การเปลี่ยนคริสตัลลีนเซลลูโลสไปเป็นตัวทำละลายโดยแบคทีเรียสกุลคลอสทริเดียที่ได้จากการคัดแยกจากธรรมชาติและการตัดต่อทางพันธุกรรม

นางสาว ชมภูนุช วิรุฒานนท์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาวิทยาศาสตรชีวภาพ
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2550

CONVERSION OF CRYSTALLINE CELLULOSE TO SOLVENT BY NATURALLY SELECTED AND RECOMBINANT CLOSTRIDIA

Miss Chompunuch Virunanon

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in Biological Sciences
Faculty of Science
Chulalongkorn University
Academic Year 2007

Thesis Title	CONVERSION OF CRYSTALLINE CELLULOSE TO SOLVENT BY NATURALLY SELECTED AND RECOMBINANT CLOSTRIDIA
By	Miss Chompunuch Virunanon
Field of Study	Biological Sciences
Thesis Advisor	Associate Professor Warawut Chulalaksananukul, Ph. D.
Thesis Co-advisor	Professor Philippe Soucaille, Ph. D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

.....Dean of the Faculty of Sciences
(Professor Supot Hannongbua, Ph. D.)

THESIS COMMITTEE

.....Chairman
(Associate Professor Preeda Boon-long, Ph.D.)

.....Thesis Advisor
(Associate Professor Warawut Chulalaksananukul, Ph. D.)

.....Thesis Co-advisor
(Professor Philippe Soucaille, Ph. D.)

.....Member
(Assistant Professor Alisa Vangnai, Ph. D.)

.....Member
(Assistant Professor Jittra Piapukiew, Ph. D.)

.....Member
(Assistant Professor Jessada Dendoungbaripant, Ph. D.)

.....External member
(Jittima Charoenpanich, Ph. D.)

ชมภูนุช วิรุฒานนท์ : การเปลี่ยนคริสตัลลีนเซลลูโลสไปเป็นตัวทำละลายโดยแบคทีเรียสกุลคลอสทริเดียที่ได้จากการคัดแยกจากธรรมชาติและการตัดต่อทางพันธุกรรม. (CONVERSION OF CRYSTALLINE CELLULOSE TO SOLVENT BY NATURALLY SELECTED AND RECOMBINANT CLOSTRIDIA) อ .ที่ปรึกษา : รศ. ดร .วรวุฒิ จุฬาลักษณานุกูล อ. ที่ปรึกษาร่วม: Prof. Philippe Soucaille, Ph.D., 182 หน้า.

มีโซฟิลิกเซลลูโลโซมของ *Clostridium cellulolyticum* มีคลัสเตอร์ของยีนขนาดใหญ่ซึ่งคล้ายคลึงมากกับใน *Clostridium acetobutylicum* ATCC824 โมเดลของมีโซฟิลิกเซลลูโลโซมนี้ได้รับการรายงานเมื่อไม่นานมานี้และทั้งนี้คลัสเตอร์ของยีนยังได้รับการจำแนกลักษณะเฉพาะ Cel48F เป็นหนึ่งในเซลลูเลสหลักของเซลลูโลโซมของ *C. cellulolyticum* โดยมีแอคทิวิตีต่อ avicel ที่ถูกย่อยโดยกรด carboxymethyl cellulose และแม้กระทั่งใน crystalline cellulose ในทางตรงกันข้าม *C. acetobutylicum* ผลิตเซลลูโลโซมที่ไม่อยู่ในสภาวะที่ทำงานได้ซึ่งเกี่ยวข้องกับการกลายของกรดอะมิโนในหน่วยย่อย Cel48A หนึ่งในสามเซลลูเลสหลักในเซลลูโลโซมของ *C. acetobutylicum* ลำดับของยีน *cel*48A มีการถอดรหัสได้เป็นสายโพลีเปปไทด์ประกอบด้วยกรดอะมิโนซึ่งมีมวลโมเลกุลจากการคำนวน 84 000 ดาลตัน โปรตีนประกอบด้วยส่วนเร่งปฏิกริยาอยู่ในสกุลที่ 48 ด้วยเหตุนี้ยีน *cel*48A จึงถูกดัดแปลงด้วยวิธีการมิวเตชันแบบจำเพาะเจาะจงเพื่อที่จะคืนความสามารถของ Cel48A ในขณะเดียวกันโปรตีนลูกผสมใหม่ระหว่าง *cel*48A dockerin domain และ *cel*48F catalytic domain ได้ถูกประกอบขึ้น ผลการทดลองพบว่าโปรตีนบริสุทธิ์ลูกผสมระหว่าง *cel*48A-*cel*48F มีความสามารถในการเร่งปฏิกิริยากับเซลลูโลสที่นำมาทดสอบทุกชนิด เป็นที่น่าสนใจว่าโปรตีนลูกผสมมีแอคทิวิตีต่อคริสตัลลีนเซลลูโลสดีเทียบเท่ากับใน Cel48F ดั้งเดิมของ *C. cellulolyticum* (11.882 and 13.4 IU/μmol ตามลำดับ) โดยภาพรวมสามารถสรุปได้ว่าโปรตีนลูกผสมมีประสิทธิภาพต่อเซลลูโลสสูงกว่าในส่วนเร่งปฏิกริยาของ *C. acetobutylicum* ดั้งเดิม ประสิทธิภาพของโปรตีนลูกผสมต่อ carboxymethyl cellulose (CMC) ซึ่งพัฒนาขึ้น 30.8 เท่าจากดั้งเดิม (Cel48A ดั้งเดิม 0.011 IU/μM ลูกผสม 0.339 IU/μM). แอคทิวิตีต่อ Phosphoric acid swollen cellulose (PASC) ปรับปรุงขึ้น 2771 เท่า (Cel48A ดั้งเดิม 0.002 IU/μM ลูกผสม 5.542 IU/μM) โปรตีนรีคอมบิแนนท์ที่สั้นลงถูกตรวจพบในสายพันธุ์ *C. acetobutylicum* p952-SA-FA ในขั้นตอนการคัดเลือกสายแบคทีเรียจากธรรมชาติ ไอโซเลททั้ง 15 ถูกจำแนกเข้าในลำดับ Clostridia อย่างรวดเร็วด้วยมาตรฐาน 4 ขั้นตอน: การสร้างสปอร์, ความสามารถในการรีดิวซ์ซัลไฟท์, ผลผลิตจากเมตาโบลิซึมและลำดับเบสของ 16S rDNA ระบบการคัดเลือกในงานวิจัยนี้เหมาะสมต่อการคัดเลือก *Clostridiaceae* ในระดับความคล้ายคลึง 83-100% การวิเคราะห์เปรียบเทียบแสดงให้เห็นว่ารีคอมบิแนนท์มีระดับของแอคทิวิตีของเอกโซกลูคาเนสได้อย่างเท่าเทียมกับสายพันธุ์ที่มาจากการคัดเลือกในธรรมชาติ

สาขาวิชา วิทยาศาสตร์ชีวภาพ
ปีการศึกษา 2550

ลายมือชื่อนิสิต.....
ลายมือชื่ออาจารย์ที่ปรึกษา.....
ลายมือชื่ออาจารย์ที่ปรึกษาร่วม.....

4573869723 MAJOR: BIOLOGICAL SCIENCES
KEY WORD: *Clostridium acetobutylicum*/ CELLULOSOME/ CELLULOSE/ *CEL*48F/ *CEL*48A

MISS CHOMPUNUCH VIRUNANON: CONVERSION OF CRYSTALLINE CELLULOSE TO SOLVENT BY NATURALLY SELECTED AND RECOMBINANT CLOSTRIDIA. THESIS ADVISOR: ASSOC. PROF. WARAWUT CHULALAKSANANUKUL, Ph. D. CO-ADVISOR: PROF. PHILIPPE SOUCAILLE, Ph. D., 182 pp.

Clostridium cellulolyticum, a mesophilic clostridial cellulosomes have a large cluster which is very similar to that of *Clostridium acetobutylicum* ATCC824. This model mesophilic cellulosome has recently been reported and cellulose gene cluster also characterized. Cel48F, one of the major cellulase of *C. cellulolyticum*, is active on degraded swallen avicel, carboxymethyl cellulose and even crystalline cellulose. In contrast, *C. acetobutylicum* produces inactive cellulosome which dues to the mutated amino acid in Cel48A subunit, one of three major cellulases in *C. acetobutylicum* cellulosome. The *cel*48A sequences encode polypeptides containing amino acids with calculated molecular mass of 84 000 Da. Protein contain a catalytic domain belonging to the family 48. For this reason, *cel*48A gene was modified by site-directed mutagenesis method to restore the activity of Cel48A. In parallel, new hybrid molecules of *cel*48A dockerin domain and *cel*48F catalytic domain were constructed. The results show that purified hybrid *cel*48A-*cel*48F has proficient activity on all cellulolytic substrates inspected. Interestingly, the hybrid protein shows activity on crystalline cellulose as well as efficient native *C. cellulolyticum* Cel48F (11.882 and 13.4 IU/μmol respectively). Overall results can be concluded that the chimera protein exhibited enhanced action on cellulose more than in native *C. acetobutylicum* catalytic domain of cellulose. The action of hybrid enzyme over carboxymethylcellulose (CMC) developed for 30.8 folds from native (native Cel48A: 0.011 IU/μM, hybrid: 0.339 IU/μM). Activity over Phosphoric acid swollen cellulose (PASC) was improved for 2771 folds (native Cel48A: 0.002 IU/μM, hybrid: 5.542 IU/μM). Truncate recombinant hybrid protein can be detected in p952-SA-FA *C. acetobutylicum* strain. In the screening from natural process, Fifteen isolates were rapidly classified as in the class Clostridia by four selected criterias: endospore formation, sulfite reducing ability, metabolic products and 16S rDNA sequence. The selective system in this research was appropriate for the screening of *Clostridiaceae* in a similarity range between 83-100%. Comparative analysis shows that recombinant also expresses exoglucanase activity as well as naturally selected Clostridia.

Field of Study: Biological Sciences
Acadamic Year 2007

Student's signature....Chompunuch Virunanon
Adivisor's signature....Warawut Chulalaksananukul
Co-advisor's signature....

ACKNOWLEDGEMENTS

I'd really appreciate my advisor Associate Professor Warawut Chulalaksananukul thesis director for priceless suggestion, discussion and proof all thesis. This research also supported by Biofuel Production by Biocatalyst Research Unit, Chulalongkorn University, Thailand.

My thesis co-advisor, Professor Philippe Soucaille and Professor Christian Croux for discussion, proof and advises thesis results and experiments in detail. Thanks for Sophie Mondeil and Maria Gonzales-Pajuelo for molecular technique initiation. I also deeply appreciate for Prokaryote team, Centre de Bioingénierie Gilbert Durand, Département de Génie Biochimique et Alimentaire, INSA, Toulouse, France.

I am grateful to Professor Henry-Piere Fierobe from Marseille, France for generous substrates (BMCC) and antibody for cellulosome detection.

Also acknowledge to Fabrice Sabathé and Laurence Girbal for kindly helpful suggestion and methodology especially in Cel9X and Cel9C activity assay.

This is also grateful to my thesis defense committee. Associate Professor Preeda Boon-long, Assistant Professor Alisa Vangnai, Assistant Professor Jessada Dendoungbaripant, Assistant Professor Jittra Piapukiew, Jittima Charoenpanich, Ph. D. for profitable comments on scientific points. I'd like to express my graditude to Department of Botany, Faculty of science, Chulalongkorn University.

Assistant Professor Jittra Piapukiew, Ph.D. for helpful 16SrDNA technique and discussion. Associate Professor Jessada Dendoungbaripant from Department of Biology, Chulalongkorn University, Thailand for MEGA4 and Neibour joining tree analysis of cellulolytic Clostridia screening.

This research was supported by Thailand Research Fund through the RGJ (The Royal Golden Jubilee Ph. D. Program grant PHD/0194/2545, Thailand), Commission on Higher Education, Thailand and The French Embassy in Thailand.

Finally, give thanks for my family and my friends for encourages.

CONTENTS

Page

Page

Page

Page

Page

Page

Page

LIST OF TABLES

Table **Page**

Table **Page**

LIST OF FIGURES

Figure **Page**

Figure **Page**

Figure **Page**

Figure **Page**

LIST OF ABBREVIATIONS

°C: Degree Celcius

BMCC: bacterial microcrystalline cellulose

CBD: cellulose-binding domain.

CBM: carbohydratebinding domain

CbpA: cellulose binding protein A

Cc: *Clostridium cellulolyticum*

Cel: cellulase

CipA: cellulosome integrating protein A

CipC: cellulosome integrating protein C

CMC: carboxymethylcellulose

Coh: cohesin domain

Da: Dalton

Doc: dockerin domain.

EngE, endoglucanase E

FRT: recombinase recognition site

GH: glycoside hydrolase

HLD: hydrophilic domain

Ig: immunoglobulin-like module

kD: kilodalton

M: Molar

Man: mannanase

MLSR: erythromycin resistant gene

OlpB: outer layer protein B

ORF: open-reading frame

ORF2p: ORF 2 polypeptide

ORFXp: ORF X polypeptide

PASC: phosphoric acid swallen cellulose

Pel: pectate lyase

PL: polysaccharide Lyase

PTS: proline–threonine–serine

Rlg: rhamnogalacturonase

SdbA; scaffoldin dockerin binding A

SLH: S-layer homology

Xyn: xylanase

CHAPTER I

INTRODUCTION

1. Rational

Due to the programmed depletion in fossil fuel, there is an increasing interest in the production of fuels and solvents from renewable resources. Butanol and butyrate can be produced by bacteria from the genus *Clostridium*. The most studied and efficient strain for this purpose is *Clostridium acetobutylicum*. It is an anaerobic bacterium that converts various carbohydrate substrates and polysaccharides into acids and solvents. Degradation of crystalline cellulose by cellulolytic Clostridia is essentially due to the action of a cellulosome. A recent and unpublished study from our group has shown that in *C. acetobutylicum,* the Cel48A enzyme, which has been reported to be the major component of the cellulosome is inactive on all cellulosic substrates tested. Three mutations have been identified in the *Cel48A* gene that might be responsible for the non-functionality of the Cel48A enzyme. Moreover, other two main subunits of *C. acetobutylicum* cellulosome were tested on binding ability and activity over cellulose. The aim of our study is to modify *C. acetobutylicum* cellulosome by restoring the activity of Cel48A using protein engineering techniques. Then, the cellulolytic, solvent producing Clostridia from soil samples in Thailand will be screened in order to compare their cellulolytic activity and solvent-producing ability between the recombinant and naturally selected Clostridia.

2. Objective

2.1. To modify inactive cellulosome of *C. acetobutylicum* by using genetic engineering technique

2.2. To isolate and characterize cellulolytic, solvent producing clostridia from natural samples in Thailand

2.3. To compare cellulolytic activity and solvent producing ability of the naturally selected and recombinant Clostridia

3. Content of thesis

Thesis scopes are listed.

3.1 Study on activity of three major cellulases in *C. acetobutylicum* cellulosome (Cel9X, Cel9C and Cel48A).

3.2 Construct hybrid and modify Cel48A subunit of *C. acetobutylicum* cellulosome. Modified Cel48A will be tested on expression and activity over crystalline cellulose.

3.3 Isolate and characterize cellulolytic, solvent producing clostridia from cow intestinal tract and decomposed samples in Thailand.

4. Expected outcomes

Natural strains from screening and developed strain appropriate for biofuel production from crystalline cellulose are expected outcome.

These developed strains will be applied in a biological process for the production of solvent by fermentation from biomass**.**

5. Strategy

5.1 Literature review

5.2 Isolation and characterization of cellulolytic, solvent producing, Clostridia from soil samples in Thailand

5.3 Culture *C. acetobutylicum* for engineer Cel48A

5.4 Gene modifying and construction of corrected pET-Cel48A plasmid for expression of Cel48A protein

5.5 Expression of corrected Cel48A protein in *E. coli* strain

5.6 Targeted inactivation of malfunction Cel48A gene in *C. acetobutylicum*

5.7 Insertion of the corrected gene into an expression vector for *C. acetobutylicum*

5.8 Study of the expression of corrected Cel48A gene in *C. acetobutylicum*

5.9 Compare their cellulolytic activity and solvent producing ability between the recombinant and natural selected Clostridia.

5.10 Writing thesis

CHAPTER II

LITERATURE REVIEWS

Plant Biomass is the only foreseeable sustainable source of fuel and materials available to humanity. Cellulosic materials are particulary attractive in this context because of their relatively low cost and plentiful supply. The central technological impediment to more widespread utilization of this important resource is the general absence of low-cost technology for overcoming the recalcitrance of cellulosic biomass. A promising strategy to overcome this impediment involves the production of cellulotic enzymes, hydrolysis of biomass, and fermentation of resulting sugars to desired products in a single process step via a cellulolytic microorganism or consortium. Such “consolidated bioprocessing”(CBP) offer very large cost reductions if microorganisms can be developed that process the required combination of substrate utilization and product formation properties (Lynd *et al*., 2002).

2.1 Structure and Composition of cellulosic Biomass

Cellulose, the most abundant component of plant biomass, in found in nature almost exclusively in plant cell walls, although it is produced by some animals (e.g., tunicates) ans a few bacteria. Despite great differences in composition and in the anatomical structure of cell walls across plant taxa, high cellulose content-typically in the range of approximately 35 to 50% of plant dry weight-is a unifying feature. In a few cases notably cotton bolls, cellulose is present in a nearly pure state. However, in most cases the cellulose fibers are embedded in a matrix of other structural biopolymers, primarily hemicelluloses and lignin, which comprise 20 to 35 and a 5 to 30% of plant dry weight. As shown in table1, cellulose, hemicelluloses and lignin composition are different among each type of lignocellulosic biomass. Although these matrix interactions vary with plant cell type and with maturity, they are a dominant structural feature limiting the rate and extent of utilization of whole, untreated biomass materials. A detailed description of these interactions and the mechanisms by which they limit hydrolysis and utilization is beyond the scope of this paper and is the topic of several recent reviews.

Table 1. Proportion of cellulose, hemicellulose and lignin in each type of lignocellulosic biomass

Biomass	Cellulose	Hemicellulose	Lignin
Herb	30-50%	15-40%	5-20%
Bagasse	40-55%	25-40%	5-25%
Hard wood	40-50%	20-30%	15-30%
Soft wood	40-55%	10-15%	25-30%
Agricultural waste	30-40%	10-40%	10-30%
Average	40-60%	20-40%	10-25%

2.1.1 Cellulosic biomass: plant cell wall structure

One of the most important distinguishing features of plant cells is the presence of cell wall. The relatively rigidity of the cell wall renders plant secondary, unlike animal which has structure allows their cells more flexibility. Cell walls are significantly thicker than plasma membranes and were visible even without microscope. The tree primary polymers that make up plant cell walls consist of about 35 to 50% cellulose, 20 to 35% hemicellulose and 10-25% lignin. Lignin fills the space in the cell wall between cellulose, hemicellulose and pectin components.

Plant cells walls also incorporate a number of proteins; the most abundant include hydroxyproline-rich glycoproteins (HRGP), also called the extensions, the arabinogalactan proteins (AGP), the glycine-rich proteins (PRPs). With the exception of glycine-rich proteins, all the previously mentioned proteins are glycosylated and contain hydroxyproline (Hyp). Each class of protein is defined by characteristic, highly repetitive protein sequence. Chimeric proteins contain two or more different domains, each with a sequence from different class of glycoprotein. Most cell wall proteins are cross-linked to the cell wall and may have structural function.

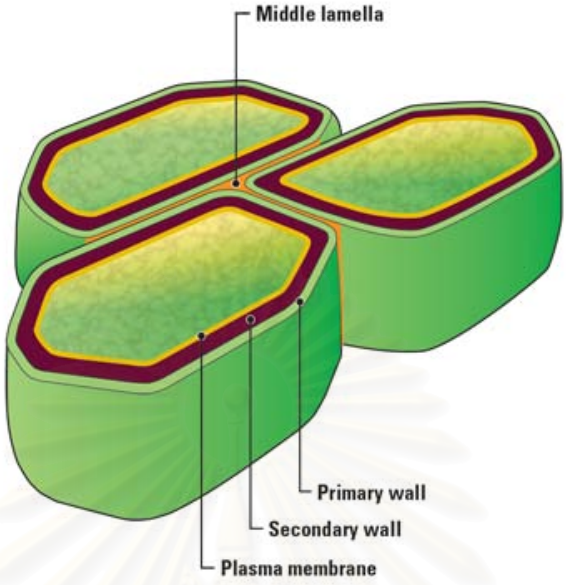


Figure 1. Schematic diagram of plant cell (Taiz and Zeiger, 1991)

The main ingredient in cell walls are polysaccharides or complex carbohydrates or complex sugars) which are built from monosaccharides (or simple sugars). There are three major regions of the wall:

1. Middle lamella-outermost layer, glue that binds adjacent cells, composed primarily of pectic polysaccharides.

2. Primary wall- wall deposited by cells before and during active growth. The primary wall of cultured sycamore cells is comprised of pectic polysaccharides (ca. 30%), cross-linking glycans (hemicellulose; 25%), cellulose (15-30%) and protein (ca. 20%) (Darvill *et al*., 1980). The actual content of the wall components varies with species and age. All plant cells have a middle lamella and primary wall.

3. Secondary wall- some cells deposit additional layer inside the primary wall. This occurs after growth stops or when the cells begin to differentiate (specialize). The secondary wall is mainly for support and is comprised

primarily of cellulose and lignin. Often can distinguish distinct layer, S1, S2 and S3-which differ in the orientation, or direction, of the cellulose microfibrils.

Eleven sugars are common in these polysaccharides including like glucose and galactose. Carbohydrates are good building blocks because they can produce a nearly infinite variety of structures.

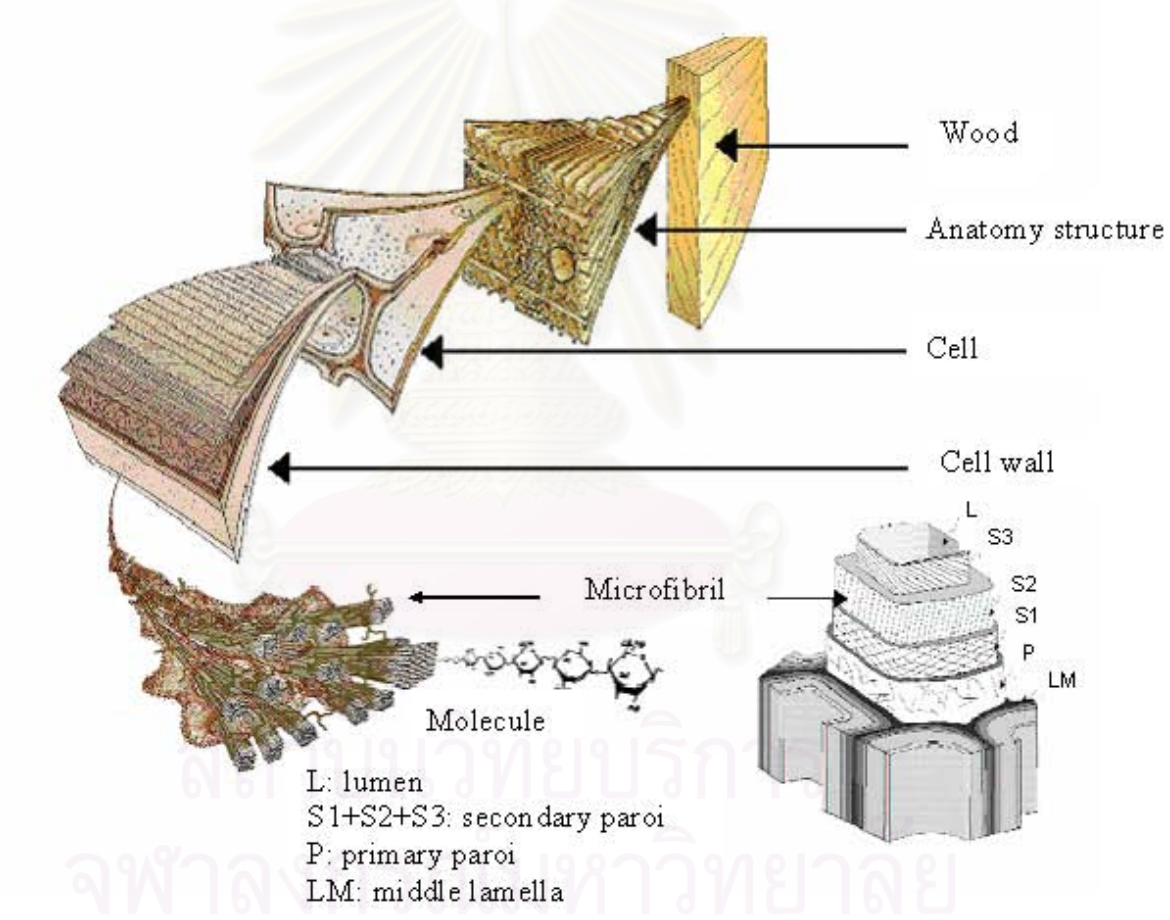


Figure 2. Cellulose structure in plant cell wall

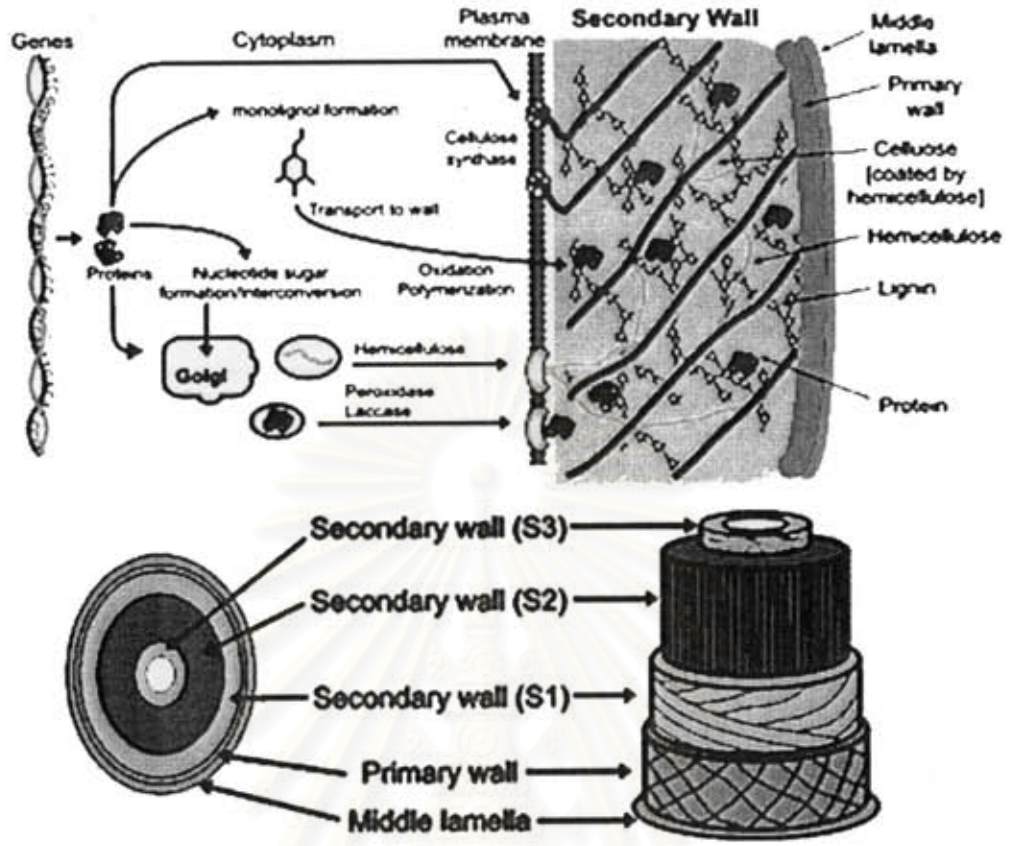


Figure 3. Structure in plant cell wall (Complex Carbohydrate Research Center, The University of Georgia, 2007)

2.1.1.1 Cellulose

Cellulose ($C_6H_{10}O_5$) is a polysaccharide of ß-glucosen with a ß-1,4-backbone. It forms the primary structure component of green plants. The primary cell wall of plant is made of cellulose. The secondary wall contains cellulose with variable amounts of lignin. Lignin and cellulose, considered together, are termed lignocellulose, which (as wood) is argued to be one of the most common biopolymers on earth. It was discovered and isolated by the French chemist Anselme Payen since 1986. Cellulose is the most abundant form of living terrestrial biomass with an estimated annual production of 1.5×10^9 tonnes (Kim and Yun, 2006).

