binding site architecture displays a cleft or groove arrangement and comprise several subsites able to accommodate the individual sugar units of the polysaccharide. As in type A CBMs, the aromatic side chains play a critical role in ligand binding, and the orientation of the aromatic residues are crucial determinants of binding specificity. The binding proficiency of the type B CBMs is assessed by the degree of polymerization of the carbohydrate ligand which high affinity to hexa-saccharides and negligible interaction with oligosaccharides with degree of polymerization of three or less. Then, type B CBMs are often called “chain binders” (Boraston et al., 2004; Filonova et al., 2007; Tomme et al., 1998; Oliveira et al, 2015).

Type C CBMs or small-sugar-binding CBMs (Figure 2.11-C, E, G, I), have the lectin-like feature of binding optimally to small sugars such as mono-, di-, or tri-saccharides due to steric restriction in the binding site, lacking the extended binding site grooves of type B CBMs. For this reason, type C CBMs are also called “lectin-like CBMs”. In general, the identification and characterization of type C CBMs falls behind type A and B classification, because their limited existence in plant cell wall of active glycoside hydrolases (Abbott et al., 2008; Boraston et al., 2003, 2004; Gregg et al., 2008; Oliveira et al, 2015).

The CBM functions to concentrate the enzyme on the surface of the substrate and this is thought to lead to improved degradation of the substrate which has been shown experimentally (Boraston et al., 2004). However, it has been shown in some instances that removal of the CBM from an enzyme only affected activity on the substrate where such substrate was insoluble (Boraston et al., 2004). The CBM is also able to cause disruption of cellulose fibers without any corresponding hydrolytic activity (Boraston et al., 2004; Gilbert, 2010; Shoseyov et al. 2006). However, the ability of a CBM to disrupt a substrate such as crystalline cellulose has only been shown in a few cases and does not apply in general to all CBMs (Boraston et al., 2004; Din et al., 1994; Hilden and Johansson, 2004).

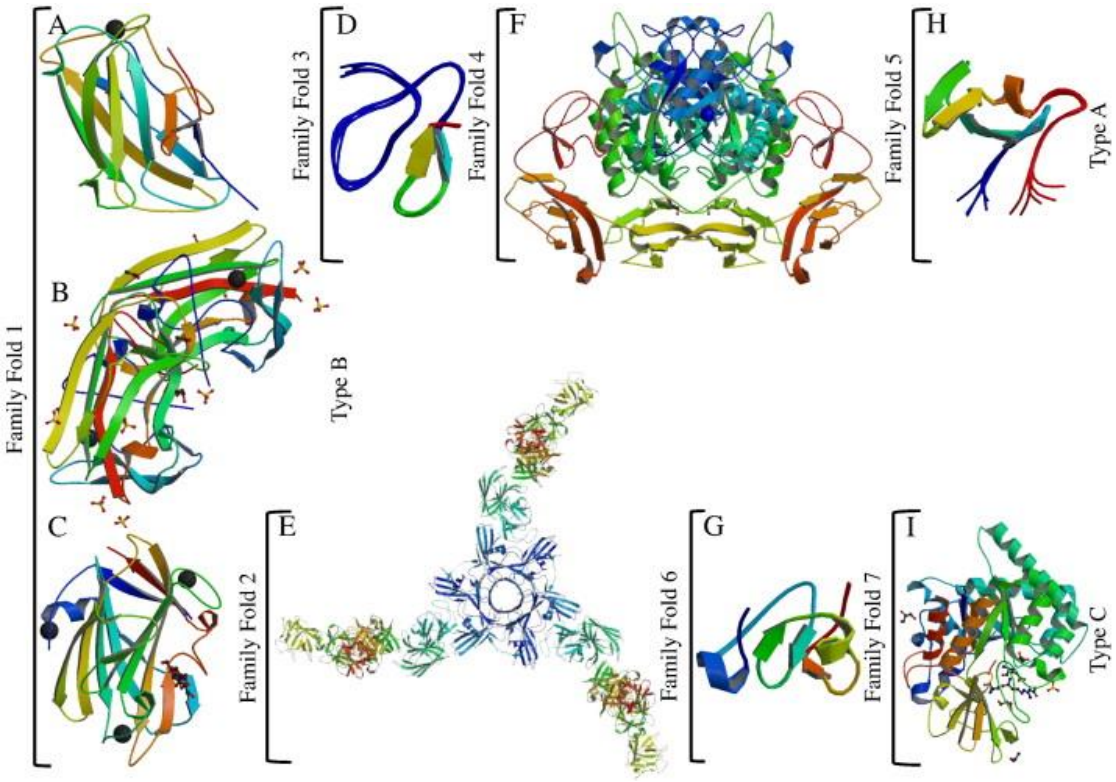


Figure 2.11 CBM functional types. CBMs shown are as follows: (A) family 3 CBM from *Clostridium thermocellum*; (B) family 4 from *C. thermocellum*; (C) family 9 from *Thermotoga maritima*; (D) family 1 CBM from *Trichoderma reesei*; (E) family 13 from *Clostridium botulinum*; (F) family 5 from *Moritella marina*; (G) family 18 CBM from *Triticum kiharae*; (H) family 10 CBM from *Cellvibrio japonicas* Ueda107; (I) family 14 CBM from *Homo sapiens* (Oliveira et al, 2015).

The CBM plays a very important role in substrate binding and targets the enzyme towards specific substrates, conferring selectivity in binding. It was illustrated that where the CBM of an enzyme was changed, the substrate specificity of the enzyme was changed (Araki et al., 2004; Boraston et al., 2004). CBMs have also been shown to selectively bind to certain areas of a polysaccharide, such as the reducing end, allowing the enzyme access to sections that may be preferentially degraded. By targeting different regions of a substrate through different CBMs, the same enzyme may be able to hydrolyse a substrate more efficiently (Boraston et al., 2004).In addition, it has been established that there is a clear relationship between adsorption of enzymes to the substrate and the efficiency of substrate hydrolysis which is closely linked to the presence of the CBM (Kotiranta et al., 1999; Varnai et al., 2010). On the other hand, it has also been established that non-productive adsorption of enzymes on lignocellulosic biomasses takes place

and affects hydrolysis efficiency. Therefore, CBMs are important as they allow the enzymes to remain bound to the substrate while they hydrolyse bonds (Boraston et al., 2004; Dyk and Pletschke, 2012). The role of the CBM has been termed an “intramolecular synergism” between the CBM and the active site by Din et al. (1994) as the activity of the enzyme is enhanced by the presence of a CBM domain.

Chapter 3 Materials and methods

3.1 Sources of polysaccharides were used

Birchwood xylan and oat-spelt xylan were from Fluka. Carboxymethylcellulose, β-D-glucan from barley, guar (galactomannan), laminarin from *Laminaria digitata* and lichenan from *Cetraria islandica* were from Sigma-Aldrich. Arabinan from sugar beet (Ara:Gal:Rha:GalUA = 88:3:2:7), xyloglucan from tamarin seed were from Megazyme international Ireland Ltd. Pectin from citrus and methylcellulose were from Nacalai Tesque, Kyoto. Soluble starch was from Kanto Chemical, Tokyo.

Corn hull was prepared as described by Tachaapaikoon et al., (2006). The dried corn hull was cut by scissors to small sizes, then ground in a blender and sieved to 80-100 mesh size. The ground corn husk was dried at 50 °C and kept for further experiments.

Ball-milled cellulose was prepared by ball mill processing of the KC flock (Nippon Paper Chemicals, Tokyo) with distilled water for 72 h.

The insoluble xylan was prepared as described by Kaewintajuk et al. (2006). A suspension of 0.5% (w/v) oat spelt xylan was stirred in deionized water for 1 h at room temperature. After that, the mixture was centrifuged at 3,000 × g for 10 min, and the supernatant containing the soluble fraction was removed. The pellet containing the insoluble fraction was washed twice with 20 volumes of deionized water and freeze dried.

3.2 Bacterial strains and plasmids

Paenibacillus xylaniclasticus stain TW1, the multienzyme producing strain, had been isolated from sludge in an anaerobic digester (Tachaapaikoon et al., 2012a, 2012b) was used for enzyme preparation.

Escherichia coli INVα (Invitrogen, Carlsbad, CA) and BL21(DE3) (Toyobo, Kyoto) were cultivated in Luria-Bertani (LB) medium supplemented with ampicillin (50 μg/mL) for use as plasmid construction and expression host for derivative constructs of plasmid pCR2.1 and pRSET/CFP (Invitrogen), respectively.

3.3 Isolation and identification of carbohydrate-binding proteins

3.3.1 Enzyme preparation

P. xylaniclasticus stain TW1 was grown in Berg's mineral salt medium (pH 7.0) (Berg et al. 1972) containing 0.2% (w/v) $NaNO_3$, 0.05% (w/v) K_2HPO_4, 0.02% (w/v) $MgSO_4 \cdot 7H_2O$, 0.002% (w/v) $MnSO_4 \cdot H_2O$, 0.002% (w/v) $FeSO_4 \cdot 7H_2O$, 0.002% (w/v) $CaCl_2 \cdot 2H_2O$, and 0.5% (w/v) corn hull as the sole carbon source. The culture was incubated in a rotary shaker at 200 rpm and 37 °C for 3 days. After centrifugation at 8,000 × g for 20 min at 4 °C, the culture supernatant was concentrated about 50-fold by freeze dryer.

3.3.2 Isolation of carbohydrate-binding proteins

The isolation method was conducted by adding 250 µl protein solution prepared from the 50-fold concentrated culture supernatant to 15 mg of insoluble polysaccharide (ball-milled cellulose or insoluble oat-spelt xylan) in 10 mM potassium phosphate buffer (KPB, pH 7.0) in 1.5-ml microtube. Sample was incubation at intervals at 4 °C for 30 min before centrifugation. The precipitate was washed 5 times with same buffer, and the precipitate and supernatant were analyzed by SDS-PAGE.

3.3.3 Gel electrophoresis

SDS-PAGE was performed on a 8% (w/v) polyacrylamide gel by the method of Laemmli (1970). After electrophoresis, the proteins were stained with Coomassie brilliant blue R-250. The molecular weight standards used was from a molecular-weight calibration kit (Bio-Rad).

3.3.4 Protein identification using MALDI-TOF/TOF MS

Selected protein bands were excised from SDS-PAGE gel and subjected to in-gel digestion with trypsin as described previously (Shin et al., 2006). MS/MS analyses to identify proteins were performed on a JP4800 Plus MALDI-TOF/TOF Analyzer (Applied Biosystems) using a 200 Hz ND: YAG laser operating at 337 nm. The 20 most intense ions per MALDI spot, with signal/noise ratios >25, were selected for subsequent MS/MS analysis in the 1 kV mode using 100 consecutive laser shots. Air served as the collision gas. Data was subjected to analysis using the Mass Standard Kit for the 4700 Proteomics Analyzer (Applied Biosystems). The NCBInr database was searched for matches to the MS/MS spectra using ProteinPilot v. 3.01 (with MASCOT as the database search engine), with peptide and fragment ion mass tolerances of 200

ppm. One missed cleavage, carbamidomethylation of cysteines and oxidation of methionines were allowed during peptide searches.

3.3.5 Polymerase chain reaction (PCR) amplification and sequencing for carbohydrate-binding proteins

Based on MALDI TOF/TOF MS data, PCR primers were designed to amplify one gene with reference to have high homology regions in the xylanase gene (*xyn11*A) of *Paenibacillus curdlanolyticus* B-6. (GenBank no. GI:283970945), identified using the Blast search tool (http://www.ncbi.nlm.nih.gov/BLAST/). The two pairs of synthetic oligonucleotide primers for this gene were *Pc*xyn11F: 5′-ACCTTGTTGTTAGCTGGGACAAACGT-3′ and *Pc*xyn11R: 5′-TTGTTCGATGTGAAGCCGGGCA-3′.

PCR reactions were performed in a Takara PCR Thermal Cycler PERSONAL. For each reaction, a PCR mixture was prepared as follows: 5.0 µl of 10× PCR buffer (Takara), 5.0 µl of dNTP mixture (Takara; 51.25 nM), 1.0 µl of each primer (50 µM each) , 0.5 µl of r-Taq DNA polymerase (Takara; 5 U/µl), 2.0 µl of DNA template, and 35.5 µl of MilliQ water, to attain 50 µl. PCR amplification conditions were 1 cycle at 95 °C for 4 min, followed by 30 cycles of denaturation at 95 °C for 20 s, annealing at 47 °C for 40 s, extension at 72 °C for 2 min, and a final elongation step at 72 °C for 3 min. Amplification products were run in a 0.8% (w/v) agarose gel, immersed in TAE buffer (40 mM Tris–acetate with 1 mM EDTA), stained with 0.5 µg/ml ethidium bromide, and visualized using a UV transilluminator Bioinstrument (ATTO). Amplified band was excised from the gels and purified prior to cloning employing a TA Cloning Kit (Invitrogen). A inserted gene was purified and sequenced on an ABI 3130X DNA Analyzer (Applied Biosystems) for seeking for carbohydrate-binding module (CBM).

3.4 Production and characterization of the recombinant CBM36

3.4.1 DNA amplification and cloning

DNA fragment encoding for the CBM36 obtained from *P. xylaniclasticus* stain TW1 was amplified by polymerase chain reaction (PCR). The primers used contained either an artificial *Bam*HI or *Eco*RI restriction endonuclease sites (underlined) in order to facilitate the cloning of the amplified DNA fragments into pRSET/CFP. The following oligonucleotide primers were used to amplify the coding regions by PCR: *Px*CBM36F-CFP: GGATCCGGTAATACAGGTGGAAC and *Px*CBM36R-CFP: GAATTCAGGAGTCGAATTGATTTCC. The amplified PCR products were then

cloned into pCR2.1. The inserted amplified DNA fragments were sequenced to confirm the absence of mutations, digested with *Bam*HI and *Eco*RI, and finally ligated into pRSET/CFP that had been digested with the same restriction enzymes.

3.4.2 Protein expression and purification

E. coli BL21 (DE3) cells harboring *Px*CBM36-pRSET/CFP was aerobically cultured at 37 °C in 1 liter of LB broth supplemented with amplicillin (50 μg/ml), the cells were allowed to grow to an optical density of about 0.4 at 600 nm. After the addition of isopropyl-1-thio-β-D-galactopyranoside (IPTG) to a final concentration of 1 mM, the cells were incubated for an additional 3 h. They were then harvested by centrifugation (8,000 x g) for 20 min, washed, and disrupted by sonication. The cell-free extract was subjected to His-tag purification on nickel-nitrilotriacetic acid (Ni-NTA) agarose (Qiagen) according to the manufacturer' s instructions. The purity of the fractions were analyzed by SDS-PAGE. The protein concentrations were determined spectrophotometrically at 280 nm using a calculated molar extinction coefficient (Mach et al., 1992). The molar absorption coefficients calculated from the tryptophan, tyrosine and cysteine contents was 47,960.

3.4.3 Substrate-binding assays with affinity gel electrophoresis and SDS-PAGE

The capacity of CBM36-CFP to bind to a variety of soluble polysaccharides were evaluated by affinity gel electrophoresis. Continuous native polyacrylamide gels (the absence of SDS) were prepared consisting of 10% (w/v) acrylamide. To one of the gels, 0.1% (w/v) soluble polysaccharide was added prior to polymerization. Approximately 5 μg of target proteins and BSA (as a non-interacting negative control) were loaded onto the gels and subjected to electrophoresis at room temperature. Proteins were visualized by staining with Coomassie brilliant blue R-250.

CBM36-CFP was allowed to bind to the 15 mg of insoluble polysaccharides on ice for 30 min in a 50 mM potassium phosphate buffer (KPB, pH 7.0) without/with EDTA in a final volume of 250 μl. The solution was then centrifuged (8,000 x g) for 5 min. The precipitate was washed 5 times with same buffer without/with EDTA, and the precipitate and supernatant were analyzed by SDS-PAGE. The polysaccharides used were ball-milled cellulose and insoluble oat-spelt xylan.

3.4.4 Preparation of plant material and CBM-labeling

This study was carried out as described by McCartney et al. (2004) and Araki et al. (2010). Sweet potato (*Ipomoea batatas* cv. Beni Kokei) was grown from a seedling in water for appropriate times. Root apices were fixed for 5 min in 50 mM sodium phosphate buffer (pH 7.2) containing 3% (w/v) paraformaldehyde and 50 mM sodium chloride with vacuum infiltration, and then for 4 h on ice. The samples were then dehydrated in an ethanol series consisting of 30% (v/v), 50% (v/v) and 70% (v/v) ethanol (20 min each at 4 °C); 80% (v/v), 90% (v/v), 95% (v/v), 99% (v/v) and 100% (v/v) ethanol (20 min each at room temperature); and 30% (v/v), 60% (v/v) and 80% (v/v) t-buthanol in 100% (v/v) ethanol (20 min each at room temperature). They were then transferred to 55 °C, allowed to warm, transferred to 100% (v/v) t-buthanol for 20 min and then to 100% (v/v) t-buthanol and liquid paraffin (1:1) and incubated for 4 h. They were transferred 3 times to low-melt paraffin for 1 h each at 65 °C and left overnight at 65 °C. They were then placed in sample moulds filled with melted paraffin and allowed to solidify in ice water. Six-µm-thick sections were cut through the roots and collected on polylysine-coated microscope slides, deparaffined in a series consisting of xylene (twice for 5 min each); 100% (v/v), 99% (v/v), 95% (v/v), 90% (v/v), 80% (v/v), 70% (v/v), 50% (v/v) and 30% (v/v) ethanol (1 min); 20 mM KPB at pH 7.0 (10 min) and 20 mM KPB with 0.25% (w/v) BSA (30 min). For CBM labeling, the root sections were incubated in 20 mM KPB containing 0.25% (w/v) BSA and 40-100 µg/ml of CBM-CFP for 15 min. They were then washed in 50 mM KPB at least 3 times. They were observed on an Axioskop 2 Plus Zeiss microscope. The images were captured using an AxioCam HRC Plan-Neofluar camera.

3.5 Identification of Carbohydrate-active enzymes in *P. xylaniclasticus* stain TW1 genome

To identify the carbohydrate-active enzymes of *P. xylaniclasticus* stain TW1, the genomic DNA was prepared and purified according to the method of Saito and Miura (1963). The nucleotide sequence was determined using a 454 GS FLX sequencer, and assembly was performed using the GS De Novo Assembler software version 2.6.

Chapter 4 Results and discussion

4.1 Isolation of carbohydrate-binding proteins

The microbial deconstruction of the plant cell wall is a critical biological process, which also provides important substrates for environmentally sustainable industries. However, the physical complexity of plant cell walls limits the access of the hydrolytic enzymes to their target substrates. To overcome the “access problem” esterases, lyases and glycoside hydrolases that degrade plant structural polysaccharides, in general, have a modular structure in which the catalytic module is appended to non-catalytic carbohydrate binding modules (CBMs), which are grouped into sequence-based families within the CAZy database (Cantarel et al., 2009). The general function of CBMs is to direct the catalytic modules to their target substrate within the plant cell wall, thereby increasing the efficiency of catalysis. The majority of CBMs display a β-sandwich fold with the ligand binding site located in either the concave surface presented by one of the β-sheets or in the loops that connect the two sheets. This binding site can either target the internal regions or the end of glycan chains (Jamal-Talabani et al., 2004). In some CBM families, especially those that recognize crystalline polysaccharides, ligand specificity is invariant, while other families contain proteins that bind to a range of different carbohydrates. Thus, CBMs are excellent model systems for studying the mechanism of protein-carbohydrate recognition. Furthermore, this diversity in ligand specificity has the usefulness in numerous biotechnological applications (Tomme et al., 1998).

From previous study, a facultative, mesophilic, anaerobic bacterium strain, TW1 (Figure 4.1) was isolated from an anaerobic digester. The 16S rRNA gene analysis revealed that strain TW1 resembled *P. cellulosilyticus*, *P. curdlanolyticus* and *P. kobensis* (97.6-97.1% similarity). However, on the basis of DNA-DNA relatedness and phenotypic analyses, phylogenetic data and the enzymatic pattern presented by Tachaapaikoon et al. (2012b), the isolated strain TW1 was classified as a novel species of the genus *Paenibacillus*, namely *P. xylaniclasticus* strain TW1. This strain was chosen for this work because it is able to produce xylanolytic-cellulolytic multienzyme complex when grown on corn hull under aerobic conditions (Tachaapaikoon et al., 2012a). General producers of multienzyme complex are frequently found in the anaerobic microorganism. Therefore, it was easier to handle aerobe. Furthermore, this strain is also able to convert lignocelluloses to fermentable sugars efficiently.

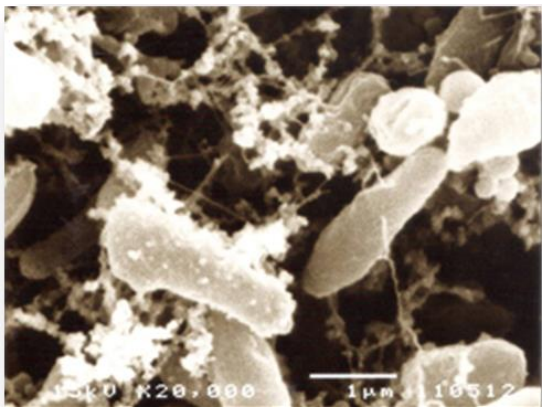


Figure 4.1 Scanning electron microscopy of a *P. xylaniclasticus* strain TW1 cell harvested at the late exponential growth phase. The scale bar in the micrograph is 1 µm.

Although, xylanolytic-cellulolytic multienzyme complexes produced from *P. xylaniclasticus* strain TW1 could effectively hydrolyze the agricultural wastes and bind substrates (Tachaapaikoon et al., 2012a), the CBMs appearing in the xylanolytic-cellulolytic multienzyme complexes from this bacterium is still unknown. Then, this work, the secreted proteins that appear to contain CBMs will be identified and the results will be useful to improve the cost of hydrolysis process, which could be reduced processing cost using immobilization of enzyme via CBMs.

In order to elucidate features of the CBMs, we carried out isolation using supernatant obtained from corn hull-aerobic grown cultures (stationary phase) of *P. xylaniclasticus* strain TW1. Based on the properties of the carbohydrate-binding ability, the carbohydrate-binding proteins were isolated from the enzyme preparation by affinity purification on ball-milled cellulose. The enzyme preparation was incubated with ball-milled cellulose at 4°C to allow the carbohydrate-binding proteins present in the enzyme preparation to bind to the cellulose. The carbohydrate-binding proteins were separated from the unbound proteins of the enzyme preparation by centrifugation, washed and then eluted from the cellulose by adding the sample application buffer, heated in a boiling-water bath for 5 min, and analyzed by SDS-PAGE. As shown in Figure 4.2, the crude enzyme preparation consisted of at least twelve proteins that were detected in the carbohydrate-binding protein preparation (namely CBP1 to CBP12). The results indicated that all these proteins probably contained the carbohydrate-binding domain. Hence, all of them were subjected to MALDI-TOF/TOF MS spectrophotometry for identification of these proteins.