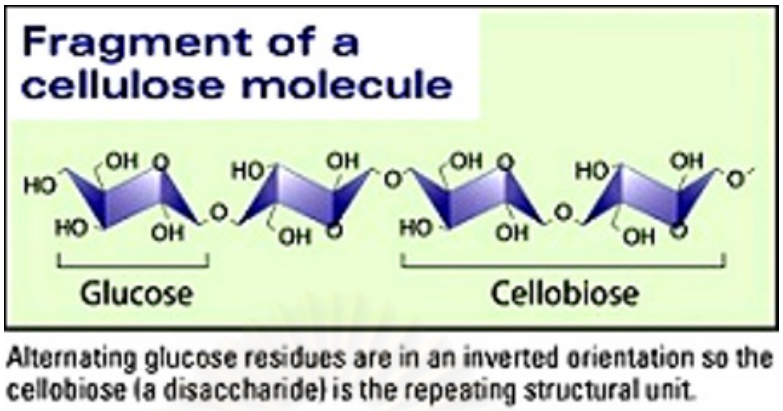


Figure 4. Cellulose is a polymer of β-D-glucose (U. S. Department of Energy Office of Science, 2007)

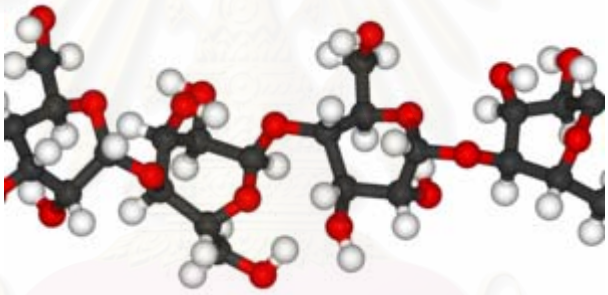


Figure 5. Cellulose in 3D structure (Brown *et al*., 2000)

An important feature of cellulose, relatively unusual in the polysaccharide world, is its crystalline structure. Cellulose is synthesized in nature as individual molecules (linear chains of glucosyl residues) which undergo self-assembly at the site of biosynthesis (Brown *et al*., 2000). There is evidence that associated hemicelluloses regulate this aggregation process (Atalla *et al*., 1993). Approximately 30 individual cellulose molecules are assembled into large units known as elementary fibrils (protofibrils), which are packed into large units called microfibrils, and these are in turn assembled into the familiar cellulose fibers.

Well known cellulose available commercial or used for research

- Highly crystalline cellulose (e.g., bacterial cellulose (BC) and bacterial microcrystalline cellulose (BMCC)) synthesized by aerobic bacterium *Acetobacter xylinum*). It made of 76-92% crystalline cellulose.

- Microcrystalline cellulose (e.g., Avicel and Sigmacell) is cellulose commercial. It can be determine crystallinity around 47% (GAL, unpublished data in thesis). They are nearly pure cellulose, and the dilute acid treatment used in their preparation removes hemicelluloses and introduces extensive amorphous regions in the cellulose fibers.

- Highly soluble cellulose (CMC: Carboxymethyl cellulose, cellodextrins, paranitrophenyl-ß-D-cellodextrins (pNP-cellodextrins) and methylumbelliferyl-β-D-cellodextrins (MUC-cellodextrins)), always used as a substrate for studies of endoglucanase production.

- Amorphous cellulose (PASC: Phosphoric acid swallen cellulose)

2.1.1.2 Cross-link glycans (Hemicellulose)

Diverse groups of carbohydrates that used to be called hemicellulose. Characterized by being soluble in strong alkaline. They are linear (straight), flat, with a ß-1,4 backbone and relatively short side chains. Two common types include glucomannans, galactoglucomannans and galactomannans. The main feature of this group is that they do not aggregate with themselves-in other words, they don't form microfibrils. However, they form hydrogen bonds with cellulose and hence the reason they are called "cross-linking glycans". There may be a fucose sugar at the end of side chains which may help keep the molecules planar by interacting with other regions of the chain.

Various agricultural residues, such as corn fiber, corn stover, wheat straw, rice straw, and sugarcane bagasse, contain about 20-40% hemicellulose, the second most abundant polysaccharide in nature. Xylans are the most abundant hemicelluloses. Hemicelluloses are heterogeneous polymers of pentoses (xylose, arabinose), hexose (mannose, glucose, galactose), and sugar acids. Unlike cellulose, hemicelluloses are

not chemically homogeneous. Hardwood hemicelluloses contain mostly xylans, whereas softwood hemicelluloses contain mostly glucomannans. Xylans of many plant materials are heteropolysaccharides with homopolymeric backbone chains of 1, 4-linked β-D-xylopyranose units. Besides xylose, xylan may contain arabinose, glucoronic acid or its 4-*O*-methyl ether, and acetic, ferulic, and *p*-coumaric acids.

Xylan from different sources, such as grasses, cereals, softwood, and hardwood, differ in composition. Rice bran neutral xylan contain 46% xylose, 44.9% arabinose, 6.1% galactose, 1.9% glucose, and 8.3% anhydrouronic acid. Corn fiber xylan is one of the complexheterxylan containing β-1,4-linked xylose residues. It contains 48-54% xylose, 33-35% arabinose, 5-11% galactose, and 3-6% glucuronic acid (Saha, 2003).

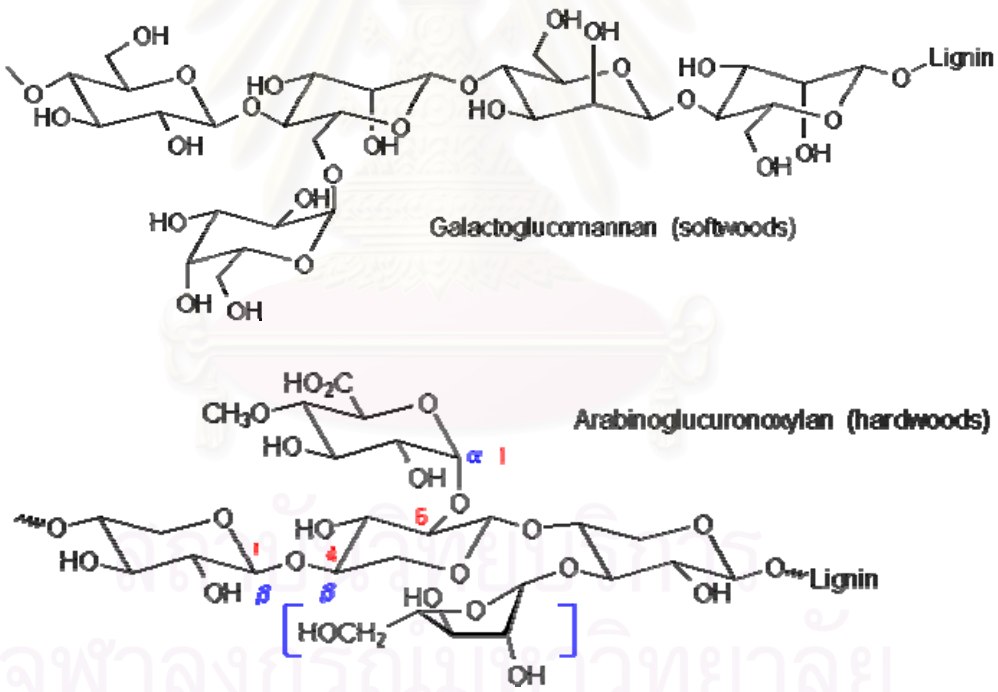


Figure 6. Structure of hemicellulose in softwood and hard wood (Saha, 2003)

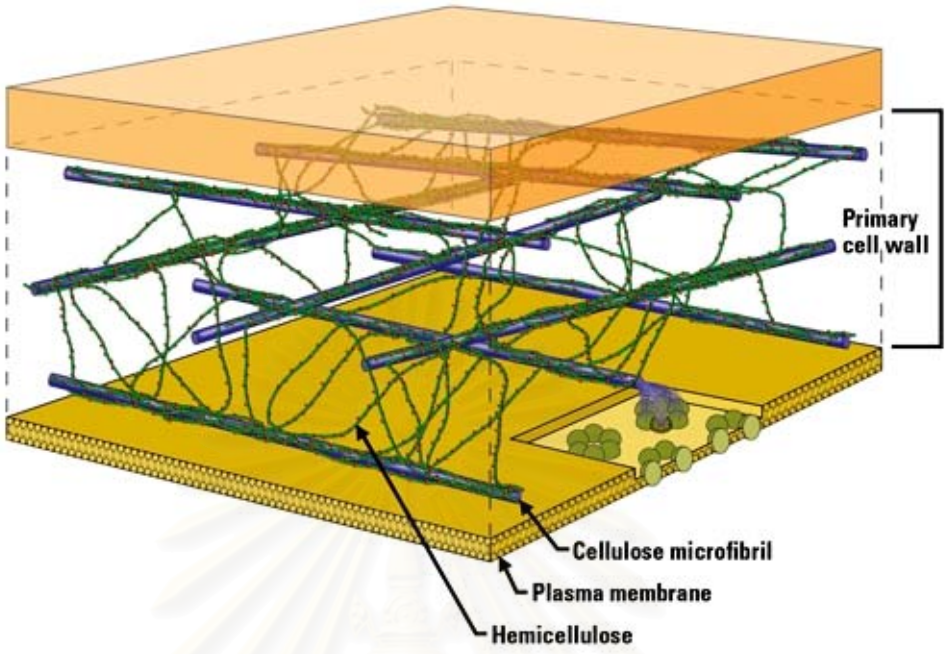


Figure 7. Location of hemicellulose in plant cell wall (Somerville, 2004)

2.1.1.3 Pectic polysaccharides

These are the most complex polysaccharides found in the cell wall and contain main chains which are composed of residue of galacturonic acid, mostly methylated, with rhamnose residues at interval in the chain. These are extracted from the wall with hot water or dilute acid or calcium chelators (like EDTA). They are the easiest constituents to remove from the wall. They form gels (i.e., used in jelly making). Another diverse group of polysaccharides that are particulary rich in galacturonic acid (galacturonan = pectic acids). Polymers of primary β 1,4 galacturonans (=polygalacturonans) are called homogalacturons (HGA) and are particularly common. These are helical in shape. Divalent cations, like calcium, also form cross-linkanges to join adjacent polymers creating a gel. Pectic polysaccharides can also be cross-linked by dihydrocinnamic or diferulic acids. The HGA6s (galacturonans) are initially secreted from the golgi as methylated polymers; the methyl groups are removed by pectin methylesterase to initiate calcium binding.

Other pectic acids include Rhamnogalacturonan II (RGII) which features rhamnose and galacturonic acid in combination with a large diversity of other sugars in varying linkages. Dimer of RGII can be cross-linked by boron atoms linked to apiose sugars in a side chain.

Although most pectic polysaccharides are acidic, other are composed of neutral sugars including arabinans and galactans. The pectic polysaccharides serve a variety of functions including determining wall porosity, providing a charged wall surface for cell-cell adhesion (middle lamella), cell-cell recognition, pathogen recognition and others.

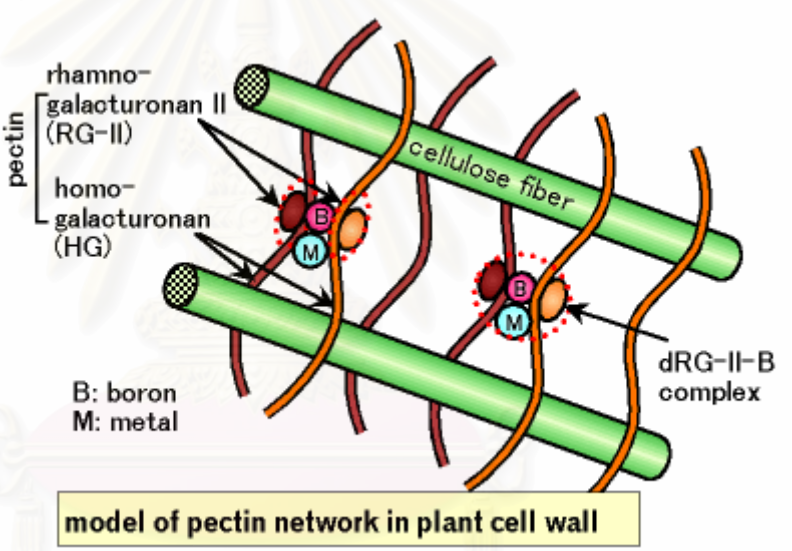


Figure 8. Model of pectin network in plant cell wall (Matsunaka, 2001)

2.1.1.4 Protein

Wall proteins are typically glycoproteins (polypeptide backbone with carbohydrate sidechains). The proteins are particularly rich in the amino acids hydroxyproline (hydroxyproline-rich glycoproteins, HPRG), proline (proline-rich protein, PRP), and glycine (glycine-rich protein, GRP). These proteins form rods

(HRGP, PRP) or beta-pleted sheets (GRP). Extensin is a well-studied HRGP. HRGP is induced by wounding and pathogen attack. The wall proteins also have a structural role since: (1) the amino acids are characteristic of other structural proteins such as collagen and gelatin; and (2) to extract the protein from the wall requires destructive conditions. Protein appears to be cross-linked to pectic substances and may have sites for lignification. The protein may serve as the scaffoldin used to construct the other wall components.

2.1.1.5 Lignin

Lignin contained polymer of phenolics, especially phenylpropanoids. Lignin is primarily a strengthening agent in the wall. It also resists fungal/pathogen attack. It is a chemical compound (complex, highly cross-linked aromatic polymer) that is most commonly derived from wood and is an integral part of the cell walls of plants, especially in tracheids, xylem fibres and sclereids.

The term lignin was introduced in 1819 and is derived from the latin word lignum. It is one of the most abundant organic compounds on earth after cellulose and chitin. Lignin makes up about one-quarter to one-third of the dry mass of wood. It has several unusual properties for being a biopolymer, such as having a network structure and lacking a defined primary structure. It fills the spaces in cell wall between cellulose, hemicellulose and pectin components.

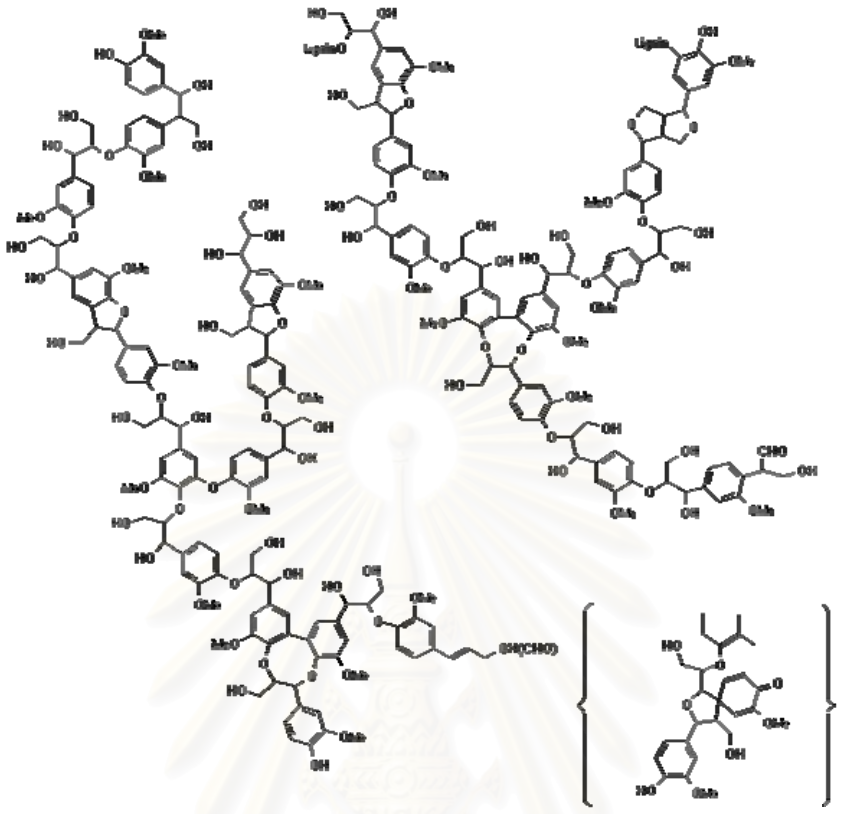


Figure 9. Structure of lignin in natural cell wall (Cole and Fort, 2007)

2.1.1.6 Suberin, wax, cutin

A variety of lipids are associated with the wall for strength and waterproofing.

2.1.1.7 Water

The wall is largely hydrated and comprised of between 75-80% water. This is responsible for some of the wall properties. For example, hydrated walls have greater flexibility and extensibility than non-hydrated wall.

2.2 Cellulolytic enzymes

2.2.1 Cellulases

For microorganisms to hydrolyze and metabolize insoluble cellulose, extracellular cellulases must be produced that are either free or cell associated.

Celluloses systems were first classified based on their mode of catalytic action and have more recently been classified based on structural properties. Three major type of enzymatic activities are found: (i) endoglucanases or 1,4-β-D-glucan-4-glucanohydrolases (EC 3.2.1.4), (ii) exoglucanases, including 1,4- β-D-glucan-glucanohydrolases (Cellodextrinases) (EC 3.2.1.74) and 1,4- β-D-glucan-cellobiohydrolases (cellobiohydrolases) (EC 3.2.1.91), and (iii) β-glucosidases (EC 3.2.1.21). Endoglucanase cut random at internal amorphous sites in the cellulose polysaccharide chain. Exoglucanase act in a processive manner on the reducing or non-reducing ends of cellulose polysaccharide chains, generate either glucose or cellobiose as major products. Exoglucanase can also act on microcrystalline cellulose, peeling cellulose chain from the microcrystalline structure.

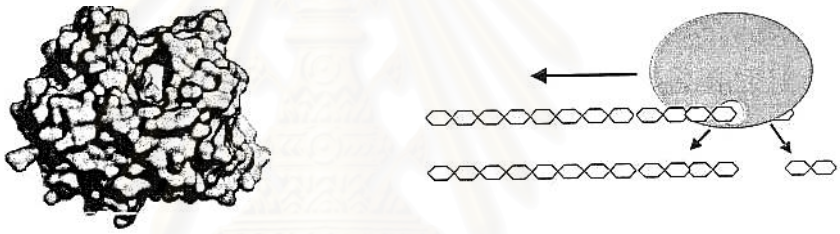


Figure 10. Mode of action type exoglucanase (Davies and Henrissat, 1995)

Figure 11. Mode of action type endoglucanase (Bayer *et al*., 2000)

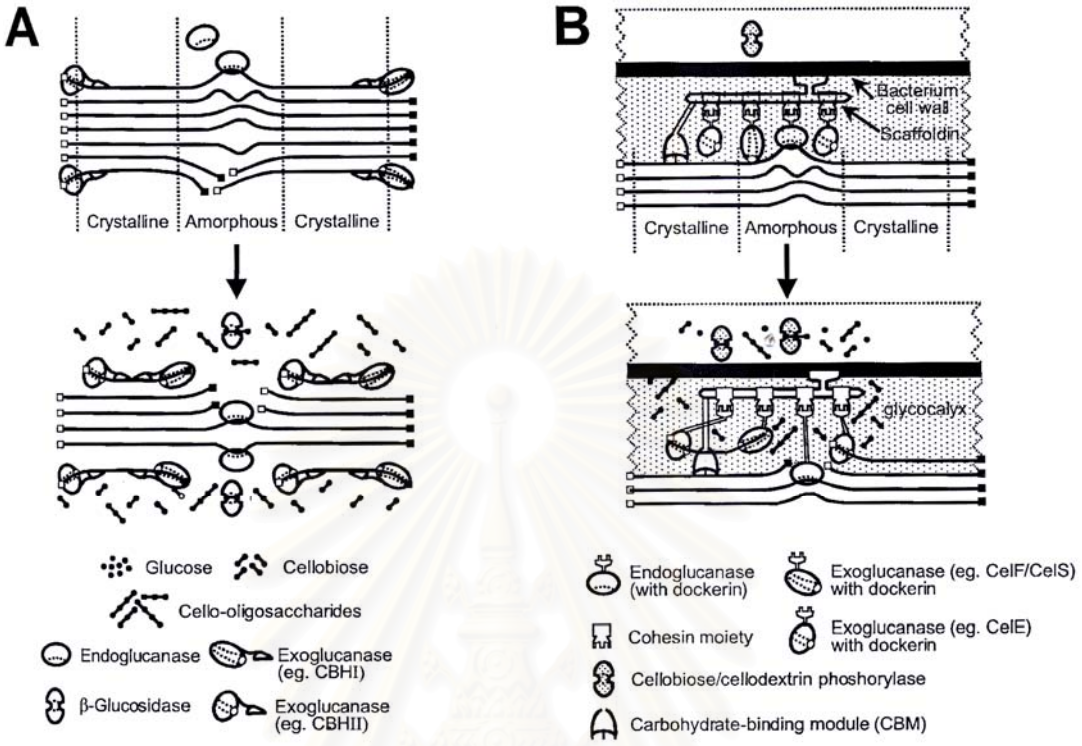


Figure 12. Hydrolysis of amorphous and microcrystalline cellulose by noncomplexed (A) and complexed (B) cellulose systems (Demain *et al*., 2005)

For cellulose degradation, a synergistic action from different components of the cellulolytic system is needed. Notice that there is no single cellulase enzyme can degrade crystalline cellulose significantly. It always required the action of combination cellulases of the cooperating for cellulose saccharification.

Examples of synergistic action of cellulases

1. Endocellulase / Exocellulase (cellulase processive) synergistic action

One of the articles of this case is the action between cellobiohydrolases (CBHI and CBHII) and endoglucanase (EGI and II) of *T. reesei* (Woodward *et al*., 1988). Initially, endocellulase always attack the amorphous zones of crystalline cellulose which generate the molecules easy to process by cellulase processive. Cellulase processive is activated by potential hydrolysis sites.

2. Synergistic action between endocellulase/endocellulase or processive cellulase / processive cellulase

Previous studied (Woodward *et al*., 1988) described synergistic action between processive cellulase one specific to the reducing end of the cellulose chain and another one specific to the nonreducing end of the cellulose chain.

3. Cellulase/β-glucosidase

It plays a major role in cellulose degradation. Cellulase catalyses reaction and form cellobiose as product. Then, β-glucosidase hydrolyses cellobiose to give 2 glucose molecules. Glucose concentration will inhibit β-glucosidase reaction (feedback inhibition). In natural, cellobiose and glucose will be consumed repidly by microorganisms. So, feedback-inhibition effect is a phenomenon occurring only *in vitro*.

4. Synergistic proximity

Bronenmeier and Staudenbauer (1988) studied *in vitro* the synergistic action different catalytic domain. Two microorganisms produce bifunctional-enzyme (presence of 2 catalytic domain), *Caldicellulosiruptor saccharobutyricum* and *A. thermophillum*. Moreover, they fuse two cellulases of *C. stercorarium* (CelZ endoglucanase and CelY exoglucanase). From the observation, it can be reported that the fusion protein express an improved activity.

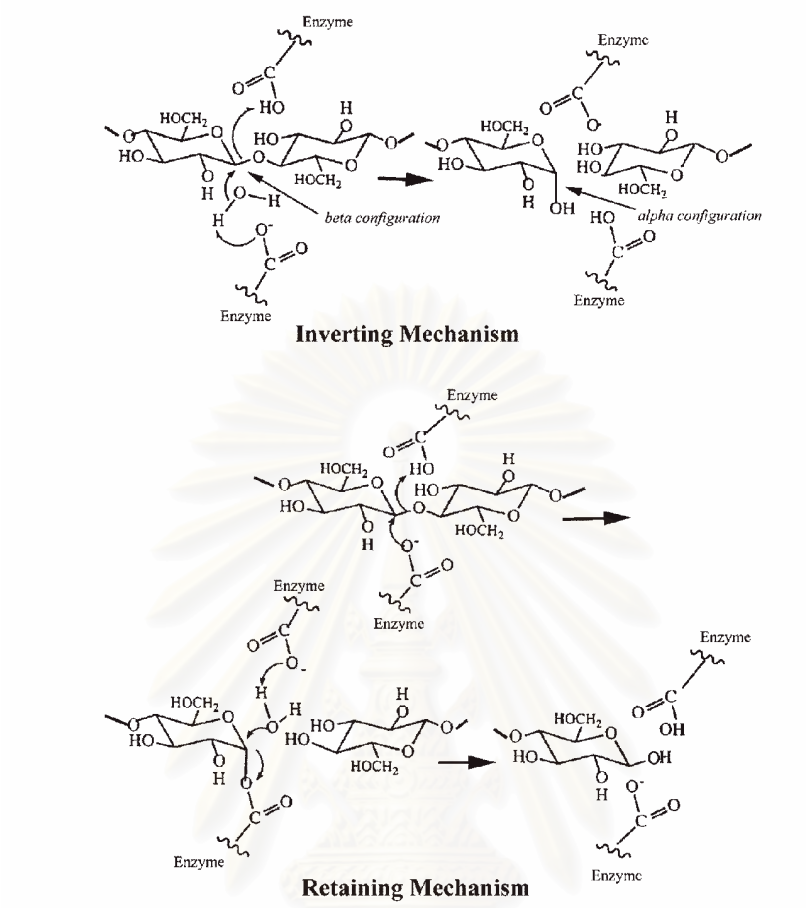


Figure 13. Stereochemically different mechanisms of hydrolysis for cellulases (Wilson *et al*., 1999)

2.2.2 Hemicellulase

Total biodegradation of xylan requires endo-β-1,4-xylanase, β-xylosidase, and several accessory enzymes. In contrast to cellulose, hemicellulose is extremely heterogeneous. Although hemicelluloses are complex heteropolysaccharides, their enzymativ degradation is relatively well known, mainly on basis of hydrolysis studies carried out with isolated soluble substrates. Several different enzymes, collectively called hemicellulases, are needed for the degradation and modification of hemicellulose. In general, the complete hydrolysis of hemicelluloses is accompanished by the combined action of endoenzymes, which cleave the main chain and produce oligosaccharides, ancillary enzymes cleaving the side-chains from the

main chain of the oligosaccharides and ß-glucosidase-enzymes leading to the final liberation of monomeric sugars.

The 1, 4-β-D-xylosidic linkages in xylan are randomly hydrolysed by endo-1,4-β-D-xylanases (EC 3.2.1.8). These enzymes have a higher hydrolysis rate when acting on polymeric xylan than on xylo-oligomers. Their main hydrolysis products are being xylobiose, xylotriose and substituted oligomers containing two to four xylosyl residues. Endoxylanases from several sources have been extensively studied, especially those produced by filamentous fungi *Tricoderma* and *Aspergillus*.

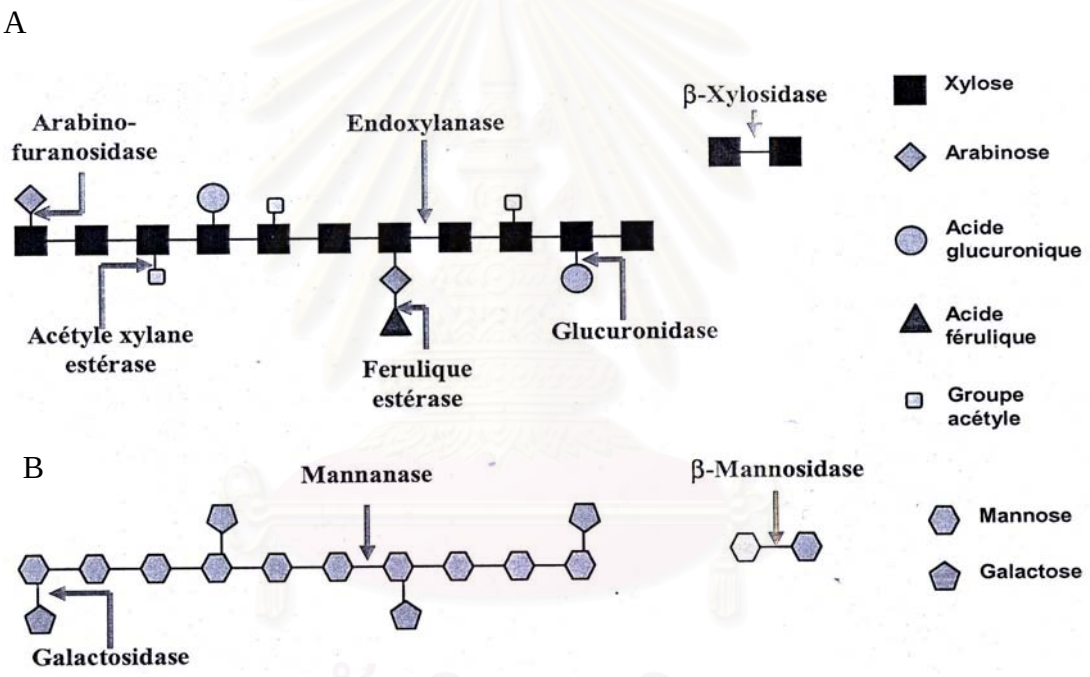


Figure 14. Hemicellulases: A-Degradation of xylan, B-degradation of galactomannan (Margolles-Clark, 1996)

The hydrolysis of the main chain of polymeric glucomannans is carried out by endo-1,4-β-mannanase (EC 3.2.1.78). These enzymes usually hydrolyse substrates with a degree of polymerization higher than three, the main products being mannobiose, mannotriose and various mixed oligosaccharides. Endo-mannanases

have been purified and characterized from several fungi, including *T.* reesei (Margolles-Clark, 1996).

In nature, the degradation of hemicellulose is carried out mainly by fungi and bacteria such as clostridia species. Many sequence-based families of glycosyl hydrolases from *Clostridium cellulovorans* are known, and their genes have been cloned, expressed and characterizd.It had been showed that cellulosomes, isolated from *C. cellulovorans* grown on xylan, were fractioned into 7 to 10 high-molecular-weight multiprotein complexes. The results indicate the greater susceptibility of substrate to hemicellulases activity which is amplified by cooperative action of xylanase and arabinofuranosidase. Hemicellulosic substrate was attacked by xylanolytic and cellulolytic enzymes cooperatively and/or sequentially in the following order: α-arabinofuranosidase→→endo-β-xylanase→→cellulases.

2.2.3 Pectinase

Pectinase catalyzes the random hydrolysis of 1,4-α-D-galactosiduronic linkages in pectin and other galacturonans. The degradation of pectin occurred by pectinolytic enzyme. It can be divided into 3 groups: glycosylhydrolases, rhamnogalacturonan-hydrolase and rhamnogalacturonan lyase. Mode of enzyme actions is shown as below.

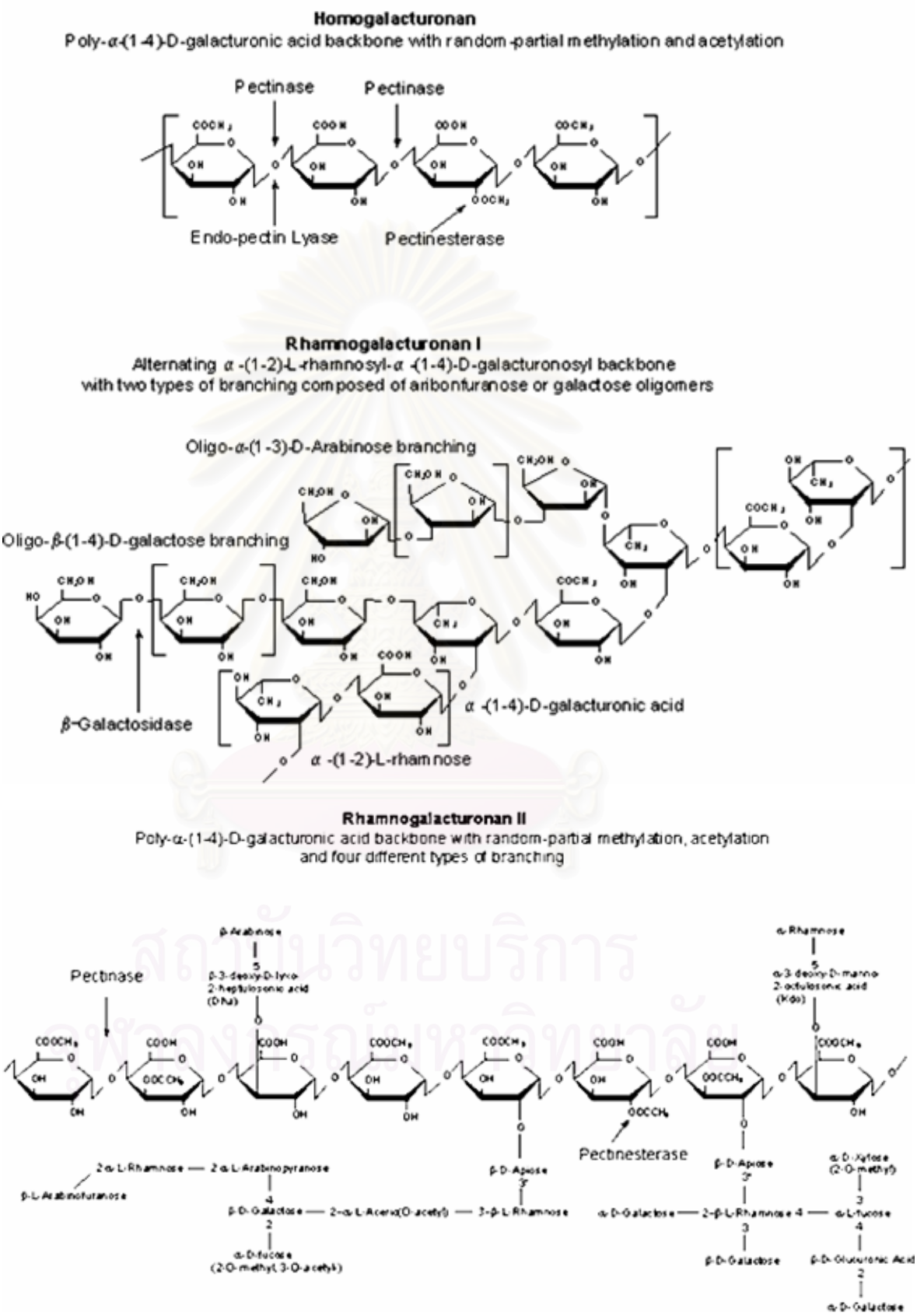


Figure 15. Pectinase and Pectinesterase specificitie (Enzyme Nomenclature, 2007)

2.3 Aerobic and anaerobic cellulolytic enzyme actions

There are two quite different ways that cellulolytic bacteria and fungi deal with the problem raised by the insolubility of cellulose and their inability to ingest cellulose particles. Most anaerobic microorganisms produce multienzyme complexes, called cellulosomes, on their cell surface. In contrast, most aerobic microorganisms secrete a set of individual cellulases into the external media where the enzymes act synergistically to degrade crystalline cellulose. In each case, the products of digestion are oligosaccharides, mostly cellobiose and glucose, are transported into the cell and metabolized. Some aerobic fungi also secrete cellobiase so that glucose is the major end product of cellulose degradation. One possible explanation for the difference in cellulase organization is that anaerobic organisms are more energy limited than aerobic organisms, and thus it is more important for them to retain the products of cellulose digestion. Some anaerobic organisms are tightly bound to cellulose by their surface cellulases so that the products of digestion are released in a confined space between the insoluble cellulose and the organism.

Cellulose utilization generally proceeds via organisms that are either aerobic or anaerobic, but not both. In addition, there is a distinct difference in cellulolytic strategy between the aerobic and anaerobic groups. Anaerobes degrade cellulose primarily via complexed cellulase systems, cellulosome. Most anaerobic species that utilize cellulose do not release measurable amounts of extracellular cellulase. They have their cellulases localized directly on the surface of the cell. Aerobic cellulose degraders utilize cellulose through the production of substantial amounts of extracellular cellulase enzymes that are freely recoverable from culture supernatants (Rapp *et al*., 1991). While many aerobic bacteria adhere to cellulose, physical contact between cells and cellulose does not appear to be necessary for cellulose hydrolysis.

2.3.1 Aerobic cellulolytic system: Free extracellular enzyme

Although a large number of microorganisms are capable of degrading cellulose, only a few of these produce significant quantities of cell free enzymes capable of completely hydrolysing crystalline cellulose *in vitro*. Fungi are the main cellulase producing microorganisms, though a few bacteria and actinomycetes have also been reported to yield cellulase activity. Fungal genera like *Trichoderma* and *Aspergillus* are taught to be cellulase producers and crude enzymes produced by these

microorganisms are commercially available for agricultural use. In general, bacterial cellulases are constitutively produced, whereas fungal cellulase is produced only in the presence of cellulose. Filamentous fungi particularly *Aspergillus* and *Trichoderma* spp. are well known efficient producers of cellulases.

2.3.2 Anaerobic cellulolytic system: Cellulosome

Protein-protein recognition plays a pivotal role in an array of biological processes. One fundamental example is the degradation of the most abundant reservoir of organic carbon in the biosphere, the plant cell wall, by anaerobic organisms. These organisms utilize a high molecular mass (megadalton) cellulose-hemicellulase complex termed the “cellulosome” in which an extensive repertoire of glycoside hydrolases are grafted on a macromolecular scaffold. It is generally believed that assembly of the catalytic components into a complex enhances the synergistic interactions between enzymes with complementary activities, leading to more efficient plant cell wall degradation (Carvalho *et al*., 2003)

Very little is known about cellulosome assembly and what control the exact composition of each individual complex. All of the cellulosomal components are secreted outside the cell and possess typical leader peptides, which are cleaved during the export process. The number of known dockerin-bearing enzymes in *C. thermocellum* is at least double the number of cohesion-dockerin pairs is therefore unlikely. Early observations on the cellulosome indicated that the complex might assume different forms. Cellulosomes isolated at early stages of growth appeared compact, whereas during the later stages of cultivation they take on a more relaxed conformation (Mayer *et al*., 1987, Shoham *et al*., 1999). It is a crucially important for the efficient breakdown and utilization of crystalline cellulose. The cellulosome is a macromolecular machine (multienzyme complex) which, like a ribosome, is dedicated to organized, concerted, synergistic, and efficient catalysis of cellular activities. Cellulosomes are unique in that no other extracellular protein complexes with the size and complexity of the cellulosomes have been reported. They have molecular weight of $2x10^6$ to $6x10^6$dalton, have diameters of about 18 nm, and contain 14 to 50 polypeptides ranging in size from 37 to 210 kDa. Over 95% of the endoglucanase

activity of *C. thermocellum* is associated with the cellulosome. Cellulosome contain 5 to 7% carbohydrate (Demain *et al*., 2005).

Cellulosome and scaffoldin have been found in many bacteria, such as *C. cellulovorans, C. cellulolyticum, C. josui, C. acetobutylicum, Acetovibrio cellulolyticus, Bacteroides cellulosolvens, R. albus, Ruminococcus flavefaciens, Vibrio* sp., and the anaerobic fungal genera *Neocallimastix, Piromyces*, and *Orpinomyces* (Demain *et al*.,2005). Microorganisms belonged to be models of cellulosome studied are *C. thermocellum* and *C. cellulolyticum*. The main difference between the cellulosome of *C. cellulolyticum* and that of *C. thermocellum* is the lack of calcium enhancement of the activities and the binding properties of the former. *C. cellulolyticum* is a mesophillic strain, and it is generally recognized that Ca^{2+} reinforced the thermostability of proteins. The recent study confirms that the cellulolytic system of *C. cellulolyticum* does not require any calcium (Gal *et al*., 1997).

The most important component of the cellulosome is the nonenzymatic scaffoldin. It is unique scaffolding protein subunits. The catalytic subunits, on the other hand, contain different modules (dockerins) which are responsible for their attachment to the scaffold. Important in the relationship are (i) cohesion domains on scaffoldin, (ii) dockerin domains on the enzymes, and (iii) a CBD on the scaffoldin, binding the complex to cellulose.

Cellulosomes are present on the bacterial cell surface and are dedicated to cellulose depolymerisation. For the bacterial cell, the biosynthesis of a cellulosome presents several advantages: (I) a direct and specific adhesion to the substrate of interest permitting efficient competition with other micro-organisms present in the same biota and (II) the proximity of the cell to the cellulose insures an efficient cellular uptake of the soluble cellodextrins by avoiding their diffusion in the extracellular medium (Shoham *et al*., 1999). From an enzymatic point of view, the cellulosome (A) allows optimum concerted activity and synergism of the cellulases, (B) avoids non-productive adsorption of the cellulases, (C) limits competition between cellulases for the sites of adsorption and (D) allows optimal processivity of the cellulase all along the cellulose fibre (Shwartz *et al*., 2001).

Figure 16. Schematic representation of the modular structure of a hypothetical clostridial extracellular enzyme. CBM: carbohydrate-binding module; Doc: dockerin module; Fn3: fibronectin type III module, GH: catalytic module of glycosyl hydrolase family; Ig: immunoglobulin-like module; SLH: surface-layer homology module. Numbers indicate the relative position of the module (Schwarz *et al*., 2004)

2.3.3 Cellulosome structure and assembly

2.3.3.1 Cellulose-Binding Domain/Module (CBD/CBM)

A carbohydrate-binding module (CBM) is defined as contiguous amino acid sequence within a carbohydrate-active enzyme with a discreet fold having carbohydrate-binding ability. 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.

Efficient enzymatic degradation of insoluble polysaccharides, such as cellulose, raw starch, and chitin, often requires a tight interaction between the enzymes and their substrates. Many polysaccharides comprise multiple domains. Typically, bacterial and fungal cellulases consist of at least a cellulose-binding domain and a catalytic domain, often separated by a linker rich in proline and hydroxyamino acids or in glycine (Tomme *et al*., 1995). The enhancement of activity of insoluble cellulose hydrolysis of bacteria such as *C. thermocellum* was found by the presence of

scaffoldin and CBM (Kataewa *et al*., 1997). Different CBMs appeared to target many more sites on crystalline cellulose. Normally, it can be divided into 8 main familys. CBM family I, II, III, IV, V, VI, VII and VIII. CBM three-dimension structures were studied by using NMR methods as described by previous publications (Kraulis *et al*., 1989, Tormo *et al*., 1996).

CBMs are classified into 10 families of related amino acid sequences (Tomme et al., 1995). As the number of CBMs assigned to a family increases, the number of amino acids strictly conserved in all of them decreases. These conserved amino acids are assumed to be important to the structure and function of the CBMs.

Family I (Kraulis *et al*., 1989)

Family II (Xu *et al*., 1995)

Family III (Tormo *et al*., 1996)

Family IV (Johnson *et al*., 1996)

Figure 17. Three-dimension structure of Cellulose-binding module (CBM)

Catalytic domains and CBMs are grouped into different families based on sequence similarities. At present, there are 12 cellulase and 8 different CBM families (Tomme *et al*., 1995).

2.3.3.1.1 Family I cellulose-binding module

This CBM family belonged to fungal cellulose-binding domain. It's properties action like cellobiohydrolase I and found in fungus *Tricoderma reesei*.

2.3.3.1.1.2 Family II cellulose-binding module

This CBM family belonged to bacterial cellulose-binding domain. It's properties action like exoglucanase/xylanase and found in fungus *Cellulomonas fimi*.

2.3.3.1.3 Family III cellulose-binding module

This CBM family is responsible directly to crystalline cellulose. Generally found in *Clostridium* sp. The crystal structure of a family–III cellulose-binding module from the cellulosomal scaffoldin unit of *C. thermocellum* has been determined (Tormo et al., 1996).

2.3.3.1.4 Family IV cellulose-binding module

This CBM family was associated with cellulase family 9. It was found in *Mixococcus xanthus*.

2.3.3.1.5 Other CBM

The other CBM components found less than family I, II or III. CBM of CelZ *Erwinia chrysanthemi* is one of the samples. This CBM belonged to family V. CBM of CelE *C. thermocellum* and CelA of *D. discoideum* also belonged to family VII and VIII respectively.

CBM plays an important role on cellulose degradation. It contains around 30-200 amino acid which contained C- and N- terminal. Its action can be appeared by two forms both reversible and irreversible on cellulolytic substrates.

Table 2. Properties of cellulose binding module family I-IV (Shoseyov and Warren, 1997)

	CBM FamilyI	CBM FamilyII	CBM FamilyIII	CBM FamilyIV
Average length (aa)	36	105	150	150
No. structure known	1	2	1	1
No. parallel β sheets	3	9, 9	12	10
shape	Wedge	Elongated barrels	Barrels	5-stranded clef
Active site	2 TYR, 1GLU; 2 face	3 TRP; 1 face	Planar aa's; shadow groove	Binds to amorphous cellulose, not to crystalline cellulose
Binding strength	Reversible	"irreversible"*	"irreversible"	
Binding characteristics	Bind to a surface, not to soluble cellulose	Binds to a surface, not to soluble cellulose	Binds to a surface, not to soluble cellulose	Binds to a single molecule only, to soluble cellulose derive. And cello-oligosacc.

* One structure is mobile on the surface of crystalline cellulose

2.3.3.2 Cohesin and dockerins

Genetical and biochemical data revealed, that in all cellulosomes investigated so far the components of the multienzyme complex are strongly bound to each other

by a duplicated, non-catalytic segment of 22 amino acid residues. It found to be conserved in all enzymes which are located in the cellulosome (Tokatlidis *et al*. 1991). This dockerin module binds specifically to the cohesin modules, located in a non-catalytic cellulosome component, for which the term "scaffoldin" was coined (cellulosome structure).

Two types of cohesion exist: type I cohesion bind specifically to type I dockerin domains on the catalytic subunits, and type II cohesions are on some cell surface proteins which bind the dockerin of scaffoldin to the cell wall. The interaction between cohesions and dockerins was found to be generally species specific: experiments carried out with isolated modules from the two species revealed that cohesions from the scaffoldin of one species binds to the dockerins of its own enzymatic subunits with high affinity, but fail to recognize those of the other species despite the relatively high sequence homology amoung the analogous components (Handelsman *et al*., 2004).

2.3.3.2.1 Dockerin

Dockerin domains have responsibility about interaction between catalytic domain and scaffolding protein. Dockerin domains contained 2 segments of 22-amino acid which can be separated to 8-10 residues (Tokatlidis *et al*., 1991, Pages *et al*., 1997). First 12 amino acid sequences segments belonged to EF-hand motif linked to calcium. This can be referred that dockerin-cohesin interactions are depended on calcium presence (Fierobe *et al*., 1999). Dockerin domain structures are revealed by NMR. Dockerin-cohesin complex are examined by crystallography and X-ray diffraction (Cavalho *et al*., 2003). The various cellulosome components are assembled by virtue of a high-affinity protein–protein interaction between reciprocal modules on the interacting subunits – the cohesin and the dockerin. In early studies on the cellulosomes from two clostridial species, *Clostridium thermocellum* and *C. cellulolyticum*, the interaction between cohesins and dockerins was found to be generally species specific: experiments carried out with isolated modules from the two species revealed that cohesins from the scaffoldin of one species bind to the dockerins of its own enzymatic subunits with high affinity, but fail to recognize those of the

other species despite the relatively high sequence homology among the analogous components.

Crystal structures of cohesins from the scaffoldin of *C. thermocellum* (Shimon *et al*., 1997) and *C. cellulolyticum* (Spinelli *et al*., 2000) have been published. The cohesins form a nine-stranded ß-sandwich with a jelly-roll topology. The ß-sandwich results from the association of a four-stranded antiparallel ß-sheet and a five-stranded mixed ß-sheet, stabilized by a hydrophobic core. The two β-sheets are composed of strands 8, 3, 6, 5 and strands 9, 1, 2, 7, 4, respectively. In addition, a solution structure of a dockerin from *C. thermocellum* cellulosomal cellobiohydrolase CelS has been solved by NMR analysis. The structure consists of two Ca^{2+} -binding loop-helix motifs that bear sequence homology to the EF-hand motif of eukaryotic calcium-binding proteins, such as calmodulin and troponin C. Very recently, a crystal structure of a cohesin–dockerin complex from *C. thermocellum* has also been solved (Cavalho *et al*., 2003). The complex shows that, while the cohesin module remains essentially unchanged, the dockerin undergoes conformational adjustments upon binding. The protein–protein contact between one face of the cohesin and a-helices 1 and 3 of the dockerin is mediated mainly by hydrophobic interactions and relatively few intermolecular hydrogen bonds. Although the structure of the heterodimer sheds additional light on the structural basis of the cohesin–dockerin interface, the function and importance of specific amino acids involved in recognition and binding are not entirely apparent from the structural data.

2.3.3.2.2 Cohesin domain

This module contained 150 amino acids sequence in the structure. Cohesin domains are separated by the other part of Proline-riches linkers. This linker plays the role on the flexibility of complex. In *A. cellulolyticus*, this linker varies from 3-37 residues in ScaA scaffoldin.

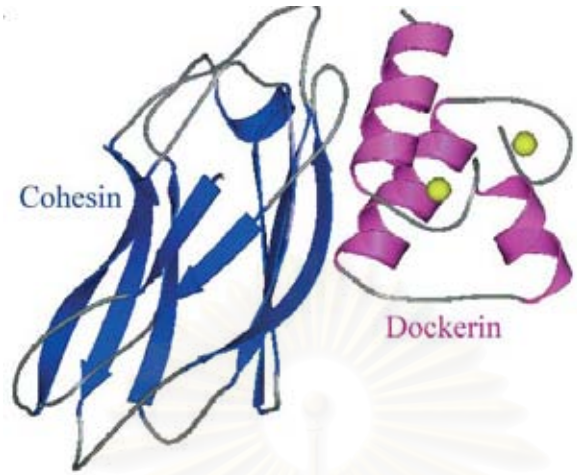


Figure 18. Contribution of cohesin residue Asp34 in the binding of the dockerin domain. Figure shows overview of the 3D structure of the cohesin–dockerin complex from *C. thermocellum* (PDB code 1OHZ). The cohesin module is rendered in blue and the dockerin in pink. Calcium ions are represented as yellow spheres (Handelsman *et al*., 2004)

2.3.3.3 Catalytic domains

Clostridia sp. cellulosome contains a variety of catalytic module. The catalytic components themselves are complex proteins consisting of catalytic and non-catalytic modules. Binding of the cellulosome to the crystalline substrate is mainly mediated by a very strongly binding CBM IIIa module of the scaffoldin. The production of the multienzyme-complex "cellulosome" may have a number of advantages for the effective hydrolysis of cellulose: a) synergism is optimized by the correct ratio between the components, which is determined by the composition of the complex; b) non-productive adsorption is avoided by the optimal spacing of components working together in synergistic fashion; c) the 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; d) a stop of hydrolysis on depletion of one structural type of cellulose at the site of adsorption is avoided by the presence of other enzymes with different specificity.

All the cellulases cloned and characterised so far hydrolysed the ß-1,4-glycosidic bonds of the cellulose by general acid catalysis (Davies *et al*., 1995). Two amino acids with carboxyl group in the side chain are always involved in the catalysis. Hydrolysis of the cellulose led either to the inversion or the retention of the anomeric carbon configuration (Schülein *et al*., 2000). These various cellulases have been classified into different glycoside hydrolases (GH) families continuously updated on the carbohydrate-active enzymes web server (CAZy): http://afmb.cnrs-mrs.fr/~pedro/CAZY/db.html. Some of those families have been regrouped under the same clan since their members share a common fold and a conserved position of their catalytic residues respective to the global fold (Bayer et. al., 1998). Three families, i.e. 6, 9 and 45, and four clans, i.e. GH-A (regrouping the families 1, 5, 10, 17, 26 and 39), GH-B (regrouping the families 7 and 16), and GH-C (regrouping the families 11 and 12) and GH-M (regrouping the families 8 and 48) have been described for cellulases (CAZy). Contrary to the IUBMB classification originally used (Felix *et al*., 1993), this classification does not define the cellulases as endoglucanase (EC 3.2.1.4) or exoglucanase (glucan 1,4-ß-glucosidase, EC3.2.1.74 and cellobiohydrolase, EC3.2.1.91) (Coutinho *et al*., 1991). In fact, this distinction is not absolute; these two catalytic activities exist in all cellulases but one mode of attack is generally predominant. Thus, the description of cellulases according to their structure and catalytic mechanism appears more definitive. Still, crystalline cellulose is generally degraded much more efficiently by the cellobiohydrolases than by endoglucanases.