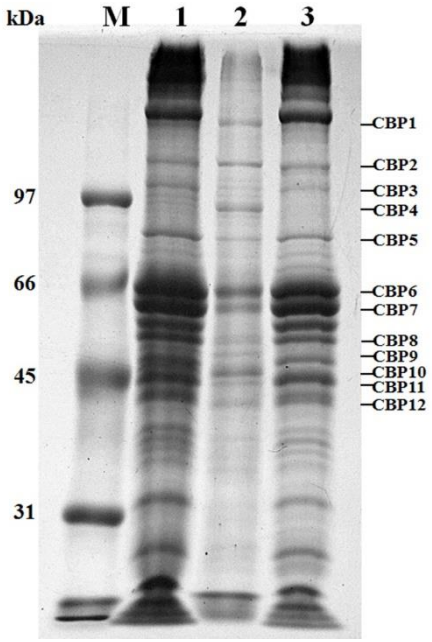


Figure 4.2 SDS-PAGE of carbohydrate-binding proteins isolated from the crude enzyme preparation of *P. xylaniclasticus* strain TW1. M, molecular weight marker (kDa); lane 1, crude enzyme preparation; lane 2, carbohydrate-binding proteins; lane 3, unbound proteins. CBP1-CBP12 in the gel indicate carbohydrate-binding proteins.

4.2 Identification of carbohydrate-binding proteins

All 12 bands of carbohydrate-binding proteins from the SDS-PAGE were digested with trypsin and the resulting peptide mixtures were analysed by MALDI-TOF/TOF MS spectrometry. The peptide mass fingerprints of the unknown components from SDS-PAGE were almost identical to unknown proteins (Table 4.1). Only 2 CBPs could be identified in this study. As shown in Table 4.1, three of the peptides from the CBP5 protein (#2, 8, 18) were obtained showing similarity to published glycoside hydrolase family 31 sequences from *P. curdlanolyticus* YK9 (GenBank no. gi|304405232). However, the known activities of glycoside hydrolase family 31 containing α-glucosidase (EC 3.2.1.20), α-1,3-glucosidase (EC 3.2.1.84), sucrase-isomaltase (EC 3.2.1.48) (EC 3.2.1.10), α-xylosidase (EC 3.2.1.177), α-glucan lyase (EC 4.2.2.13), isomaltosyltransferase (EC 2.4.1.-) and α-mannosidase (EC 3.2.1.24) were reported within this family (http://www.cazy.org/GH31_characterized.html). Additionally, no information on the CBM appearing in the glycoside hydrolase family 31 from *P. curdlanolyticus* YK9 was presented. For more clearly, PCR primer designs, sequencing, cloning and characterization of the recombinant CBP5 might be further investigated for the activity of this protein and the binding region of protein

on ball-milled cellulose. CBP12 band from SDS-PAGE gel could be identified as xylanase. An NCBI BLAST search of the peptide sequence of CBP12 (KGTITVDGGTYDIYETTRV) showed similarity to *P. curdlanolyticus* B6 xylanase belonging to glycoside hydrolase family 11 with a CBM family 36. Intriguingly, no report on the binding of CBM36 to cellulose was discovered (Jamal-Talabani et al., 2004). Hence, we might find the new binding properties of CBM36 from CBP12. Therefore, CBP12 was further studied for seeking the characteristic of this protein.

Table 4.1 Peptide sequences determined from carbohydrate-binding protein digest samples.

Protein	Peptide no. and sequence	BLAST identification
CBP1	#10: GKQLADDSSSAGR	Unknown protein
	#12: WLANIMLNGDPR	Unknown protein
	#18: EQRVLDSDGNPLPSEPR	Unknown protein
CBP2	#6: YDMYDKYFQK	Unknown protein
	#7: IKASVLGVSAAYVR	Unknown protein
	#13: IKQYPFQLSGGLCQR	Unknown protein
CBP3	#15: QISNLQQSISDAEQR	Unknown protein
	#16: NARTVFSELSQATSNK	Unknown protein
	#18: FVEFYGSGLSNISLADR	Unknown protein
CBP4	#7: GWPSTIGFLVDTSR	Unknown protein
	#9: SVPKLFSWIR	Unknown protein
	#13: IKQYPFQLSGGLCQR	Unknown protein
CBP5	#2: RKGEEWSVAFYRD	Glycoside hydrolase family 31 (*Paenibacillus curdlanolyticus* YK9)
	#8: RVPWAYDQEAVDVTRY	Glycoside hydrolase family 31 (*Paenibacillus curdlanolyticus* YK9)
	#18: KGYGVFVNQPELVSFEVGSEKV	Glycoside hydrolase family 31 (*Paenibacillus curdlanolyticus* YK9)

Table 4.1 (Cont.) Peptide sequences determined from carbohydrate-binding protein digest samples.

Protein	Peptide no. and sequence	BLAST identification
CBP6	#1: DDELLSAYR	Unknown protein
	#3: LVIENPEKPR	Unknown protein
	#14: IVGVNMGDSIPQVDIPRLIK	Unknown protein
CBP7	#1: FYLVGSKMR	Unknown protein
	#3: MMDTIESIR	Unknown protein
	#5: EQFGGVKDTAGR	Unknown protein
CBP8	#2: ASEQPTQIR	Unknown protein
	#3: QVVNPEDGSR	Unknown protein
	#6: EEFLGVIACIR	Unknown protein
CBP9	#3: EIAAAQDGIAR	Unknown protein
	#4: ENNIQFVDFR	Unknown protein
	#13: DTLCVCCMAASNVTGR	Unknown protein
CBP10	#10: QVASDFKSLSGVGSSQR	Unknown protein
	#11: NGFFAMGMAEKANSASR	Unknown protein
	#15: SLADAEGYETFVIPDDVGGR	Unknown protein
CBP11	#4: GPLTTPIGGGIR	Unknown protein
	#14: DLLNEYEFPGDDTPIVR	Unknown protein
	#16: YQLVQPQTIIVILEKIR	Unknown protein
CBP12	#15: KGTITVDGGTYDIYETTRV	Xylanase (*Paenibacillus curdlanolyticus* B6)

4.3 Properties of the CBP12

By the BLAST database search, the amino acid sequence (KGTITVDGGTYDIYETTRV) was found to have high homology with xylanase11A (Xyn11A) of *P. curdlanolyticus* B-6 (ADB54799.1). Then, to clone the gene of the CBP12 in the multienzyme complex of *P. xylaniclasticus* strain TW1, the forward and reverse primers were designed based on both amino acid sequence and nucleotide sequence of *P. curdlanolyticus* B-6. The amplified fragment was

cloned into the TA vector and sequenced (see “Materials and methods”). The ORF of the CBP12 gene consisted of 1,155 bp and encoded a protein of 385 amino acids (Figure 4.3) with a molecular weight of 41.5 kDa and pI 7.7. The alignment of predicted amino acid sequence revealed that mature CBP12 was composed of two major functional domains, a catalytic domain belonging to family-11 glycoside hydrolases and a carbohydrate-binding module classified as family-36. A glycine- and asparagine-repeated sequence existed between the two domains (Figure 4.4).

```
   1  atgacaacgaaggaggaagatgtgatgaaaccaagaaaaatgaaatggatgctggccatgctgctcagcttcaca
   1  M  T  T  K  E  E  D  V  M  K  P  R  K  M  K  W  M  L  A  M  L  L  S  F  T

  76  ctcgcattgccagcaggcgtcgcgaacgccgtaacgattacgaataatgagtccggcacccatgacggctactat
  26  L  A  L  P  A  G  V  A  N  A  V  T  I  T  N  N  E  S  G  T  H  D  G  Y  Y

 151  tatgagctatggaaggattccgggaacacaagcatgacgcttaacagcggcggcaccttcagtgcgacgtggagc
  51  Y  E  L  W  K  D  S  G  N  T  S  M  T  L  N  S  G  G  T  F  S  A  T  W  S

 226  aacattaacaatgcgttattccgtaaaggcaaaaaattcaatgaaacacagacccaccagcaaatcggcaacatc
  76  N  I  N  N  A  L  F  R  K  G  K  K  F  N  E  T  Q  T  H  Q  Q  I  G  N  I

 301  tcggtcaattacagcgctaactattctccgaacggcaactcttatctctgcatctatggatggtcggttgacccg
 101  S  V  N  Y  S  A  N  Y  S  P  N  G  N  S  Y  L  C  I  Y  G  W  S  V  D  P

 376  ctcgtcgagtattatatcgtcgacagctggggcagctggcgtcctcctggcgcctcgtccaaaggtacgatcacc
 126  L  V  E  Y  Y  I  V  D  S  W  G  S  W  R  P  P  G  A  S  S  K  G  T  I  T

 451  gttgacggcggcacctacgacatttatgagacgactcgtgtgaaccagccttccattaaaggcacggcaacgttc
 151  V  D  G  G  T  Y  D  I  Y  E  T  T  R  V  N  Q  P  S  I  K  G  T  A  T  F

 526  cagcagtattggagcgtccggacgtcgaagcgtacgagcggcacaatctcggtaaccgatcacttcaatgcttgg
 176  Q  Q  Y  W  S  V  R  T  S  K  R  T  S  G  T  I  S  V  T  D  H  F  N  A  W

 601  gcaagcagaggcatgcagctgggcaaaatgtacgaagtatcgctcacggtagaaggctaccaaagcagcggaagc
 201  A  S  R  G  M  Q  L  G  K  M  Y  E  V  S  L  T  V  E  G  Y  Q  S  S  G  S

 676  gcgacaattacaagcaatacgattacaatcggcggcaacggcggcggcggtaatacgggcggtggcaataacggc
 226  A  T  I  T  S  N  T  I  T  I  G  G  N  G  G  G  G  N  T  G  G  G  N  N  G

 751  ggcggcaacaatggaggaggcaacaatggcggcggtaatacaggtggaacgacgagaattgaagccgagagcatg
 251  G  G  N  N  G  G  G  N  N  G  G  G  N  T  G  G  T  T  R  I  E  A  E  S  M

 826  acgaagagcggccaatacactggcaatatctactctccgttcaatggtgtagcactatacgctaacaacgactcg
 276  T  K  S  G  Q  Y  T  G  N  I  Y  S  P  F  N  G  V  A  L  Y  A  N  N  D  S

 901  gtgaaattcacgcataactttacgaaaagcacgaacaacttctcgctccgcggcgcttcgaacaactcccaaatg
 301  V  K  F  T  H  N  F  T  K  S  T  N  N  F  S  L  R  G  A  S  N  N  S  Q  M

 976  gctagagtcgatctgaaggtcggcggcgtgacgaagggcaccttctattacggcggcagctatcctgcagtctat
 326  A  R  V  D  L  K  V  G  G  V  T  K  G  T  F  Y  Y  G  G  S  Y  P  A  V  Y

1051  acgataaacaacgtcaatacgggaaccggaaatcaacaggtcgagctcattcttacagccgataacgggcaatgg
 351  T  I  N  N  V  N  T  G  T  G  N  Q  Q  V  E  L  I  L  T  A  D  N  G  Q  W

1126  gatgcgttcctcgattatttggaaatcaattag
 376  D  A  F  L  D  Y  L  E  I  N  *
```

Figure 4.3 Nucleotide sequence of the *cbp12* gene and deduced amino acid sequence of the gene product.

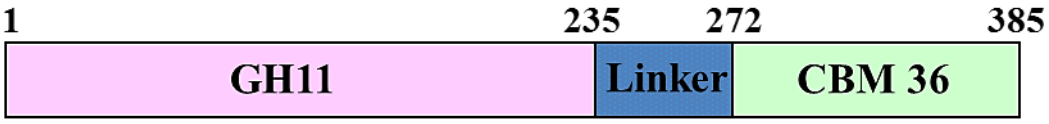


Figure 4.4 The modular structure of CBP12 from *P. xylaniclasticus* strain TW1. Domain boundaries are given by amino acid sequence number above the module schematic. GH11, family-11 catalytic module of the glycoside hydrolases; CBM36, family-36 CBM.

Family-11 glycoside hydrolase domain and family-36 carbohydrate-binding module (CBM36) of CBP12 have highly homology for xylanase11A from *P. curdlanolyticus* B-6 (99% identity, ADB54799.1), endo-β-1,4 xylanase from *Paenibacillus* sp. Aloe-11 (75% identity, ZP_09772596.1), xylanase X from *P. campinasensis* BL11 (72% identity, ABB77852.1), and CBM36 of *P. terrae* HPL-003 (75% identity, YP_005077359.1), *Bacillus* sp. YA-335 (73% identity, CAA41784.1) and *P. polymyxa* (50% identity, P45796.1), respectively. However, the only CBM36 currently characterized, that from *P. polymyxa* xylanase 43A (P45796.1), shows calcium-dependent binding of xylans and xylooligosaccharides (Jamal-Talabani et al, 2004). Therefore, CBM36 of CBP12 from *P. xylaniclasticus* strain TW1 was still attracted and, henceforth, was produced and characterized of this CBM.

4.4 Characterization of the CBM36 from *P. xylaniclasticus* stain TW1 (*Px*TW1CBM36)

In this study the *Px*TW1CBM36-cyan fluorescent protein (*Px*TW1CBM36-CFP) was constructed. The binding characteristics and specificities of *Px*TW1CBM36-CFP against isolated carbohydrates and the recognition of carbohydrates located within plant cell walls as the substrate available for cognate enzymes were investigated. Moreover, the target region of CBM36 in plants was currently unknown and the fluorescent tag was used for monitoring CBM36 binding in plants.

4.4.1 Ligand specificity of *Px*TW1CBM36-CFP

To study the binding characteristics of CBM36 from *P. xylaniclasticus* strain TW1, *Px*TW1CBM36 was ligated in pRSET/CFP for construction of CFP fusion *Px*TW1CBM36 (*Px*TW1CBM36-CFP), expressed under the induction of IPTG and purified for subsequent analysis. For the purification strategy, the recombinant protein was purified based on its N-terminal 6×His-tag by affinity purification using a Ni-NTA agarose column. The purified

*Px*TW1CBM36-CFP was analyzed by SDS-PAGE and desalted by dialysis against 10 mM KPB (pH 7.0) for further study.

In vitro binding studies using semi-quantitative affinity gel electrophoresis showed that migration of purified *Px*TW1CBM36-CFP through native polyacrylamide gels loaded with soluble birchwood xylan and oat-spelt xylan were slower than that of BSA (control) (Figure 4.5). Moreover, *Px*TW1CBM36-CFP could bind tightly to both soluble birchwood and oat-spelt xylans, moderately to sugar beet arabinan (α-linked arabinan) and weakly to laminarin (β-1,3-glucan), tamarind seed xyloglucan, methylcellulose and lichenan. However, *Px*TW1CBM36-CFP could not bind to carboxymethylcellulose and barley β-glucan. Although, barley β-glucan is a mixed-linkage β-1,3-1,4-glucan with a structure similar to that of lichenan, the structure of barley β-glucan was generally categorized as containing more β-1,3-linked cellotetraosyl units than lichenan (Wood et al., 1994). These results indicated that the frequency and composition of substituted groups might have an effect on the binding ability of CBM36. In addition, *Px*TW1CBM36-CFP did not bind to galactomannan (guar gum), pectin or starch (Table 4.2). The result revealed that *Px*TW1CBM36-CFP showed relatively tight binding for xylans from different origins, which was similar to the ligand binding of CBM36 of xylanase 43 from *P. polymyxa* ATCC 842 (Jamal-Talabani et al., 2004). However, *Px*TW1CBM36-CFP was also capable of binding to other carbohydrates in addition to xylans. Hence, CBM36 from the strain TW1 might have a broad binding ability, compared to CBM36 from those of the strain ATCC 842 (Jamal-Talabani et al., 2004).

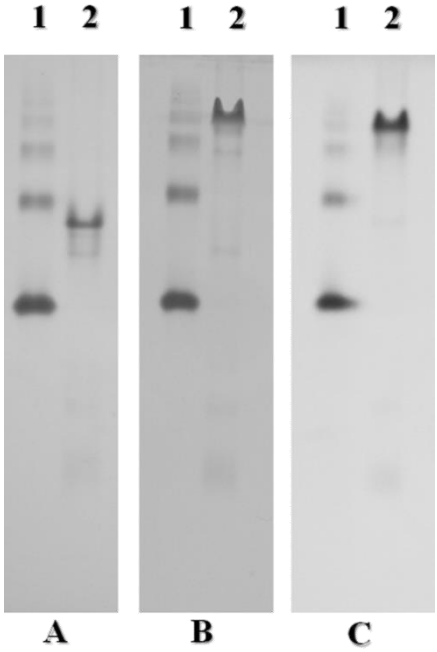


Figure 4.5 Affinity gel electrophoresis of *Px*TW1CBM36-CFP against soluble xylans. *Px*TW1CBM36-CFP in lane 2 was electrophoresed on non-denaturating polyacrylamide gels containing no ligand (A), birchwood xylan (B) or oat-spelt xylan (C). All samples contained approximately 5 µg of protein. BSA was used as a non-carbohydrate binding control (lane 1).

Table 4.2 Semi-quantitative assessment of ligand specificity of *Px*TW1CBM36-CFP determined by affinity gel electrophoresis.

Soluble carbohydrate	Binding ability*
Birchwood xylan	+++
Oat-spelt xylan	+++
Sugar beet arabinan	++
Laminarin	+
Tamarind seed xyloglucan	+
Methylcellulose	+
Carboxymethylcellulose	-
Lichenan	+
Barley β-glucan	-
Guar gum	-
Pectin	-
Starch	-

*Semi-quantitative binding is indicated as tight (+++), moderate (++), weak (+), and no significant interaction (-) depending on how far the proteins migrated on the gel.

4.4.2 Effects of EDTA on the binding ability of *Px*TW1CBM36-CFP

Jamal-Talabani et al. (2004) reported that the CBM36 of *P. polymyxa* ATCC 842 was a calcium-dependent type. This was because the binding property of CBM36 from *P. xylaniclasticus* stain TW1 differed from the strain ATCC 842. Then, to investigate whether ligand recognition was metal-dependent, the ability of *Px*TW1CBM36-CFP to bind to insoluble carbohydrates was determined by incubating the CBM, insoluble carbohydrate and EDTA at a final concentration of 10 mM. The supernatant fraction (unbound protein) and the precipitate fraction (bound protein) obtained by centrifugation were subjected to SDS-PAGE. The binding ability was determined by comparing the protein concentration in the supernatant fraction with that in the precipitate fraction. Interestingly, as shown in Figure 4.6 (I), the strong binding to ball-milled cellulose and insoluble oat-spelt xylan were observed under the presence of 10 mM EDTA. To confirm the influence of

EDTA on the binding ability of *Px*TW1CBM36-CFP, 50 mM EDTA was used for this investigation (Figure 4.6 (II)). It was apparent that, mainly *Px*TW1CBM36-CFP was still bound to both insoluble carbohydrates. These results indicated that the chelating agent (EDTA) had a few effects on the binding of *Px*TW1CBM36-CFP to both insoluble substances. Carbohydrate binding was a strong capability although a high concentration of EDTA was added to the protein. Furthermore, these observations also found that *Px*TW1CBM36-CFP presented a high affinity for both insoluble xylan and cellulose. Generally, the ligand specificity of CBMs reflects the target substrate hydrolyzed by their catalytic module (Carvalho et al., 2004). It is likely that xylan is a biologically significant ligand recognized by CBM36 because they are located predominantly in xylanases (see URL http://www.cazy.org/CBM36_all.html). Thus, it is especially unusual for a *Px*TW1CBM36 to bind to xylan and cellulose.

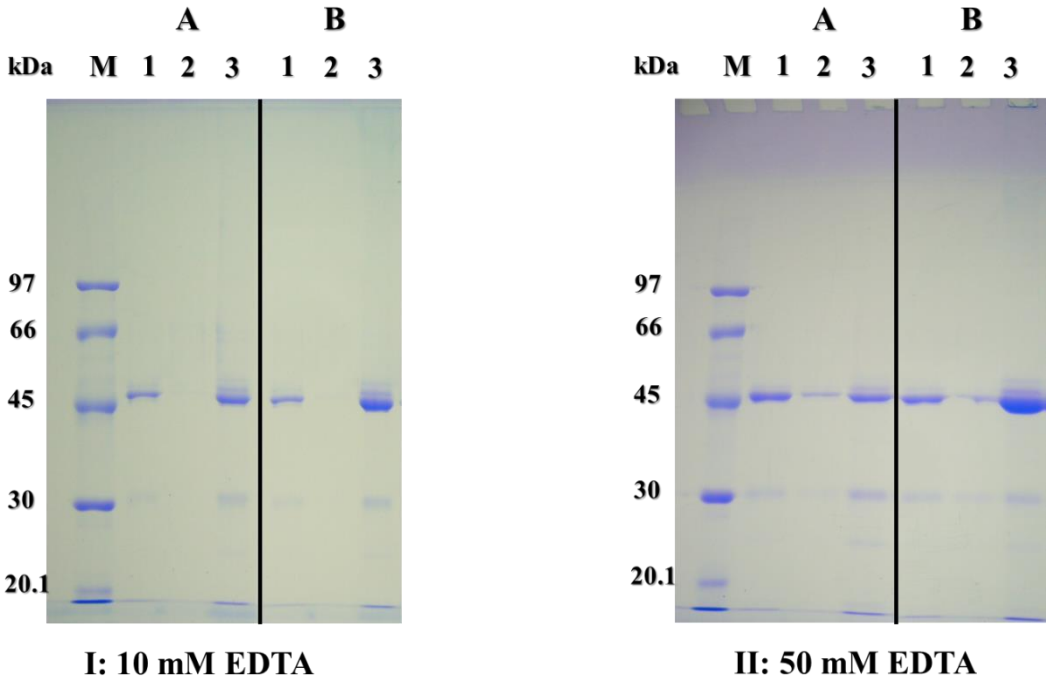


Figure 4.6 The influence of EDTA on binding ability of *Px*TW1CBM36-CFP to insoluble polysaccharides. In this experiment, the *Px*TW1CBM36-CFP was incubated with ball-milled cellulose (A) or oat-spelt xylan (B), containing 10 mM EDTA (I) or 50 mM EDTA (II). After centrifugation, the protein in the supernatant (lane 2) and precipitate (lane 3) were analyzed by SDS-PAGE. *Px*TW1CBM36-CFP was used as a positive control (lane 1).

Sequence-based relatives of *Px*TW1CBM36 using BLAST was also analyzed. The revealed 17 proteins that contain a module with significant similarity to *Px*TW1CBM36 were obtained. All of the proteins were from bacteria, and almost of proteins contained xylanase catalytic modules, revealed that the primary target for this CBM was xylan. However, one protein from *Clostridium phytofermentas* contained a carbohydrate esterase family 4 (CE4) with this family known to also present xylan-specific esterases. In addition, the observation of essentially conserved metal binding sites, derived from the alignment of CBM36 sequences, initially led us to known the residues that coordinate Ca^{2+} in CBM36 (Try29, Asp105, Trp109 and Asp110) were conserved in all proteins. Furthermore, Try15 which dominated the second sugar binding site was also comprised within the family (Figure 4.7).