2.3.3.3.1 Catalytic mechanism

In most cases, the hydrolysis of the glycosidic bond is performed by two catalytic residues of the enzyme: a general acid (proton donor) and a nucleophile/base. Depending on the spatial position of these catalytic residues, hydrolysis occurs via overall retention or overall inversion of the anomeric configuration.

2.3.3.3.2 Catalytic domain families

Family 5

This family is the largest cellulase family containing 57 genes. Of these 52 code for endocellulases and they are retaining enzymes. The other five genes code for β1–3 exoglucanases.Most of the family 5 genes are from bacteria, but some are from fungi. The three-dimensional structures of four family 5 catalytic domains have been

reported and they have an *α*/*β*-barrel fold, which is the most common found among all proteins. Many of the family 5 genes do not code for a cellulose-binding domain. Those that also lack a dockerin domain may not function in the degradation of crystalline cellulose. This has been shown to be true for the carboxymethylcellulase (CMCase) from the anaerobic bacterium *Prevotella bryantii*, where the gene appears to be required for growth on *β*-glucan, a glucose polymer with alternating 1–4, 1–3 linkages.

Family 6

It contains nine genes coding for both endo- and exocellulases from bacteria and fungi. The structures of two family 6 enzymes, *Trichoderma reesei* CBHII – a fungal exocellulase and *Thermomonospora fusca* E2 – an actinomycete endocellulase, have been determined and they are modified *α*/*β* barrels that close the barrel in a slightly different way than it is closed in standard *α*/*β* barrel proteins.When the three-dimensional structures of CBHII and E2 are overlaid there are two loops which cover the active site cleft to form an active site tunnel in CBHII. The E2 active site cleft is much more open because one loop is much shorter and the other has a different conformation. The enzymes in this family all catalyze hydrolysis with inversion of the anomeric carbon configuration.

Family 7

It contains only fungal genes, coding for both endo- and exocellulases. The enzymes in this family all utilize the retaining mechanism. The structure of the *T. reesei* CBHI catalytic domain has been determined and it is a unique structure with a *ß* sandwich forming the active site and many loops connecting the *b* strands. This structure is larger than those of the family 5 and 6 catalytic domains and has a long active site tunnel with enough room for seven glucosyl residues. Three more family 7 structures have been solved: *Fusarium oxysporum* endoglucanase I with a nonhydrolyzable thiooligosaccharide substrate analogue, *Humicola insolens* endoglucanase I, and *T. reesei* endoglucanaseI.

Family 8

This enzyme contains nine genes, all bacterial, which appear to code for endocellulases utilizing the inverting mechanism. The structure of a family 8

endocellulase, *C. thermocellum* Cel A, has been determined and it is an (*a*/*a*) 6 barrel similar to those found in families 9 and 48.

Family 9

The enzyme contains 19 cellulase genes that belong to two subfamilies distinguished by the presence or absence of a family III CBD closely attached to the catalytic domain. All of the enzymes in this family are inverting. The three-dimensional structures of the catalytic domain of *C. thermocellum* CelD and the catalytic domain plus the family III CBD of *T. fusca* E4 have been determined and the catalytic domains are $(a/a)_6$ barrels. The E4 family III CBD was aligned with the catalytic cleft so that a cellulose molecule bound in the active site could also be bound to the CBD.The enzymes without the attached CBD are all endocellulases, while E4, with the attached CBD, is a processive endoglucanase. This family does not contain any fungal genes, but includes genes from both bacteria and plants.

Family 12

This family contains nine genes, all coding for retaining endoglucanases. They include bacterial and fungal genes. The structure of an endocellulase has been determined and it is a jelly roll made up of *ß*-sheets very similar to the structure of a family 11 xylanase. Families 44, 60 and 61 are small cellulase containing families with only a few members thus far. Family 44 includes an inverting endoglucanase and a mannanase from bacteria. No structures have been determined for these families.

Family 48

It contains six cellulase genes and they all code for inverting enzymes. Some of the enzymes studied so far appear to have low specific activities on cellulose substrates and several are present in their respective organisms in relatively large amounts suggesting they are exocellulases. These genes are present in both anaerobic and aerobic bacteria as well as anaerobic fungi. The three-dimensional structure of one family 48 catalytic domain, *C. cellulolyticum* CelF, has been reported and it is an $(a/a)_6$ barrel similar to that found for family 8 and family 9 cellulases. As expected for an exocellulase part of the active site is in a tunnel.

Table 3. Cellulases catalytic domain characteristics of *C. cellulolyticum*

MW (Molecular weight); Inclusion bd. (Inclusion body); GH (Glycoside hydrolase)

	Cel5A	**Cel8C**	**Cel48F**	**Cel9G**	**Cel9E**	**Cel9M**
Domains	 GH5 -do	 GH8-do	 GH48-do	 GH9-CBM3c-do	 CBM4-Ig-GH9-do	 GH9-do
References	Fierobe 1991 Ducros 1995	Fierobe 1993	Reverbel-Leroy 1996 Parsiegla 1998, 2000	Gal 1997 Mandelman 2003	Gaudin 2000	Bélaich 2002 Parsiegla 2002
MW(kDa)	51	47	77	76	94	55
Conformation	α_6/β_8	α_6/β_8	$(\alpha/\alpha)_6$	$(\alpha/\alpha)_6$	$(\alpha/\alpha)_6$	$(\alpha/\alpha)_6$
Structure						
Inclusion bd.	No	No	Yes	Yes	Yes	No
Action mode	**Endo**	**Endo**	**Endo-processive**	**Endo**	**Endo-processive**	**Endo**
Protein activities (U/μM)						
Avicel	5.4	0.8	13.4	5	5.8	2.5
PASC	100	336	42.5	38	71.5	57
BMCC	ND	ND	3.9	9	5.5	2
Xylan	280	6.2	ND	0	2.3	ND
Lichennan	2090	528	ND	140	5.8	ND
pNP-cellobiose	1.5	0	ND	0	34.8	0

2.4. The Clostridial cellulases

2.4.1 The *Clostridium thermocellum* cellulosome

C. thermocellum is a moderately thermophillic bacterium (55-65ºC), repeatedly isolated from hot springs and wet, rooting biomass. It is highly specialized for growth on cellulose and cellodextrins as carbon and energy sources. Cellobiose and soluble cellodextrins are taken up and hydrolysed intracellularity by cellobiose- and cellodextrin-phosphorylase. The extracellular cellulolytic complex of *C. thermocellum* differs between the species somewhat in size (from 2 to 6.5 MDa) and composition. In some strains the cellulosomes aggregate to larger supercomplexes, called polycellulosomes, with a molecular mass up to 100 MDa.

The cellulosome of the anaerobic bacterium *C. thermocellum* has been extensively studied. In this complex, the enzymes are bound to a noncatalytic protein termed “scaffolding” (CipA), CipA contains nine reiterated sequences referred to as “cohesion” domains that interact with comprises a 23-residue tandemly repeated sequence. Cellulosome assembly is mediated by the interaction of dockerin domains of each enzyme with one of the complemetary cohesion domains of CipA. In *C. thermocellum*, the nine cohesion domains of CipA are unable to discriminate between the individual dockerin present in the various enzymes, thus any individual cellulosome complex may comprise a different ensemble of catalytic subunits appended to CipA (Carvalho *et al*., 2003).

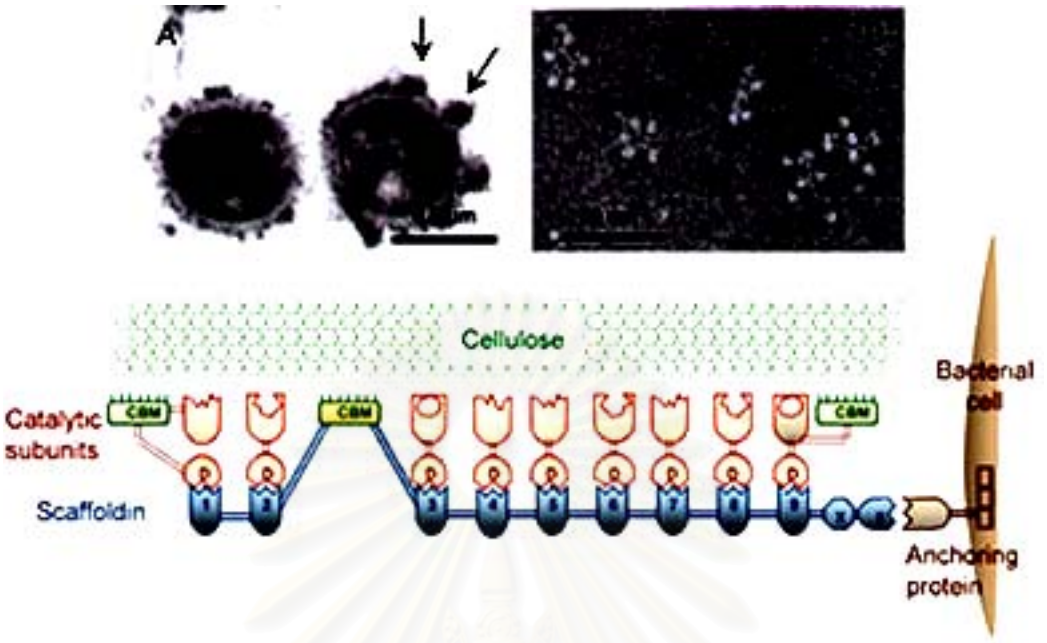


Figure 19. Cellulosomes transmission electron micrographs of *C. thermocellum* cells showing cellulosome aggregate on cell surfaces (indicated by arrows; A), and, at higher magnification, individual cellulosomes of *C. papyrosolvens* (B). Also shown a model of *C. thermocellum* cellulosome and an associated cell-surface anchorin (Leschine, 2004)

Unlike fungal cellulases, the *C. thermocellum* cellulases complex has very high activity on crystalline cellulose; i. e., it has “true cellulose activity” (also called Avicelase) which is characterized by its ability to completely solubilize crystalline forms of cellulose such as cotton and Avicel. This unique cellulose system has been studied by a number of groups biochemically, immunologically, and via molecular biological techniques. The complex is comprised of: (i) numerous endo-β-glucanases which are responsible for the random breakdown of amorphous type of cellulose, including CMC and TNP-CMC; (ii) at least four exoglucanases; (iii) a cellobiose phosphorylase that breaks down cellobiose to glucose and glucose-1-phosphate; (iv) a cellodextrin phosphorylase that phosphorylyzed β-1,4-oligoglucans; and (v) two β-glucosidases that hydrolyze cellobiose to glucose. *C. thermocellum* also processes at least six xylanases, two lichenases, two laminarases, and minor activities of β-

xylosidase, β-galactosidase, and β-mannosidse. A strain of *C. thermocellum* has been shown to degrade pectin and probably produce pectin lyase, polygalacturonate hydrolase, and pectin methylesterase.The activity responsible for hydrolysis of crystalline cellulose, unlike the endoglucanases, to be inhibited by cellobiose, glucose had no such inhibitory effect (Demain *et al*., 2005).

Hypothetical reading frames and cloned genes in the unfinished genome sequence. Only reading frames containing dockerin modules, a Shine-Dalgarno sequence and stop codon, and a recognizable module composition are listed. The protein (gene) designation and enzymatic activity or function (if known), and the ORF number from *http://genome.ornl.gov/microbial/cthe/* (Cthe) are given (as of October 2003). If no Cthe number is indicated, the gene was cloned but is missing in the genomic sequence. The components are sorted according to their GH family or their putative function, as obvious from the catalytic module family. Components with more than one catalytic module or unknown modules are listed at the end.

Table 4. Cellulosomal-encoding gene and structure of *C. thermocellum* cellulosome

	Gene	**Reading frame/function**	**Structure****	**Refernce**
1.	*cipA* +	scaffoldin, Cthe1933-1930	2(Coh1)-CBM3a-7(Coh1)-X2-Doc2	Fujino *et al*., 1992, Zverlov et al., 1999
	GH2			
2.	*cthe1580*	GH2-CBM6-Doc1		
	GH5			
3.	*celO*	cellobiohydrolase, Cthe1674	CBM3b-GH5-Doc1	Zverlov, in preparation
4.	*cthe1575*	GH5-CBM6-Fn3–Doc1		
5.	*celB*	endoglucanase, Cthe0374	GH5-Doc1	Grépinet *et al*., 1986
6.	*celG* +	endoglucanase, Cthe0885	GH5-Doc1 27	Lemaire *et al*., 1992
7.		Cthe0444	GH5-Doc1	
	GH8			
8.	*celA* +	endoglucanase, Cthe0722	GH8-Doc1	Béguin *et al*., 1985
	GH9			

9.	*cbhA* +	cellobiohydrolase	CBM4-Ig-GH9-2(Fn3)-CBD3b-Doc1	Zverlov *et al*., 1998
10.	*celK* +	cellobiohydrolase, Cthe2598	CBM4-Ig-GH9-Doc1	Zverlov *et al*., 1998
11.	*celD*	endoglucanase, Cthe0968	Ig-GH9-Doc1	Joliff *et al*., 1986
12.		Cthe1953	GH9-CBM3c-CBM3b-Doc1	
13.		Cthe0850	GH9-CBM3c-CBM3b- Doc1	
14.	*celN* +	endoglucanase, Cthe1222	GH9-CBM3c-Doc1	Zverlov, in preparation
15.	*celR* +	endoglucanase, Cthe1837	GH9-CBM3c–Doc1	Zverlov, in preparation
16.	*celQ* +	endoglucanase, Cthe0300	GH9-CBM3c-Doc1	Arai *et al*., 2001
17.	*celF*	endoglucanase, Cthe0382	GH9-CBM3c-Doc1	Navarro *et al*., 1991
18.		Cthe1308	GH9-CBM3c–Doc1	
19.		Cthe0727	GH9-Doc1	
20.	*celT* +	endoglucanase	GH9-Doc1	Kurokawa *et al*., 2002
	Xylanases			
21.	*xynD* +	xylanase, Cthe0688	CBM22-GH10–Doc1	Zverlov, in preparation
22.	*xynC* +	xylanase, Cthe0626	CBM22-GH10-Doc1	Hayashi *et al*., 1997
23.	*xynA*, *xynU* +	xylanase, Cthe1161	GH11-CBM4-Doc1-NodB	Hayashi *et al*., 1999
24.	*xynB*, *xynV* +	xylanase	GH11-CBM4-Doc1	Hayashi *et al*., 1997
	Other hemicellulases			
25.	*licB* +	lichenase	GH16-Doc1	Zverlov *et al*., 1994
26.	*chiA* +	chitinase	GH18-Doc1	Zverlov *et al*., 2002
27.	*manA* +	mannanase, Cthe0533	CBM-GH26-Doc1	Halstead *et al*., 1999
28.		Cthe2142	GH26-Doc1	
29.		Cthe1127	GH30-CBM6-Doc1	
30.		Cthe2333	GH53-Doc1	
31.		Cthe0269	GH81-Doc1	
	Putative glycosidases			

32.		Cthe1665	GH39-2(CBM6)-Doc1	
33.		Cthe1579	GH43-CBM6-Doc1	
34.		Cthe0268	GH43-CBM13-Doc1	
35.		Cthe0484	GH43-2(CBM6)-Doc1	
	GH48			
36.	*celS* +	exoglucanase, Cthe0939	GH48-Doc1	Wang *et al*., 1993
	Xyloglucanhydrolase			
37.	*xghA* +	xyloglucanase, Cthe2335	GH74-CBM2-Doc1	Zverlov, in preparation
	Putative carbohydrate esterases			
38.		Cthe0066	Fn3-CE12-Doc1-CBM6-CE	
39.		Cthe1577	CE1-CBM6-Doc1	
	Putative pectinases			
40.		Cthe2008	GH28-Doc1	
41.		Cthe2236	PL1-Doc1-CBM6	
42.		Cthe1810	<-Doc1-CBM6-PL9	
43.		Cthe2234	PL10-UN-Doc1	
44.		Cthe0702	Doc1-CBM6-PL11	
	Multifunctional components			
45.	*celJ* +	cellulase, Cthe0301	X-Ig-GH9-GH44-Doc1-X	Ahsan *et al*., 1996
46.	*celH*	endoglucanase, Cthe0837	GH26-GH5-CBD9-Doc1	Yaguee *et al*., 1990
47.		Cthe1667	GH30-GH54-GH43-Doc1	
48.		Cthe1211	GH54-Doc1-GH43	
49.		Cthe1666	GH54-GH43-Doc1	
50.	*xynZ* +	xylanase, Cthe1691	CE1-CBM6-Doc1-GH10	Guglielmi *et al*., 1998

51.	*xynY*	xylanase, Cthe2036	CBM22-GH10-CBM22-Doc1-CE1	Fontes *et al*., 1995
52.	*celE* +	endoglucanase, Cthe0940+, Cthe2702, Cthe2514	GH5-Doc1-CE2	Hazelwoodd *et al*., 1990
	Putative protease inhibitors			
53.		Cthe1412	Fn3-Doc1-serpin	
54.		Cthe1413	DOC1-SERPIN	
	Components with unknown function			
55.		Cthe0694	2(UN)-UN-UN(CelP 550-870)-Doc1	
56.		Cthe1578	UN-CBM6-Doc1	
57.	*cseP* +	Cthe1223	UN-Doc1	Zverlov *et al*., 2003
58.		Cthe1474	Doc1-UN	
59.		Cthe0287	UN1-UN2-Doc1	
60.		Cthe0416	Doc1-UN	
61.		Cthe0073	UN-Doc1	
62.		Cthe0649	UN-Doc1	

*A "+" in the reference column indicates, that the component was shown to be present in the cellulosome.

** Module classification according to (7), URL: http://afmb.cnrs-mrs.fr/CAZY/: Coh, cohesin module; Doc, dockerin module; CBM, carbohydrate binding module; X, hydrophobic module; GH, glycosyl hydrolase family; Fn3, fibronectin III module; Ig, immunoglobulin like fold; NodB, acetylxylan-esterase NodB type; CE, carbohydrate esterase; PL, pectin lyase; UN, unknown module; serpin, serine-protease inhibitor homologue.

Filter paper degradation was effected by *C. thermocellum* strain DSM 1237 (the type strain) and F7 in anaerobic flask (Zverlov, unpublished data). Normally, none of aerobic cultures did show efficient cellulose degradation at the high

temperatures, although the operator of one of the composting plants has detected temperatures up to 70 ºC in the aerated compost from which samples were taken. Consequently, the anaerobic bacterium *C. thermocellum* (and clos relatives) is an ubiquitous and easily isolated anaerobic thermophilic bacterium which seems to play a mojor role in anaerobic cellulose degradation in nature.

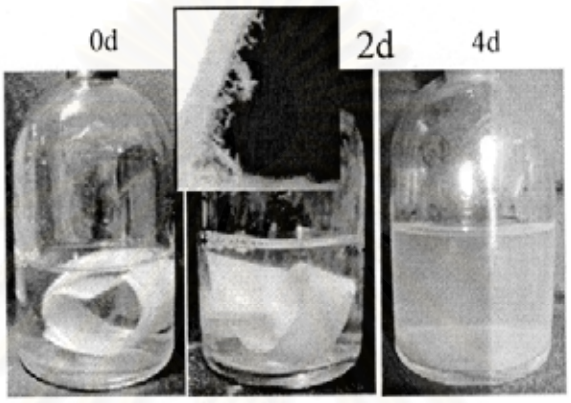


Figure 20. Digestion of Whatman filter paper No. 1 in anaerobic cultures and incubated 0, 2, 4 days at 60 °C (Zverlov *et al*., unpublished data)

2.4.2 Cel48A protein of *C. thermocellum* (CelS)

Tali and collaborator (2003) reported biochemical results confirm earlier reports that the family 48 cellobiohydrolase subunit, CelS, is the major component of the *C. thermocellum* cellulosome (Bayer *et al*., 1983**,** Lamed *et al*., 1983**,** Lamed *et al*., 1985**,** Morag *et al*., 1993**,** Morag *et al*., 1991**,** Tali *et al*., 2003**)**. It also support a previous study **(**Bayer *et al*., 1985**)**, in which it was observed that growth of cells on cellobiose led to a dramatic decrease in the content of the CelS subunit and concomitant increases in other cellulosomal components. These components were determined in this work to represent XynC, CelG, and/or CelB. On the basis of these results, it would appear that the composition of cellulosomal components is regulated by the growth conditions. However, for numerous reasons, studying gene regulation of cellulosomal subunits at the protein level is problematic. For example, the various cellulosomal subunits exhibit similar levels and types of activities on similar substrates. In many cases, the molecular sizes are very similar and the protein subunits often migrate similarly on SDS-polyacrylamide gels. Many of the subunits share

conserved regions in their primary structures (notably in their dockerin sequences) and thus exhibit a high degree of immunogenic cross-reactivity

This family 48 glycoside hydrolase is the most abundant cellulosomal enzyme subunit from this and other cellulosome-producing clostridia. CelS exhibits exoglucanase activity, and its activity is strongly inhibited by the presence of cellobiose**.**

2.4.3 The *Clostridium cellulolyticum* cellulosome

Clostridium cellulolyticum, a mesophillic cellulolytic bacterium, has been extensively studied to determine its growth parameters when cultivated on glucose, cellobiose and cellulose. The gene coding for the scaffoldin protein (or CipC, which stands for cellulosome integrating protein C) has been identified and partially sequenced and found to be part of a large cluster composed of at least six genes, two of which (*celC* and *celG*) from an operon. As enzymes involved in cellulose breakdown by *C. cellulolyticum* are firmly embedded into the cellulosome, the now classical strategy of cloning and sequencing the genes and overproducing the proteins in *Escherichia coli*, has been used for resolving the cellulolytic system of this bacterium. The ‘congo red plate method’ allowed characterize 5 endoglucanases (CelA, CelD, CelC, CelG, and CelE). All the deduced endoglucanase sequences contain a characteristic N terminus signal sequence typical of secreted proteins. At their C terminus, they also contain a duplicated sequence of about 22 residues (DS), typical of clostridia cellulases, which play a major role in the cellulosome structuring. Some cellulose (CelD, CelE and CelG) contain, in addition to the catalytic and the reiterated, as a cellulose binding domain on the basis of sequence comparison.

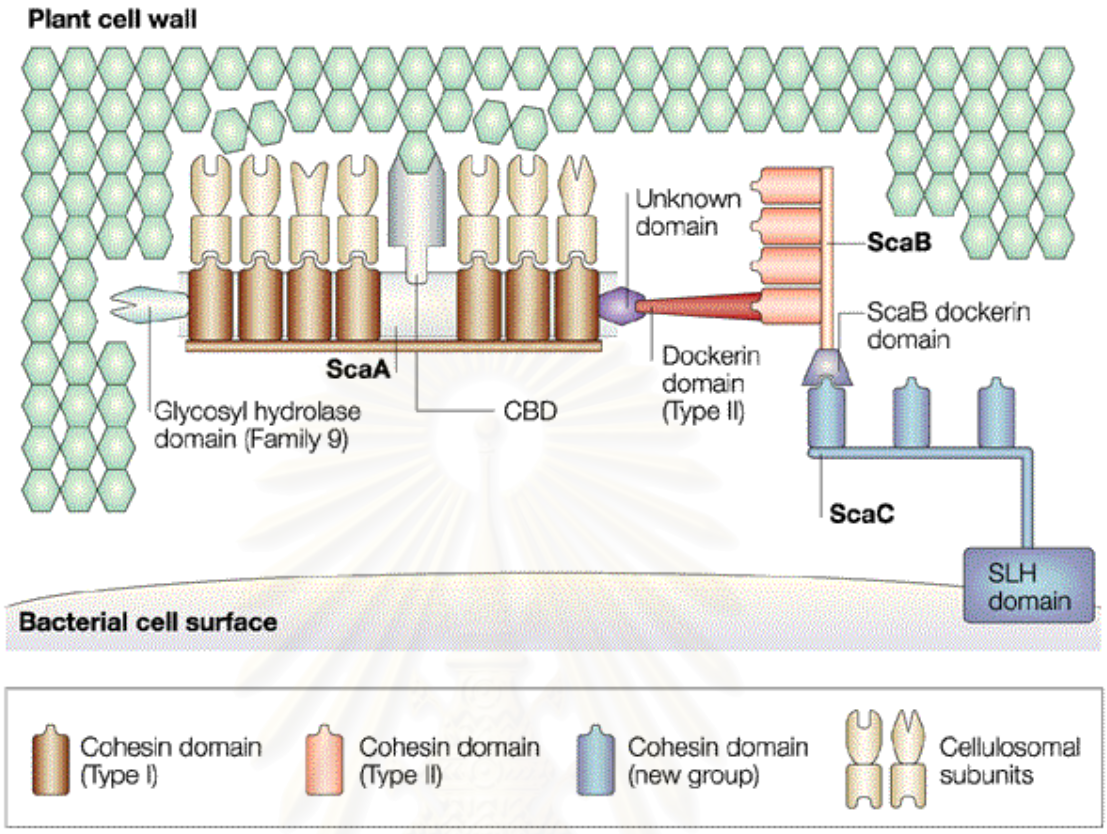


Figure 21. Scematic representation depicting a functional model of a cellulosome and the interaction of its component part with the cellulose substrate (Doi and Kosugi, 2004)

The scaffoldin sub-unit (based on the scaffoldin from *C. cellulolyticum*) and its complement of enzymes is bound to the cellulose component of the plant cell wall by virtual potent family-3a CBM. In the presence of cellulose, the inter-cohesin modules of the scaffoldin undergo large-scale motion to adjust the respective positions of the complexed catalytic sub-units include a CBM that mediates a relatively weak interaction with the substrate. The name of the cellulosomal modules are given in the legend (Hammel *et al*., 2005).

Table 5. Cellulosomal enzymes of *C. cellulolyticum*

Catalytic subunit	Chromosomic location	Signal peptide (a.a.)	Mature protein (a.a.; kDa)	Catalytic activity	Modular structure[a]	Accession number Swissprot/GenBank
Cc-Cel48F	*cel* cluster	29	693; 77.6	Cellobiohydrolase	GH_{48}–Doc_I	**P37698**
Cc-Cel8C	*cel* cluster	32	428; 47.2	Endoglucanase	GH_8–Doc_I	**P37699**
Cc-Cel9G	*cel* cluster	35	690; 76.1	Endoglucanase	GH_9–CBM_{IIIc}–Doc_I	**P37700**
Cc-Cel9E	*cel* cluster	35	857; 93.8	Cellobiohydrolase	CBM_{IV}–Ig–GH_9–Doc_I	**AAA73869**
Cc-Cel9H	*cel* cluster	31	706; 80.0	Putative cellulase	GH_9–CBM_{IIIc}–Doc_I	**AAG45157**
Cc-Cel9J	*cel* cluster	31	722; 81.2	Putative cellulase	GH_9–CBM_{IIIc}–Doc_I	**AAG45158**
Cc-Man5K	*cel* cluster	25	399; 45.1	Putative mannanase	Doc_I–GH_5	**AAG45159**
Cc-Cel9M	*cel* cluster	30	496; 54.5	Endoglucanase	GH_9–Doc_I	**AAG45160**
Cc-Rgl11Y	*cel* cluster	27	650; 70.8	Pectinase	PL_{11}–Doc_I	**AAG45161**
Cc-Cel5N	*cel* cluster	29	505; 56.6	Putative cellulase	GH_5–Doc_I	**AAG45162**
Cc-Cel5A	Out of *cel* cluster	26	449; 50.8	Endoglucanase	GH_5–Doc_I	**P17901**
Cc-Cel5D	Out of *cel* cluster	24	560; 63.4	Endoglucanase	GH_5–CBM_{XI}–Doc_I	**P25472**

[a] GH, glycoside hydrolase; PL, polysaccharide lyase; Doc, dockerin domain; CBM, carbohydrate binding domain; Ig, immunoglobulin-like module. Subscript denotes the family number.
[b] nd, not determined.

2.4.3.1 Cel48F protein of *C. cellulolyticum*

Cc-Cel48F and Cc-Cel9E are the most abundant catalytic components of the *C. cellulolyticum* cellulosome (Gal *et al*., 1997). While GH belonging to family 48 was originally considered as exoglucanases, the work of Reverbel-Leroy *et al*. clearly demonstrated that Cc-Cel48F is a processive cellulase with an endo activity (Reverbel-Leroy *et al*., 1997). Cc-Cel48F is a perfect example of a cellulosomal cellulase possessing the dual activities of endoglucanase and cellobiohydrolase. The active site of Cc-Cel48F is composed of a long tunnel followed by an open cleft (Parsiegla *et al*., 1998). The tunnel part will be initially open like a cleft to fix a single cellulose chain. After closing of the cleft, the chain of cellulose is progressively slid through the tunnel and thus continuously supplies the leaving group site with glucose residues after the catalytic cleavage. The specific arrangement of the aromatic residues present along the tunnel reduces the sliding resistance of the substrate chain. While the initial hydrolysis liberates soluble cellodextrins from cellobiose to cellohexaose with the cellotetraose as the major end-product, longer incubation (more than 1 h) leads mainly to the formation of cellobiose. Cc-Cel48F was also reported to exhibit an overall capacity to act in synergism with other cellulases (Gaudin *et al*., 2000).

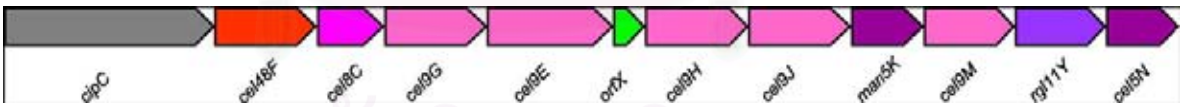


Figure 22. Schema of the *cel* cluster in *C. cellulolyticum* (Desvaux, 2005**)**

Abbreviations:

CBM, carbohydratebinding domain; CbpA, cellulose binding protein A; Cc, *Clostridium cellulolyticum*; Cel, cellulase ; CipA, cellulosome integrating protein A ; CipC, cellulosome integrating protein C ; Coh, cohesin domain; Doc, dockerin domain; EngE, endoglucanase E ; GH, glycoside hydrolase ; HLD, hydrophilic domain ; Ig, immunoglobulin-like module

Man, mannanase; OlpB, outer layer protein B; ORF, open-reading frame ; ORFXp, ORF X polypeptide; ORF2p, ORF 2 polypeptide; Pel, pectate lyase; PL,

polysaccharide Lyase, PTS, proline–threonine–serine; Rlg, rhamnogalacturonase; SdbA, scaffoldin dockerin binding A; SLH, S-layer homology

Interestingly, for *C. cellulovoran*sand *C. acetobutylicum,* which belong to cluster I, genetically analysis have shown that the synteny and sequence homology of the cellulosomal genes are very similar to those of *C. cellulolyticum* (Desvaux, 2005, Bélaïch *et al*., 1997, Lee *et al*., 1985, Nolling *et al*., 2001). Taking into account the phylogenetic distance between these three species, horizontal gene transfer rather than a common bacterial ancestor seemsa most probable hypothesis explaining the close taxonomic relatedness of those cellulosomal genes.

2.4.4 The *C. acetobutylicum* cellulosome

The type strain, *C. acetobutylicum* ATCC824, was isolated in 1924 from garden soil in Connecticut12, and is one of the best-studied solventogenic clostridia. Strain relationships among solventogenic Clostridia have been analyzed, the ATCC 824 strain was shown to be closely related to the historical Weizmann strain. The ATCC824 strain has been characterized from a physiological point of view and used in a variety of molecular biology and metabolic engineering studies in the US and in EU. This strain is known to utilize a broad range of monosaccharides, disaccharides, starches, and other substrates such as inulin, pectin, whey and xylan but not crystalline cellulose, Physical mapping of the genome demonstrated that this strain has a 4MB chromosome with 11 ribosomal operons and harbors a large plasmid, about 200 kb in size, which carries the genes involved in solvent formation, hence the name **pS0Ll3O.** Much work has been done to elucidate the metabolic pathways by which solvents are produced, and to isolate solvent-tolerant, or solventoverproducing strains ***C. acetobutylicum*** ATCC824 and related organisms, and these have been used to develop modified strains with altered solventogenic properties.

Its ability to ferment pentoses has long been recognized. However, its ability to degrade and ferment hemicellulose was not adequately documented until recently. Lee et al. reported that xylanolytic activity was widely distributed among the solvent-producing Clostridium strains. In particular, *C. acetobutylicum* ATCC 824 possessed xylanase, xylosidase, and arabinofuranosidase activities. Under defined conditions in a chemostat, strain ATCC 824 was able to utilize 50% of the oat spelt xylan for growth.

More recently, Lemmel et al. showed that *C. acetobutylicum* ATCC 39236 was also able to grow on xylan.

The paper by Sabathé and Soucaille reports the existing expression in vivo of designer minicellulosomes by *Clostridium acetobutylicum*. This organism, whose genome has been completely sequenced, is very important for the industrial production of solvents. Ironically, *c. acetobutylicum* produced >665 kDa cellulosome that does not degrade crystalline cellulose, although it contains large cellulosome synthesis cluster gene. If the cellulosome genes could be activated to produce a cellulose-degrading cellulosome or if cellulosome genes could be transformed into *C. acetobutylicum* and expressed to form an active cellulosome, then the organism might be able not only to degrade cellulose but also produce solvents as a single organism system (Doi, 2003).

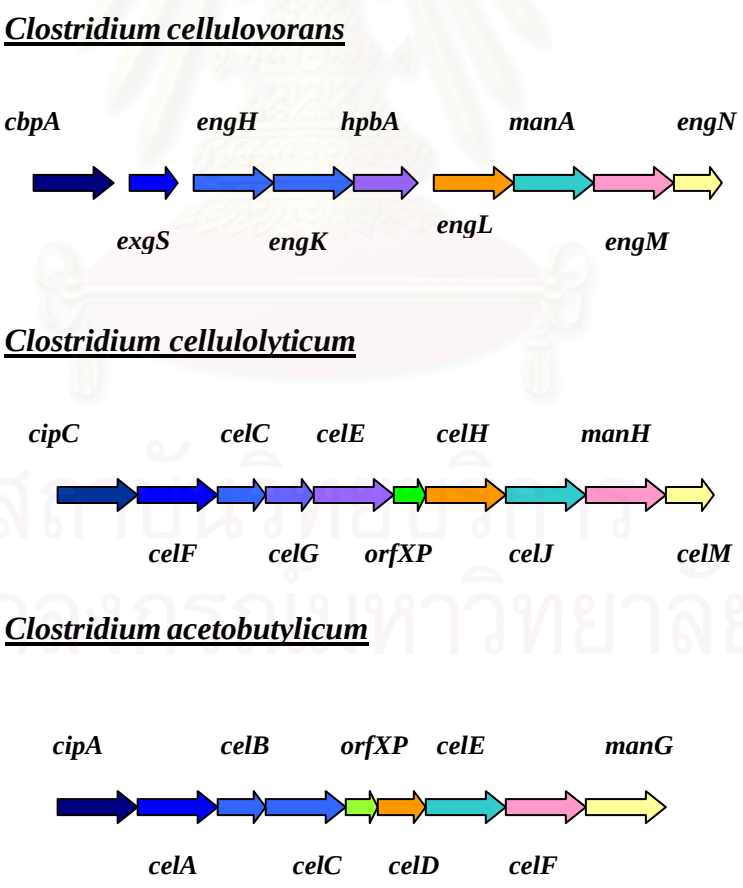
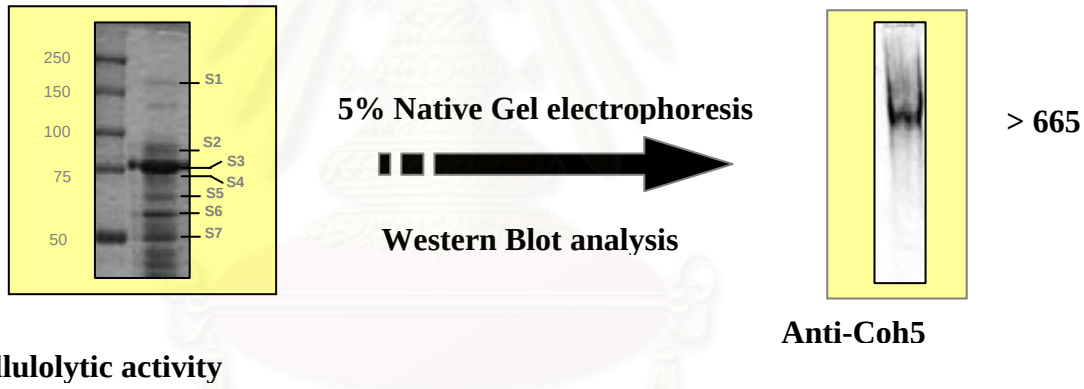


Figure 23. Component of cellulolytic gene cluster of Clostridia

2.4.4.1 Cel48A of *C. acetobutylicum*

Sabathé et al., 2003 reported and demonstrated that *C. acetobutylicum* produced an inactive cellulosome with an apparent molecular mass of 665 kDa. Biochemical and Western analysis revealed that the *C. acetobutylicum* cellulosome comprised four major subunits, including the scaffolding protein CipA and the cellobiohydrolases Cel48A, Cel9X, and Cel9C or Cel9E. However, lack of cellulolytic activity is hypothesized based on protein sequence analysis of main catalytic subunits (Cel48A, Cel9X and Cel9C). Unpublished data from our group shows that there is no mutation point in Cel9X and Cel9C protein sequences. In contratst, Cel48A protein sequence has three different points of mutation occurred in part of linker protein sequences. This may be related to the absence of cellulolytic activity on crystalline cellulose in this organism.



	Specific activity (UI/mg)			
	Avicel	BMCC	CMC	PASC
Culture supernatant	-	-	0.115	0.037
Cellulosome fraction	-	-	0.011	0.0020

Figure 24. Cellulosome fractions and specific activity (IU/mg protein) (Sabathé, unpublished data)

Design protein engineer process for malfunctional Cel48A replacement in *C. acetobutylicum* ATCC824. First, abnormal Cel48A gene will be deleted. After that, chimeric protein molecule will be designed by use gene for express Cel48F catalytic

domain from plasmid pETFc (Reverbel-leroy *et al*., 1996) with modification for correct protein expression. We modify this plasmid by replace the Cel48F dockerin domain with Cel48A dockerin domain. This dockerin domain is appropriate for *C. acetobutylicum* ATCC824 cellulosome assembly. The replacement is very important because the construct of cellulosome in *Clostridium* sp. needs species-specific dockerin domain.

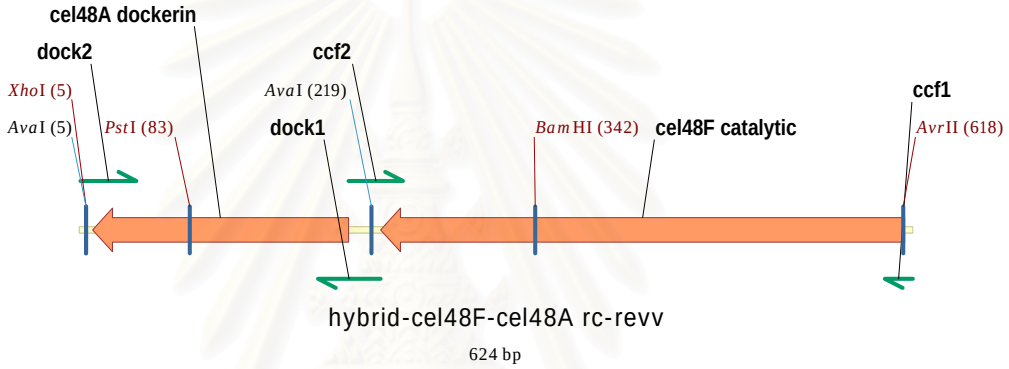


Figure 25. Design of hybrid Cel48F catalytic module fused with Cel48A dockerin module

2.4.4.2 Chimeric cellulosome/ Minicellulosome

Cellulolytic *Clostridia* species, such as *Clostridium cellulolyticum*, produce and secrete large cellulolytic complexes called cellulosomes that efficiently degrade cellulose and related plant cell wall polysaccharides. The catalytic subunits participating with cellulosomes, and therefore appended with a dockerin domain, comprise mainly cellulases but also hemicellulases such as xylanases, a mannanase, and at least one pectinase. As do most species of cellulolytic bacteria, *C. cellulolyticum* grows rather slowly, even on cellobiose, the major product released by cellulosomes on cellulose. Its catabolic pathways are adapted to low carbon flow, which is characteristic of growth on cellulose, and the main products of the metabolism are acetate, lactate, and ethanol.

On the contrary, *Clostridium acetobutylicum* is unable to grow on crystalline cellulose, although its genome contains a large cluster of genes encoding cellulolytic

enzymes and a scaffoldin (Nolling *et al*., 2001, Sabathé *et al*., 2002, Mingardon *et al*., 2005). Interestingly, it has been shown that the bacterium produces small amounts of a 665-kDa cellulosome devoid of activity towards crystalline cellulose and is poorly active on carboxymethyl cellulose or phosphoric acid-swollen cellulose (Sabathé *et al*., 2002, Mingardon *et al*., 2005). On the other hand, *C. acetobutylicum* grows much faster than *C. cellulolyticum*, rapidly consumes cellobiose, and produces substantial amounts of solvents of industrial interest such as butanol (Fischer *et al*., 1993, Mingardon *et al*., 2005).

Fierobe *et al*. (2001) attempt to design and produce active cellulosome chimeras (Figure 26). Defined chimera cellulosomes were produced in which selected enzymes were incorporated in specific locations within a multicomponent complex. The molecular building blocks of this approach are based on complementary protein modules from the cellulosomes of two clostridia, *Clostridium thermocellum* and *Clostridium cellulolyticum*, wherein cellulolytic enzymes are incorporated into the complexes by means of high-affinity species-specific cohesion-dockerin interactions. Several chimeric scaffoldins and hybrid enzymes were designed. The chimeric scaffoldins comprised an optional CBD and two cohesion domains of unlike specificity, one from each clostridial species. Recombinant enzyme constructs contained a catalytic module together in the same polypeptide chain with a dockerin domain from either species. The cellulosome chimeras exhibited enhanced synergy on a microcrystalline cellulose substrate.

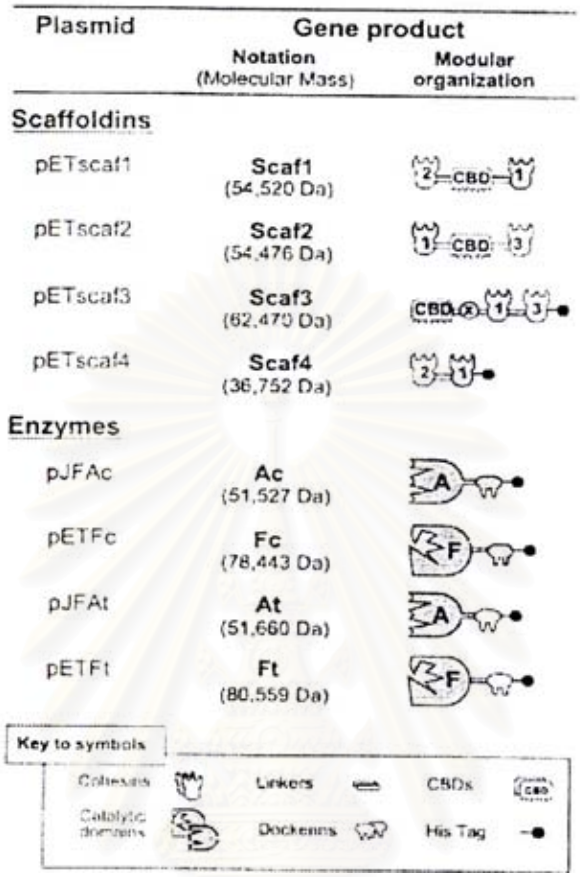


Figure 26. Schematic representation of the recombinant proteins use in Fierobe studied. *White* (*C. thermocellum*) and *grey* (*C. cellulolyticum*) *symbols* denoted the source of the respective domains (see "*Key to symbols*") from Fierobe *et al*. (2001)

The cohesion domains are numbered according to their original position in the respective native cellulosome scaffoldin. A hydrophilic domain (*X*) of unknown function is part of the *C. cellulolyticum* scaffoldin. In the shorthand notation for the enzymes, *A* and *F* represent the catalytic domins from *C. cellulolyticum* cellulosome family-5A and family-48 CelF, respectively; *c* and *t* refer to the dockerin domains derived from *C. cellulolyticum* or *C. thermocellum*, respectively.

The engineered protein were produced in *E. coli* and affinity-purified in one step on either cellulose or nickel-nitrilotriacetic acid according to the presence of a CBD or His tag, respectively. The chimeric scaffoldin s were found to be very stable upon storage for several days at 4ºC, whereas low levels of spontaneous cleavage

between the catalytic and the dock modules were detectable for both wild-type and hybrid enzymes.

Figure 27. Schematic representation of chimeric cellulosome (Fierobe *et al*., 2001)

A chimeric scaffoldin is produced containing an optional carbohydrate-binding module (CBM) and multiple (n) cohesion modules of different dockerin specificities. The dockerin counterparts comprise distinct modules as part of the polypeptide chains of the desired protein component (eg. enzyme A, B, C and N). The chimeric cellulosome complex is constructed by simply mixing in solution the chimeric scaffoldin and dockerin containing components. The resultant complex exhibit enhanced synergistic functions due to the close proximity of the interacting components.

The data indicate that the activity levels of the chimeric cellulosomes were significantly higher than those of the combined free enzyme systems, thereby demonstrating that proximity of the different enzymes within the complex indeed appears critical to the observed enhancement of synergistic action. The incorporation of the enzymes into defined chimeric cellulosomes provided a further enhancement for of 2-3 folds.

2.5 Application in bioprocess

2.5.1 Regulation of solvent production in *Clostridium acetobutylicum*

Clostridium acetobutylicum, a strictly anaerobic spore-forming bacterium, usually shows a biphasic batch fermentation pattern. After producing acetate and butyrate during exponential growth, the organism switches to the formation of acetone, butanol, and ethanol shortly before entering the stationary phase. The mechanisms responsible for the onset of solventogenesis are currently the focus of much scientific research. In batch cultures, the initiation and sustained production of solvents are associated with a low extracellular and intracellular pH and a high undissociated butyric acid concentration. In continuous culture, ATP and NADPH availabilities appear to play a key role in product selectivity. High ATP concentration related to low ATP demand or high efficiency of ATP generation would lead to enhanced solvent production (i) for glucose-sufficient cultures at a low pH with biomass recycling; (ii) for iron-, nitrogen-, or phosphate-limited cultures; and (iii) during shifts induced on phosphatelimited cultures by lowering the pH or adding organic acids. Ethanol and butanol productions were associated with increased availability of reducing power (i) when the *in vivo* activity of the hydrogenase was decreased by CO gassing or by adding methyl viologen; (ii) during a shift in solvent production induced by lowering the pH when acetyl coenzyme A (CoA) was first converted to acetone (a pathway consuming no reducing energy), creating a redox imbalance; (iii) when an NADH pressure was provided by culturing the microorganism on glucose and a more reduced substrate (Girbal *et al*., 1994).

In the central pathway, thiolase catalyses the condensation of two acety-CoA molecules to form one acetoacetyl-CoA molecule, the precursor of the fourcarbon solvents. This reaction plays an important role in determining the ratio between the two-carbon (acetate, ethanol) and the three- (acetone) and four- (butyrate and butanol) -carbon products.

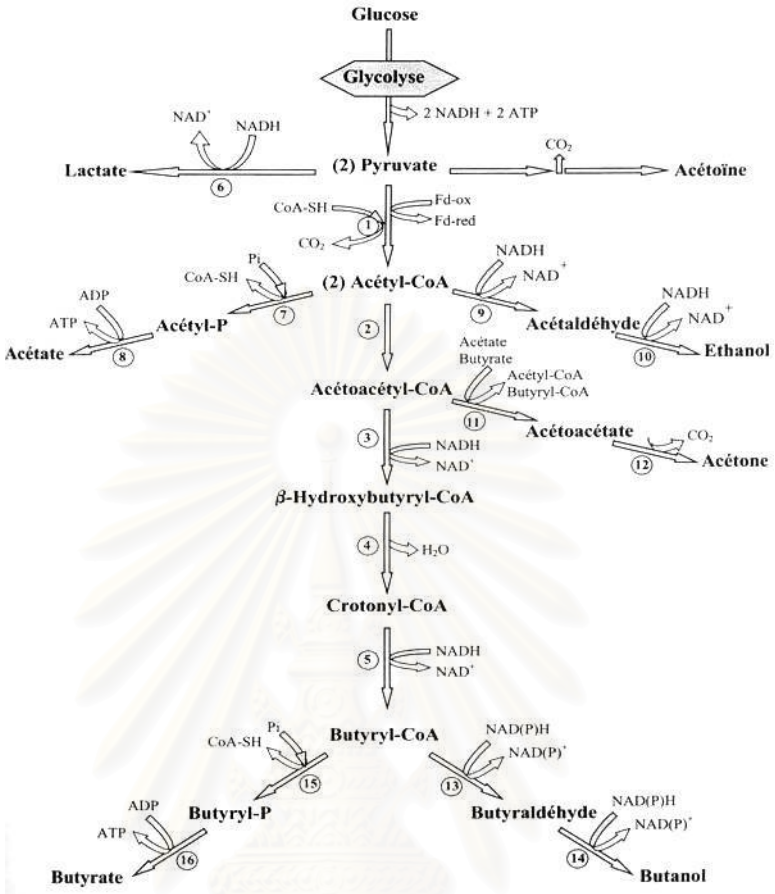


Figure 28. Metabolic pathway of *Clostridium acetobutylicum* ATCC824 (Roger *et al*., 1993)

During exponential growth of *C. acetobutylicum* on a variety of sugars and starch at pH values greater than about 5.6. (depend upond the strain) the major fermentation products are butyrate, acetate, CO_2 and H_2. As acids accumulation occurred in batch culture, growth becomes linear and gradually stops. When, pH slips down to 4.5-4.0, a classical shift in the fermentation occurs and solvents formation is trigged. Solventogenesis are induced by lowering the pH of continuouse culture. Butanol and ethanol, without acetone) were associated with an increase in the NAD(P)H availability inside the cell.

Table 6. Induction of Solventogenesis (Rogers and Gottschalk, 1993)

Conditions	pH		Effects
	Decrease below5.5 during fermentation		Increase of internal concentration of acids
Butyrate Addition	pH 7.0 pH 4.5	100 mM 20 mM	Increase of internal concentration of acids
Acetoacetate Butyrate uncouplers	pH 4.5	8-12 mM 20 mM 5 μM	Induction (1-3 h addition)
CO gassing addition	pH 6.8 pH 5.0	0.1 g/l 0.2 g/l	Inhibition of ethanol production, induction, increase of acetone production
H_2	Increase of partial pressure		Fast induction

Clostridia and cellulosic bioethanol production

Cellulose feedstocks were already in use during the second world-war. The cellulose materials in the form of micro-fibrils are very closely associated with lignin. So, the cellulosic materials are first made free from lignin and then the cellulose is broken down into fermentable glucose prior to fermentation.