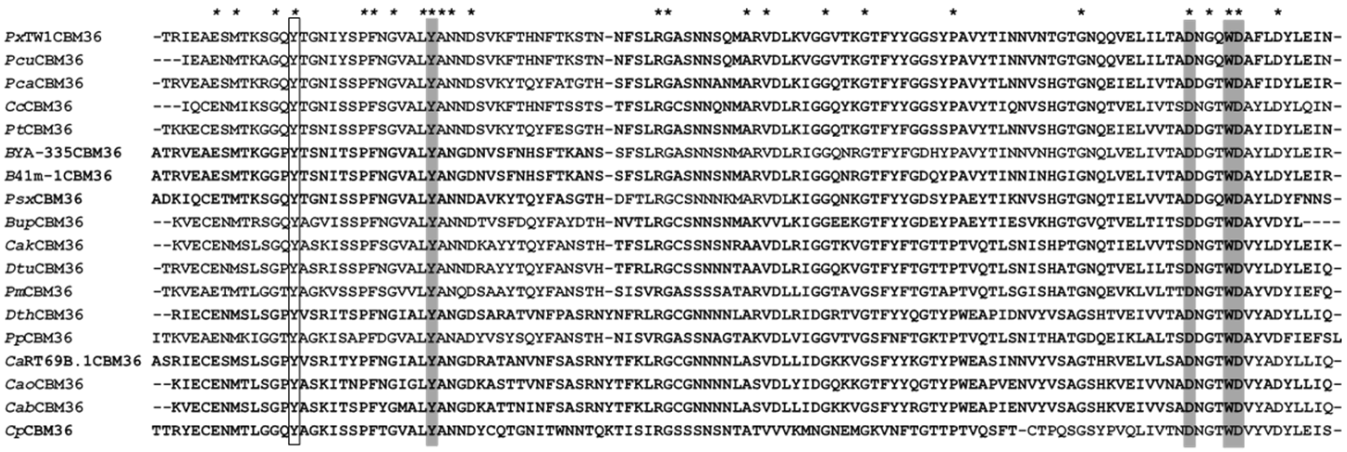


Figure 4.7 Alignment of CBM36 sequences. The sequence alignment was derived from a search of the NCBI database using *Px*TW1CBM36 as the query sequence and the BLASTp search engine. Residues that were invariant within the family were indicated by an asterisk. Amino acids that were the ligand-binding calcium, were highlighted in grey. The aromatic residue that dominated the second sugar binding subsite was highlighted in the open box.

It have been reported that the role of calcium ion, only contributed to carbohydrate recognition in those CBMs that interact with one or, at most, two sugars, but not in those modules that bind to four or more sugars. Moreover, the calcium ion requirement was well explained in CBMs that bind to the internal regions of polysaccharides and calcium could mediated particularly tight binding to sugars, consistent with the view that charged dipole-dipole interactions were stronger than uncharged ones (Fersht et al., 1985). As seem in CBM36 currently characterized,

that from *P. polymyxa* xylanase 43A (Jamal-Talabani et al., 2004), which interact with only two or three internal sugars and the calcium ion plays a dominant role in polysaccharide binding. However, *Px*TW1CBM36 was still binding to insoluble polysaccharides in the presence of EDTA at 50 mM. Thus, it is possible that calcium-dependency might not be the only mode of substrate binding by *Px*TW1CBM36 but the conserved second sugar binding site might also be the functional mode of binding to polysaccharides in the presence of EDTA.

4.4.3 Binding of *Px*TW1CBM36-CFP to the cell wall of sweet potato roots

The capacity of the xylan-binding CBM36 to bind to cell walls in sections of roots from sweet potato was assessed by using Cyan Fluorescent Protein (CFP)-labeled recombinant carbohydrate-binding module (*Px*TW1CBM36-CFP). This analysis shows that the *Px*TW1CBM36-CFP bound extensively to epidermal tissue of root (Figure 4.8, A2). It showed effectively and specifically to bind to epidermal tissue of sweet potato roots this likely reflects the variable structure of xylan within plant tissue. Due to xylans are known to be a taxonomically variable set of polymers, and biochemical analysis have indicated that a β-1,4-linked xylose polymer, which could be decorated with glucuronosyl or 4-O-methyl-glucuronosyl residues at O2 and α-arabinofuranosyl or acetyl groups at O2 or O3, with the nature and extent of substitution varying between species, cell type and plants age (O'Neill and York, 2003; Ebringerova´ and Heinze, 2000). Even more intriguing was the observation that although *Px*TW1CBM36-CFP exhibits higher recognition binding to epidermal tissue than other tissues, its capacity to recognize the polysaccharide within plant cell walls was limited by the older roots (Figure 4.8, B2). The differences in binding capacity profile of *Px*TW1CBM36-CFP with 1-week-root olds and 1-month-root olds was likely to reflect a combination of the accessibility of the target ligands within the cell wall composites and fine details of xylan structure. In nature, the older plant, xylan chains were highly intimately associated with other cell wall components to form the hard and rigid structure, providing an explanation for why the CBM displayed the weak binding to xylan when it is embedded in the plant cell wall.

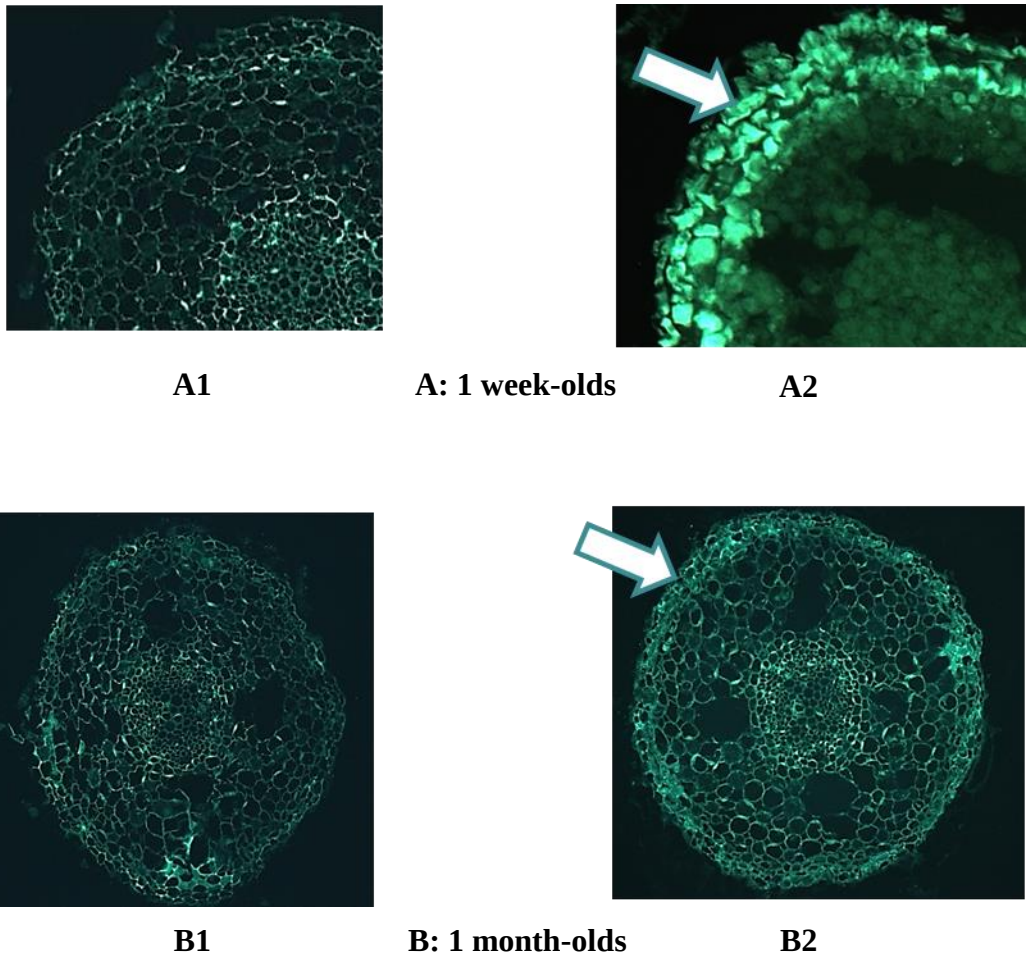


Figure 4.8 Binding of *Px*TW1CBM36-CFP to cell walls of 1-week-olds (A) and 1-month-olds (B) sweet potato roots. In (1) the plant cell walls were non-labeling, whereas in (2) the cell walls were labeling with *Px*TW1CBM36-CFP. Arrows indicate recognition of epidermal tissues.

4.5 Carbohydrate-active enzymes in *P. xylaniclasticus* strain TW1 genome

In this report not only carbohydrate-binding module but carbohydrate-active enzymes in *P. xylaniclasticus* stain TW1 genome were also investigated. The nucleotide sequence was determined using a 454 GS FLX sequencer, and assembly was performed using the GS De Novo Assembler software version 2.6. A total of 348,957 reads including up to 139,127,032 bp were obtained. Finally, we obtained the *P. xylaniclasticus* strain TW1 draft genome of 5,081,541 bp distributed in 112 contigs. From nucleotide sequence, the data exhibited many carbohydrate-active enzymes in *P. xylaniclasticus* strain TW1 genome. It contained 28 xylanolytic enzymes, 9 cellulolytic enzymes and 3 other plant cell wall-degrading enzymes (Table 4.3). In addition, not only CBM36 but *P. xylaniclasticus* strain TW1 contained many CBM which different family. As introduction, the complete and rapid hydrolysis of polysaccharides in agricultural residues requires

not only cellulolytic enzymes but also the cooperation of many enzymes such as xylanolytic enzymes and side chain-cleaving enzymes such as deacetylase and α-L-arabinofuranosidase. Thus, this result explained that why crude enzyme produced from *P. xylaniclasticus* strain TW1 displayed the capable of binding to insoluble xylan and cellulose and efficiently hydrolyzed raw agricultural residues.

Table 4.3 List of carbohydrate-active enzymes identified in the annotated genome dataset of *P. xylaniclasticus* strain TW1.

ORF	Xylanolytic enzyme
701	GH11 xylanase-CBM36
2480	GH74 xyloglucanase-CBM3-unknown
5677	GH10 xylanase-Fn3-CBM3-unknown
5821	GH43 beta-xylosidase-unknown2
7692	GH10 xylanase
8099	GH43 alpha-arabinofruranosidase
9962	GH11 xylanase
10482	GH43 beta-xylosidase-CBM6
10270	Esterase (acetylxylan esterase)-Fn3-Fn3-CBM3-unknown
10740	Esterase
11776	Unknown7-CE4 deacetylase
12456	GH8 xylanase
14115	CE4 deacetylase-unknown4
14429	CE4 deacetylase
14607	GH43 beta-xylosidase-CBM6
14618	GH43 beta-xylosidase

Table 4.3 (Cont.) List of carbohydrate-active enzymes identified in the annotated genome dataset of *P. xylaniclasticus* strain TW1.

ORF	Xylanolytic enzyme
15422	GH11 xylanase-CBM36
15471	Esterase
19642	Esterase-Fn3-CBM3-unknown
22604	CE4 deacetylase
25386	GH43 beta-xylosidase-CBM6-CBM36
25489	CE4 deacetylase
26091	GH10 xylanase
26123	Unknown5-CE4 deacetylase-unknown6
26629	GH43 xylosidase
27382	Esterase
27461	GH10 xylanase-CBM9-Fn3-Fn3-CBM3-unknown
27501	GH10 xylanase-CBM9-SLH
ORF	**Cellulolytic enzyme**
6016	GH9 cellulase
6648	GH5 cellulase-CBM17-CBM28-Fn3-SLH
7862	GH9 cellulase-Fn3-Fn3-CBM3-unknown
7869	GH48 cellobiohydrolase-Fn3-Fn3-Fn3-CBM3-unknown
7875	GH6 cellulase-Fn3-Fn3-Fn3-CBM3-unknown
7881	GH5 cellulase-Fn3-CBM3-unknown

Table 4.3 (Cont.) List of carbohydrate-active enzymes identified in the annotated genome dataset of *P. xylaniclasticus* strain TW1.

ORF	Cellulolytic enzyme
7884	GH5 cellulase-Fn3-CBM3-unknown
10492	E-Cellulase domain 2-GH9
10167	GH5 cellulase-CBM11-CBM3-unknown
10188	GH5 cellulase-Fn3-Fn3-CBM3-unknown
ORF	**Other plant cell wall-degrading enzyme**
2857	PL11 Polysaccharide lyase (Rhamnogalactan lyase)-unknown
16478	GH26 mannanase-unknown
24568	GH26 mannanase-unknown3

References

Abbott, D.W., Eirin-Lopez, J.M. and Boraston, A.B. (2008) Insight into ligand diversity and novel biological roles for family 32 carbohydrate-binding modules. *Molecular Biology and Evolution*, Vol. 25, pp.155-167.

Araki, R., Ali, K., Sakka, M., Kimura, T., Sakka, K. and Ohmiya, K. (2004) Essential role of the family-22 carbohydrate-binding modules for β-1,3-1,4-glucanase activity of *Clostridium stercorarium* Xyn10B. *FEBS Letters*, Vol. 561, pp. 155-158.

Araki, Y., Karita, S., Tsuchiya, T., Kondo, M. and Goto, M. (2010) Family 17 and 28 carbohydrate-binding modules discriminated different cell-wall sites in sweet potato roots. *Bioscience Biotechnology and Biochemistry*, Vol. 74, pp. 802-805.

Banerjee, G., Scott-Craig, J.S. and Walton, J.D. (2010) Improving enzymes for biomass conversion: a basic research perspective. *BioEnergy research*, Vol. 3, pp. 82-92.

Bayer, E.A., Kenig, R. and Lamed, R. (1983) Adherence of *Clostridium thermocellum* to cellulose. *Journal of Bacteriology*, Vol. 156, pp. 818-827.

Bayer, E.A., Morag, E. and Lamed, R. (1994) The cellulosome a treasure-trove for biotechnology. *Trends in Biotechnology*, Vol. 12, pp. 379-386.

Bayer, E.A., Chanzy, H., Lamed, R. and Shoham, Y. (1998) Cellulose, cellulase and cellulosome. *Journal of Structural Biology*, Vol. 8, pp. 548-557.

Bayer, E.A., Belaich, J.-P., Shoham, Y. and Lamed, R. (2004) The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annual Review of Microbiology*, Vol. 58, pp. 521-554.

Beg, Q.K., Kapoor, M., Mahajan, L. and Hoondal, G.S. (2001) Microbial xylanases and their industrial applications: a review. *Applied Microbiology and Biotechnology*, Vol. 56, pp. 326-338.

Beguin, P. and Lemaire, M. (1996) The cellulosome: An exocellular, multiprotein complex specialized in cellulose degradation. *Critical Reviews in Biochemistry and Molecular Biology*, Vol. 31, pp. 201-236.

Berg, B., Hofsten, B.V. and Pettersson, G. (1972) Growth and cellulase formation by *Cellvibrio fulvus*. *Journal of Applied Bacteriology*, Vol. 35, pp. 201-214.

Bhat, M.K. and Bhat, S. (1997) Cellulose degrading enzymes and their potential industrial applications. *Biotechnology advance*, Vol. 15, pp. 583-620.

Biely, M. (1985) Microbial xylanolytic systems. *Trends in Biotechnology*, Vol. 3, pp. 286-290.

Bolam, D.N., Ciruela, A., McQueen-Mason, S., Simpson, P., Williamson, M.P., Rixon, J.E., Boraston, A., Hazlewood, G.P. and Gilbert, H.J. (1998) *Pseudomonas* cellulose-binding domains mediate their effects by increasing enzyme substrate proximity. *Biochemical Journal*, Vol. 331, pp. 775-781.

Boraston, A.B., Kwan, E., Chiu, P., Warren, R.A. and Kilburn, D.G. (2003) Recognition and hydrolysis of noncrystalline cellulose. *Journal of Biological Chemistry*, Vol. 278, pp. 6120-6127.

Boraston, A.B., Bolam, D.N., Gilbert, H.J. and Davies, G.J. (2004) Carbohydrate-binding modules: finetuning polysaccharide recognition. *Biochemical Journal*, Vol. 382, pp. 769-781.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V. and Henrissat, B. (2009) The carbohydrate-active enzymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Research*, Vol. 37, pp. D233-238.

Carvalheiro, F., Duarte L.C. and Gırio, F.M. (2008) Hemicelluloses biorefineries: a review on biomass pretreatments. *Journal of Scientific and Industrial Research*, Vol. 67, pp. 849-864.

Carvalho, A.L., Goyal, A., Prates, J.A.M., Bolam, D.N., Gilbert, H.J., Pires, V.M.R., Ferreira, L.M.A., Planas, A., Romão, M.J. and Fontes, C.M.G.A. (2004). The family 11 carbohydrate-binding module of *Clostridium thermocellum* Lic26A-Cel5E accommodates β-1,4- and β-1,3-1,4-mixed linked glucans at a single binding site. *Journal of Biological Chemistry*, Vol. 279, pp. 34785-34793.

Coughlan, M.P. and Hazlewood, G.P. (1993) β-1,4-D-xylan degrading enzyme systems: Biochemistry, molecular biology and applications. *Biotechnology and Applied Biochemistry*, Vol. 17, pp. 259-289.

Davies, G. and Henrissat, B. (1995) Structures and mechanisms of glycosyl hydrolases. *Structure*, Vol. 3, pp. 853-859.

Din, N., Damude, H.G., Gilkes, N.R., Miller, R.C., Warren, R.A. and Kilburn, D.G. (1994) C1-Cx revisited: intramolecular synergism in a cellulase. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 91, pp. 11383-11387.

Dyk, J.S.V. and Pletschke, B.I. (2012) A review of lignocellulose bioconversion using enzymatic hydrolysis and synergistic cooperation between enzymes-Factors affecting enzymes, conversion and synergy. *Biotechnology Advances*, Vol. 30, pp. 1458-1480.

Ebringerova´, A. and Heinze, T. (2000) Xylan and xylan derivatives-biopolymers with valuable properties, 1. Naturally occurring xylans structures, isolation procedures and properties. *Macromolecular Rapid Communications*, Vol. 21, pp. 542-556.

Fan, L.T., Gharpuray, M.M. and Lee, Y.H. (1987) Cellulose hydrolysis, in *Biotechnology Monographs*, Aiba, S., Fan, L.T., Fiechter, A., Klein, J. and Schugerl, K. (Eds), Vol. 3, Springer-Verlag, Berlin Heidelberg, pp.1-16.

Fersht, A.R., Shi, J.-P., Knill-Jones, J., Lowe, D.M., Wilkinson, A.J., Blow, D.M., Brick, P., Carter, P., Waye, M.M.Y. and Winter, G. (1985) Hydrogen bonding and biological specificity analysed by protein engineering. *Nature*, Vol 314, pp. 235-238.

Fierobe, H.-P., Pages, S., Belaich, A., Champ, S., Lexa, D. and Belaich, J.P. (1999) Cellulosome from *Clostridium cellulolyticum*: Molecular study of the dokerin/cohesin interaction. *Biochemistry*, Vol. 38, pp.12822-12832.

Filonova, L., Gunnarsson, L.C., Daniel, G. and Ohlin, M. (2007) Synthetic xylan-binding modules for mapping of pulp fibres and wood sections. *BMC Plant Biology*, Vol. 7, pp. 54.

Gilbert, H.J. (2010) The biochemistry and structural biology of plant cell wall deconstruction. *Plant Physiology*, Vol. 153, pp. 444-455.

Gilkes, N.R., Henrissat, B., Kilburn, D.G., Jr, M.R.C. and Warren, R.A. (1991) Domains in microbial beta-1,4-glycanases: Sequence conservation, function, and enzyme families. *Microbiological Reviews*, Vol. 55, pp. 303-315.

Gregg, K.J., Finn, R., Abbott, D.W. and Boraston, A.B. (2008) Divergent modes of glycan recognition by a new family of carbohydrate-binding modules. *Journal of Biological Chemistry*, Vol. 283, pp. 12604-12613.

Han, Y. and Chen, H. (2007) Synergism between corn stover protein and cellulase. *Enzyme and Microbial Technology*, Vol. 41, pp. 638-645.

Harris, P.V, Welner, D., McFarland, K.C., Re, E., Poulsen, J-C.N., Brown, K., Salbo, R., Ding, H., Vlasenko, E., Merino, S., Xu, F., Cherry, J., Larsen, S. and Leggio, L.L. (2010) Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase Family 61: structure and function of a large, enigmatic family. *Biochemistry*, Vol. 49, pp. 3305-16.

Henrissat, B. and Bairoch, A. (1996) Updating the sequence-based classification of glycosyl hydrolases. *Biochemical Journal*, Vol. 316, pp. 695-696.

Higuchi, T. (1978) Lignin structure and morphological distribution in plant cell walls, in *Lignin Biodegradation: Microbiology, Chemistry, and Potential Applications*, Kirk, K.T., Higuchi, T. and Chang, H.M. (Eds), Vol. 1, CRC Press, Boca Raton, pp. 1-17.

Higuchi, T. (1997) Biochemistry and molecular biology of wood, Springer, Berlin.

Hilden, L. and Johansson, G. (2004) Recent developments on cellulases and carbohydrate-binding modules with cellulose affinity. *Biotechnology Letters*, Vol. 26, pp. 1683-1693.

Jamal-Talabani, S., Boraston, A.B., Turkenburg, J.P., Tarbouriech, N., Ducros, V.M.-A. and Davies, G.J. (2004) Ab initio structure determination and functional characterization of CBM36: a new family of calcium-dependent carbohydrate binding modules. *Structure*, Vol. 12, pp. 1177-1187.

Jiang, Z.Q., Deng, W., Li, X.T., Ai, Z.L., Li, L.T. and Kusakabe, I. (2005) Characterization of a novel, ultra-large xylanolytic complex (xylanosome) from *Streptomyces olivaceoviridis* E-86. *Enzyme and Microbial Technology*, Vol. 36, pp. 923-929.

Juturu, V. and Wu, J.C. (2014) Microbial cellulases: Engineering, production and applications. *Renewable and Sustainable Energy Reviews*, Vol. 33, pp. 188-203.

Kaewintajuk, K., Chon, G.H., Lee, J.-S., Kongkiattikajorn, J., Rattanakhanokchai, K., Kyu, K.L., Lee, J.H., Roh, M.S., Choi, Y.Y., Park, H. and Lee, Y.S. (2006) Hydrolysis of agricultural residues and kraft pulps by xylanolytic enzymes from alkaliphilic *Bacillus* sp. strain BK. *Journal of Microbiology and Biotechnology*, Vol. 16, pp. 1255-1261.

Kotiranta, P., Karlsson, J., Siika-aho, M., Medve, J., Viikari, L., Tjerneld, F. and Tenkanen, M. (1999) Adsorption and activity of *Trichoderma reesei* cellobiohydrolase I, endoglucanase II, and the corresponding core proteins on steam pretreated willow. *Applied Biochemistry and Biotechnology*, Vol. 81, pp. 82-90.

Kulkarni, N., Shendye, A. and Rao, M. (1999) Molecular and biotechnological aspects of xylanases. *FEMS Microbiology Reviews*, Vol. 23, pp. 411-456.

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

Lamed, R., Setter, E. and Bayer, E.A. (1983) Characterization of cellulose-binding, cellulase-containing complex in *Clostridium thermocellum*. *Journal of Bacteriology*, Vol. 156, pp. 828-836.

Lin, L.L. and Thomson, J.A. (1991) An analysis of the extracellular xylanases and cellulases of *Butyrivibrio fibrisolvens* h17c. *FEMS Microbiology Letters*, Vol. 68, pp. 197-204.

Lynd, L.R., Weimer, P.J., Van Zyl, W.H. and Pretorius, I.S. (2002) Microbial cellulose utilization: Fundamentals and biotechnology. *Microbiology and Molecular Biology Reviews*, Vol. 66, pp. 506-577.

Mach, H., Middaugh, C.R. and Lewis, R.V. (1992) Statistical determination of the average values of the extinction coefficients of tryptophan and tyrosine in native proteins. *Analytical Biochemistry*, Vol. 200, pp. 74-80.

McCartney, L., Gilbert, H.J., Bolam, D.N., Boraston, A.B. and Knox, J.P. (2004) Glycoside hydrolase carbohydrate-binding modules as molecular probes for the analysis of plant cell wall polymers. *Analytical Biochemistry*, Vol. 326, pp. 49-54.