The cost of cellulose in the *tricoderma-yeast* process is still prohibitively high, wherease for the direct clostridia coculture process, the enzyme cost very little because they are produced by the fermenting organism in the course of ethanol production. The direct fermentation of cellulose to ethanol could save 50 cent per gallon compared to a state-of-the-art *Tricoderma-yeast* simultaneous saccharification and fermentation process, since the former process combines cellulose production, hydrolysis, and fermentation in a single bioreactor (Hogsett *et al*., 1992). Conversion of mixed hardwood flour to ethanol in a continuouse fermetor was 2.5 times higher with *C. thermocellum* than with the simultaneous saccharification and fermentation

process using *Tricoderma* cellulose, β-glucosidase, and *S. cerevisiae* (South *et al*., 1993)

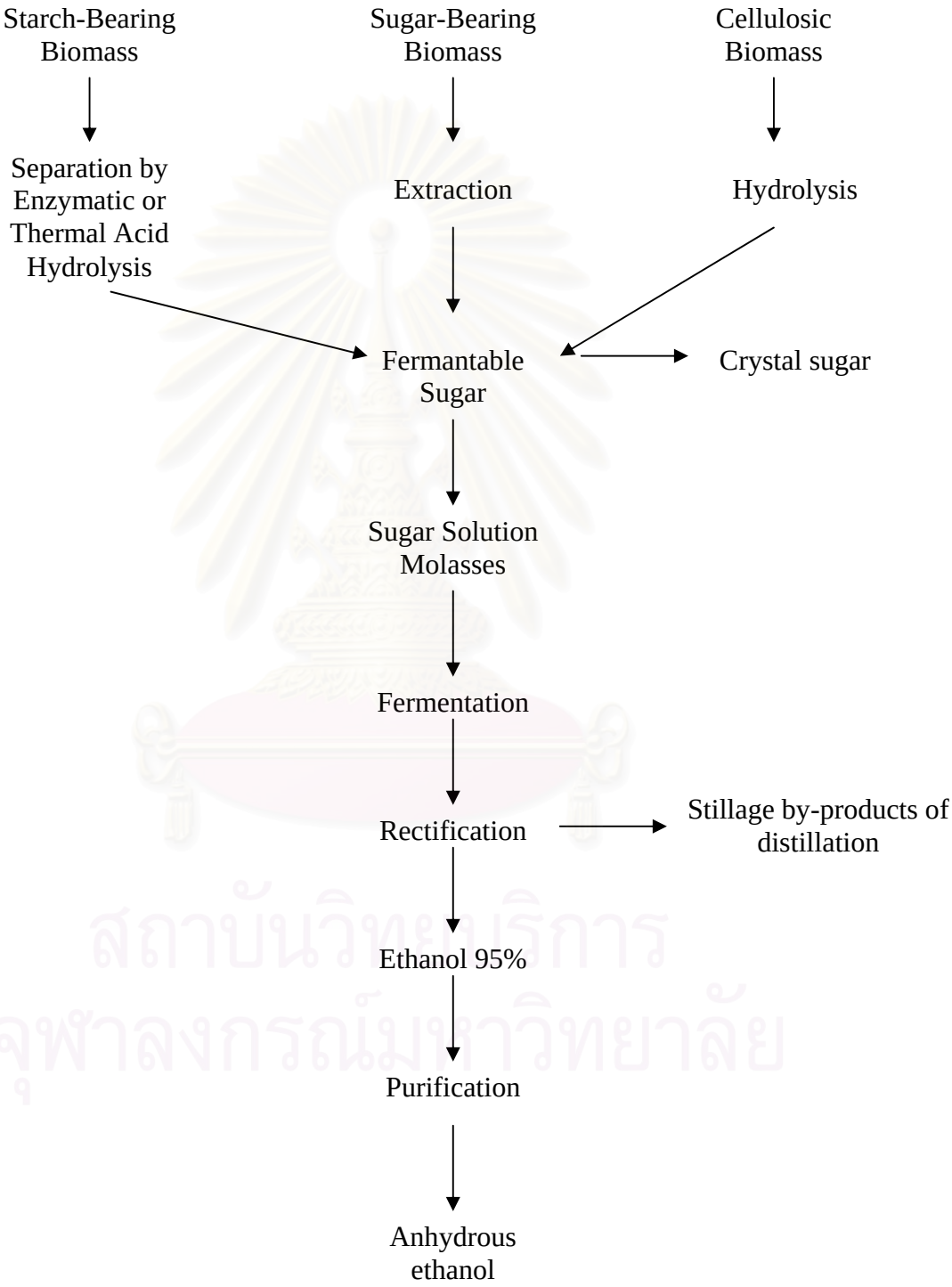


Figure 29. Flow diagram of ethanol production from cellulosic, sugar bearing and starch-bearing raw materials

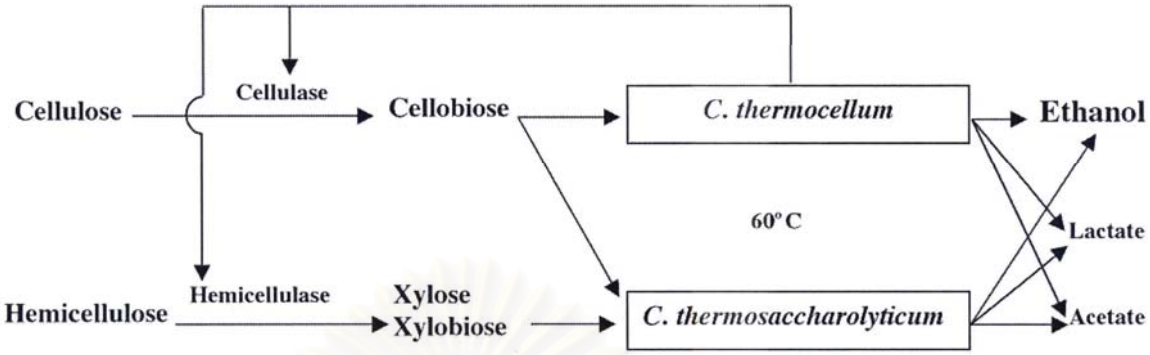


Figure 30. *Clostridia* coculture process in which *C. thermocellum* serves as the cellulose and hemicellulase producer. The hemicellulose derived pentoses can be utilized by *C. thermosaccharolyticum* but not *C. thermocellum*. *C. thermosaccharolyticum* uses cellobiose faster and is a better ethanol producer. In addition to cellobiose, cellodextrins are also produces from cellulose and can be utilized directly (Demain *et al*., 2005)

Advantages of using *C. thermocellum* for ethanol fermentation from biomass (Demain *et al*., 2005) were shown as figure30. Firstly, the cellulolytic and ethanogenic nature, allowing saccharification and fermentation in a single step. Secondly, the anaerobic nature, negating the need for expensive oxygen transfer. Thirdly, low cell growth yield, favoring ethanol conversion. Next, the thermophillic nature, facilitating ethanol removal and recovery. The thermophillic nature, reducing cooling cost. Furthermore, thermophillic fermentation being less prone to contamination. Finally,thermophillic biomass-degrading enzyme enhancing protein stability. Amenability to cocultire with other ethanol-producing and pentose-fermenting organisms.

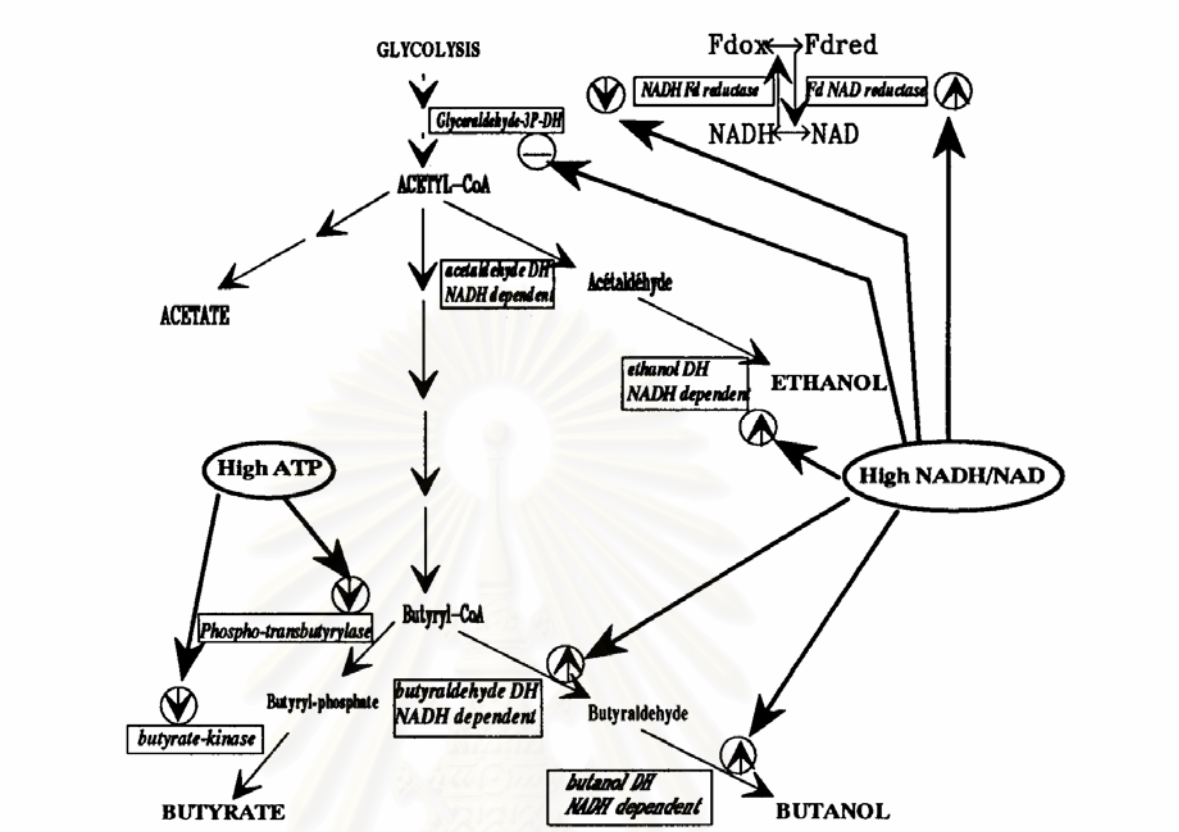


Figure 31. Scheme of the regulatory mechanisms involved in carbon and electron flow orientation during growth of *C. acetobutylicum* in continuous steady-state, phosphate-limited cultures maintained at pH 6.5 (Girbal *et al*., 1994)

2.5.2 Organic acid formation

Breakdown of sugars to pyruvate proceeds via Embden-Meyerhof-Parnas pathway. Acetate, lactate and butyrate are categorie in products of this acid formation. From the pyruvate molecules form, two acetates are produced by the enzyme pyruvate: ferredoxin (Fd)-oxidoreductase, phosphotransacetylase, and acetate kinase. CO_2 act as the hydrogen acceptor for the reducing equivalents generate during sugar breakdown (Rogers and Gottschalk, 1993).

To provide a hydrogen acceptor, CO_2 is reduced to formate and then linked to tetrahydrofolate (H_4F) in the formyl-H_4F synthetase reaction and it is reduced stepwise to methyl-H_4F. The final reactions resulting in the formation of third acetate are complex and require the enzyme CO-dehydrogenase, which has three active sites. The methyl group is transferred from methyl-H_4F via a corrinoid enzyme to the CO-

dehydrogenase. Independently, the second CO_2 molecule from glycolysis is reduced to yield enzyme-bound CO. An acetyl enzyme is formed which subsequently is cleaved by the action of coenzyme A generating acetyl-CoA, which again can be converted into acetate via acetyl phosphate. Thus, three molecules of acetate are formed from one molecule of hexose, and there are practically no by-products.

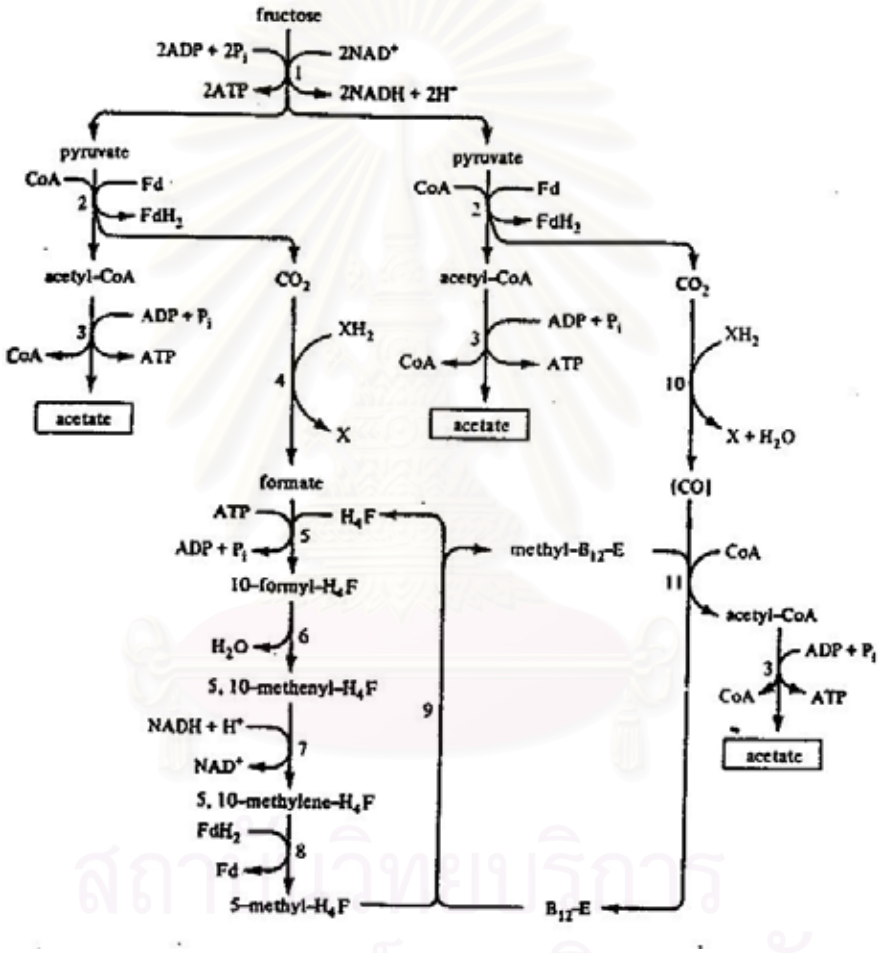
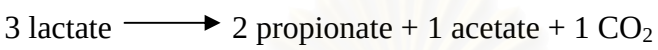


Figure 32. Pathway of the acetate formation (Gottschalk, 1986)

(1) Degradation of fructose via the Embden-Meyerhof-Parnas pathway; (2) pyruvate-ferredoxin (Fd) oxidoreductase; (3) phosphotransacetylase plus acetate kinase; (4) formate dehydrogenase; (5) formyltetrahydrofolate (H_4F) synthetase; (6) methenyl-H_4F cyclohydrolase; (7) methylene H_4F dehydrogenase; (8) methylene-H_4F reductase; (9) H_4F: B12 methyltransferase; (10) CO dehydrogenase; (11) acetyl-CoA-synthesizing coenzyme; (CO), enzyme-bound

Only a few clostridia species are known which produce lactate in a lactate/sugar ratio comparable to the one of homofermentative lactic acid bacteria. A species to be mentioned in this context is *C. thermolacticum*. It ferments a great variety of sugar including starch, xylan, lactose and various pentoses. *C. propionicum* converts lactate to propionate, acetate and CO_2 according to the following equation



2.5.3 Hydrogen production

The technology for large scale hydrogen production from biological active organisms or from waste biomass is not well developed. However, hydrogen can be obtained from water by means of many physico-chemical methods. But the production cost of hydrogen is very high. Various biotechnological mean of hydrogen production through photosynthetic bacteria and anaerobic fermentation of biological waste material.

Table 7. Hydrogen producing clostridia (Verma and Behera, 2003)

Clostridium species	Type of substrate	Products
C. acetobutylicum	Glucose, glycerol	Butyrate, Acetate, Butanol
C. butylicum	Pyruvate	Acetone, Ethanol, CO_2, H_2
C. butylicum	Glucose	Butyrate, Acetate, Butanol, 2-propane, CO_2, H_2O
C. kluyveri	Ethanol, Acetate, CO_2	Butyrate, H_2
C. botulinum	Protein, Amino-acids	Acetate, Lactate
C. histolyticum		NH_3, H_2
C. sticklandii		
C. sporogenes		
C. tetanomorphum	Glutamate, Histidine	Butyrate, Acetate, NH_3, CO_2, H_2O
C. butyricum	Glucose, Starchdextrin	
C. tyrobutyricum	Glucose or Lactate Glucose + Acetate	Butyrate, Acetate, CO_2, H_2O

The genus *Clostridium* belongs to the family *Bacilliaceae*. They are gram-positive bacteria and very motile by virtue of processing petrichous flagella. The

vegetative cells are rods, but their shape is variable and influenced by environmental factors. The clostridia are characterized physiologically by their intensely fermentative metabolism and their relation to oxygen. The optimum temperature for growth of most known clostridia is in the range of 30-40 ºC. The main carbon sources of these bacteria are glucose, starch, lactate, butyrate, CO_2, H_2 etc. The bacterium *Clostridium kluyveri* metabolises ethanol and acetate to butyrate, caproate and hydrogen. The fermentation of glucose by *Clostridium kutricum* and *C. acetobutylicum* can be regards as also prototype clostridia fermentation. Final products are acetate, butyrate, butanol, ethanol, acetone, 2-propanol, CO_2 and H_2.

2.6 Basic methods for the isolation of Clostridia

No single method may be relied upon to isolate all the varieties of clostridia that might be present in a given sample, still less all the species that might be present in samples from widely different environments. Knowledge of the varying habits of this group of bacteria is, therefore, a prerequisite for their successful isolation. In the present communication, no attempt is made to describe all the specialized methods that have been devised to facilitate isolation of individual clostridia or group of clostridia; rather, the general principles of isolation methods are described and illustrated by media and methods that we have found successful in examining specimens for the commoner mesophilic clostridia from a variety of environments including clinical material, fishery products, marine and fresh-water sediments.

Sucessfully isolation of many bacteria depends on exploitation of specific properties of the organism sought. In the case of clostridia, anaerobic incubation in a liquid medium is itself often sufficient for effective enrichment, and isolation may be achieved simply by subsequent plating on to solid media. The relatively high resistance of spores to heat and chemicals may be exploited, provided the organism is present in the spore form; some clostridial spores, however are no more resistant than many vegetative cells. Finally, various selective agents may be incorporated into media to inhibit the growth of other organisms. The success of individual selective agents depends, however, on the nature of associated microflora, and no single inhibitor is universally effective.

Carbon dioxide often improves the growth of clostridia; it will certainly improve the germination rate of some spores (Wynne and Foster, 1948). The carbon dioxide may be provided either in the atmosphere, for example as a mixture of hydrogen with 5% (v/v) carbon dioxide, or by the inclusion of 0.1% (w/v) of sodium bicarbonate in the culture medium (Hobbs *et al*., 1971).

It is common practice to treats samples with heat or chemical before primary culture. The spores of many mesophilic clostridia survive heat treatment at 80°C for 10 min, although survival is a function both of a number of spores present as well as of temperature and time of exposure. This treatment destroys all vegetative cells present. Various chemicals may also be used to kill vegetative organisms; indeed use of chemicals such as ethyl alcohol may be useful when heating can not be applied because the spores of some clostridia are thermolabile. The majority of clostridia grow at an optimumtemperature of 37ºC, and unless thermophilic or psychophilic sought, this incubation temperature should be used.

2.6.1 The use of differential Reinforced Clostridial medium for the isolation and enumeration of Clostridia

Many of selective media that are routinely used to enumerate clostridia were designed primarily for the isolation of Clostridium. These media inhibit many other clostridia (Gibbs and Freame, 1965). Differential Reinforced Clostridial Medium (DRCM) is based on the reinforced Clostridial Medium of Hirsch and Grinsted (1954) and Gibbs and Hirsch (1956), it is a rich medium that supports good growth of most clostridia, but also permits growth of the other anaerobes and facultative anaerobes. In order to detect growth of clostridia in RCM, sodium sulfide and iron citrate were added. The sulfite is reduced by clostridia to form sulfide which, when combined to give iron-sulfide, blacken the medium. It is important to realize that tha blackening of DRCM is not specific for clostridia, but can be caused by other bacteria such as salmonellae, *Proteus* spp., bacteroides and some strains of *Escherichia coli*.

Gibbs and Freame (1965) listed the advantage of using DRCM as: (a) good recovery of low number of clostridia; (b) no need to use anaerobic jar; (c) incubation can be prolonged without dehydration of the medium; (d) gas formation is no problem as it can be in agar cultures; (e) cultures can be examined daily without disturbing the growth conditions. The disadvantages were listed as; (a) use of the MPN count, which

is less accurate and more tedious in preparation than a plate count; (b) the need to make aseptic additions of the sulfide/citrate solution ineach bottle of basal medium immediately before use, and (c) the growth and blackening in non-pasteurized counts by some bacteria other than *Clostridia*.

2.6.2 Cellulose-decomposing clostridia

Because of the small number of cellulolytic anaerobes in many soils, isolation is often easier when proceded by enrichment. The media and methods to be described are specific for cellulolytic anaerobes rather than cellulolytic clostridia. Cellulose suspension prepared by ball-milling (Hungate, 1950) is convenient to use. Growth factors always added by yeast extract or it can be compensated by rumen fluid. Cellulolytic clostridia can be growth in both form of colony on the top of agar or sometimes no colony appeared. Sharp and smear clear zones depend on the location of colony in agar. Localized digestion of cellulose always appeared and can be observed around small colony. In case no colony appeared, generalized cellulose digestion zone will appeared with no colony visible. This effect is probably caused by the accumulation of cellulolytic enzymes in the condensation water.

Since plant cell wall has different composition, hard to digest or break chemical bond by any specific enzyme alone. Enzyme cocktail or cellulosome is needed. In Clostridia species, its cellulolytic enzyme system is macromolecule extracellular enzyme. Cellulase, lichenase, xylanase and hemicellulase subunit bind to the core protein by species specific recognition. All subunits act synergistically to degrade cellulose efficiently. Screening of cellulolytic bacteria is one of the simply ways to obtain new cellulolytic enzyme from natural.

CHAPTER III

MATERIALS AND METHODS

3.1 To measure Cel9X and Cel9C enzyme subunit activity

3.1.1 Bacterial strains and plasmids

C. acetobutylicum ATCC 824 was used as the source of genomic DNA. *Escherichia coli* DH5α was used as a host for the pGEM-T derivative plasmids and the pET-22b(+) (Novagen) derivates. *E. coli* BL21(DE3) was used as the host for expression vectors.

3.1.2 Growth conditions

C. acetobutylicum was grown anaerobically at 37°C in a synthetic medium as previously described (Vasconcelos *et al*., 1994). *E. coli* DH5α was grown at 37°C in Luria-Bertani medium supplemented with ampicillin (100 µg/ml) when required. *E. coli* BL21(DE3) was grown under the same conditions of *E. coli* DH5α, except that isopropyl-β-D-thiogalactopyranoside (IPTG) 50 µM was added to provide optimal conditions for the heterologous protein production.

3.1.3 DNA isolation and manipulation

Total genomic DNA from *C. acetobutylicum* ATCC 824 was isolated as previously described (Mermelstein *et al*., 1993). Plasmid DNA was extracted from *E. coli* with the QIAprep kit (Qiagen, Courtaboeuf, France). Restriction enzymes and T4 DNA ligase were obtained from New England BioLabs (Beverly, USA) and GIBCO/BRL (Life Technologies, Cergy Pontoise, France), respectively and used according to the manufacturer's instructions. DNA fragments were purified from agarose gels with the QIAquick gel purification kit (Qiagen). PCR was carried out using chromosomal DNA as a template with the Expand High Fidelity System (Boehringer Manheim).

3.1.4 Expression of recombinant CelC and Cel9X proteins

The region of *celC* gene that encodes the mature protein was amplified by PCR with the primers $CelC_{dir}$ (5' AAGTAACCATGGATGGA TGAAACTAATACAAATTTTAATTATGGTGAAGCAC-3') and $CelC_{rev}$ (5'-

ATAACTCGAGAGGATTCACACTAGAAATCTTTCTGATTAACTCC-3'). The PCR oligonucleotide primers were designed to amplify the gene without its leader peptide. These primers create a *Nco*I site and a *Xho*I site upstream and downstream of the coding sequence, respectively. The amplified fragment was subcloned into the pGEM-T easy vector (Promega, Charbonnières, France) and digested with *Nco*I and *Xho*I and then ligated to *Nco*I/*Xho*I-digested pET-22. The coding sequence of *celC* was fused in frame with a downstream sequence of the vector encoding six histidine residues (His Tag). The resulting plasmid, pET-C, was used to transform *E. coli* BL21(DE3) cells for the expression of the recombinant protein.The pET-X constructed in the same way.

3.1.5 Purification of the CelC and Cel9X recombinant proteins

Recombinant cells were grown in 2000 ml Luria-Bertani medium supplemented with glycerol (12 g/l) and chloramphenical (10 µg/ml) and carbenicillin (50 µg/ml) with shaking to an optical density of 1.8 at 600 nm, and then cooled to 16°C. IPTG was added to a final concentration of 50 µM, and the culture was incubated with shaking at 16°C for 15 h. The final optical density of the culture was about 2.5-3. The cells were cooled to 5°C, harvested by centrifugation, resuspended in cold 30 mMTrisHCl, pH9; 0.5 mM$CaCl_2$ and sonicated at 4°C in five cycles of 10 s with a 30 s interval between each cycle. The crude extract was centrifuged at 10.000 × g for 15 min, and the supernatant was loaded on a 3 ml Ni-nitrilotriacetic acid column previously equilibrated with cold 30 mMTrisHCl, pH9; 0.5 mM$CaCl_2$ buffer. After that, washes with cold 30 mMTrisHCl, pH8; 0.5 mM$CaCl_2$ buffer supplemented with 10 mM imidazole, the protein was eluted in cold 30 mMTrisHCl, pH8; 0.5 mM$CaCl_2$ buffer supplemented with 60 mM imidazole. The eluted fraction was dialyzed immediately against 25 mM Sodium acetate, pH 5.0, 6.0, 7.0 buffer and concentrated in an Amicon concentrator. The protein purification was monitored by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (Laemmli et al., 1970). Aliquots of the pure protein were stored at -20°C.

According to the same purification protocol, the full-length CipA was purified from a cell-free extract of *E. coli* BL21(DE3) harbouring pET-CipA (Sabathé and Soucaille, 2002).

3.1.6 Enzyme assays

Cellulolytic activities were assayed by mixing appropriate concentrations of pure Cel9C with carboxymethyl cellulose (CMC), Avicel, Bacterial Cellulose (BC) and phosphoric acid swollen cellulose (PASC) at a final concentration of 0.8%, in Sodium Acetate Buffer (PPB, 25 mM, pH 7.0) at 37°C. BC was a generous gift from H.P. Fierobe (BIP, Marseille, France) and PASC was prepared from Avicel as described by Walseth (1952). Aliquots were collected at specific intervals and centrifuged at 10,000 RPM for 5 min at 4°C to remove insoluble substrate and the reducing sugar contents were determined by the Bicinchoninic acid method (Wood, and Bhat, 1988), assuming that 1 U enzyme liberates 1 µmol of glucose equivalent min^{-1} (mg protein)$^{-1}$. Specific activities were determined in the linear range of the reaction. The protein concentration was determined as previously described by Lowry *et al*. (1951).

3.1.7 Surface plasmon resonance (Biacore) analysis

Real-time detection of the binding of Cel9C to CipA was performed at 25°C with a biomolecular interaction analysis biosensor-based analytical system. All experiments were performed on a Biacore 3000 instrument (Biacore AB, Uppsala, Sweden) using certified running buffer containing 10 mM Hepes, (pH 7.4), 150 mM NaCl, 3 mM EDTA, 0.005% polysorbate 20 (v/v) (HBS-EP). Sensorchips CM5 (sensorchips with a carboxymethylated dextran matrix) were obtained from Biacore AB.

Reagents for amine coupling are supplied by Biacore AB as Amine coupling kit containing N-hydroxysuccinimide (NHS) and N-ethyl-N’-(dimethyl-aminopropyl)-carbodiimide (EDC) and 1 M ethanolamine hydrochloride pH 8.5.

Immobilisation of histidine-tagged proteins. Histidine-tagged CipA fusion protein was immobilised on a CM5 sensorchip by amine coupling chemistry. Briefly, amine coupling introduces N-hydroxysuccinimide esters into the surface matrix by modification of the carboxymethyl groups with a mixture of N-hydroxysuccinimide (NHS) and N-ethyl-N’-(dimethyl-aminopropyl)-carbodiimide (EDC). These esters then react spontaneously with amines on the ligand to form covalent links. Deactivation of excess reactive groups on the sensor surface is performed using 1 M ethanolamine hydrochloride pH 8.5.