Merino, S.T. and Cherry, J. (2007) Progress and challenges in enzyme development for biomass utilization. *Advances in Biochemical Engineering/Biotechnology*, Vol. 108, pp. 95-120.

Murashima, K., Kosugi, A. and Doi, R.H. (2002) Synergistic effects on crystalline cellulose degradation between cellulosomal cellulases from *Clostridium cellulovorans*. *Journal of Bacteriology*, Vol. 184, pp. 5088-5095.

Ohara, H., Karita, S., Kimura, T., Sakka, K. and Ohmiya, K. (2000) Characterization of the cellulolytic complex (cellulosome) from *Ruminococcus albus*. *Bioscience Biotechnology and Biochemistry*, Vol. 64, pp. 254-260.

Oliveira, C., Carvalho, V., Domingues, L. and Gama, F.M. (2015) Recombinant CBM-fusion technology-Applications overview. *Biotechnology Advances*, Vol. 33, pp. 358-369.

O'Neill, M.A. and York, W.S. (2003). The composition and structure of plant primary cell walls. In The Plant Cell Wall, pp. 1-54. Edited by J.K.C. Rose. Ithaca, New York, Blackwell Publishing/CRC Press.

Perlack, R.D., Wright, L.L., Turhollow, A.F., Graham, R.L.; Sotcks, B.J. and Erbach, D.C. (2005) Biomass as feedstock for a bioenergy and bioproducts industry: the technical feasibility of a billion-ton annual supply, DOE/GO-102005-2135, April. U. S. Department of Energy and U.S. Department of Agriculture.

Puls, J. and Schuseil, J. (1993) Chemistry of hemicelluloses: relationship between hemicellulose structure and enzymes required for hydrolysis, in *Hemicellulose and Hemicellulases*, Coughlan, M.P., and Hazlewood, G.P. (Eds), Portland Press Ltd, London, pp.1-28.

Saito, H. and Miura, K.I. (1963) Preparation of transforming deoxyribonucleic acid by phenol treatment. *Biochimica et Biophysica Acta*, Vol. 72, pp. 619-629.

Shin, G.-W., Palaksha, K.J., Yang, H.-H., Shin, Y.-S., Kim, Y.-R., Lee, E.-Y., Oh, M.J. and Jung, T.S. (2006) Partial two-dimensional gel electrophoresis (2-DE) maps of *Streptococcus iniae* ATCC29178 and *Lactococcus garvieae* KG9408. *Diseases of Aquatic Organisms*, Vol. 70, pp. 71-79.

Shoham, Y., Lamed, R. and Bayer, E.A. (1999) The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides. *Trends in Microbiology*, Vol. 7, pp. 275-281.

Shoseyov, O., Shani, Z. and Levy, I. (2006) Carbohydrate binding modules: biochemical properties and novel applications. *Microbiology and Molecular Biology Reviews*, Vol. 70, pp. 283-295.

Sunna, A. and Antranikian, G. (1997) Xylanolytic enzymes from fungi and bacteria. *Critical Reviews in Biotechnology*, Vol. 17, pp. 39-67.

Tachaapaikoon, C., Kyu, K.L. and Ratanakhanokchai, K. (2006) Purification of xylanase from alkaliphilic *Bacillus* sp. K-8 by using corn husk column. *Process Biochemistry*. Vol. 41, pp. 2441-2445.

Tachaapaikoon, C., Kyu, K.L., Pason, P. and Ratanakhanokchai, K. (2012a) A novel multienzyme complex from a newly isolated facultative anaerobic bacterium. *Paenibacillus* sp. TW1, *Acta Biologica Hungarica*, Vol. 63, pp. 288-300.

Tachaapaikoon, C., Tanasupawat, S., Pason, P., Sornyotha, S., Waeonukul, R., Kyu, K.L. and Ratanakhanokchai, K. (2012b) *Paenibacillus xylaniclasticus* sp. nov., a xylanolytic-cellulolytic bacterium isolated from sludge in an anaerobic digester. *The Journal of Microbiology*, Vol. 50, pp. 394-400.

Tomme, P., Boraston, A., McLean, B., Kormos, J., Creagh, A.L., Sturch, K., Gilkes, N.R., Haynes, C.A., Warren, R.A. and Kilburn, D.G. (1998) Characterization and affinity applications of cellulose-binding domains. *Journal of Chromatography B: Biomedical Sciences and Applications*, Vol. 715, pp. 283-296.

Varnai, A., Siika-aho, M. and Viikari, L. (2010) Restriction of the enzymatic hydrolysis of steam pretreated spruce by lignin and hemicellulose. *Enzyme and Microbial Technology*, Vol. 46, pp. 185-193.

Wood, P.J., Weisz, J. and Blackwell, B.A. (1994) Structural studies of (1⟶3),(1⟶4)-β-D-glucans by ^{13}C-nuclear magnetic resonance spectroscopy and by rapid analysis of cellulose-like regions using high-performance anion-exchange chromatography of oligosaccharides released by lichenase. *Cereal Chemistry*, Vol. 71, 301-307.

Output

1. Proceeding/Conferences /Symposiums

1.1 **Sornyotha, S.**, Karita, S., Tachaapaikoon, C. and Ratanakhanokchai, K., 2013, “Affinity isolation and rapid identification of carbohydrate-binding proteins of xylanolytic-cellulolytic multienzyme complexs from *Paenibacillus xylaniclasticus* TW1 by MALDI-TOF/TOF MS”, The 1st International Symposium on Microbial Technology for Food and Energy Security, November 25-27, The Rama Gardens Hotel, Bangkok, Thailand, pp. 281-287.

1.2 **Sornyotha, S.**, Karita, S., Tsuchiya, T., Tachaapaikoon, C. and Ratanakhanokchai, K., 2014, “A novel xylan-binding CBM family 36 from multienzyme complex-producing bacterium, *Paenibacillus xylaniclasticus* TW1: A high affinity for insoluble cellulose in addition to xylans and specificity for target carbohydrate in epidermis cell walls”, MIE BIOFORUM 2014-Lignocellulose Degradation and Biorefinery, November 18-21, Shima City, Mie, Japan, pp. 111-118.

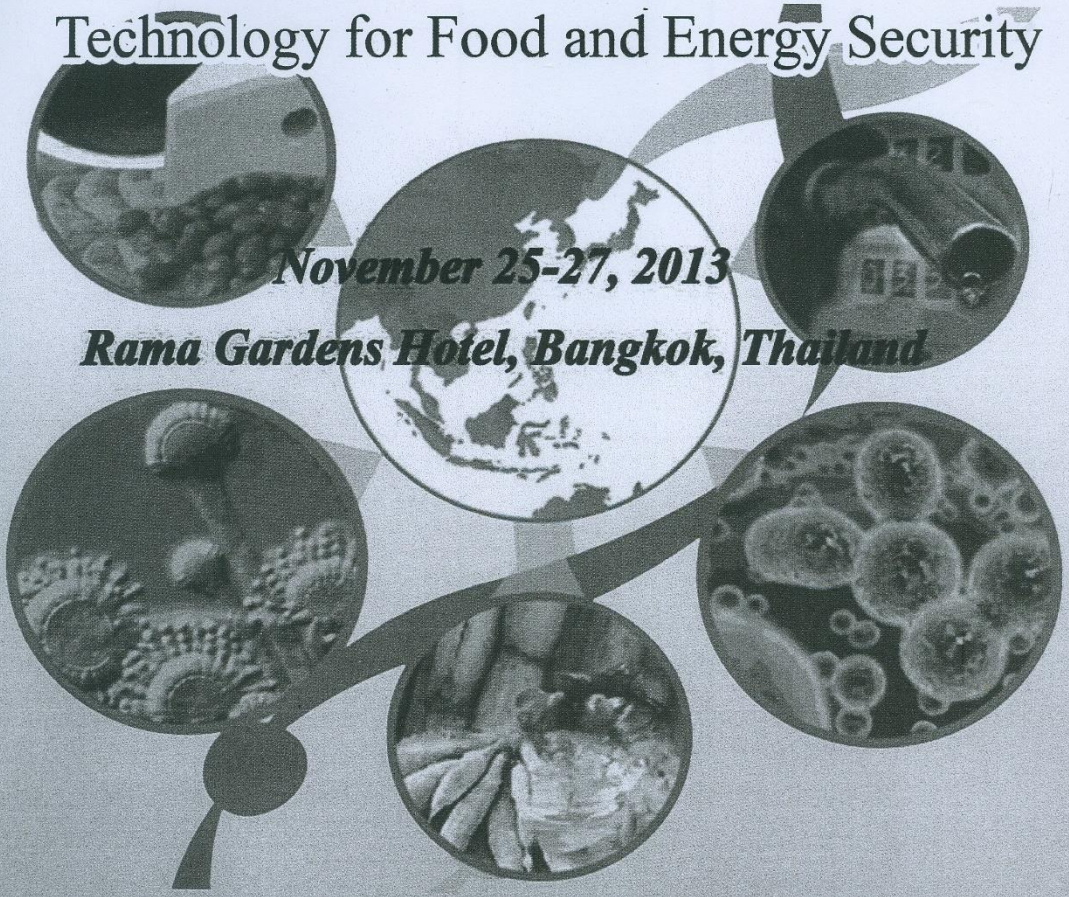
Appendix

Appendix A

Proceeding under title: Affinity isolation and rapid identification of carbohydrate-binding proteins of xylanolytic-cellulolytic multienzyme complexs from *Paenibacillus xylaniclasticus* TW1 by MALDI-TOF/TOF MS.

Sornyotha, S., Karita, S., Tachaapaikoon, C. and Ratanakhanokchai, K., 2013, “Affinity isolation and rapid identification of carbohydrate-binding proteins of xylanolytic-cellulolytic multienzyme complexs from *Paenibacillus xylaniclasticus* TW1 by MALDI-TOF/TOF MS”, The 1st International Symposium on Microbial Technology for Food and Energy Security, November 25-27, The Rama Gardens Hotel, Bangkok, Thailand, pp. 281-287.

The 1st International Symposium on Microbial Technology for Food and Energy Security



November 25-27, 2013

Rama Gardens Hotel, Bangkok, Thailand

Kasetsart University

and

Japan International Research Center for Agricultural Sciences (JIRCAS), Japan

General Organizing Committee

Masayoshi Saito (Chairperson)	Japan International Research Center for Agricultural Sciences (JIRCAS)
Patcharee Tungtrakul (Thailand)	Institute of Food Research and Product Development (IFRPD)
Warunee Thanapsase (Thailand)	KAPI, Kasetsart University
Gassinee Trakoontivakorn (Thailand)	IFRPD, Kasetsart University
Chakrit Tachaapaikoon (Thailand)	King Mongkut's University of Technology Thonburi
Achara Nuntagij (Thailand)	Department of Agriculture
Akihiko Kosugi (Japan)	Japan International Research Center for Agricultural Sciences (JIRCAS)
Yoshinori Murata (Japan)	Japan International Research Center for Agricultural Sciences (JIRCAS)
Takamitsu Arai (Japan)	Japan International Research Center for Agricultural Sciences (JIRCAS)
K Sudesh Kumar (Malaysia)	University Science Malaysia
Cheng Yongqiang (China)	China Agricultural University
Joko Sulistyo (Indonesia)	Indonesian Institute of Science
Wolfgang Schwarz (Germany)	Technische Universität München

SECRETARIES:

Kazuhiko Nakahara (Japan)	JIRCAS
Pilanee Vaithanomsat (Thailand)	KAPI, Kasetsart University

Advisory Committee

Vudtechai Kapilakanchana (President), Kasetsart University, Thailand
Masa Iwanaga (President), JIRCAS, Japan
Thira Sutabutra (Former Minister of Agriculture and Cooperatives), Thailand
Yupa Pankaew (Former Director of KAPI), Kasetsart University
Klanarong Sriroth, Kasetsart University

Local Organizing Committee (Thailand)

Patcharee Tungtrakul (Chairperson)	Kasetsart University
Khanok Ratanakhanokchai	King Mongkutt's University of Technology Thonburi
Warunee Thanapase	Kasetsart University
Sarote Sirisansaneeyakul	Kasetsart University
Poonsuk Prasertsan	Prince of Songkla University
Suwayd Ningsanond	Suranaree University of Technology
Chakrit Tachaapaikoon	King Mongkut's University of Technology Thonburi
Sumol Nilratnisakorn	University of Phayao
Achara Nuntagi	Department of Agriculture
Phatchayaphon Meunchang	Department of Agriculture

SECRETARIES:

Pilanee Vaithanomsat	Kasetsart University
Gassinee Trakoontivakorn	Kasetsart University

The 1st International Symposium on Microbial Technology for Food and Energy Security

EP05: Affinity Isolation and Rapid Identification of Carbohydrate-binding Proteins of Xylanolytic-cellulolytic Multienzyme Complexs from *Paenibacillus xylaniclasticus* TW1 by MALDI-TOF/TOF MS

Somphit Sornyotha,[1] Shuichi Karita,[2] Chakrit Tachaapaikoon[3] and Khanok Ratanakhanokchai[3]

[1]King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520 Thailand,
E-mail: kssomphi@kmitl.ac.th

[2]Mie University, Tsu 514-8507 Japan

[3]King Mongkut's University of Technology Thonburi, Bangkok 10150 Thailand

Abstract

A facultative anaerobic bacterium, *Paenibacillus xylaniclasticus* TW1, isolated from an anaerobic digester, produces an extracellular xylanolytic-cellulolytic multienzymes system when grown on corn hull under aerobic conditions. The enzyme was found to be capable of binding to insoluble substrates and efficiently hydrolyzed raw agricultural residues. The carbohydrate-binding proteins (CBPs) were isolated from the crude enzyme preparation by affinity purification on ball-milled cellulose. This showed 12 proteins by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Moreover, the results from matrix-assisted laser desorption/ionization time-of-flight/time-of-flight mass spectrometry (MALDI-TOF/TOF MS) found that the peptide mass fingerprints of the unknown components from SDS-PAGE were almost identical as unknown proteins. This result suggested that the other proteins were possibly present as the new glycoside hydrolase family. The only two bands from SDS-PAGE could be identified as the glycoside hydrolase family 31 (CBP5) and family 11 (CBP12). Surprisingly, both of identical proteins showed CMCase activity on zymogram. However, no cellulase activity was reported within family 31 and family 11. Moreover, although CBP5 showed similarity to published glycoside hydrolase family 31 sequences from *P. curdlanolyticus* YK9 but no information on the carbohydrate-binding module (CBM) appearing in the glycoside hydrolase family 31 from *P. curdlanolyticus* YK9 was presented. In case of CBP12, the peptide sequence of CBP12 showed similarity to *P. curdlanolyticus* B6 xylanase with a CBM family 36. Intriguingly, no report on the binding of CBM36 to cellulose was discovered. These data indicated that both of identical proteins possibly displayed the new hydrolase family with the new CBM family or the new binding properties. Hence, CBP5 and CBP12 should be further investigated for elucidation of the different properties of these proteins as compared to previous report.

Keywords: Carbohydrate-binding proteins, MALDI-TOF/TOF MS, *Paenibacillus xylaniclasticus TW1, Xylanolytic-cellulolytic multienzyme complexes*

Introduction

Efficient conversion of lignocellulosic materials such as agricultural wastes into fuel ethanol has become a world priority for producing environmentally friendly renewable energy at reasonable price for the transportation sector. Fuel ethanol can be utilized as oxygenate of gasoline elevating its oxygen content, allowing a better oxidation of hydrocarbons and reducing the amounts of greenhouse gas emissions into the atmosphere. Moreover, lignocellulosic materials such as agricultural wastes are attractive feedstocks for bioethanol production, due to agricultural wastes are cost effective, renewable and abundant.

Many cellulolytic microorganisms and their cellulase systems have been studied extensively for degradation of naturally abundant lignocelluloses to valuable products such as fermentable sugars, chemicals, and liquid fuel. Efficient enzymatic degradation of insoluble lignocelluloses requires a tight interaction between the enzymes and their substrates and the cooperation of multiple enzymes to enhance the hydrolysis due to the complex structure. Cellulases and xylanases containing cellulose-binding domains (CBDs) and/or xylan-binding domains (XBDs) are associated into a high molecular weight

multienzyme complex, cellulosome, and have been found in many anaerobic cellulolytic microorganisms and multienzyme complexes (Schwarz, 2001). These complexes are dedicated to hydrolyzing lignocelluloses because of their ability to bind to insoluble cellulose and/or xylan via CBDs and XBDs, respectively. The CBMs affect binding to the substrate surface, presumably to facilitate hydrolysis by bringing the catalytic domain in close proximity to the insoluble substrate. Thus, the arrangement of enzymes into a multienzyme complex has advantage over single enzyme systems. The most intensively studied multienzyme complex is the cellulosome produced by anaerobic thermophilic *Clostridium thermocellum* (Bayer and Lamed, 1986). However, the few studies concerning multienzyme complex production under aerobic conditions were reported (Pason et al., 2006). Therefore, we have undertaken a study on the xylanolytic-cellulolytic multienzyme complex production under aerobic conditions for the degradation of lignocellulosic materials.

From previous study, a facultative, mesophilic, anaerobic bacterium strain, TW1 (**Figure 1**) was isolated from an anaerobic digester. Strain TW1 produced xylanolytic-cellulolytic multienzyme complexes when grown on corn hull under aerobic conditions (Tachaapaikoon et al., 2012a). The 16S rRNA gene analysis revealed that strain TW1 resembled *P. cellulosilyticus*, *P. curdlanolyticus* and *P. kobensis* (97.6-97.1% similarity). However, on the basis of DNA-DNA relatedness and phenotypic analyses, phylogenetic data and the enzymatic pattern presented by Tachaapaikoon et al. (2012b), the isolated strain TW1 was classified as a novel species of the genus *Paenibacillus*, namely *P. xylaniclasticus* TW1.

Although, xylanolytic-cellulolytic multienzyme complexes produced from *P. xylaniclasticus* TW1 could effectively hydrolyze the agricultural wastes and bind substrates (Tachaapaikoon et al., 2012a), the CBMs appearing in the xylanolytic-cellulolytic multienzyme complexes from this bacterium is still unknown. Then, this work, the secreted proteins that appear to contain CBMs will be identified and the results will be useful to improve the cost of hydrolysis process, which could be reduced processing cost using immobilization of enzyme via CBMs.

Materials and methods

Enzyme preparation

Paenibacillus xylaniclasticus TW1, the multienzyme producing strain, was isolated from an anaerobic digester. It was grown in Berg's mineral salt medium (pH 7.0) (Berg et al., 1972) containing 0.2% $NaNO_3$, 0.05% K_2HPO_4, 0.02% $MgSO_4 \cdot 7H_2O$, 0.002% $MnSO_4 \cdot H_2O$, 0.002% $FeSO_4 \cdot 7H_2O$, 0.002% $CaCl_2 \cdot 2H_2O$, and 0.5% corn hull as the sole carbon source. The culture was incubated in a rotary shaker at 200 rpm and 37 °C for 3 days. After centrifugation at 8,000 × g for 10 min at 4 °C, the culture supernatant was concentrated about 50-fold by freeze dryer.

Isolation of carbohydrate-binding proteins

The carbohydrate-binding proteins, which are membered in the multienzyme complex, were isolated using affinity purification on cellulose. The isolation method was conducted by adding 250 µl of protein from the 50-fold concentrated culture supernatant to 0.015 g of ball-milled cellulose in 10 mM potassium phosphate buffer (KPB, pH 7.0) in 1.5-ml Eppendorf tubes. Sample was incubated at 4 °C for 30 min before centrifugation. The precipitate was washed 5 times with same buffer, and the precipitate and supernatant were analyzed by SDS-PAGE.

Gel electrophoresis analysis and zymogram

SDS-PAGE was performed on a 7% polyacrylamide gel by the method of Laemmli (1970). After that, the proteins were stained with Coomassie brilliant blue R-250. The molecular weight standards used were from a molecular-weight calibration kit (Bio-Rad). CMCase zymogram was prepared from SDS-7% polyacrylamide gels containing 0.1% carboxymethylcellulose (CMC), as described previously (Ratanakhanokchai et al., 1999).

Protein identification using MALDI-TOF/TOF MS

Selected protein bands (12 bands) were excised from SDS-PAGE gels and subjected to in-gel digestion with trypsin as described previously (Shin et al., 2006). MS/MS analyses to identify proteins were performed on a JP4800 Plus MALDI-TOF/TOF Analyzer (Applied Biosystems) using a 200 Hz ND: YAG laser operating at 337 nm. The 20 most intense ions per MALDI spot, with signal/noise ratios >25, were selected for subsequent MS/MS analysis in the 1 kV mode using 100 consecutive laser shots. Air served as the collision gas. Data were subjected to analysis using the Mass Standard Kit for the 4700 Proteomics Analyzer (Applied Biosystems). The NCBInr database was searched for matches to the MS/MS spectra using ProteinPilot v. 3.01 (with MASCOT as the database search engine), with peptide and fragment ion mass tolerances of 200 ppm. One missed cleavage, carbamidomethylation of cysteines and oxidation of methionines were allowed during peptide searches.

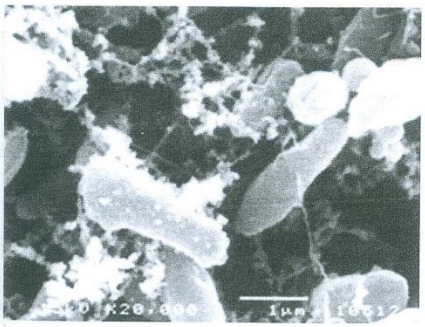


Figure 1. Scanning electron microscopy of a *P. xylaniclasticus* TW1 cell harvested at the late exponential growth phase. The scale bar in the micrograph is 1 μm.

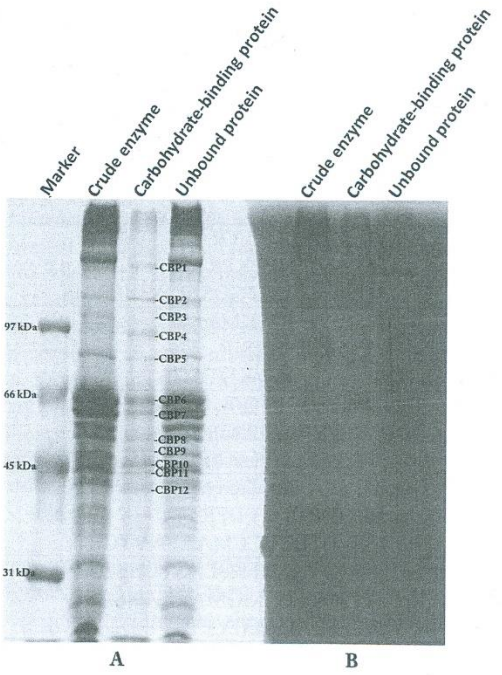


Figure 2. Patterns of proteins and enzyme activities by gel electrophoresis. (A) SDS-PAGE. (B) Zymogram for cellulase activity.

Table 1. Peptide sequences determined from carbohydrate-binding protein digest samples.

Protein	Peptide no. and sequence	BLAST identification
CBP1	#10: GKQLADDSSSAGR	Unknown protein
	#12: WLANIMLNGDPR	Unknown protein
	#18: EQRVLDSDGNPLPSEPR	Unknown protein
CBP2	#6: YDMYDKYFQK	Unknown protein
	#7: IKASVLGVSAAYVR	Unknown protein
	#13: IKQYPFQLSGGLCQR	Unknown protein
CBP3	#15: QISNLQQSISDAEQR	Unknown protein
	#16: NARTVFSELSQATSNK	Unknown protein
	#18: FVEFYGSGLSNISLADR	Unknown protein
CBP4	#7: GWPSTIGFLVDTSR	Unknown protein
	#9: SVPKLFSWIR	Unknown protein
	#13: IKQYPFQLSGGLCQR	Unknown protein
CBP5	#2: RKGEEWSVAFYRD	Glycoside hydrolase family 31 (*Paenibacillus curdlanolyticus* YK9)
	#8: RVPWAYDQEAVDVTRY	Glycoside hydrolase family 31 (*Paenibacillus curdlanolyticus* YK9)
	#18: KGYGVFVNQPELVSFEVGSEKV	Glycoside hydrolase family 31 (*Paenibacillus curdlanolyticus* YK9)
CBP6	#1: DDELLSAYR	Unknown protein
	#3: LVIENPEKPR	Unknown protein
	#14: IVGVNMGDSIPQVDIPRLIK	Unknown protein
CBP7	#1: FYLVGSKMR	Unknown protein
	#3: MMDTIESIR	Unknown protein
	#5: EQFGGVKDTAGR	Unknown protein
CBP8	#2: ASEQPTQIR	Unknown protein
	#3: QVVNPEDGSR	Unknown protein
	#6: EEFLGVIACIR	Unknown protein
CBP9	#3: EIAAAQDGIAR	Unknown protein
	#4: ENNIQFVDFR	Unknown protein
	#13: DTLCVCCMAASNVTGR	Unknown protein
CBP10	#10: QVASDFKSLSGVGSSQR	Unknown protein
	#11: NGFFAMGMAEKANSASR	Unknown protein
	#15: SLADAEGYETFVIPDDVGGR	Unknown protein
CBP11	#4: GPLTTPIGGGIR	Unknown protein
	#14: DLLNEYEFPGDDTPIVR	Unknown protein
	#16: YQLVQPQTIIVILEKIR	Unknown protein
CBP12	#15: KGTITVDGGTYDIYETTRV	Xylanase (*Paenibacillus curdlanolyticus* B6)

Results and discussion

Isolation of carbohydrate-binding proteins

The microbial deconstruction of the plant cell wall is a critical biological process, which also provides important substrates for environmentally sustainable industries. However, the physical complexity of plant cell walls limits the access of the hydrolytic enzymes to their target substrates. To overcome the "access problem" glycoside hydrolases, esterases, and lyases that degrade plant structural polysaccharides, in general, have a modular structure in which the catalytic module is appended to non-catalytic carbohydrate binding modules (CBMs), which are grouped into sequence-based families within the CAZy database (Cantarel et al., 2009). The general function of CBMs is to direct the cognate catalytic modules to their target substrate within the plant cell wall, thereby increasing the efficiency of catalysis. CBMs, in addition to their family assignment, have also been defined as type A, B, and C modules

reflecting their mode of binding (Boraston et al., 2004). Type A CBMs bind to crystalline surfaces such as cellulose and chitin, type B modules recognize the internal regions of single glycan chains, whereas type C proteins typically recognize no more than two sugars and often target the end of glycan chains (Boraston et al., 2004). The majority of CBMs display a β-sandwich fold with the ligand binding site located in either the concave surface presented by one of the β-sheets (a topography that facilitates the targeting of the internal regions of glycan chains) or in the loops that connect the two sheets. This latter binding site can either target the end or, less frequently, the internal regions of glycan chains (Jamal-Talabani et al., 2004). In some CBM families, typically those that recognize crystalline polysaccharides, ligand specificity is invariant, while other families contain proteins that bind to a range of different carbohydrates. Thus, CBMs are excellent model systems for studying the mechanism of protein-carbohydrate recognition. Furthermore, this diversity in ligand specificity underpins the exploitation of these protein modules in numerous biotechnological applications (Tomme et al., 1998).

The *P. xylaniclasticus* TW1 strain used in this study was identified by Tachaapaikoon et al. (2012b). This strain was chosen for this work because it is able to produce xylanolytic-cellulolytic multienzyme complex when grown on corn hull under aerobic conditions. General producers of multienzyme complex are frequently found in the anaerobic microorganism. Therefore, it was easier to handle aerobe. Furthermore, this strain is also able to convert lignocelluloses to fermentable sugars efficiently.

Due to that the ability to bind to insoluble polysaccharides has suggested a feature of multienzyme complexes. Therefore, carbohydrate-binding proteins (CBPs) which are composite in the xylanolytic-cellulolytic multienzyme complex were interesting. The CBPs were isolated from the crude enzyme preparation by affinity purification on ball-milled cellulose. The crude enzyme preparation was incubated with ball-milled cellulose for 30 min to allow the CBPs present in the crude enzyme preparation to bind the cellulose. The CBPs were separated from the unbound proteins of the crude enzyme preparation by centrifugation, eluted from the cellulose, and analyzed by SDS-PAGE. As shown in Figure 2, although the crude enzyme preparation consisted of many proteins, only 12 proteins were detected in the ball-milled cellulose-binding fraction, showing that most proteins could not bind to ball-milled cellulose. In addition, the carbohydrate-binding proteins were analyzed by zymogram using CMC as substrates to determine the enzyme activities of the catalytic components. The result found that all 12 proteins showed the activity on CMCase zymogram (Figure 2). Consequently, all 12 carbohydrate-binding proteins should belong to the family of glycoside hydrolases, especially cellulase family.

Identification of carbohydrate-binding proteins

All 12 bands of carbohydrate-binding proteins from the SDS-PAGE were digested with trypsin and the resulting peptide mixtures were analysed by MALDI-TOF/TOF MS spectrometry. The peptide mass fingerprints of the unknown components from SDS-PAGE were almost identical to unknown proteins (Table 1). This result suggested that these proteins possibly presented as the new glycoside hydrolase family. Only 2 CBPs could be identified in this study. As shown in Table 1, three of the peptides from the CBP5 protein (#2, 8, 18) were obtained showing similarity to published glycoside hydrolase family 31 sequences from *P. curdlanolyticus* YK9 (GenBank no. gi|304405232). However, the known activities of glycoside hydrolase family 31 containing α-glucosidase (EC 3.2.1.20), α-1,3-glucosidase (EC 3.2.1.84), sucrase-isomaltase (EC 3.2.1.48) (EC 3.2.1.10), α-xylosidase (EC 3.2.1.177), α-glucan lyase (EC 4.2.2.13), isomaltosyltransferase (EC 2.4.1.-) and α-mannosidase (EC 3.2.1.24) (http://www.cazy.org/GH31_characterized.html) which no cellulase activity was reported within this family. Additionally, no information on the CBM appearing in the glycoside hydrolase family 31 from *P. curdlanolyticus* YK9 was presented. For more clearly, PCR primer designs, sequencing, cloning and characterization of the recombinant CBP5 should be further investigated for the activity of this protein and the binding region of protein on ball-milled cellulose. CBP12 band from SDS-PAGE gel could be identified as xylanase. An NCBI BLAST search of the peptide sequence of CBP12 (KGTITVDGGTYDIYETTRV) showed similarity to *P. curdlanolyticus* B6 xylanase belonging to glycoside hydrolase family 11 with a CBM family 36. Intriguingly, CBP12 showed CMC hydrolysis activity on zymogram (Figure 2) but it could be identified as xylanase. Furthermore, no report on the

binding of CBM36 to cellulose was discovered (Jamal-Talabani et al., 2004). Hence, we might find the new bi-functional activity and the new binding properties of CBM36 from CBP12. However, CBP12 should be also further studied for seeking the characteristic of this protein.

Conclusion

SDS-PAGE of the crude enzyme preparation exhibited that only 12 proteins could bind to ball-milled cellulose. Using the in-gel tryptic digestion followed by MALDI-TOF/TOF analysis, it was found that only two of CBPs represented as the glycoside hydrolase family 31 (CBP5) and family 11 (CBP12), while other CBPs shown non-similarity to known function proteins. This result suggested that other proteins possibly presented as new glycoside hydrolase family. However, both identical proteins should be further investigated for elucidation of the different properties of these proteins as compared to the previous reports.

Acknowledgements

This study was supported through Project on Invitation to Japan for Environmental Research (PIER) by the 2005 World Exposition in Aichi, Japan and implemented by the International Center for Environmental Technology Transfer (ICETT). This study was also supported by the Thailand Research Fund (TRF Grant) and King Mongkut's Institute of Technology Ladkrabang.

References

Bayer, E.A., Lamed, R. (1986) Ultrastructure of the cell surface cellulosome of *Clostridium thermocellum* and its interaction with cellulose. Journal of Bacteriology 167: 828-836.

Berg, B., von Hofsten, B., Pettersson, G. (1972) Growth and cellulase formation by *Celluvibrio fulvus*. Journal of Applied Bacteriology 35: 201-214.

Boraston, A.B., Bolam, D.N., Gilbert, H.J., Davies, G.J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochemical Journal 382: 769-781.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V. Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Research 37: D233-238.

Jamal-Talabani, S., Boraston, A.B., Turkenburg, J.P., Tarbouriech, N., Ducros, V.M., Davies, G.J. (2004) Ab initio structure determination and functional characterization of CBM36: a new family of calcium dependent carbohydrate-binding modules. Structure 12: 1177-1187.

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

Pason, P., Kyu, K.L., Ratanakhanokchai, K. (2006) *Paenibacillus curdlanolyticus* strain B-6 xylanolytic-cellulolytic enzyme system that degrades insoluble polysaccharides. Applied and Environmental Microbiology 72: 2483-2490.

Ratanakhanokchai, K., Kyu, K.L., Tanticharoen, M. (1999) Purification and properties of a xylan-binding endoxylanase from alkaliphilic *Bacillus* sp. strain K-1. Applied and Environmental Microbiology 65: 694-697.

Schwarz, W.H. (2001) The cellulosome and cellulose degradation by anaerobic bacteria. Applied Microbiology and Biotechnology 56: 634-649.

Shin, G.W., Palaksha, K.J., Yang, H.H., Shin, Y.S., Kim, Y.R., Lee, E.Y. (2006) Partial two-dimensional gel electrophoresis (2-DE) maps of *Streptococcus iniae* ATCC29178 and *Lactococcus garvieae* KG9408. Diseases of Aquatic Organisms 70(1-2): 71-79.

Tachaapaikoon, C., Kyu, K.L., Pason, P., Ratanakhanokchai, K. (2012a) A novel multienzyme complex from a newly isolated facultative anaerobic bacterium, *Paenibacillus* sp. TW1. Acta Biologica Hungarica 63: 288-300.

Tachaapaikoon, C., Tanasupawat, S., Pason, P., Sornyotha, S., Waeonukul, R., Kyu, K.L., Ratanakhanokchai, K. (2012b) *Paenibacillus xylaniclasticus* sp. nov., a xylanolytic-cellulolytic bacterium isolated from sludge in an anaerobic digester. Journal of Microbiology 50: 394-400.

Tomme, P., Boraston, A., McLean, B., Kormos, J., Creagh, A.L., Sturch, K., Gilkes, N.R., Haynes, C.A., Warren R.A.J., Kilburn, D.G. (1998) Characterization and affinity applications of cellulose-binding domains. Journal of Chromatography B 715: 283-296.

Appendix B

Proceeding under title: A novel xylan-binding CBM family 36 from multienzyme complex-producing bacterium, *Paenibacillus xylaniclasticus* TW1: A high affinity for insoluble cellulose in addition to xylans and specificity for target carbohydrate in epidermis cell walls.

Sornyotha, S., Karita, S., Tsuchiya, T., Tachaapaikoon, C. and Ratanakhanokchai, K., 2014, “A novel xylan-binding CBM family 36 from multienzyme complex-producing bacterium, *Paenibacillus xylaniclasticus* TW1: A high affinity for insoluble cellulose in addition to xylans and specificity for target carbohydrate in epidermis cell walls”, MIE BIOFORUM 2014-Lignocellulose Degradation and Biorefinery, November 18-21, Shima City, Mie, Japan, pp. 111-118.

Lignocellulose Degradation and Biorefinery

EDITED BY
Kazuo Sakka
Tetsuya Kimura
Yutaka Tamaru
Shuichi Karita
Makiko Sakka
Sadanari Jindou

UNI PUBLISHERS CO., LTD

A novel xylan-binding CBM family 36 from multienzyme complex-producing bacterium, *Paenibacillus xylaniclasticus* TW1 : a high affinity for insoluble cellulose in addition to xylans and specificity for target carbohydrate in epidermis cell walls

Somphit Sornyotha[1], Shuichi Karita[2], Tohru Tsuchiya[3], Chakrit Tachaapaikoon[4], and Khanok Ratanakhanokchai[5]

[1]*Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand*

[2]*Graduate School of Regional Innovation Studies, Mie University, Tsu 514-8507, Japan*

[3]*Plant Functional Genomics Institute, Life Science Research Center, Mie University, Tsu 514-8507, Japan*

[4]*Pilot Plant Development and Training Institute, King Mongkut's University of Technology Thonburi, Bangkok 10150, Thailand*

[5]*Enzyme Technology Laboratory, School of Bioresources and Technology, King Mongkut's University of Technology Thonburi, Bangkok 10150, Thailand*

A xylan-binding carbohydrate-binding module family 36 (xylan-binding CBM36) of xylanase Xyn11A from *Paenibacillus xylaniclasticus* TW1 was cloned, expressed, purified and studied for binding characteristics. The results found that *P. xylaniclasticus* TW1 xylan-binding CBM36 displayed broad binding ability to carbohydrates which had a high affinity for xylans and insoluble cellulose. Interestingly, this is the first study that indicated that xylan-binding CBM36 had an affinity for insoluble cellulose. In addition, the binding characteristic of xylan-binding CBM36 on carbohydrates embedded within plant cell walls was also elucidated. It displayed the strong recognition for ligands located in the epidermal tissue of sweet potato roots. Based on these results, *P. xylaniclasticus* TW1 xylan-binding CBM36 is a new tool and could be applied for targeting enzymes on the surface of plants.

A facultative, mesophilic, anaerobic bacterium, *Paenibacillus xylaniclasticus* TW1, isolated from a waste treatment reactor was a true cellulolytic/xylanolytic organism, as it could grow on xylan, filter paper, or lignocellulosic substances as a sole source of carbon and produced the xylanolytic-cellulolytic multienzyme complex under aerobic conditions. This enzyme system could be degraded the lignocellulosic materials effectively and it was able to bind to substrate via CBM. (1)

Carbohydrate-binding modules (CBMs) are non-catalytic modules that bind to poly- or oligosaccharides. One or more CBMs are often attached to a catalytic module of plant cell wall polysaccharide-degrading enzymes (2). The general function of CBMs is to facilitate the association of the parent enzyme with its substrate, thereby increasing the local concentration of the enzyme and consequently enhancing its activity. (3, 4) There are currently 69 sequence-based families (http://www.cazy.org/Carbohydrate-Binding-Modules.html). In addition, CBMs have also been classified as types A, B and C modules reflecting their topography and binding properties. Type A CBMs have a flat hydrophobic surface that interacts with crystalline surfaces, type B CBMs have a cleft to accommodate single glycan chains, and type C CBMs typically bind no more than two sugars and often target the end of glycan chains (5).

The xylanase Xyn11A from *P. xylaniclasticus* TW1 consists of 385 amino acids and contains a catalytic glycoside hydrolase (GH) family 11 module, a CBM family 36 and the glycine-, asparagine- and threonine-rich sequence existed between the two modules (GenBank Accession# JX069975). Family 36 CBM of Xyn11A has high homology with CBM36 from *P. curdlanolyticus* B-6 (98% identity, ADB54799.1), *P. campinasensis* G1-1 (79% identity, AEI54132.1) and *Paenibacillus* sp. Aloe-11 (78% identity, ZP_09772596.1), respectively. However, only CBM36 was characterized, that from *P. polymyxa* ATCC 842 xylanase 43A (P45796.1), displayed a xylans-binding CBM (6). In addition, CBM36 from *P. xylaniclasticus* TW1 showed only 50% sequence identity with CBM36 of *P. polymyxa* ATCC 842. Therefore, xylan-binding CBM36 of Xyn11A from *P. xylaniclasticus* TW1 was still attractive. In this study, we sought to elucidate the binding characteristic of the *P. xylaniclasticus* TW1 xylan-binding CBM36 on isolated carbohydrates. Moreover, we also identified the specificity for target carbohydrates embedded within plant cell walls. The specificity of the xylan-binding CBM36 to bind to plant cell walls was assessed by using cyan fluorescent protein (CFP)-labeled recombinant CBM. Therefore, the *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP was constructed, produced and studied for binding characteristics.

MATERIALS AND METHODS

Bacterial strains and plasmids. *Escherichia coli* INVα (Invitrogen, Carlsbad, CA) and BL21(DE3) (Toyobo, Kyoto) were cultivated in Luria-Bertani (LB) medium supplemented with ampicillin (50 μg/ml) for use as plasmid construction and expression host for derivative constructs of plasmid pCR2.1 and pRSET/CFP (Invitrogen), respectively.

DNA amplification and cloning. DNA fragment encoding for the xylan-binding CBM36 obtained from *P. xylaniclasticus* TW1 was amplified by polymerase chain reaction (PCR). The primers used contained either an artificial *Bam*HI or *Eco*RI restriction endonuclease sites (underlined) in order to facilitate the cloning of the amplified DNA fragments into pRSET/CFP. The following oligonucleotide primers were used to amplify the coding regions by PCR: *P. xylaniclasticus* TW1 xylan-binding CBM36F-CFP: GGATCCGGTAATACAGGTGGAAC and *P. xylaniclasticus* TW1 xylan-binding CBM36R-CFP: GAATTCAGGAGTCGAATTGATTTCC. The amplified PCR products were then cloned into pCR2.1. The inserted amplified DNA fragments were sequenced to confirm the absence of mutations, digested with *Bam*HI and *Eco*RI, and finally ligated into pRSET/CFP that had been digested with the same restriction enzymes.

Protein expression and purification. *E. coli* BL21(DE3) cells harboring *P. xylaniclasticus* TW1 xylan-binding CBM36-pRSET/CFP was aerobically cultured at 37 °C in 1 liter of LB broth supplemented with amplicillin (50 μg/ml), the cells were allowed to grow to an optical density of about 0.4 at 600 nm. After the addition of isopropyl-1-thio-β-D-galactopyranoside (IPTG) to a final concentration of 1 mM, the cells were incubated for an additional 3 h. They were then harvested by centrifugation (8,000 x *g*) for 20 min, washed, and disrupted by sonication. The cell-free extract was subjected to His-tag purification on nickel-nitrilotriacetic acid (Ni-NTA) agarose (Qiagen) according to the manufacturer' s instructions. The purity of the fractions was analyzed by SDS-PAGE. The protein concentrations were determined spectrophotometrically at 280 nm. The molar absorption coefficients calculated from the tryptophan, tyrosine and cysteine contents was 47,960 $M^{-1}cm^{-1}$ (7).

Qualitative carbohydrate-binding assays. The affinities of *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP for soluble carbohydrates were evaluated by native affinity polyacrylamide gel electrophoresis (native affinity PAGE) as described by Montanier et al. (4). The separating gel contained 10% (w/v) polyacrylamide gels. The carbohydrate ligands

at a concentration of 0.1% (w/v) were incorporated into the gel prior to polymerization. The non-binding negative control reference protein was bovine serum albumin (BSA). Electrophoresis was carried out for 1 h at room temperature. Proteins were visualized by Coomassie blue staining.

Binding of *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP to insoluble carbohydrates was examined as follows: *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP (50 μg) was mixed with the insoluble polysaccharides (15 mg) on ice for 30 min in a 50 mM KPB (pH 7.0) at a final volume of 250 μl. The solution was then centrifuged (8,000 × *g*) for 5 min. The precipitate was washed 5 times with the same buffer, and the precipitate and the supernatant were analyzed by SDS-PAGE. The carbohydrates used were ball-milled cellulose and oat-spelt xylan.

Preparation of plant material and labeling for *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP and fluorescence microscopy. This study was carried out as described by McCartney et al. (8) and Araki et al. (9). Sweet potato (*Ipomoea batatas* cv. Beni Kokei) was grown from a seedling in water for 1 week. Root apices were fixed for 5 min in 50 mM sodium phosphate buffer (pH 7.2) containing 3% (w/v) paraformaldehyde and 50 mM sodium chloride with vacuum infiltration, and then for 4 h on ice. The samples were then dehydrated in an ethanol series consisting of 30% (v/v), 50% (v/v) and 70% (v/v) ethanol (20 min each at 4 °C); 80% (v/v), 90% (v/v), 95% (v/v), 99% (v/v) and 100% ethanol (20 min each at room temperature); and 30% (v/v), 60% (v/v) and 80% (v/v) *t*-buthanol in 100% ethanol (20 min each at room temperature). They were then transferred to 55 °C, allowed to warm, transferred to 100% *t*-buthanol for 20 min and then to 100% *t*-buthanol and liquid paraffin (1:1) and incubated for 4 h. They were transferred 3 times to low-melt paraffin for 1 h each at 65 °C and left overnight at 65 °C. They were then placed in sample moulds filled with melted paraffin and allowed to solidify in ice water. Six-μm-thick longitudinal sections were cut through the roots and collected on polylysine-coated microscope slides (Frontia Matsunami Glass, Tokyo), deparaffined in a series consisting of xylene (twice for 5 min each); 100%, 99% (v/v), 95% (v/v), 90% (v/v), 80% (v/v), 70% (v/v), 50% (v/v) and 30% (v/v) ethanol (1 min); 20 mM KPB at pH 7.0 (10 min) and 20 mM KPB with 0.25% (w/v) BSA (30 min). For CBM labeling, the root sections were incubated in 20 mM KPB containing 0.25% (w/v) BSA and 40-100 μg/ml of *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP for 15 min. They were then washed in 50 mM KPB at least 3 times. They were

observed on an Axioskop 2 Plus Zeiss microscope. The images were captured using an AxioCam HRC Plan-Neofluar camera.

RESULTS AND DISCUSSION

Carbohydrate specificity of *P. xylaniclasticus* TW1 xylan-binding CBM36. In this study, we used native affinity PAGE to investigate the binding specificities of *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP. As shown in Table 1, *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP could thus be shown to bind strongly to both of oat-spelt xylan and birchwood xylan, moderately to arabinan from sugar beet (α-linked arabinan) and weakly to xyloglucan from tamarind seed, methylcellulose (derivatized cellulose), lichenan (mixed-linkage β-1,3-1,4-glucan) and laminarin (β-1,3-glucan). However, *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP could not bind to derivatized cellulolose such as carboxymethylcellulose which substituted with carboxymethyl groups (-CH_2COOH) on some of the hydroxyl groups (-OH) and also not bind to the mixed-linkage β-1,3-1,4-glucan as β-glucan from barley. These results suggested that the frequency and composition of substitution might have effect on binding ability of CBM36. In addition, *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP did not bind to pectin and galactomannan (guar). It was apparent, therefore, that *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP showed relatively strong binding for xylans from different origins, which was similar to the ligand binding of CBM36 of xylanase 43 from *P. polymyxa* ATCC 842 (6). However, *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP was also capable of binding to other carbohydrates in addition to xylans. Hence, CBM36 from the strain TW1 had a broad binding ability, compared to CBM36 from those of the strain ATCC 842 (6).

The ability of *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP to bind insoluble carbohydrates was also qualitatively evaluated by incubating the proteins with the carbohydrates and comparing the protein concentrations in the supernatant fraction (unbound protein) and in the precipitate fraction (bound protein) by SDS-PAGE. As shown in Fig. 1, *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP showed high affinity for insoluble fraction of oat-spelt xylan. Moreover, it also has a strong affinity for the insoluble cellulose as ball-milled cellulose, suggesting that this xylan-binding CBM36 prefers insoluble cellulose in addition to xylans. Interestingly, this is the first study that indicated that xylan-binding CBM36 had an affinity for insoluble cellulose.