Flow cell 2 of a CM5 sensorchip was coated with histidine-tagged CipA fusion protein. Briefly, histidine-tagged CipA fusion protein was injected at 10µg/ml over the sensor surface after NHS-EDC activation. Finally, 2250 RU of CipA protein were immobilized on flow cell 2. Then deactivation with ethanolamine was performed.

3.1.8 BIA analysis

The biosensor assay was run at 25°C in HBS-EP as running buffer. All samples were injected in HBS-EP buffer. Kinetic constants were determined using Biaevolution version 3.1 software.

3.1.9 Nucleotide sequence accession number

The Genbank accession number for the nucleotide sequences of *cel9C* and *cel9X* are AE001437 and AE001437.1 AE001437.1:649182.651818 respectively.

3.2 To modify the cellulosome of *C. acetobutylicum* by using genetic engineering technique

3.2.1 Bacterial strains for engineer Cel48A

The bacterium used in this investigation is *Clostridium acetobutylicum* ATCC 824 (American Type Culture Collection). This strain was used as a source for genomic DNA. It was grown anaerobically at 37°C (85% N_2 10% H_2; Coy Labolatory Products, Inc., Ann Arbor, Mich.). *E. coli* strains used in this study are *E. coli* DH5α and *E. coli* BL21 (DE3). *E. coli* DH5α was used as a host for all cloning purposed. *E. coli* BL21(DE3) was use as a host for plasmid pET-Cel48A and pET-FH3 expression vector which provide protein containing a C-terminal His-tag under the control of T7 RNApolymerase (Sabathé and Soucaille, 2002). *E. coli* DH5α was grown at Luria-Bertani medium supplemented with appropriated antibiotics when required. *E. coli* BL21(DE3) was grown at 30°C on a rotary shaker or on agar plates with Luria-Bertani medium, $MgSO_4$ (10 mM), isopropylthiogalactoside (IPTG) (1 mM), 5-bromo-4-chloro-3-indoyl-β- galactoside (50 µg/ml) as required. *E. coli* strains were stored in glycerol solution at –70°c (Wang *et al*., 1993).

3.2.2 Gene modifying and construction of pET-Cel48A plasmid for expression of Cel48A protein

Total genomic DNA from *C. acetobutylicum* ATCC824 was isolated as

previously described by Hansen *et al*. (1992). Plasmid DNA was extracted from *E. coli* with the QIAprep kit (Qiagen, Courtaboeuf, France). All plasmids were constructed in *E. coli* DH5α first and then transformed into *C. acetobutylicum*. DNA restriction and cloning were performed according to standard procedures. Restriction enzymes and T_4 DNA ligase were obtained from New England Biolabs (NEB, Beverly, Mass.) and GIBCO/BRL (Life technologies, Cergy Pontoise, France), respectively, and used according to the manufacturer's instructions. DNA fragments were purified from agarose gels with QIAquick gel purification kit (Qiagen). Transformations were established. Procedures for transformation of *E. coli* (Sambrook and Russel, 2001), preparation of competent cells of *E. coli*, and their transformations were done. Plasmid pET-22b (+) containing the DNA fragment encoding the Cel48A will serve as the template for all PCRs. PCRs with mutagenesis primers were done according to the “Quick change site directed by Mutagenesis Kit” (Stratagene). Primers were be designed for corrected Cel48A protein sequence (The corrected amino acid sequence between position 567 to 572 is -Glu-Gln-Arg-Gly-Asp-Tyr-). The PCR products were used to transformed into DH5α competent cells. Plasmid DNA in each case was isolated and sequenced. Mutants possessing the correct nucleotide changes will be used for further study.

3.2.3 Expression of corrected Cel48A protein in *E. coli* BL21 CodonPlus strain

Recombinant cells were grown in 2000 ml Luria-Bertani medium supplemented with glycerol (12 g/l), ampicillin (200 μg/ml) and carbenicillin (50 μg/ml). IPTG was added to induce recombinant protein production. The final optical density at 600 nm of the culture is about 2-2.5. The cells were cooled to 4°C, harvested by centrifugation, resuspended in Tris-HCl pH 9.0 buffer, and sonicated at 0°C in five cycles of 30 s with a 60 s interval between each cycle. Aliquots of pure protein were stored at 4°C. Samples of cell-free extract of *E.coli* BL21(DE3) (25μl) containing plasmid for expression in *E.coli* and *C. acetobutylicum* proteins were subjected to SDS-10% PAGE and transferred electrophoretically onto nitrocellulose membranes. The blots were blocked with blocking buffer containing 2% bovine serum albumin, 25 mM $CaCl_2$, in Tris-buffered saline (50 mM Tris-HCl [pH7.5]-150 mM NaCl). Secondary antibody alkaline phosphatase antibody (Invitrogen, Carlsbad,

Calif.) were used for detection according to the manufacturer's instructions (Ding *et al*., 2000).

3.2.4 Hybrid *C. cellulolyticum*-48Fcatalytic/*C. acetobutylicum*-dockerin fragment construction

C. cellulolyticum-48Fcatalytic and *C. acetobutylicum*-dockerin fragment were synthesized separately by *Taq* polymerase high-fidelity. PCR amplification was done to link two fragments by use Ccf1 and Dock2 as primer. Ligate synthesize hybrid molecule with pETFc plasmid as described in topic 2.6. Transform into *E. coli* BL21 codonplus. Select the right clones by use 50 µg/ml ampicillin and chloramphenical.

3.2.5 Bacterial strain, plasmid and culture condition for chimeric Cel48

The bacterium used in this investigation is *Clostridium acetobutylicum* ATCC 824(American Type Culture Collection). This bacterial strain was used as a source for genomic DNA. It is grown anaerobically at 37°c (85% N_2 10% H_2; Coy Labolatory Products, Inc., Ann Arbor, Mich.). For plasmid isolation, *C. acetobutylicum* was grown in Clostridial Basal medium and harvested at late exponential phase (Optical density at 620 nm of about 0.9-1.0). Growth was monitored as the optical density at 620 nm by using a spectronic20 spectrophotometer (Bausch&Lomb, Inc., Rochester N.Y.) (Kim *et al*., 1990).

3.2.6 Introduction of chimeric *C. cellulolyticum*-48Fcatalytic/*C. acetobutylicum*-dockerin fragment into expression plasmid

Plasmid-pET-FH3 was constructed by design specific primer to amplified hybrid of Cel48F catalytic domain and then link with Cel48A dockerin domain gene. After synthesis the hybrid molecule, then replace dockerin domain of Cel48F in the plasmid pET-Fc (Leverbel-Leroy *et al*., 1996) with Cel48A dockerin domain. This plasmid obtained by overlap-extension polymerase chain reaction. Hybrid Cel48F-catalytic domain was amplified by primer forward 5' ga**cctagg**ttgtgcttcttca3' (*Avr*II site in bold) and reverse 5' tactttatatgtcatgctcgggaa*gag tattgcataaactc* 3'. Linker part and Cel48A dockerin domain was amplified by forward primer 5' *ttcccgagcatgacatataaagta*ttcgctaatacagctacaccagg 3' (overlap linker part in italic) and reverse 5' ttt**ctcgag**ttttgcaattaatttagtaagttccataatatc 3' (introduce *Xho*I site in bold).

3.2.7 Nucleotide accession number

The DNA sequence for Cel48A gene is deposited in the GenBank Database

under accession number AE007607 and AE001437.

3.2.8 Construction of corrected pET-Cel48A, pET-FH3 plasmid and expression of Cel48A protein

Plasmid DNA which is corrected by PCR nucleotide change was used. The Cel48A protein was produced from the cytoplasmic fraction from *E.coli* BL21(DE3) CodonPlus cell harboring the resulting plasmid, pET-Cel48A and pET-FH3. Expression of the Cel48A and hybrid-cel48A-F protein was obtained after induction by 40 μM IPTG (isopropyl-β-D-thiogalactopyranoside).

3.2.9 Competent cells

E. coli BL21(DE3) CodonPlus (Stratagene) about 15-20 ml was grown overnight. Then, 100 μl of the culture will be inoculated into LB medium and incubate on a table-top shaker at 37°c until an OD at 600 nm reaches 0.3-0.5. Then cells were spun down and resuspend in 0.5 volume of cold 0.1 M $RbCl_2$. Competent cells were kept at -70ºC.

3.2.10 Expression and purification of chimeric Cel48A recombinant protein

Cells were grown in 2000-3000 ml Luria-Bertani medium supplemented with glycerol (12 g/l), chloramphenical and carbennicillin (50 μg/ml) in case using pET-Cel48A and using chloramphenical and ampicillin (50 μg/ml), shaking to an optical density of 0.8 at 600 nm, and then cool to 16°C. IPTG (isopropyl-β-D-thiogalactopyranoside) was added to a final concentration of 2 mM, and the culture was incubated while shaking at 16°C for 12-16 h. The final optical density at 600 nm of the culture is about 2.5-3. The cells were cooled to 4°C, harvested by centrifugation, resuspended in cold 30 mM TrisHCl , pH 9.0; 0.5 mM $CaCl_2$ buffer (equilibrate buffer), and sonicated at 4°C in five cycles of 30 s with a 60 s interval between each cycle. The crude extract was centrifuged at 10,000 x g for 10 min., and the supernatant was loaded on a 4 ml HIS-Select^{TM} Nickel Affinity column previously equilibrated with equilibrate buffer. The eluted fraction was dialyzed immediately against 30 mM TrisHCl; 0.5 mM $CaCl_2$, pH 8.0 buffer and concentrated in an Microsep 50K Omega concentrator. The protein purification was monitored by sodium dodecyl sulfate-polyacrylamind gel electrophoresis. Aliquots of pure protein was stored at 4°C and used within one week.

3.2.11 Western blotting

Samples of cell-free extract of *E. coli* BL21(DE3) CodonPlus (2ug) containing plasmid for expression in *E. coli* proteins (2µg) were subjected to SDS-10% PAGE and transferred electrophoretically onto nitrocellulose membranes. The blots were blocked with blocking buffer containing 2% bovine serum albumin, in PBS complete buffer (pH7.3). Secondary antibody alkaline phosphatase antibody (Invitrogen, Carlsbad, Calif.) was used for detection according to the manufacturer's instructions (Ding *et al*., 2000).

3.2.12 Activity assay

Rapid 2,2'-bicinchoninic-acid assay was used for assay reducing sugar released from enzyme reaction as described previously (Doner et al, 1988). Reducing-sugar standards were prepared freshly. Assay solution A and B were prepared as described previously (Doner *et al*, 1991). The assay reagents were prepared daily by mixing equal volumes of assay solution A and B before assay. Standard reducing glucose solutions were prepared by diluting aqueous 0.5 mM solutions of glucose. Standard curves were plotted with glucose concentrations from 0 to 100 µmol/0.2 ml. To test tube containing 0.2 ml of reducing carbohydrate, 0.8 ml of assay reagent will be added, and the mixtures were heated in a dry bath at 100°C for 15 min. The tube is then cooled down before the reading (Perkin-Elmer, Norwalk, CT, Model LC-55 variable wavelength uv-vis detector) at 560 nm.

3.2.13 p952-SA-FA vector transformation into *C. acetobutylicum* ATCC824

Plasmid DNA which is corrected by PCR nucleotide change was used. Electroporation were performed by using a Gene Pulser connected to a pulse controller (Bio-Rad Laboratories, Richmond, Calif.). Two cultures of *C. acetobutylicum* were prepared in CGM medium by add 1 ml of *C. acetobutylicum* spore in MS medium. Heat-shock was performed at 80 °C for 12 minutes. Incubate at 37 °C overnight until the OD_{620} reach to 2. Then, Cells were grown to late exponential phase. The cultures were grown in 50 ml 2YTG, incubated at 37 °C until the OD_{620} reach to 1. Place the culture on ice at least 30 min. Culture and buffer were transferred into anaerobic condition. Centrifuged for 10 min at 6,000 rpm by using an RC-2

Centrifuge (Ivan Sorvall, Inc., Norwalk, Conn.) to recover the pellet and resuspended in 4 ml electroporation buffer. Pellets were washed and centrifuge again, discared the supernatant. All pellets were resuspended into 500 ml electroporation buffer pH 7.4 and transferred into 4 mm electroporation cuvette (EUROGENTEC, Inc.). *C. acetobutylicum* competent cells were kept on ice at last 20 min before transformation. Purified concentrated plasmids (methylated or non-methylated) were prepared and incubated with competent cells before transformation. Electrotransformation were done at the condition following standard procedure: U=1.8kV, C=50μF, R=600Ω. The electroporated cell suspension were incubated for 4 hr at 37°C in an Incu-Block (Denville Scientific, Inc., Denville, N.J.) and subsequently purified by using phenol and chloroform extraction (Kim *et al*., 1990).

3.2.14 Targeted inactivation of malfunction Cel48A gene in *C. acetobutylicum*

pREPcel48::upp was constructed. Cel48A deletion fragment was synthesized by *C. acetobutylicum* genomic DNA amplification by CelA1, CelA2, CelA3 and CelA4 primer. The, complete fragment was ligated to pConS2-1::upp. Plasmid was introduced into *C. acetobutylicum* ATCC 824 by electroporation. A liquid culture of strain 824 (pREPcel48::upp) will then be used to confluently streak agar (15 g/l) plates of Reinforced Clostridial medium containing 40 mg of Erythromycin per liter. These isolates will be grown in liquid medium containing 40 mg of Erythromycin per liter, selected for Erythromycin resistance and Chloramphenical sensitive. The culture will be used for plasmid DNA isolation, transformation of pET-Cel48A plasmid, and frozen stock preparation. The criteria for further screening are the presence of both inactivation plasmid and pREPcel48::upp, resistance to Erythromycin.

3.2.15 Study of corrected Cel48A gene presence in *C. acetobutylicum*

After transformation, clostridia transformants were cultured on Reinforced Clostridial agar supplemented with 40 μg/ml ampicillin and 200 μg/ml erythromycin. Colonies were appeared within 24-48 hours after incubate at 37°C in strictly anaerobic condition. Corrected phenotypic expression colonies were picked and restreak on antibiotic plates to check plasmid stability. Selected clones were inoculated into 2YTG medium supplemented with 40 μg/ml ampicillin and 200 μg/ml erythromycin and incubated overnight. Harvest cells by centrifugation and use as template for PCR

reaction directly. PCR reactions were performed to check transformed p952-SA-FA appearance and correction. Primer pREP-D and DockI were used for amplification. The tube were incubated for 2 min at 94°C for initiation and then subjected to 30 cycles consisting of denaturation at 94°C for 30 second, annealing at 52°C for 30 second, and primer extension at 72°C for 1.15 min. The tubes were then incubated for 7 min at 72°C. Aliquots (4µl) of the amplification products were analyzed by electrophoresis in 0.8 % (wt/vol) agarose gels and then ethidium bromide staining. Expected PCR product from corrected plasmid amplification should be appeared by 7 kbp DNA fragment.

Table 8. Bacterial strain, plasmid and primer

Strain, plasmid or primer	Relevant characteristics	Source or reference
Strain		
E. coli DH5α	F-Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *deo*R *rec*A1 *ton*A	Invitrogen
E. coli BL21 CodonPlus(DE3)-RIL strain*	*E. coli* BF *omp*T *hsd*S (r_B^- m_B^-) *dcm*$^+$ Tet^r *gal*λ (DE3) *end*A The [*argU ileY leuW* Cam^r]	Stratagene
C. acetobutylicum	Wild type	ATCC
Plasmid		
pET-Fc	F1 ORI, Amp^r, *lac* promoter, T7 promotor	Reverbel-Leroy, 1996
pET-Cel48A	F1 origin, Carb^r, , *lac* promoter, T7 promotor	In this study
pET-FH3	F1 origin, Amp^r, *lac* promoter, T7 promotor	In this study
P952-SA-FA	F1 ORI, Amp^r, Ery^r, Pthl- *lac* operator	In this study
pET22-Cel9C-1	F1 ORI, Cam^r, *lac* promoter, T7 promotor	In this study
pET-Cel9X-C1	F1 ORI, Cam^r, *lac* promoter, T7 promotor	In this study
Primer		
Ccf1	5’GACCTAGGTTGTGCTTCTTCA3’	Reverbel-Leroy, 1996
Ccf2	5’TACTTTATATGTCATGCTCGGGAAGAGTATTGCATAAACTC3’	Reverbel-Leroy, 1996
DockI	5’TTCCCGAGCATGACATATAAAGTATTCGCTAATACAGCTACACCAGG3’	In this study
Dock II	5’TTTCTCGAGTTTTGCAATTAATTTAGTAAGTTCCATAATATC3’	In this study
CelA1	AAAAGGGATCCCAGATGGAAAAATAACAATCACAGGTTCAGCACCAG	In this study
CelA2	**GGGG***AGGCCT***AAAAAGGGGG**AAATATAACTGTAGATGTAAGAGCTACAGCC	In this study
CelA3	**CCCCCTTTTT***AGGCCT***CCCC**TTAGCTGCAGCAGATGTTAATAATGATGGTG	In this study
CelA4	AAAAGGATCCTCTTGTATATAGAACCAGCTTGGACGCC	In this study

*These strain, derivative of *E. coli* B, are general protein expression strains that lack

both the L_{on} protease and the OmpT protease, which can degrade protein during purification. Dcm methylase, naturally lacking in *E. coli* B, is inserted into the genomes (Stratagene). *Bam*HI site = GGATCCC, *Stu*I site = *AGGCCT*, Bold letter = linker fragment between CelA2 and CelA3 primer.

3.3 To isolate and characterize cellulolytic, solvent producing clostridia from soil samples in Thailand

3.3.1 Isolation of cellulolytic bacteria

Thirty isolates of butanol-tolerant anaerobic cellulolytic bacteria were isolated. The samples collected from soil, dry decomposer and cow dunk. Fifteen isolates were identified in genus *Clostridium*. Genus isolation was performed by using 4 criteria for Clostridia identification Gram's staining, sulfite reducing ability, anaerobic fermentation products, and spore-forming process were done. Thienability for dissimilatory sulfate reduction were performed by using differential Clostridial agar (Difco Inc.). Anaerobic fermentation product in each samples were analysed with Gas Chromatography (Shimadsu Model GC 7AG), Porapak Q80-100 mesh column.

Inoculate 1 g of soil sample into MS media with modified. Avicel was added to the media as the carbon source of selective media. MS media, used in the isolation of the organisms, contained: distilled water, 1000ml; $MgSO_4$, 0.22g,; KH_2PO_4, 0.55g; glacial acetic acid, 2.3 ml; $FeSO_4$, 0.011g; Para-amino benzoic acid, 5 ml; Biotin, 4 ml; Resazurin, 1 ml; Avicel PH101 (Fluka), 20g. After preparation all of the composition, N_2 gas always added into the bottle of media. The pH of MS media was adjusted to 6.5. The media were heated under an N_2 atmosphere until the resazurin indicator became colorless and the distributed into bottles (30 ml of medium per anaerobic culture bottle; Belco Glass) that were being flushed with N_2. The bottles were sealed with Neoprene stopper and autoclaved. After cooling either to room temperature broth) or 45 to 50 °C (agar medium), the steriled media were supplemented with a filter-sterilized 5% (wt/vol) solution of cysteine until final concentration reach to 0.5 g/l and 1 mM $CaCl_2$. For select butanol-tolerant anarobic bacteria, butanol was added until final concentration in each bottle reach to 5 g/l.

Differential Clostridia agar was prepared by Weenk method (1991). This media contained: distilled water, 100 ml; Peptone from Casein, 5.0 g; Peptone from meat, 5.0 g; meat extract, 8.0 g; starch, 1 g; D(+)glucose, 1.0 g; yeast extract, 1.0 g;

cysteinium chloride, 0.5 g; resazurin, 0.002 g; agar-agar, 20 g. Final pH should be 7.6. After sterilization, media were supplemented with 5 ml/l ferric (III) ammonium citrate solution and 7.5 ml/litre medium sodium sulfite.

The isolates were grown on MS selective media for select butanol-tolerant anaerobic bacteria. MS broth supplemented with Avicel as carbon source were use for this purpose. After inoculation, the bottles were incubated at 37°C for 7-14 days under an atmosphere of 5% CO_2, 10% H_2, and 85% N_2 in an anaerobic chamber (Coy Laboratory Products, Inc., Ann Arbor, Mich.). Subculture processes were done for 3 times to confirm bacterial selection under selective condition.

3.3.2 Nutritional Characteristics

The ability of the isolates to utilize various soluble compounds as fermentable substrates for growth was determined by estimating visually the turbidity and gas production after three transfers in liquid medium containing the potential substrate. Seven-day-old MS broth cultures were used to inoculate both MS broth supplemented with Avicel and MS broth supplemented with glucose. Both type of medias were not add any butanol. Two additional successive transfers in the same medium were performed with 1 ml of culture as inoculum.

Analysis of fermentation products were performed by Gas Chromatography. Cultures (30 ml) were inoculated with 1 ml of 7-day-old culture in reinforced clostridial media and then incubated at 37°C for 96 hrs. Fermentation products formed in these cultures were quantitated as described below. Ethanol, acetic acid, acetone, butanol and butyric acid standard were prepared by range 0, 2, 4, 6, 8, and 10 g/l. Product were determined by the Shimadzu Model GC 7AG, Porapak Q80-100 mesh column. N2 in column 50 ml/min. Butanol, acetone, acetic acid, butyric acid and ethanol retention time are 4.09, 1.77, 2.3, 12.36 and 1.31 respectively.

3.3.3 PCR Amplification and sequencing

The extraction of DNA samples were used in PCR. 16S rDNA fragment were amplified with primer 27F and 1492R. About 1.6 kb fragments of 16S rDNA were amplified by PCR with *Taq* polymerase using the primer of Eubac 2F (AGAGTTTGATCCTGGCTCAG) and1492R (GGTTACCTTGTTACGACTT) (Gillan, 1998). The PCR amplification procedure was performed with Omnigene temperature cycle as described. The tube were incubated for 3 min at 94°C and then

subjected to 25 cycles consisting of denaturation at 94°C for 1 min, annealing at 55°C for 1 min, and primer extension at 72°C for 2 min. The tubes were then incubated for 10 min at 72°C. Aliquots (4µl) of the amplification products were analyzed by electrophoresis in 1 % (wt/vol) agarose gels and then ethidium bromide staining. PCR products were then purified for 16S rDNA gene sequencing by automate sequencer (Macrogen Inc., Korea).

3.3.4 Bioinformatic tools for Sequence analysis

The sequence obtained were compaired to known sequencing by using BLAST program and the SIMILARITY_RANK tool of the Ribosomal Database Project (RDP). The sequence was checked for chimeric molecules by using the CHECK_CHIMERA tool of the RDP. Nucleotide sequence of close evolutionary relatives of our sequences were retrieved from the National Center for Biotechnology Information World Wide Web ENTREZ browser that maintains and distributes the GenBank sequence database.

3.3.5 Enzyme assay

CMCase activity was tested on PC/agar plates containing 0.1% CM cellulose. Culture waere spotted and incubated for 72 hr at 37°C, then the plate were stained with 1% Congo red solution for 5 min and destained several times by washing with 1M NaCl. Orange halos on a red background indicated CMC activity. Cellulolytic assay were prepared. Cells were culture in MS supplemented with avicel then incubate 37 ºC 72 hrs, centrifuge to collect supernatant. Clear supernatant was added with avicel final 1% (wt/vol). Final volume will be adjusted to be 600 ml. Reactions were incubated 37 ºC with slight shaking (190 rpm). 0.5 ml samples were taken at specific intervals, centrifuged twice. The reducing sugar content of 200 µl of the supernatant was estimated by BCA assay as described previous by Doner *et al*. (1992).

3.3.6 Cellulosome expression and purification

Selected strains and *C. acetobutylicum* wild type were grown for 16 hours in 50 ml serum bottles with shaking at low speed (100 rpm). Cellulose was harvested by decantation of the culture fluids and washed two times in 50 ml of 100 mM Potassium Phosphate Buffer, pH 7.0 (PBB) to remove cells. The cellulose pellet was filtered through a 3-µm-pore-size glass filter and extensively washed two times with 50 ml of

100 mM PPB, and one time with 50 ml of 25 mM PPB. The cellulosome was then eluted with 10 ml 1% triethylamine. The eluted fraction was centrifuged at 15,000 × g for 15 min to remove all of the insoluble material. The 10 ml eluted fraction that was obtained was concentrated to a final volume of 1 ml by centrifugation on a 50KD polysulfone membrane (PALL bioscience).

3.3.7 Western Blottiing

Western blotting was carried out by standard procedures on nitrocellulose hybond-C extra (Amersham) when polyclonal antibodies were used. Primary antibody anti-celF was a gift from Fierobe, H. P. The dilution used with the primary antibodies was 1/1,000. Anti-rabbit immunoglobulin G conjugated with horseradish peroxidase (Amersham) was used as the secondary antibody.

3.4 To compare cellulolytic activity and solvent producing ability of the naturally selected and recombinant Clostridia

The cellulase activity on carboxymethyl cellulose and crystalline cellulose (Cellulone; Weyer-haeuser) were determined as described below. Disodium-2,2'-bicinchonate was purchased from Pierce. Glucose, cellobiose, and cellotriose will be purchased from Sigma, purified oligogalacturonic acid was isolated. Aqueous stock solutions of each carbohydrate were prepared at 0.25 mM. Assay solution and protocol will be prepared as described previously (Doner *et al*., 1991). The tube was then cooled down before the reading (Perkin-Elmer, Norwalk, CT, Model LC-55 variable wavelength uv-vis detector) at 560 nm.

Solvent producing ability will be studied by inoculate 1 ml of bacterial culture into modified MS media. Avicel was added to the media as the carbon source. They were grown anaerobically at 37°c (85% N_2 10% H_2; Coy Labolatory Products, Inc., Ann Arbor, Mich.) for 96 hrs. Solvent products in each species were kept and then analyzed with GC (Shimadzu Model Gc 7AG), Porapak Q80-100 mesh column.

CHAPTER IV
RESULTS AND DISCUSSION

Results were divided into 4 parts. Each part has relevant topic together. Discussion is included in each part.