Table 1 Ligand specificity of *P. xylaniclasticus* TW1 xylan-binding CBM36 determined by native affinity polyacrylamide gel electrophoresis.

Ligand[a]	Binding[b]
Oat-spelt xylan	+++
Birchwood xylan	+++
Methylcellulose	+
Carboxymethylcellulose	-
Arabinan (Sugar beet)	++
Xyloglucan (Tamarind seed)	+
Lichenan (Icelandic moss)	+
Laminarin (*Laminaria digitata*)	+
β-D-glucan (Barley)	-
Galactomannan (Gal:Man, 38:40) (Guar)	-
Pectin (Citrus)	-

[a] Ligands were screened at a concentration of 1 mg/ml.

[b] No detectable binding, −; weak binding, +; moderate binding, ++; strong binding, +++.

Fig. 1. Adsorption of *P. xylaniclasticus* TW1 xylan-binding CBM36 to insoluble carbohydrates.

P. xylaniclasticas TW1 xylan-binding CBM36 was incubated with insoluble carbohydrates including ball-milled cellulose (1) and oat-spelt xylan (2). After centrifugation, proteins in the precipitate (C) and in the supernatant (B) were analyzed by SDS-PAGE. Lane A contains the purified *P. xylaniclasticas* TW1 xylan-binding CBM36 as control.

Binding of *P. xylaniclasticus* TW1 xylan-binding CBM36 to cell walls of sweet potato root. *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP was allowed to bind to the sections of roots from sweet potato. This study found that the *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP bound extensively to epidermal tissue of root (Fig. 2). It showed effectively and specifically to bind to epidermal tissue of sweet potato roots this likely reflects the variable structure of xylan within plant tissue. Due to xylans are known to be a taxonomically variable set of polymers, and biochemical analysis have indicated that a β-1,4-linked xylose polymer, which could be decorated with glucuronosyl or 4-O-methyl-glucuronosyl residues at O2 and α-arabinofuranosyl or acetyl groups at O2 or O3, with the nature and extent of substitution varying between species, cell type and plants age (10, 11). However, ligand specificity studies indicated that *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP had high affinity for both xylan and cellulose (Fig. 1). Hence, not only xylan but also cellulose within cell walls might be recognized by *P. xylaniclasticus* TW1 xylan-binding CBM36-CFP, since these carbohydrates are the major components of plant cell walls. Furthermore, the recognition profile of *P. xylaniclasticus* TW1 xylan-binding CBM36 in plant cell walls is likely to reflect the region of the accessibility of target substrates for their cognate enzyme. The specificity of *P. xylaniclasticus* TW1 xylan-binding CBM36 to bind to epidermal tissue of sweet potato roots mirrors a very specific region of substrate available for the cognate enzyme. Finally, as wall degradation progresses, the *P. xylaniclasticus* TW1 xylan-binding CBM36 will become effective with a wider range of substrates as indicated by its recognition of isolated carbohydrates. Thus, this CBM also represents valuable tools with which to probe the intricate architecture of plant cell walls, and possible as a new tool for targeting enzymes to surface of plant tissue.

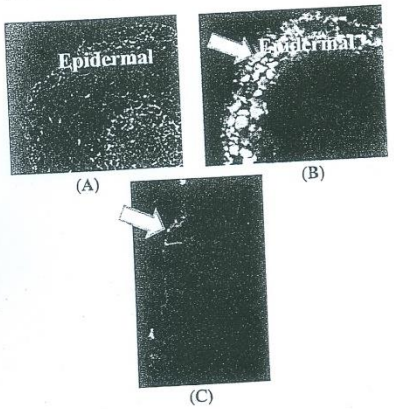


Fig. 2. Binding of *P. xylaniclasticus* TW1 xylan-binding CBM36 to cell walls of 1-week-olds sweet potato roots.

The binding of *P. xylaniclasticas* TW1 xylan-binding CBM36 to the section was detected by fluorescence microscopy. In (A) the plant cell walls were non-labeling, whereas in (B&C) the cell walls were labeling with *P. xylaniclasticas* TW1 xylan-binding CBM36-CFP. Arrows indicate recognition of epidermal tissues. In (C) the root were cut and used as non-fixation samples.

ACKNOWLEDGMENTS

This study was supported through Project on Invitation to Japan for Environmental Research (PIER) by the 2005 World Exposition in Aichi, Japan and implemented by the International Center for Environmental Technology Transfer (ICETT). This study was also supported by the Thailand Research Fund (TRF Grant) and King Mongkut's Institute of Technology Ladkrabang.

REFERENCES

1) Tachaapaikoon, C., et al., Acta Biol. Hung., 63, 288-300 (2012).
2) Tsukimoto, K., et al., FEBS Lett., 584, 1205-1211 (2010).
3) Bolam, D.N., et al., Biochem. J., 331, 775-781 (1998)
4) Montanier, C., et al., J. Biol. Chem., 285, 31742-31754 (2010).
5) Boraston, A.B., et al., Biochem. J., 382, 769-781 (2004).
6) Jamal-Talabani, S., et al., Structure, 12, 1177-1187 (2004).
7) Mach, H., et al., Anal. Biochem., 200, 74-80 (1992).
8) McCartney, L., et al., Anal. Biochem., 326, 49-54 (2004).
9) Araki, Y., et al., Biosci. Biotechnol. Biochem., 74, 802-805 (2010).
10) Ebringerova´, A., Heinze, T., Macromol. Rapid Commun., 21, 542-556 (2000).
11) O'Neill, M.A., York, W.S., in Rose J.K.C. (ed.) The plant cell wall, the composition and structure of plant primary cell walls, p. 1-54. CRC Press, Boca Raton, FL.

Appendix C

Manuscript under title: Atypical characteristics of family 36 carbohydrate-binding module of *Paenibacillus xylaniclasticus* TW1 xylanase and their recognition in epidermal tissue of sweet potato roots.

Running title: CBM36 from *Paenibacillus xylaniclasticus* TW1

Somphit Sornyotha,[1] Shuichi Karita,[2] Tohru Tsuchiya,[3] Chakrit Tachaapaikoon[4] and Khanok Ratanakhanokchai[5]

[1]Faculty of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok 10520, Thailand

[2]Graduate School of Regional Innovation Studies, Mie University, Tsu 514-8507, Japan

[3]Plant Functional Genomics Institute, Life Science Research Center, Mie University, Tsu 514-8507, Japan

[4]Pilot Plant Development and Training Institute, King Mongkut's University of Technology Thonburi, Bangkok 10150, Thailand

[5]Enzyme Technology Laboratory, School of Bioresources and Technology, King Mongkut's University of Technology Thonburi, Bangkok 10150, Thailand

Correspondence.

Khanok Ratanakhanokchai

E-mail: khanok.rat@kmutt.ac.th (K. Ratanakhanokchai)

Tel: +66-2470-7755. Fax: +66-2470-7760

Shuichi Karita

E-mail: karita@bio.mie-u.ac.jp (S. Karita)

Tel: +81-59-231-9619. Fax: +81-59-231-9684

Abstract

Cellulose-bound protein 12 (CBP12) produced from *Paenibacillus xylaniclasticus* TW1 as xylanase family 11 with a carbohydrate-binding module family 36 (CBM36) was identified by using affinity purification and matrix-assisted laser desorption/ionization time-of-flight/time-of-flight mass spectrometry (MALDI-TOF/TOF MS). The CBM36 of the strain TW1 (*Px*TW1CBM36) showed only 50% homology with the characterized CBM36 from *P. polymyxa* xylanase 43A. The recombinant *Px*TW1CBM36 fused with cyan fluorescent protein (r*Px*TW1CBM36-CFP) displayed broad binding ability to carbohydrates which had a high affinity for xylan and insoluble cellulose. Although the essentially metal binding sites were conserved, the binding capacity of r*Px*TW1CBM36-CFP was not perturbed by the addition of ethylenediaminetetraacetic acid (EDTA), suggesting that it might be a metal-independent CBM. In addition, the binding characteristic of r*Px*TW1CBM36-CFP was specific for carbohydrate ligands only in the young epidermal tissue of sweet potato roots. This study may provide a new tool for targeting enzymes on the surface of plant cell walls.

Keywords: carbohydrate-binding module family 36, xylanase family 11, metal-independent binding, *Paenibacillus xylaniclasticus*, sweet potato roots

INTRODUCTION

Plant cell walls are the most abundant and renewable source of valuable products such as fermentable sugars, chemicals and liquid fuel (Hasunuma et al., 2013). The cell wall of plants is comprised of cellulose, hemicelluloses (mainly xylan), lignin, and other components. These components are tightly connected and intertwined via covalent bonds, hydrogen bonds and van der Waals to form a highly complex structure (Pérez et al., 2002) The complete and rapid hydrolysis of polysaccharides in plant cell walls requires cellulolytic (Hasunuma et al., 2013) and xylanolytic enzymes (Gírio et al., 2010).

The previous reports indicated that a few microorganisms could produce multienzyme complexes under aerobic conditions (Jiang et al., 2005; Pason et al., 2006; van Dyk et al., 2009; Tachaapaikoon et al., 2012a). Among them, *P. xylaniclasticus* TW1 was a true cellulolytic/xylanolytic organism as it

could grow on xylan, filter paper, or agricultural wastes as a sole source of carbon and produced the xylanolytic-cellulolytic multienzyme complex under aerobic conditions (Tachaapaikoon et al., 2012a, b). The multienzyme complexes have been reported to dedicate to hydrolyze plant cell wall substances because of their ability to bind to insoluble cellulose and/or xylan via CBMs such as cellulose-binding and xylan-binding modules, respectively (Bayer et al., 1994). The general function of CBMs is to direct the cognate catalytic modules to their target substrates within the plant cell wall, thereby increasing the efficiency of catalysis (Shoseyov et al., 2006). CBMs are divided into families based on amino acid sequence similarity. Currently, there are 69 defined CBM families (see URL http://www.cazy.org/Carbohydrate-Binding-Modules.html). In addition, CBMs have also been classified as types A, B and C modules reflecting their topography and binding properties. Type A CBMs bind to crystalline surfaces and type B CBMs bind to the internal regions of single glycan chains, while type C CBMs typically bind no more than two sugars and often target the end of glycan chains (Boraston et al., 2004). Thus, they are considered to be an important factor in the efficient degradation of plant cell wall materials (Ohmiya et al., 1997). Intriguingly, no information on the CBMs appearing in the xylanolytic-cellulolytic multienzyme complex from *P. xylaniclasticus* TW1 is reported. Therefore, in this study, we sought to identify the secreted enzymes that appear to contain CBMs by using affinity purification and applying MALDI-TOF/TOF MS. Based on our strategies, we have found CBM36 with the xylanase family 11 produced from the strain TW1. To date, only CBM36 from *P. polymyxa* ATCC 842 xylanase 43A was characterized, while CBM36 from the strain TW1 showed low amino acid sequence similarities (50% identities) to CBM36 from the strain ATCC 842. Thus, a CBM36 from the strain TW1 was still attractive.

In this report, the r*Px*TW1CBM36-CFP was constructed. The binding characteristics and specificities of r*Px*TW1CBM36-CFP against isolated carbohydrates and the recognition of carbohydrates located within plant cell walls as the substrate available for cognate enzymes were investigated. Moreover, the target region of CBM36 in plants was currently unknown and the fluorescent tag was used for monitoring CBM36 binding in plants.

METHODS

Sources of carbohydrates. Birchwood and oat-spelt xylans were from Fluka, Buchs, Switzerland. Carboxymethylcellulose, β-D-glucan from barley, guar gum (galactomannan), laminarin from Laminariadigitata and lichenan from Cetrariaislandica were from Sigma-Aldrich, St. Louis, USA. Arabinan from sugar beet (Ara:Gal:Rha:GalUA = 88:3:2:7), and xyloglucan from tamarind seed were from Megazyme International Ireland Ltd. Pectin from citrus and methylcellulose were from Nacalai Tesque, Kyoto, Japan. Soluble starch was from Kanto Chemical, Tokyo, Japan. Corn hull was prepared as described by Tachaapaikoon et al. (2006).

Enzyme preparation. *P. xylaniclasticus* TW1, the multienzyme producing strain, (Tachaapaikoon et al., 2012b) was grown in Berg's mineral salt medium (pH 7.0) (Berg et al., 1972) containing 0.2% (w/v) $NaNO_3$ (Wako Pure Chemical Industries Ltd., Osaka, Japan), 0.05% (w/v) K_2HPO_4 (Wako), 0.02% (w/v) $MgSO_4 \cdot 7H_2O$ (Wako), 0.002% (w/v) $MnSO_4 \cdot H_2O$ (Wako), 0.002% (w/v) $FeSO_4 \cdot 7H_2O$ (Wako), 0.002% (w/v) $CaCl_2 \cdot 2H_2O$ (Wako), and 0.5% (w/v) corn hull as a sole carbon source. The culture was incubated in a rotary shaker at 200 rpm and 37 °C for 3 days. After centrifugation at 8,000 g for 20 min at 4 °C, the culture supernatant was used as crude enzyme.

Isolation of cellulose-bound proteins. The isolation of cellulose-bound protein was conducted by adding 250 µl protein solution (25 mg ml^{-1}) to 15 mg of ball-milled cellulose, KC flock (Nippon Paper Chemicals, Tokyo, Japan) in 10 mM potassium phosphate buffer (KPB, pH 7.0) in 1.5-ml Eppendorf tubes. The sample was incubated at intervals at 4 °C for 30 min before centrifugation. The precipitate was washed 5 times with the same buffer. After that the precipitate (cellulose-bound proteins) and the supernatant (unbound proteins) were analyzed by SDS-PAGE.

Gel electrophoresis. SDS-PAGE was performed on an 8% (w/v) polyacrylamide gel by the method of Laemmli (1970). After electrophoresis, the proteins were stained with Coomassie brilliant blue R-250 (Nacalai Tesque, Kyoto, Japan). The molecular mass standards used were from a molecular-mass calibration kit (GE Healthcare, Tokyo, Japan).

Protein identification using MALDI-TOF/TOF MS. Selected protein bands (12 bands) were excised from SDS-PAGE gel and subjected to in-gel digestion with trypsin as described previously (Hanna

et al., 2000). MS/MS analyses to identify proteins were performed on a 4800 Plus MALDI-TOF/TOF Analyzer (Applied Biosystems, Tokyo, Japan) using a 200 Hz ND: YAG laser operating at 337 nm. The 20 most intense ions per MALDI spot, with signal/noise ratios >25, were selected for subsequent MS/MS analysis in the 1 kV mode using 100 consecutive laser shots. Air was served as the collision gas. Data were subjected to analysis using the Mass Standard Kit for the 4700 Proteomics Analyzer (Applied Biosystems). The NCBInr database was searched for matches to the MS/MS spectra using Protein Pilot v. 3.01 (with MASCOT as the database search engine), with peptide and fragment ion mass tolerances of 200 ppm. One missed cleavage, carbamidomethylation of cysteines and oxidation of methionines were allowed during peptide searches.

Polymerase chain reaction (PCR) amplification and sequencing for CBP12. Based on the MALDI TOF/TOF MS data, PCR primers were designed to amplify one gene with reference to have high homology regions in the xylanase gene (xyn11A) of *P. curdlanolyticus* B-6 (GenBank no. gi|283970946) and identified using the Blast search tool (http://www.ncbi.nlm.nih.gov/BLAST/). The two pairs of synthetic oligonucleotide primers for this gene were *Pb*xyn11F: 5´-ACCTTGTTGTTAGCTGGGACAAACGT-3´ and *Pb*xyn11R: 5´-TTGTTCGATGTGAAGCCGGGCA-3´.

PCR reactions were performed in a Takara PCR Thermal Cycler PERSONAL. For each reaction, a PCR mixture was prepared as follows: 5.0 µl of 10× PCR buffer (Takara, Shiga, Japan), 5.0 µl of dNTP mixture (Takara; 51.25 nM), 1.0 µl of each primer (50 µM each), 0.5 µl of r-Taq DNA polymerase (Takara; 5 U µl-1), 2.0 µl of DNA template, and 35.5 µl of MilliQ water, to attain 50 µl. PCR amplification conditions were 1 cycle at 95 °C for 4 min, followed by 30 cycles of denaturation at 95 °C for 20 s, annealing at 47 °C for 40 s, extension at 72 °C for 2 min, and a final elongation step at 72 °C for 3 min. Amplification products were run in an 0.8% (w/v) agarose gel, immersed in TAE buffer (40 mM Tris-acetate with 1 mM EDTA), stained with 0.5 µg ethidium bromide ml-1, and visualized using a UV transilluminator Bioinstrument (ATTO). The amplified band was excised from the gels and purified prior to cloning using a TA Cloning Kit (Invitrogen, CA, USA). An inserted gene was purified and sequenced on an ABI 3130X DNA Analyzer (Applied Biosystems). The encoding amino acid sequence of cbp12 gene was predicted the molecular mass and pI by using Compute pI/Mr tool on the ExPASy Server (http://web.expasy.org/compute_pi/).

Cloning, expression and purification of r*Px*TW1CBM36-CFP. To construct the plasmid of CFP fusion r*Px*TW1CBM36, the coding region was first amplified from *P. xylaniclasticus* TW1 genomic DNA using PCR. The following oligonucleotide primers were designed for amplifying *Px*TW1*cbm36* gene based on the sequence deposited in GenBank with accession number JX069975, while *Bam*HI and *Eco*RI restriction enzymatic cutting sites were introduced into the termination of the primers. The forward primer was *Px*TW1*cbm36*F: 5´-GGATCCGGTAATACAGGTGGAAC-3´, and the reverse primer was *Px*TW1*cbm36*R: 5´-GAATTCAGGAGTCGAATTGATTTCC-3´. The underlined letters refer to the restriction enzymatic cutting sites of *Bam*HI and *Eco*RI. The PCR conditions were as follows: 95 °C for 4 min for 1 cycle, 95 °C for 20 s, 48 °C for 40 s, and 72 °C for 50 s for 30 cycles. The amplified PCR products were then cloned into pCR2.1 (Invitrogen). The inserted amplified DNA fragments were sequenced to confirm the absence of mutations, digested with *Bam*HI and *Eco*RI, and finally ligated into the expression vector pRSET/CFP (Invitrogen) that had been digested with the same restriction enzymes, yielding plasmid pRSET/CFP-*Px*TW1CBM36, which produces a recombinant protein with a 6 × His-tag at the N terminus.

Recombinant *Escherichia coli* BL21(DE3) (Toyobo, Kyoto, Japan) harboring the pRSET/CFP-*Px*TW1CBM36 was aerobically cultured at 37 °C in 1 liter of Luria-Bertani (LB) broth supplemented with ampicillin (50 μg ml^{-1}) and the cells were allowed to grow to an optical density of about 0.4 at 600 nm. After the addition of isopropyl-1-thio-β-D-galactopyranoside (IPTG) to a final concentration of 1 mM, the cells were incubated for an additional 3 h. They were then harvested by centrifugation (8,000 g) for 20 min, washed, and disrupted by sonication. The cell-free extract was subjected to His-tag purification on nickel-nitrilotriacetic acid (Ni-NTA) agarose (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The purity of the fractions was analyzed by SDS-PAGE. The concentration of purified r*Px*TW1CBM36-CFP was determined by UV absorbance (280 nm) using a calculated molar extinction coefficient of 47,960 $M^{-1}cm^{-1}$ (Mach et al., 1992).

Nucleotide sequence accession number. The nucleotide sequence for the *P. xylaniclasticus* TW1 xylanase gene was deposited in the GenBank database under accession number JX069975.

Ligand-binding assays with affinity gel electrophoresis. The capacity of r*Px*TW1CBM36-CFP to bind to a variety of soluble carbohydrates was evaluated by affinity gel electrophoresis. Continuous native polyacrylamide gels (the absence of β-mercaptoethanol, SDS and heating) were prepared consisting of 10% (w/v) acrylamide. The soluble carbohydrates were incorporated into the gel at a concentration of 0.1% (w/v) prior to polymerization. Approximately 5 μg of target proteins or bovine serum albumin (BSA) from Sigma-Aldrich, as a non-interacting negative control, was loaded onto the gels and subjected to electrophoresis at room temperature. Proteins were visualized by staining with Coomassie brilliant blue R-250.

Effect of EDTA on the binding of r*Px*TW1CBM36-CFP to insoluble carbohydrates. r*Px*TW1CBM36-CFP (40 μg) was allowed to bind to the insoluble carbohydrates (15 mg) on ice for 30 min in a 50 mM KPB (pH 7.0) with/without EDTA in a final volume of 250 μl. The solution was then centrifuged (8,000 g) for 5 min. The precipitate was washed 5 times with the same buffer with/without EDTA and the precipitate and the supernatant were analyzed by SDS-PAGE. The insoluble carbohydrates used were ball-milled cellulose and insoluble fractions of oat-spelt xylan which were prepared as described previously (Zhang et al., 2012).

Preparation of plant material and CBM-labeling. This study was carried out as described by McCartney et al. (2004) and Araki et al. (2010). Sweet potato (Ipomoea batatas cv. Beni Kokei) was grown from a seedling in water for the appropriate times. Roots were fixed for 5 min in 50 mM sodium phosphate buffer (pH 7.2) containing 3% (w/v) paraformaldehyde and 50 mM sodium chloride with vacuum infiltration and then for 4 h on ice. The samples were then dehydrated in an ethanol series consisting of 30% (v/v), 50% (v/v) and 70% (v/v) ethanol (20 min each at 4 °C), 80% (v/v), 90% (v/v), 95% (v/v), 99% (v/v) and 100% ethanol (20 min each at room temperature), and 30% (v/v), 60% (v/v) and 80% (v/v) t-buthanol in 100% ethanol (20 min each at room temperature). They were then transferred to 55 °C, allowed to be warm, transferred to 100% t-buthanol for 20 min and then to 100% t-buthanol and liquid paraffin (1:1) and incubated for 4 h. They were transferred 3 times to low-melt paraffin for 1 h each at 65 °C and left overnight at 65 °C. They were then placed in sample moulds filled with melted paraffin and allowed to solidify in ice water. Six-μm-thick transverse sections were cut through the roots and collected on polylysine-coated microscope slides (Frontia Matsunami Glass, Tokyo, Japan), deparaffined in a series consisting of xylene (twice for 5 min each), 100%,

99% (v/v), 95% (v/v), 90% (v/v), 80% (v/v), 70% (v/v), 50% (v/v) and 30% (v/v) ethanol (1 min), 20 mM KPB at pH 7.0 (10 min) and 20 mM KPB with 0.25% (w/v) BSA (30 min). After that, the root sections were incubated with 50 µg ml-1 of r*Px*TW1CBM36-CFP in 20 mM KPB for 30 min. They were then washed in 50 mM KPB at least 3 times and observed on an Axioskop 2 Plus Zeiss microscope. The images were captured using an AxioCam HRC Plan-Neofluar camera.

RESULTS AND DISCUSSION

Isolation of cellulose-bound proteins

In order to elucidate features of the CBMs, we carried out isolation using supernatant obtained from corn hull-aerobic grown cultures (stationary phase) of *P. xylaniclasticus* TW1. Based on the properties of the cellulose-binding ability, the cellulose-bound proteins were isolated from the crude enzyme by affinity purification on ball-milled cellulose. The crude enzyme was incubated with ball-milled cellulose at 4 °C to allow the cellulose-bound proteins present in the enzyme preparation to bind to the cellulose. The cellulose-bound proteins were separated from the unbound proteins of the enzyme preparation by centrifugation, washed and then eluted from the cellulose by adding the sample application buffer, heated in a boiling-water bath for 5 min, and analyzed by SDS-PAGE. As shown in Fig. 1, the crude enzyme consisted of at least twelve proteins that were detected in the cellulose-bound proteins (namely CBP1 to CBP12). The results indicated that these proteins probably contained the cellulose-binding domain. Hence, all of proteins were subjected to MALDI-TOF/TOF MS spectrophotometry for identification of these proteins.

Identification of cellulose-bound proteins

All twelve bands of cellulose-bound proteins from the SDS-PAGE were digested with trypsin and the resulting peptide mixtures were analyzed by MALDI-TOF/TOF MS spectrometry. The peptide mass fingerprints of the unknown components from SDS-PAGE were almost identical to the unknown function proteins (data not shown). This result suggested that these proteins are possibly present as new cellulose-bound protein families. It is expected that the use of the apparent results in this study, together with the further works, will contribute to an understanding of these proteins. However, only

the band of CBP12 from SDS-PAGE gel could be identified as xylanase. An NCBI BLAST search of the amino acid sequence of CBP12 (KGTITVDGGTYDIYETTRV) shows similarity to *P. curdlanolyticus* B-6 xylanase belonging to glycoside hydrolase family 11 (GenBank no. gi|283970946; http://www.ncbi.nlm.nih.gov/BLAST/). Therefore, the gene encoding protein CBP12 of *P. xylaniclasticus* TW1 was further identified by cloning and sequencing.

By the BLAST database search, the amino acid sequence (KGTITVDGGTYDIYETTRV) was found to have high homology with xylanase11A (Xyn11A) of *P. curdlanolyticus* B-6. Then, to clone the gene of the CBP12 from *P. xylaniclasticus* TW1, the forward and reverse primers were designed based on both the amino acid sequence and nucleotide sequence of the strain B-6 xylanase11A. The amplified fragment was cloned into the TA vector and sequenced as described in Methods. The *cbp12* gene consisted of 1155 bp and encoded 385 amino acid residues (Fig. 2a) with a predicted molecular mass of 42 kDa and pI 7.7. The alignment of the predicted amino acid sequence revealed that mature CBP12 was composed of two major functional domains, a catalytic domain belonging to family11 glycoside hydrolases and a carbohydrate-binding module classified as family 36. The glycine-, asparagine- and threonine-rich sequence existed between the two domains (Fig. 2a and 2b).

Family 36 CBM of CBP12 has high homology with CBM36 from *P. curdlanolyticus* B-6 (98% identity, GenBankID: ADB54799.1), *P. campinasensis* G1-1 (79% identity, GenBank ID: AEI54132.1) and *Paenibacillus* sp. Aloe-11 (76% identity, GenBank ID: ZP_09772596.1), respectively. However, no characteristics of CBM36 from these bacteria were reported. Until the present, only CBM36 from *P. polymyxa* ATCC 842 xylanase 43A (GenBank ID: P45796.1) was characterized, showing the calcium-dependent binding of xylans and xylooligosaccharides (Jamal-Talabani et al., 2004). In addition, CBM36 from *P. xylaniclasticus* TW1 showed only 50% amino acid sequence identity with CBM36 of the strain ATCC 842. Consequently, CBM36 of CBP12 from the strain TW1 was still attractive and, thus, this CBM was further characterized.

Ligand specificity of r*Px*TW1CBM36-CFP

To study the binding characteristics of CBM36 from *P. xylaniclasticus* TW1, r*Px*TW1CBM36 was ligated in pRSET/CFP for construction of CFP fusion r*Px*TW1CBM36 (r*Px*TW1CBM36-CFP), expressed under the induction of IPTG and purified for subsequent analysis. For the purification

strategy, the recombinant protein was purified based on its N-terminal 6 × His-tag by affinity purification using a Ni-NTA agarose column. The purified r*Px*TW1CBM36-CFP was analyzed by SDS-PAGE and dialyzed against 10 mM KPB (pH 7.0) for further study.

In vitro binding studies using semi-quantitative affinity gel electrophoresis showed that migration of purified r*Px*TW1CBM36-CFP through native polyacrylamide gels loaded with soluble birchwood and oat-spelt xylans were slower than that of BSA (control) (Fig. 3). Moreover, r*Px*TW1CBM36-CFP could bind tightly to ball-milled cellulose and both soluble birchwood and oat-spelt xylans, moderately to sugar beet arabinan (α-linked arabinan) and weakly to laminarin (β-1,3-glucan), tamarind seed xyloglucan, methylcellulose and lichenan. However, r*Px*TW1CBM36-CFP could not bind to carboxymethylcellulose and barley β-glucan. Although, barley β-glucan is a mixed-linkage β-1,3-1,4-glucan with a structure similar to that of lichenan, the structure of barley β-glucan was generally categorized as containing more β-1,3-linked cellotetraosyl units than lichenan (Wood et al., 1994). These results indicated that the frequency and composition of substituted groups might have an effect on the binding ability of CBM36. In addition, r*Px*TW1CBM36-CFP did not bind to galactomannan (guar gum), pectin or starch (Table 1). The result revealed that r*Px*TW1CBM36-CFP showed relatively tight binding for xylans from different origins, which was similar to the ligand binding of CBM36 of xylanase 43 from *P. polymyxa* ATCC 842 (Jamal-Talabani et al., 2004). However, r*Px*TW1CBM36-CFP was also capable of binding to cellulose and other carbohydrates in addition to xylans. Hence, CBM36 from the strain TW1 had a broad binding ability, compared to CBM36 from those of the strain ATCC 842 (Jamal-Talabani et al., 2004).

Binding of r*Px*TW1CBM36-CFP to insoluble carbohydrates in the presence of EDTA

Jamal-Talabani and coworkers (2004) reported that the CBM36 of *P. polymyxa* ATCC 842 was a calcium-dependent type. Because of the low amino acid sequence similarity (50%), binding property of CBM36 from *P. xylaniclasticus* TW1 probably differs from the strain ATCC 842. To investigate ligand recognition, the ability of r*Px*TW1CBM36-CFP to bind to insoluble carbohydrates was determined by incubating the CBM, insoluble carbohydrates, and EDTA at a final concentration of 10 mM. The supernatant fraction (unbound protein) and the precipitate fraction (bound protein) obtained by centrifugation were subjected to SDS-PAGE. The binding ability was determined by

comparing the protein concentration in the supernatant fraction with that in the precipitate fraction. Interestingly, as shown in Fig. 4a, the strong binding to ball-milled cellulose and insoluble oat-spelt xylan were observed under the presence of 10 mM EDTA. To confirm the influence of EDTA on the binding ability of r*Px*TW1CBM36-CFP, 50 mM EDTA was used for this investigation (Fig. 4b). It was apparent that, mainly r*Px*TW1CBM36-CFP was still bound to both insoluble carbohydrates. These results indicated that the chelating agent had a few effects on the binding of r*Px*TW1CBM36-CFP to both insoluble substances. Carbohydrate binding was a strong capability although a high concentration of EDTA was added to the protein. These results revealed that *Px*TW1CBM36 possibly displayed a calcium-independent carbohydrate-binding module. Furthermore, these observations also found that r*Px*TW1CBM36-CFP presented a highly bound to both insoluble xylan and cellulose. Generally, the ligand specificity of CBMs reflects the target substrate hydrolyzed by their cognate enzyme (Carvalho et al., 2004). It is likely that xylan is a biologically significant ligand recognized by CBM36 because they are located predominantly in xylanases (see URL http://www.cazy.org/CBM36_all.html). Thus, it is especially unusual for a *Px*TW1CBM36 to bind to xylan and cellulose. Present studies, however, showed that a *Px*TW1CBM36 displayed broad ligand binding than the previous report (Jamal-Talabani et al., 2004).

Sequence-based relatives of *Px*TW1CBM36 using BLAST were analyzed. The revealed 17 proteins that contained a module with significant similarity to *Px*TW1CBM36 were obtained (Fig. 5). All the proteins were from bacteria and almost proteins contained xylanase catalytic modules. This data indicated that the primary target for CBM family 36 was xylan. Proteins from microorganisms e.g. *Clostridium phytofermentas*, *Ruminococcus albus* and *Streptomyces thermoviolaceus* contained a carbohydrate esterase family 4 (CE4) which family is also known to reveal xylan-specific esterases (see URL http://www.cazy.org/CE4_characterized.html; Montanier et al., 2010). In addition, the observation of essentially conserved metal binding sites, derived from the alignment of CBM36 sequences, initially led us to know the residues that coordinate Ca^{2+} in CBM36 (Tyr29, Asp105, Trp109 and Asp110) (Jamal-Talabani et al., 2004; Montanier et al., 2010) were invariant in all members. Furthermore, Tyr15 which dominated the second sugar binding site (Montanier et al., 2010) was also comprised and was thus functionally conserved within the family.

It has been reported that the role of calcium in the function of CBMs was restricted to CBMs that display a β-sandwich fold, where the ligand binding site comprised the loops that link the two β-sheets (Abou-Hachem et al., 2002). The metal ion, currently, only contributes to carbohydrate recognition in those CBMs that interact with one or, at most, two sugars, but not in those modules that bind to more extensive ligands. This suggests that calcium could mediate particularly tight binding to sugars. This feature of metal ion requirement was well illustrated in CBMs that bind to the internal regions of carbohydrates. CBMs that bind to four or more sugars in these polymers did not utilize calcium in ligand recognition (Montanier et al., 2010). As seem in the module, CBM36 from *P. polymyxa* ATCC 842 xylanase 43A (Abou-Hachem et al., 2002) and CBM60 from *Cellvibrio japonicas* (Montanier et al., 2010), which interact with two or three internal sugars and the metal ion, play a dominant role in carbohydrate binding, consistent with the view that charged dipole-dipole interactions were stronger than uncharged ones (Fersht et al., 1985). In sharp contrast, however, r*Px*TW1CBM36-CFP was still binding to insoluble carbohydrates in the presence of EDTA at 50 mM. Thus, it is possible that the ligand recognition site between CBM36 from *P. xylaniclasticus* TW1 and the strain ATCC 842 might be different. Hence, we propose that the ligand recognition of *Px*TW1CBM36 might be presented as metal-independent. The metal-independent CBM is beneficial to biotechnological process since additional of required metal ion is negligible, and metal ions contaminated in process have impact to smaller extent.

Binding of r*Px*TW1CBM36-CFP to the cell wall of sweet potato roots

The capacity of the CBM36 to bind to cell walls in sections of roots from sweet potatoes was assessed by using the CFP-labeled recombinant carbohydrate-binding module (r*Px*TW1CBM36-CFP). This analysis showed that the r*Px*TW1CBM36-CFP displayed a highly distinctive pattern of cell recognition with strong binding to the epidermal tissue of sweet potato roots (Fig. 6). In vitro studies indicated that r*Px*TW1CBM36-CFP had high affinity for both xylan and cellulose (Fig. 4). Hence, not only xylan but also cellulose within cell walls might be recognized by *Px*TW1CBM36, since these carbohydrates are the major components of plant cell walls. In addition, r*Px*TW1CBM36-CFP exhibited to bind specifically to young epidermal tissue (Fig. 6a-2); this likely reflects the variability of cell wall components and the structure of carbohydrates within plant tissue (O'Neill & York, 2003) It has been reported that CBMs in different families display significant variation in specificity for target ligands in

cell walls, providing a biological rationale for this structural diversity. Different CBMs therefore have the capacity to target appended catalytic modules to specific cell walls (McCartney et al., 2006). Furthermore, in view of the variation in specificity displayed by the different CBMs, it is interesting that very few xylanases contain multiple xylan-binding CBMs from distinct families (McCartney et al., 2006). For example, xylanases/arabinofuranosidases from *Caldicelluorupter* sp. Rt69B.1 contain a CBM6 and a CBM22 (Morris et al., 1999), or *P. polymyxa* ATCC 842 xylanase contains a CBM6 and a CBM36 (Jamal-Talabani et al., 2004). In contrast, numerous xylanases contain distinct xylan- and cellulose-binding CBMs (Waeonukul et al., 2009). Thus, in nature, individual xylanases may target specific cell walls. Even more intriguing was the observation that although r*Px*TW1CBM36-CFP exhibits higher recognition binding to epidermal tissue than other tissues, its capacity to recognize the polysaccharide within plant cell walls was limited by the older roots (Fig. 6b-2). The differences in the binding capacity profile of r*Px*TW1CBM36-CFP with a 1-week-old root and 1-month-old root was likely to reflect a combination of the accessibility of the target ligands within the cell wall composites and fine details of the carbohydrate structure. In the cell wall, the older plant, cellulose and xylan chains were highly intimately associated with other cell wall components, especially lignin to form the hard and rigid structure, (O'Neill & York, 2003) which were therefore limited to make contacts with r*Px*TW1CBM36-CFP. The weak binding to the epidermis of the older roots suggested that probable extensive bonds, e.g. hydrogen bonds, covalent bonds, as well as van der Waals, between each component to form a complex structure within the plant cell wall (Pérez et al., 2002) might not accessible to CBM. From these reasons, an explanation was provided for why the CBM displayed the weak binding to target ligands when it is embedded in the older plant cell wall.

The recognition profile of r*Px*TW1CBM36-CFP in plant cell walls is likely to reflect the region of the accessibility of target substrates for their cognate enzyme. The specificity of r*Px*TW1CBM36-CFP to bind to epidermal tissue of sweet potato roots mirrors a very specific region of substrate available for the cognate enzyme. However, the binding specificity of r*Px*TW1CBM36-CFP to plant cell walls could also potentiate for enzymatic hydrolysis by localizing the cognate enzyme onto the surface of the target substrate. In addition, some CBMs which bind to cellulose have been reported to disrupt the hydrogen bonding network between cellulose chains of the crystalline regions, thereby increasing the substrate access (Din et al., 1994). Thus, the hypothesis of the attractive function of

*Px*TW1CBM36 not only locks the cognate xylanase onto the intimate plant cell walls through their xylan and cellulose binding abilities but may also increase the substrate access through their cellulose binding and disrupting connection networks of cellulose fibers as described by Din et al. (1994) and Zverlov et al. (2002). From our observation, therefore, CBM36 of *P. xylaniclasticus* TW1 is one of the key technologies for the biomass degradative process.

In conclusion, this work describes a novel metal-independent family 36 of CBM by targeting; the primary with xylans likely directs the cognate enzyme to highly decorated regions of the carbohydrate. It was significantly displayed in specificity for target ligands only on the epidermal tissue of young sweet potato roots, providing the capacity to target the appended catalytic module to specific cell walls. Thus, the portfolio of specificity presented by this CBM can represent a valuable tool for probing the intricate architecture of plant cell walls and the possible use as a new tool for targeting enzymes on the surface of plant tissue.

ACKNOWLEDGEMENTS

This report is an outcome of the program supported by the 2005 World Exposition in Aichi, Japan and implemented by the International Center for Environmental Technology Transfer (ICETT). The program was hosted by the Advanced Fusion Technology Unit, Graduate School of Regional Innovation Studies, Mie University. The program titled as “Project on Invitation to Japan for Environmental Research (PIER)” invites research fellows from developing countries to extend their studies more in cooperation with advice of Japanese Institutes in the area of waste management and recycling issues. Acknowledgements also go to the King Mongkut’s University of Technology Thonburi, Thailand under the Higher Education Research Promotion and National Research University Project of Thailand, Office of Higher Education Commission, the Thailand Research Fund (TRF Grant) and King Mongkut’s Institute of Technology Ladkrabang.

The authors wish to express their gratitude to Ms. Julie Koldewyn for carefully proofreading the manuscript of this work.

REFERENCES

Abou-Hachem, M., Karlsson, E. N., Simpson, P. J., Linse, S., Sellers, P., Williamson, M. P., Jamieson, S. J., Gilbert, H. J., Bolam, D. N. & Holst, O. (2002). Calcium binding and thermostability of carbohydrate binding module CBM4-2 of Xyn10A from *Rhodothermus marinus*. Biochemistry 41, 5720-5729.

Araki, Y., Karita, S., Tsuchiya, T., Kondo, M. & Goto, M. (2010). Family 17 and 28 carbohydrate-binding modules discriminated different cell-wall sites in sweet potato roots. Biosci Biotechnol Biochem 74, 802-805.

Bayer, E. A., Morag, E. & Lamed, R. (1994). The cellulosome: a treasure trove for biotechnology. Trends Biotechnol 12, 379-386.

Berg, B., Hofsten, B. V. & Pettersson, G. (1972). Growth and cellulase formation by *Cellvibrio fulvus*. J Appl Bacteriol 35, 201-214.

Boraston, A. B., Bolam, D. N., Gilbert, H. J. & Davies, G. J. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem J 382, 769-781.

Carvalho, A. L., Goyal, A., Prates, J. A. M., Bolam, D. N., Gilbert, H. J., Pires, V. M. R., Ferreira, L. M. A., Planas, A., Romão, M. J. & Fontes, C. M. G. A. (2004). The family 11 carbohydrate-binding module of *Clostridium thermocellum* Lic26A-Cel5E accommodates β-1,4- and β-1,3-1,4-mixed linked glucans at a single binding site. J Biol Chem 279, 34785-34793.

Din, N., Damude, H. G., Gilkes, N. R., Miller, Jr. R. C., Warren, R. A. J. & Kilburn, D. G. (1994). C1-Cx revisited: intramolecular synergism in a cellulase. Proc Natl Acad Sci U S A 91, 11383-11387.

Fersht, A. R., Shi, J. -P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y. & Winter, G. (1985). Hydrogen bonding and biological specificity analysed by protein engineering. Nature 314, 235-238.

Gírio, F. M., Fonseca, C., Carvalheiro, F., Duarte, L. C., Marques, S. & Bogel-Łukasik, R. (2010). Hemicelluloses for fuel ethanol: a review. Bioresour Technol 101, 4775-4800.

Hanna, S. L., Sherman, N. E., Kinter, M. T. & Goldberg, J. B. (2000). Comparison of proteins expressed by *Pseudomonas aeruginosa* strains representing initial and chronic isolates from a cystic fibrosis patient: an analysis by 2-D gel electrophoresis and capillary column liquid chromatography-tandem mass spectrometry. Microbiology 146, 2495-2508.

Hasunuma, T., Okazaki, F., Okai, N., Hara, K. Y., Ishii, J. & Kondo, A. (2013). A review of enzymes and microbes for lignocellulosic biorefinery and the possibility of their application to consolidated bioprocessing technology. Bioresour Technol 135, 513-522.

Jamal-Talabani, S., Boraston, A. B., Turkenburg, J. P., Tarbouriech, N., Ducros, V. M. -A. & Davies, G. J. (2004). Ab initio structure determination and functional characterization of CBM36: a new family of calcium-dependent carbohydrate binding modules. Structure 12, 1177-1187.

Jiang, Z. Q., Deng, W., Li, X. T., Ai, Z. L., Li, L. T. & Kusakabe, I. (2005). Characterization of a novel, ultra-large xylanolytic complex (xylanosome) from *Streptomyces olivaceoviridis* E-86. Enzyme Microb Technol 36, 923-929.

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

Mach, H., Middaugh, C. R. & Lewis, R. V. (1992). Statistical determination of the average values of the extinction coefficients of tryptophan and tyrosine in native proteins. Anal Biochem 200, 74-80.

McCartney, L., Gilbert, H. J., Bolam, D. N., Boraston, A. B. & Knox, J. P. (2004). Glycoside hydrolase carbohydrate-binding modules as molecular probes for the analysis of plant cell wall polymers. Anal Biochem 326, 49-54.

McCartney, L., Blake, A. W., Flint, J., Bolam, D. N., Boraston, A. B., Gilbert, H. J. & Knox, J. P. (2006). Differential recognition of plant cell walls by microbial xylan-specific carbohydrate-binding modules. Proc Natl Acad Sci U S A 103, 4765-4770.

Montanier, C., Flint, J. E., Bolam, D. N., Xie, H., Liu, Z., Rogowski, A., Weiner, D. P., Ratnaparkhe, S., Nurizzo, D. & other authors (2010). Circular permutation provides an evolutionary link between two families of calcium-dependent carbohydrate binding modules. J Biol Chem 285, 31742-31754.

Morris, D. D., Gibbs, M. D., Ford, M., Thomas, J. & Bergquist, P. L. (1999). Family 10 and 11 xylanase genes from *Caldicellulosiruptor* sp. strain Rt69B.1. Extremophiles 3, 103-111.

Ohmiya, K., Sakka, K., Karita, S. & Kimura, T. (1997). Structure of cellulases and their applications. Biotechnol Genet Eng 14, 365-414.

O'Neill, M. A. & York, W. S. (2003). The composition and structure of plant primary cell walls. In The Plant Cell Wall, pp. 1-54. Edited by J. K. C. Rose. Ithaca, New York, Blackwell Publishing/CRC Press.

Pason, P., Kyu, K. L. & Ratanakhanokchai, K. (2006). *Paenibacillus curdlanolyticus* strain B-6 xylanolytic-cellulolytic enzyme system that degrades insoluble polysaccharides. Appl Environ Microbiol 72, 2483-2490.

Pérez, J., Muñoz-Dorado, J., De-la-Rubia, T. & Martínez, J. (2002). Biodegradation and biological treatments of cellulose, hemicellulose and lignin: an overview. Int Microbiol 5, 53-63.

Shoseyov, O., Shani, Z. & Levy, I. (2006). Carbohydrate binding modules: biochemical properties and novel applications. Microbiol Mol Biol Rev 70, 283-295.

Tachaapaikoon, C., Kyu, K. L. & Ratanakhanokchai, K. (2006). Purification of xylanase from alkaliphilic *Bacillus* sp. K-8 by using corn husk column. Process Biochem 41, 2441-2445.

Tachaapaikoon, C., Kyu, K. L., Pason, P. & Ratanakhanokchai, K. (2012a). A novel multienzyme complex from a newly isolated facultative anaerobic bacterium, *Paenibacillus* sp. TW1. Acta Biol Hung 63, 288-300.

Tachaapaikoon, C., Tanasupawat, S., Pason, P., Sornyotha, S., Waeonukul, R., Kyu, K. L. & Ratanakhanokchai, K. (2012b). *Paenibacillus xylaniclasticus* sp. nov., a xylanolytic-cellulolytic bacterium isolated from sludge in an anaerobic digester. J Microbiol 50, 394-400.

van Dyk, J. S., Sakka, M., Sakka, K. & Pletschke, B. I. (2009). The cellulolytic and hemi-cellulolytic system of *Bacillus licheniformis* SVD1 and the evidence for production of a large multi-enzyme complex. Enzyme Microb Technol 45, 372-378.

Waeonukul, R., Pason, P., Kyu, K. L., Sakka, K., Kosugi, A., Mori, Y. & Ratanakhanokchai, K. (2009). Cloning, sequencing, and expression of the gene encoding a multidomain endo-β-1,4-xylanase from *Paenibacillus curdlanolyticus* B-6, and characterization of the recombinant enzyme. J Microbiol Biotechnol 19, 277-285.

Wood, P. J., Weisz, J. & Blackwell, B. A. (1994). Structural studies of (1⟶3),(1⟶4)-β-D-glucans by 13C-nuclear magnetic resonance spectroscopy and by rapid analysis of cellulose-like regions using high-performance anion-exchange chromatography of oligosaccharides released by lichenase. Cereal Chem 71, 301-307.

Zhang, J., Tang, M. & Viikari, L. (2012). Xylans inhibit enzymatic hydrolysis of lignocellulosic materials by cellulases. Bioresour Technol 121, 8-12.

Zverlov, V. V., Velikodvorskaya, G. A. & Wolfgang H. Schwarz, W. H. (2002). A newly described cellulosomal cellobiohydrolase, CelO, from *Clostridium thermocellum*: investigation of the exo-mode of hydrolysis, and binding capacity to crystalline cellulose. Microbiology 148, 247-255.

FIGURE LEGENDS

Fig. 1. SDS-PAGE of cellulose-bound proteins isolated from the crude enzyme of *P. xylaniclasticus* TW1. M, molecular weight marker (kDa); lane 1, crude enzyme; lane 2, cellulose-bound proteins; lane 3, unbound proteins. CBP1-CBP12 in the gel indicated cellulose-bound proteins.

Fig. 2. Nucleotide sequence of the *cbp12* gene and deduced amino acid sequence of the gene product (a) and schematic of CBP12 (b). In b, domain boundaries are given by amino acid sequence number above the module schematic. GH11, family-11 catalytic module of the glycoside hydrolases; CBM36, family-36 CBM.

Fig. 3. Affinity gel electrophoresis of r*Px*TW1CBM36-CFP against soluble xylans. r*Px*TW1CBM36-CFP in lane 2 was electrophoresed on non-denaturating polyacrylamide gels containing no ligand (a), birchwood xylan (b) or oat-spelt xylan (c). BSA was used as a non-carbohydrate binding control (lane 1).

Fig. 4. The influence of EDTA on the binding ability of r*Px*TW1CBM36-CFP to insoluble carbohydrates. In this experiment, the r*Px*TW1CBM36-CFP was incubated with ball-milled cellulose (i) or insoluble oat-spelt xylan (ii), containing 10 mM EDTA (a) or 50 mM EDTA (b). After centrifugation, the protein in supernatant (lane 2) and precipitate (lane 3) were analyzed by SDS-PAGE. r*Px*TW1CBM36-CFP was used as a positive control (lane 1).

Fig. 5. Alignment of CBM36 sequences. The sequence alignment was derived from a search of the NCBI database using *Px*TW1CBM36 as the query sequence and the BLASTp search engine. Residues that were invariant within the family were indicated by an asterisk. Amino acids that were the ligand-binding calcium were highlighted in grey. The aromatic residue that dominated the second sugar binding subsite was showed in the open box. The sequences are as follows: *Px*TW1CBM36 from *P. xylaniclasticus* TW1 xylanase, *Pcu*CBM36 from *P. curdlanolyticus* B-6 xylanase 11A (GenBank ID: ADB54799.1), *Pca*CBM36 from *P. campinasensis* G1-1 xylanase (GenBank ID: AEI54132.1), *Cc*CBM36 from *Clostridium clariflavum* DSM19732 GH11 (GenBank ID: YP_005047554.1), *Pt*CBM36 from *P. terrae* HPL-003 xylanase/chitin deacetylase (GenBank ID: YP_005077359.1), *B*YA-335CBM36 from *Bacillus* sp. YA-335 endo-1,4-β-xylanase (GenBank ID: CAA41784.1), *B*41m-1CBM36 from *Bacillus* sp. 41m-1 XynJ (GenBank ID:2DCJ¬_A), *Psx*CBM36 from *Pseudobutyrivibrio xylanivorans* xylanase/deacetylase (GenBank ID: P83513.2), *Bup*CBM36 from *Butyrivibrio proteoclasticus* B316 endo-1,4-β-xylanase (GenBank ID: YP_003830062.1), *Cak*CBM36 from *Caldicellulosiruptor kronotskyensis* 2002 endo-1,4-β-xylanase (GenBank ID: YP_004022816.1), *Dtu*CBM36 from *Dictyoglomus turgidum* DSM 6724 endo-1,4-β-xylanase (GenBank ID: YP_002352183.1), *Pm*CBM36 from *P. mucilaginosus* KNP414 XynD (GenBank ID: YP_004642759.1), *Dth*CBM36 from *D. thermophilum* H-6-12 endo-1,4-β-xylanaseA (GenBank ID: YP_002251776.1), *Pp*CBM36 from *P. polymyxa* ATCC 842 Xyn43A (GenBank ID: P45796.1), *Ca*RT69B.1CBM36 from *Caldicellulosiruptor* sp. Rt69B.1 XynD (GenBank ID: AAB95327.1), *Cao*CBM36 from *Ca. owensensis* OL endo-1,4-β-xylanase (GenBank ID: YP_004001463.1), *Cab*CBM36 from *Ca. bescii* DSM 6725 endo-1,4-β-xylanase (GenBank ID: YP_002572022.1) and *Cp*CBM36 from *C. phytofermentans* ISDg polysaccharide deacetylase (GenBank ID: YP_001558798.1).

Fig. 6. Binding of r*Px*TW1CBM36-CFP to cell walls of 1-week-old (a) and 1-month-old (b) sweet potato roots. (1) The plant cell walls were non-labeled, whereas (2) the cell walls were labeled with r*Px*TW1CBM36-CFP. Arrows indicate recognition of epidermal tissues. The regions corresponding to epidermis (e), cortex (c) and vascular cylinder (vc) were indicated.

Table 1 Semi-quantitative assessment of ligand specificity of r*Px*TW-1CBM36-CFP determined by affinity gel electrophoresis.

Soluble carbohydrate	Binding ability*
Birchwood xylan	+++
Oat-spelt xylan	+++
Sugar beet arabinan	++
Laminarin	+
Tamarind seed xyloglucan	+
Methylcellulose	+
Carboxymethylcellulose	-
Lichenan	+
Barley β-glucan	-
Guar gum	-
Pectin	-
Starch	-

*Semi-quantitative binding is indicated as tight (+++), moderate (++), weak (+), and no significant interaction (-) depending on how far the proteins migrated on the gel.

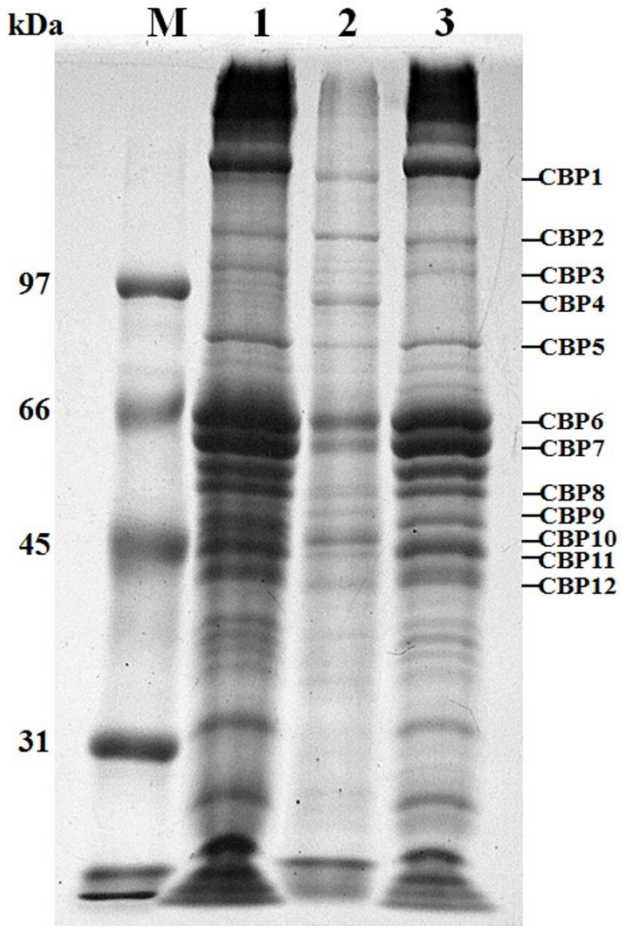


Fig. 1. Sornyotha *et al*.

(a)

```
   1  atgacaacgaaggaggaagatgtgatgaaaccaagaaaaatgaaatggatgctggccatgctgctcagcttcaca
   1  M  T  T  K  E  E  D  V  M  K  P  R  K  M  K  W  M  L  A  M  L  L  S  F  T

  76  ctcgcattgccagcaggcgtcgcgaacgccgtaacgattacgaataatgagtccggcacccatgacggctactat
  26  L  A  L  P  A  G  V  A  N  A  V  T  I  T  N  N  E  S  G  T  H  D  G  Y  Y

 151  tatgagctatggaaggattccgggaacacaagcatgacgcttaacagcggcggcaccttcagtgcgacgtggagc
  51  Y  E  L  W  K  D  S  G  N  T  S  M  T  L  N  S  G  G  T  F  S  A  T  W  S

 226  aacattaacaatgcgttattccgtaaaggcaaaaaattcaatgaaacacagacccaccagcaaatcggcaacatc
  76  N  I  N  N  A  L  F  R  K  G  K  K  F  N  E  T  Q  T  H  Q  Q  I  G  N  I

 301  tcggtcaattacagcgctaactattctccgaacggcaactcttatctctgcatctatggatggtcggttgacccg
 101  S  V  N  Y  S  A  N  Y  S  P  N  G  N  S  Y  L  C  I  Y  G  W  S  V  D  P

 376  ctcgtcgagtattatatcgtcgacagctggggcagctggcgtcctcctggcgcctcgtccaaaggtacgatcacc
 126  L  V  E  Y  Y  I  V  D  S  W  G  S  W  R  P  P  G  A  S  S  K  G  T  I  T

 451  gttgacggcggcacctacgacatttatgagacgactcgtgtgaaccagccttccattaaaggcacggcaacgttc
 151  V  D  G  G  T  Y  D  I  Y  E  T  T  R  V  N  Q  P  S  I  K  G  T  A  T  F

 526  cagcagtattggagcgtccggacgtcgaagcgtacgagcggcacaatctcggtaaccgatcacttcaatgcttgg
 176  Q  Q  Y  W  S  V  R  T  S  K  R  T  S  G  T  I  S  V  T  D  H  F  N  A  W

 601  gcaagcagaggcatgcagctgggcaaaatgtacgaagtatcgctcacggtagaaggctaccaaagcagcggaagc
 201  A  S  R  G  M  Q  L  G  K  M  Y  E  V  S  L  T  V  E  G  Y  Q  S  S  G  S

 676  gcgacaattacaagcaatacgattacaatcggcggcaacggcggcggcggtaatacgggcggtggcaataacggc
 226  A  T  I  T  S  N  T  I  T  I  G  G  N  G  G  G  G  N  T  G  G  G  N  N  G

 751  ggcggcaacaatggaggaggcaacaatggcggcggtaatacaggtggaacgacgagaattgaagccgagagcatg
 251  G  G  N  N  G  G  G  N  N  G  G  G  N  T  G  G  T  T  R  I  E  A  E  S  M

 826  acgaagagcggccaatacactggcaatatctactctccgttcaatggtgtagcactatacgctaacaacgactcg
 276  T  K  S  G  Q  Y  T  G  N  I  Y  S  P  F  N  G  V  A  L  Y  A  N  N  D  S

 901  gtgaaattcacgcataactttacgaaaagcacgaacaacttctcgctccgcggcgcttcgaacaactcccaaatg
 301  V  K  F  T  H  N  F  T  K  S  T  N  N  F  S  L  R  G  A  S  N  N  S  Q  M

 976  gctagagtcgatctgaaggtcggcggcgtgacgaagggcaccttctattacggcggcagctatcctgcagtctat
 326  A  R  V  D  L  K  V  G  G  V  T  K  G  T  F  Y  Y  G  G  S  Y  P  A  V  Y

1051  acgataaacaacgtcaatacgggaaccggaaatcaacaggtcgagctcattcttacagccgataacgggcaatgg
 351  T  I  N  N  V  N  T  G  T  G  N  Q  Q  V  E  L  I  L  T  A  D  N  G  Q  W

1126  gatgcgttcctcgattatttggaaatcaattag
 376  D  A  F  L  D  Y  L  E  I  N  *
```

(b)

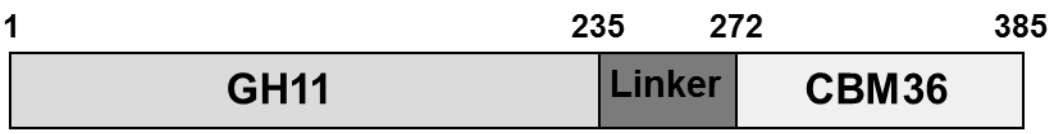


Fig. 2. Sornyotha *et al*.

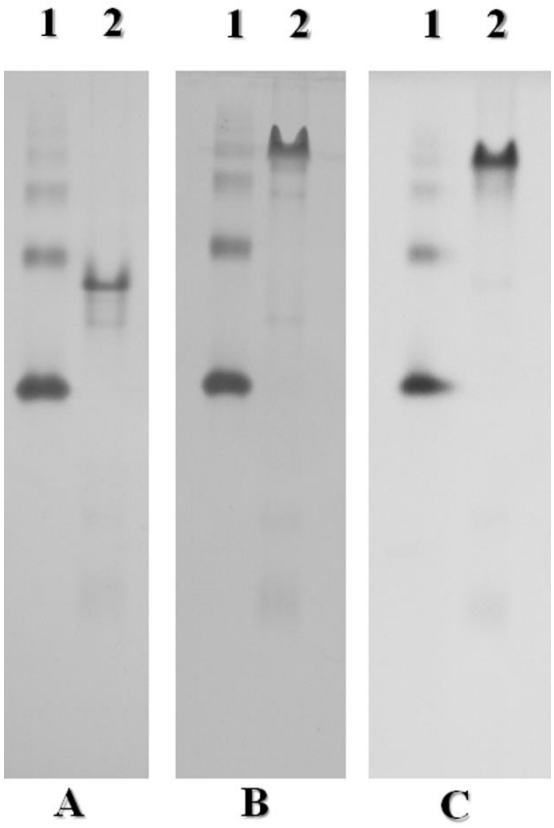


Fig. 3. Sornyotha *et al*.

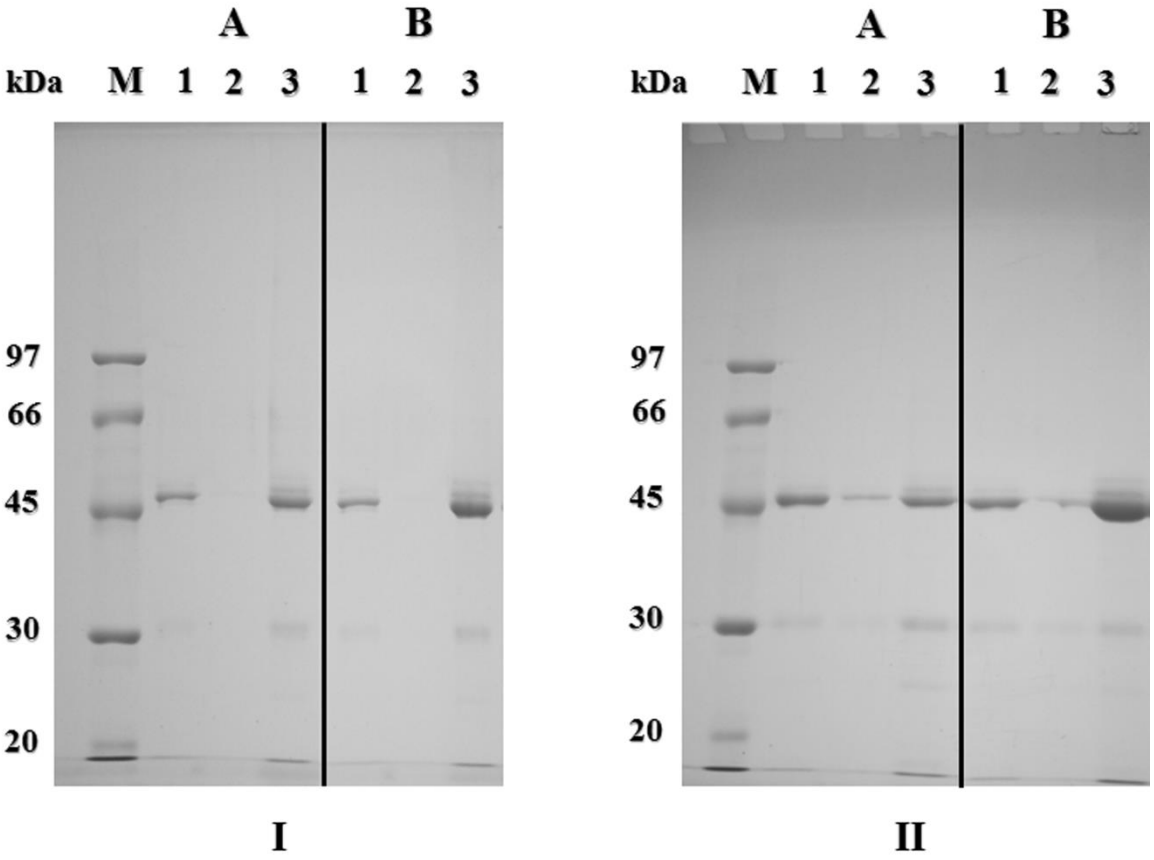


Fig. 4. Sornyotha *et al*.

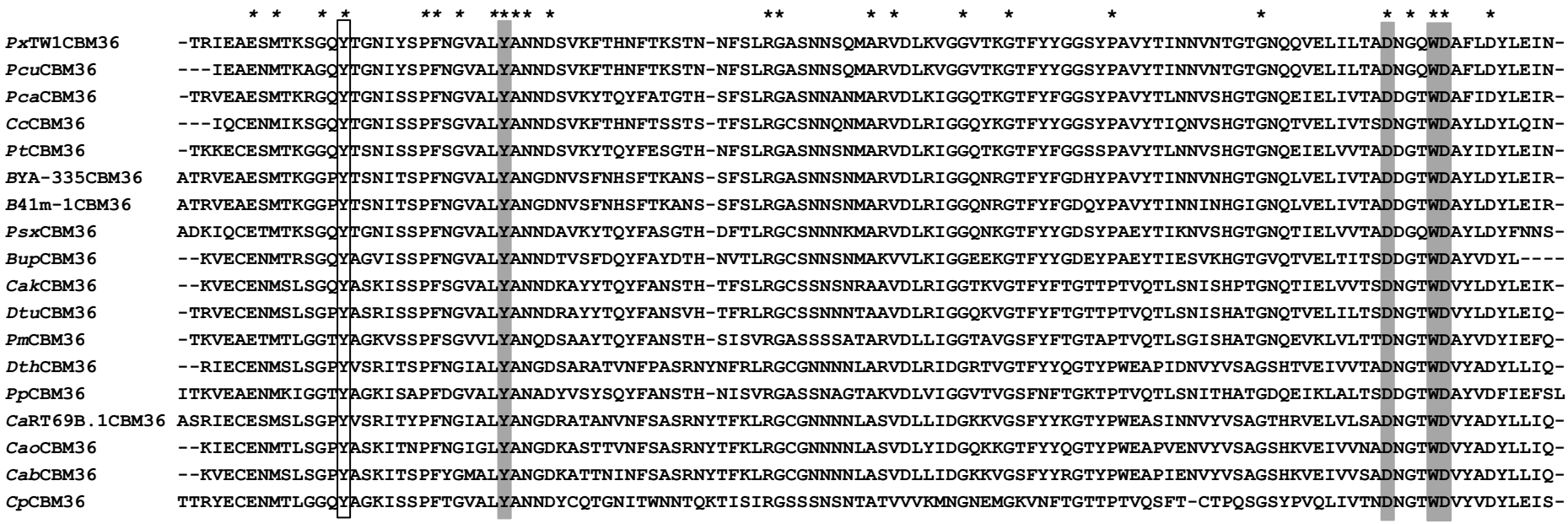


Fig. 5. Sornyotha *et al*.

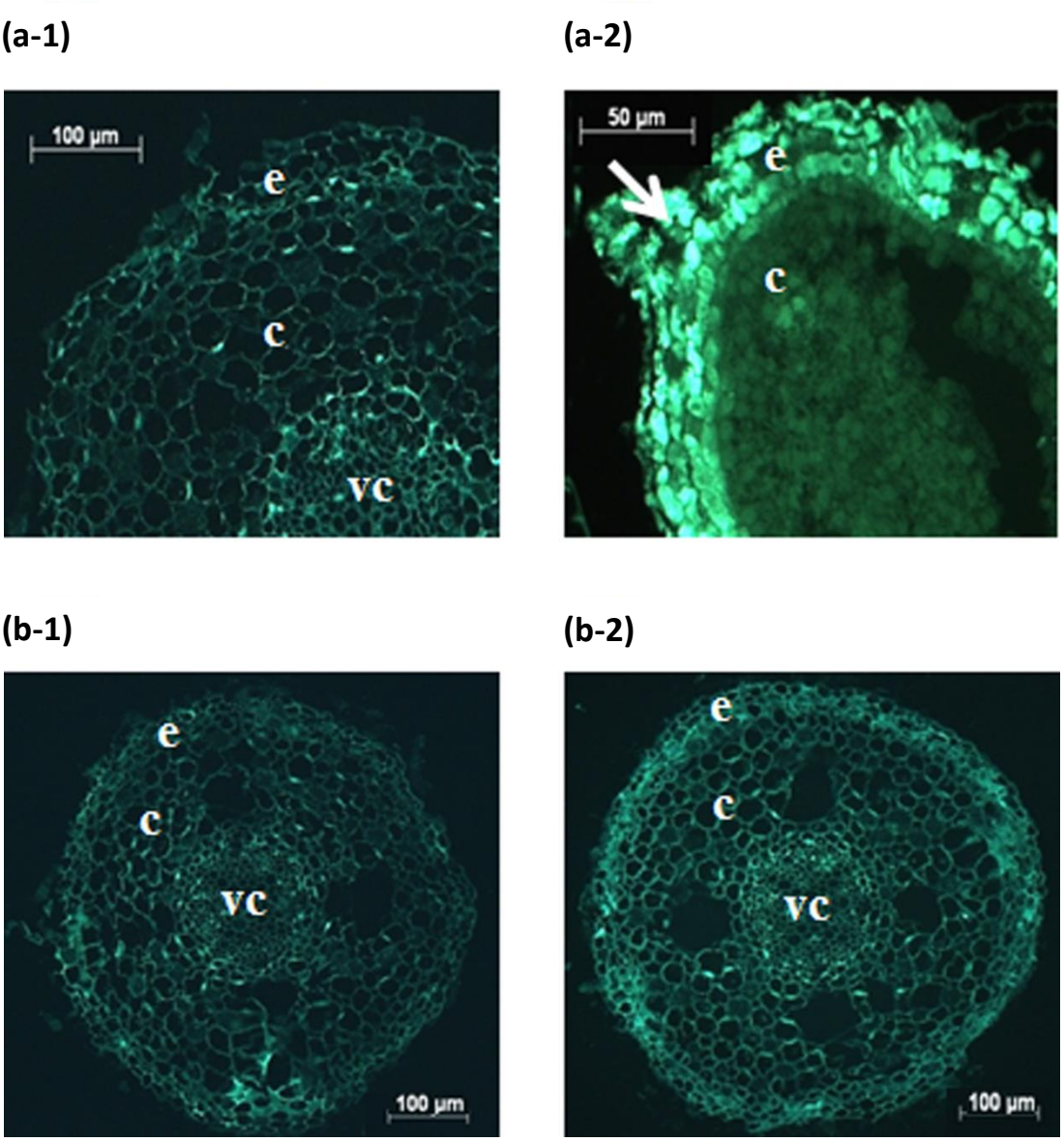


Fig. 6. Sornyotha *et al*.