4.1 First two main enzyme domains: Cel9X and Cel9C and relation to *C. acetobutylicum* ATCC824 inactive cellulosome

4.2 Cel48A, suspect and evidences for malfunction cellulosome

4.3 Ineffective *cel*48A gene deletion

4.4 Screening: alternative ways for improve sustainable strains in process

4.1 First two main enzyme domains: Cel9X and Cel9C and relation to *C. acetobutylicum* inactive cellulosome.

Clostridium acetobutylicum ATCC 824 produces a multicomponent protein complex (with an apparent molecular weight of 665 kDa), or cellulosome, unable to hydrolyze crystalline cellulose. To understand why this cellulosome is inactive, two of the major components of the cellulosome, Cel9C and Cel9X, were characterized from a biochemical point of view. The *celC* and *celX* sequences encode polypeptides containing respectively 713 and 878 amino acids with calculated molecular mass of 79700 and 95902 Da. Both proteins contain a catalytic domain belonging to the family 9 and a dockerin module but while cel9C has a carbohydrate-binding module (CBM) of the family 3c, the CBM of cel9X is of the family 4. The recombinant form of the Cel9C and Cel9X cellulases, tagged by a C-terminal histidine tail, were overexpressed in *Escherichia coli* and purified by affinity chromatography on a Ni-nitrilotriacetic acid column. The recombinant Cel9C and Cel9X were active on carboxymethylcellulose, amorphous cellulose but also on crystalline cellulose. Furthermore, the ability of Cel9C and Cel9X to interact with the full size CipA scaffolding protein was demonstrated by binding assays using surface plasmon resonance (SPR). Cel9C and cel9X bound to the CipA cohesin domains with a K_D of respectively 1.24×10^{-7} M and 0.97×10^{-7} M. These results strongly suggest that the lack of activity of the *C. acetobutylicum* cellulosome on crystalline cellulose might be due to a deficiency in Cel48A, the most abundant cellulase of the cellulosome.

4.1.1 Two of the three major cellulases of the cellulosome of *Clostridium acetobutylicum* are active on crystalline cellulose

Although the solvent producing strain *Clostridium acetobutylicum* can utilize all the sugars resulting from cellulose (Compere *et al*., 1979) and hemicellulose (Dunning *et al*., 1945; Nakhamanovich *et al*., 1959) hydrolysis and degrades polymers such as starch or xylan, it is not able to hydrolyze and grow on crystalline cellulose (Lee *et al*., 1985).

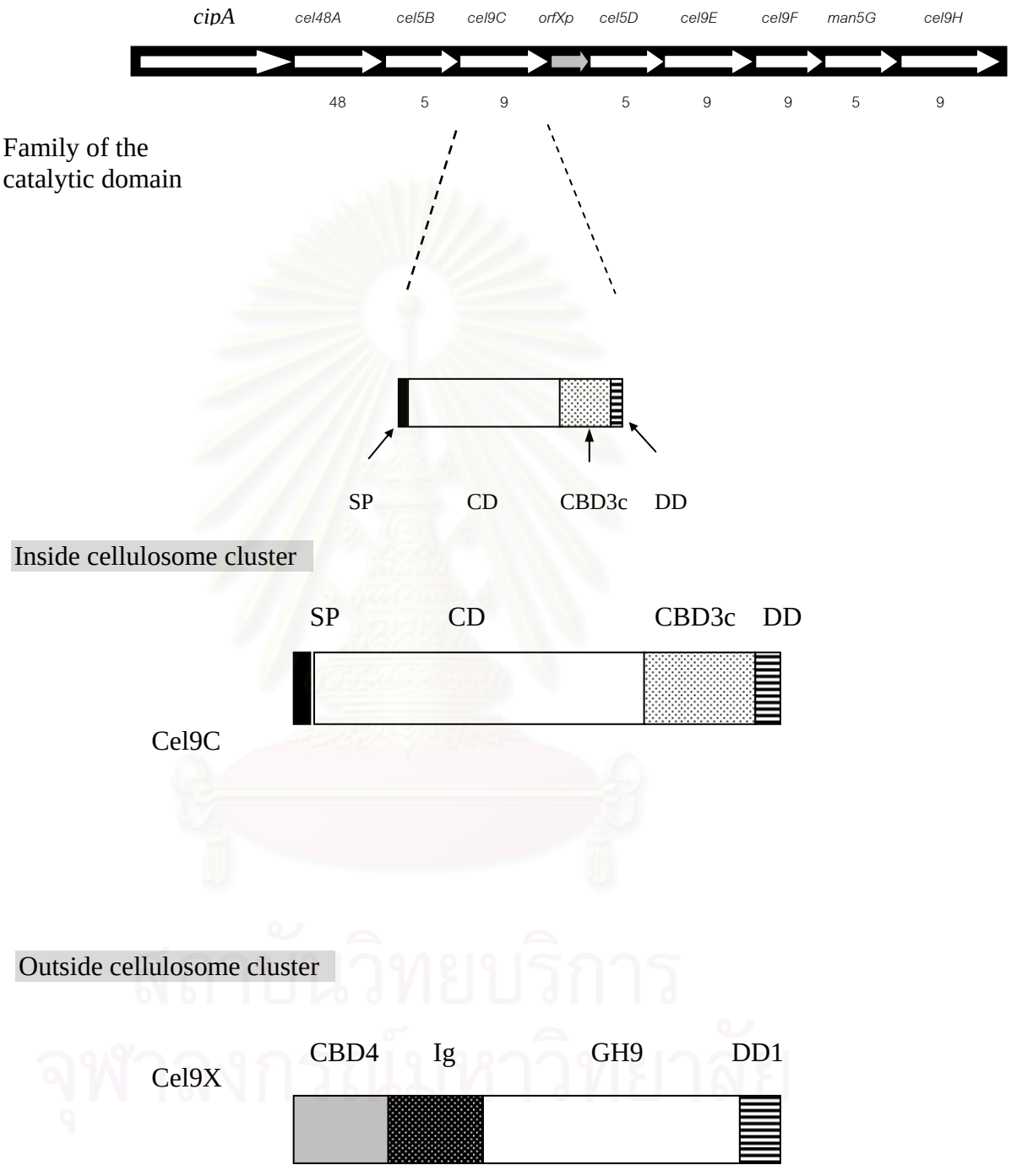


Figure 33. Cellulosomal gene cluster of *C. acetobutylicum*, and the domain structure of Cel9C. Symbols : ■, signal peptide (SP); □, catalytic domain (CD); ⊠ , CBD3c; ⊟ , dockerin domain (DD). Domain structure of Cel9X ▣ , CBD4; ▩, Ig; □ , Glycosyl hydrolases family 9.

Cellulosome, a macromolecular complex for cellulose degradation, has been genetically and biochemically characterized in four *Clostridium* species including *C. cellulolyticum* (Gal *et al*., 1997), *C. cellulovorans* (Doi *et al*., 1994), *C. josui* (Kakiuchi, 1998) and *C. thermocellum* (Lamed *et al*., 1988), but not in *C. acetobutylicum*. Surprisingly, the genome sequencing of *C. acetobutylicum* ATCC 824 (Nölling *et al*., 2001) has revealed the presence of a large cellulosomal gene cluster. Sequence analysis revealed that this cluster contains the genes for the CipA scaffolding protein, the Cel48A cellulase, several cellulases of family 5 and 9, the Man5G mannanase, and a hydrophobic protein, OrfXp. Genetic organization of this large cluster is very similar to those of the mesophilic *C. cellulolyticum* and *C. cellulovorans* (Bélaïch *et al*., 1997, Tamura *et al*., 2000). As *C. acetobutylicum*, is unable to grow on cellulosic substrates, the existence of a cellulosomal gene cluster in the genome raises questions about its expression, function and evolution. In an earlier work, we have established that *C. acetobutylicum* can produce a cellulosome with an apparent molecular weight of about 665 kDa (Sabathé *et al*., 2002). Biochemical and immunochemical analyses of the cellulosomal components revealed the existence of at least four subunits including the CipA scaffolding protein and three major cellulases, Cel48A, Cel9X, and Cel9C cellulases.

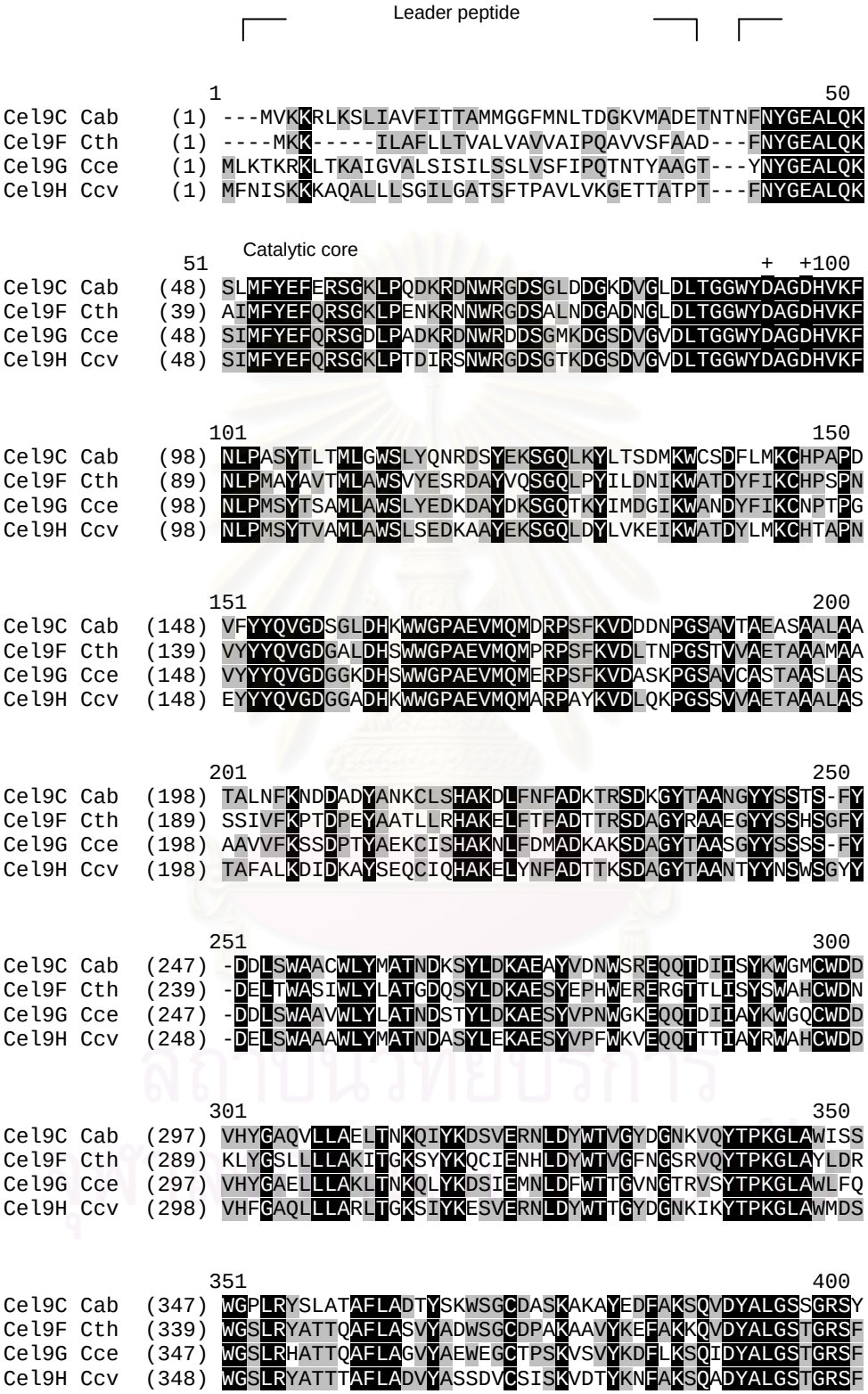
Our ultimate goal is to elucidate the relationships between the gene sequence data and the structure/function of the cellulosome system from *C. acetobutylicum*. We previously characterized the CipA scaffolding protein containing five cohesin domains (Sabathé and Soucaille, 2002). The role of cohesin domains as receptors for catalytic subunits in the assembly of the cellulosome has been clearly defined by several authors (Kruus *et al*., 1995; Pagès *et al*., 1996; Salamitou *et al*., 1994). In the present study, we described the purification and the biochemical characterization of Cel9C and Cel9X, two of the three major cellulases of the cellulosome of *C. acetobutylicum*. We demonstrated that they are both active on crystalline cellulose and that they can form a complex with the CipA scaffolding protein, suggesting that the lack of activity of the *C. acetobutylicum* cellulosome is linked to Cel48A the major cellulase of the cellulosome.

4.1.2 Sequence analysis of Cel9C and Cel9X

The *celC* sequence (CAC0913) identified in the large cellulosomal gene cluster (Fig. 33) (Sakon *et al*., 1997), encodes a polypeptide containing 713 amino acids with a calculated molecular mass of 79700 Da. The N-terminal region contained a typical prokaryotic leader peptide. A putative signal-peptide-processing site was detected between the A-32 and D-33, residues leading to a mature protein of 681 aa with a theoretical molecular mass of 76122 Da.

Analysis of the deduced amino acid sequence of Cel9C revealed a multidomain architecture consisting of a glycosyl hydrolase family 9 catalytic domain (GH9) followed by a family 3 carbohydrate-binding module (CBM) and a duplicated sequence corresponding to a type I dockerin module . The presence of the dockerin module is a clear indication that Cel9C can dock to the cohesin modules of the CipA scaffolding protein, and explain why Cel9C was found to be one of the component of the cellulosome complex (Sabathé *et al*., 2002). Comparison with sequences in the EMBL-GenBank database showed a high degree of identity between the Cel9C catalytic domain and those of the *C. cellulolyticum* Cel9G (69%), *C. cellulovorans* Cel9H (68%) and *C. thermocellum* Cel9F (64%) cellulases. Moreover, the residues involved in the catalysis of family 9 glycosyl hydrolases were conserved in Cel9C. We can speculate that the conserved Glu462 residue is the catalytic acid/base and the conserved Asp90 or Asp93 is the base.

The catalytic core is followed by a CBM3. This domain displays 70, 55 and 53% of identity with the CBM3 of respectively *C. cellulovorans* Cel9H, *C. cellulolyticum* Cel9G and *C. thermocellum* Cel9F cellulases. Structure-based sequence alignment shown in Fig. 33, revealed that the structure of the CBM3 from the *C. acetobutylicum* Cel9C cellulase seems to be very similar to the previously elucidated family 3 CBM, including a nine-stranded jelly-roll topology which exhibits distinctive structural element consistent with family 3 CBM (Shimon *et al*., 2001).


Leader peptide
Catalytic core
Cel9C Cab (1) ---MVKKRLKSLIAVFITTAMMGGFMNLTDGKVMADETNTNFNYGEALQK
Cel9F Cth (1) ----MKK-----ILAFLLTVALVAVVAIPQAVVSFAAD---FNYGEALQK
Cel9G Cce (1) MLKTKRKLTKAIGVALSISILSSLVSFIPQTNTYAAGT---YNYGEALQK
Cel9H Ccv (1) MFNISKKKAQALLLSGILGATSFTPAVLVKGETTATPT---FNYGEALQK
Cel9C Cab (48) SLMFYEFERSGKLPQDKRDNWRGDSGLDDGKDVGLDLTGGWYDAGDHVKF
Cel9F Cth (39) AIMFYEFQRSGKLPENKRNNWRGDSALNDGADNGLDLTGGWYDAGDHVKF
Cel9G Cce (48) SIMFYEFQRSGDLPADKRDNWRDDSGMKDGSDVGVDLTGGWYDAGDHVKF
Cel9H Ccv (48) SIMFYEFQRSGKLPTDIRSNWRGDSGTKDGSDVGVDLTGGWYDAGDHVKF
Cel9C Cab (98) NLPASYTLTMLGWSLYQNRDSYEKSGQLKYLTSDMKWCSDFLMKCHPAPD
Cel9F Cth (89) NLPMAYAVTMLAWSVYESRDAYVQSGQLPYILDNIKWATDYFIKCHPSPN
Cel9G Cce (98) NLPMSYTSAMLAWSLYEDKDAYDKSGQTKYIMDGIKWANDYFIKCNPTPG
Cel9H Ccv (98) NLPMSYTVAMLAWSLSEDKAAYEKSGQLDYLVKEIKWATDYLMKCHTAPN
Cel9C Cab (148) VFYYQVGDSGLDHKWWGPAEVMQMDRPSFKVDDDNPGSAVTAEASAALAA
Cel9F Cth (139) VYYYQVGDGALDHSWWGPAEVMQMPRPSFKVDLTNPGSTVVAETAAAMAA
Cel9G Cce (148) VYYYQVGDGGKDHSWWGPAEVMQMERPSFKVDASKPGSAVCASTAASLAS
Cel9H Ccv (148) EYYYQVGDGGADHKWWGPAEVMQMARPAYKVDLQKPGSSVVAETAAALAS
Cel9C Cab (198) TALNFKNDDADYANKCLSHAKDLFNFADKTRSDKGYTAANGYYSSTS-FY
Cel9F Cth (189) SSIVFKPTDPEYAATLLRHAKELFTFADTTRSDAGYRAAEGYYSSHSGFY
Cel9G Cce (198) AAVVFKSSDPTYAEKCISHAKNLFDMADKAKSDAGYTAASGYYSSSS-FY
Cel9H Ccv (198) TAFALKDIDKAYSEQCIQHAKELYNFADTTKSDAGYTAANTYYNSWSGYY
Cel9C Cab (247) -DDLSWAACWLYMATNDKSYLDKAEAYVDNWSREQQTDIISYKWGMCWDD
Cel9F Cth (239) -DELTWASIWLYLATGDQSYLDKAESYEPHWERERGTTLISYSWAHCWDN
Cel9G Cce (247) -DDLSWAAVWLYLATNDSTYLDKAESYVPNWGKEQQTDIIAYKWGQCWDD
Cel9H Ccv (248) -DELSWAAAWLYMATNDASYLEKAESYVPFWKVEQQTTTIAYRWAHCWDD
Cel9C Cab (297) VHYGAQVLLAELTNKQIYKDSVERNLDYWTVGYDGNKVQYTPKGLAWISS
Cel9F Cth (289) KLYGSLLLLAKITGKSYYKQCIENHLDYWTVGFNGSRVQYTPKGLAYLDR
Cel9G Cce (297) VHYGAELLLAKLTNKQLYKDSIEMNLDFWTTGVNGTRVSYTPKGLAWLFQ
Cel9H Ccv (298) VHFGAQLLLARLTGKSIYKESVERNLDYWTTGYDGNKIKYTPKGLAWMDS
Cel9C Cab (347) WGPLRYSLATAFLADTYSKWSGCDASKAKAYEDFAKSQVDYALGSSGRSY
Cel9F Cth (339) WGSLRYATTQAFLASVYADWSGCDPAKAAVYKEFAKKQVDYALGSTGRSF
Cel9G Cce (347) WGSLRHATTQAFLAGVYAEWEGCTPSKVSVYKDFLKSQIDYALGSTGRSF
Cel9H Ccv (348) WGSLRYATTTAFLADVYASSDVCSISKVDTYKNFAKSQADYALGSTGRSF

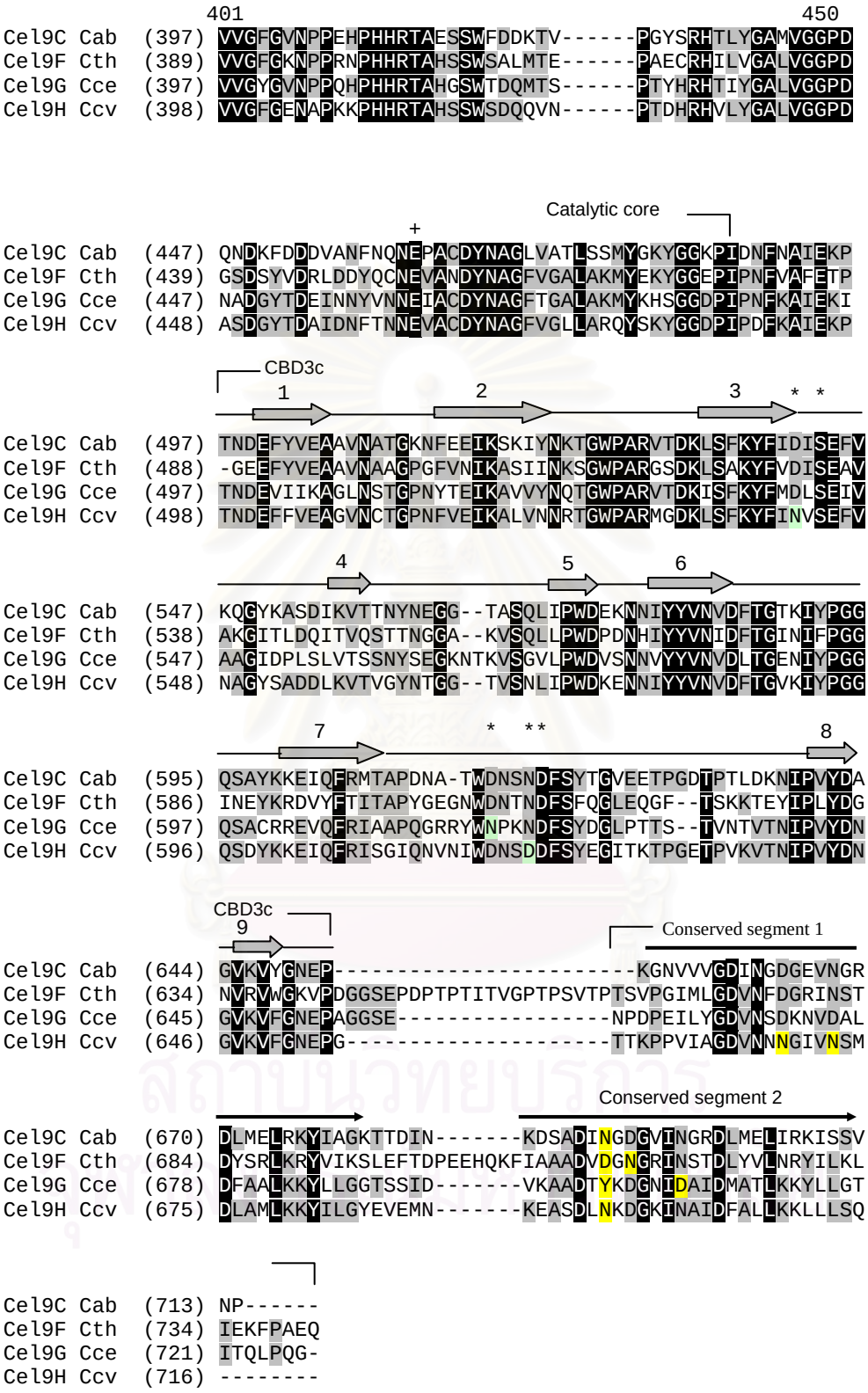


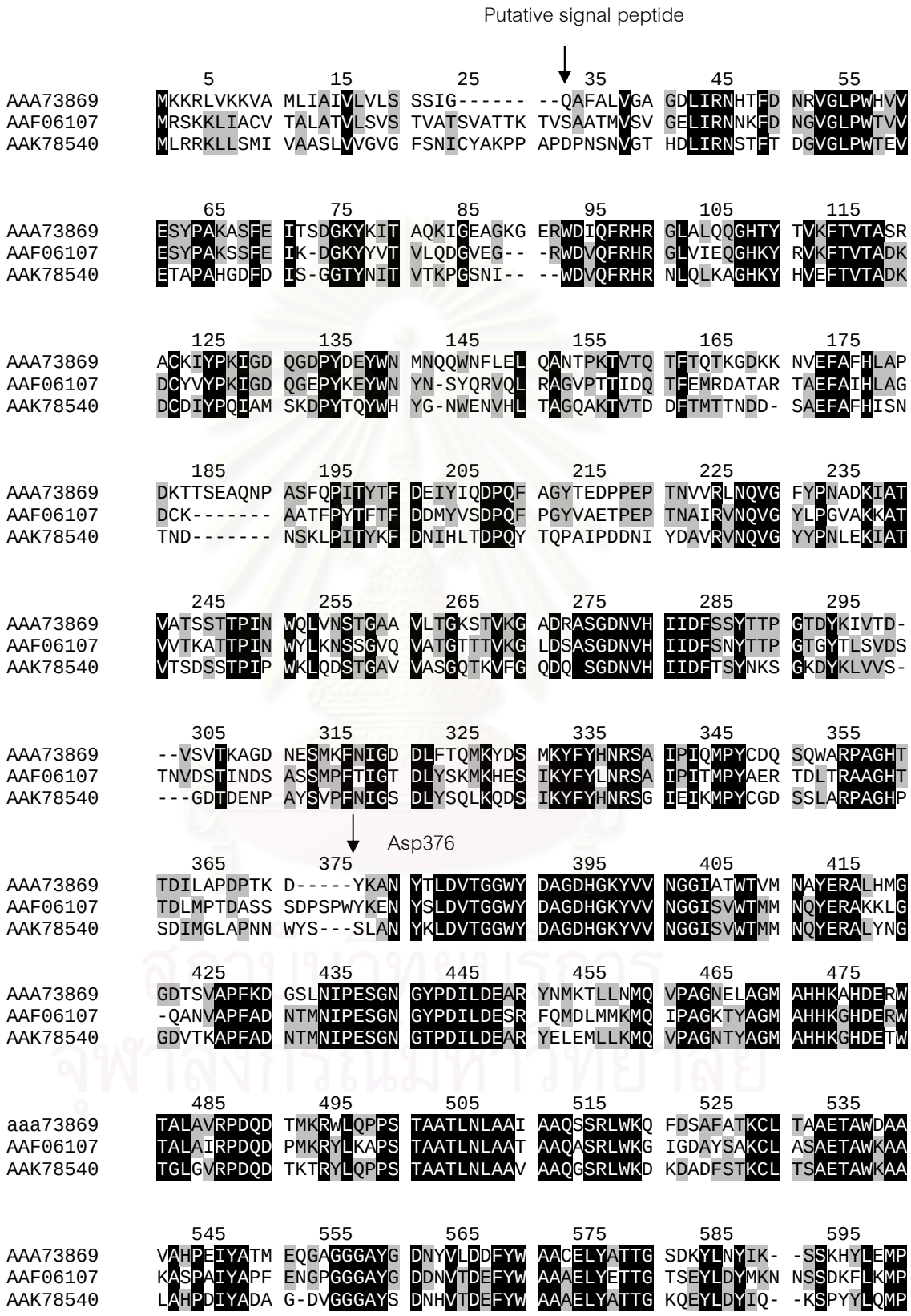
Figure 34. Multiple alignments of selected cellulases with Cel9C from *C. acetobutylicum*. The sequences shown in the figure are from the following proteins:

Cel9F from *C. thermocellum*, Cel9G from *C. cellulolyticum* and Cel9H from *C. cellulovorans*. Residues which are identical are on a solid background, whereas residues which are similar or identical in at least 50% of the proteins are dark shaded background. (*) Residues involves in calcium binding. (+) residues involves in the putative catalytic. EF-Hand motifs are shown in yellow. Horizontal arrows indicated duplicated segments. Secondary-structural elements are indicated and enumerated.

The *cel9X* sequence (CAC 0561) identified outside from the cellulosomal gene cluster (Fig. 34) (Nölling *et al*., 2001), encodes a polypeptide containing 878 amino acids with a calculated molecular mass of 95902 Da. The N-terminal region contained a typical prokaryotic leader peptide. A putative signal-peptide-processing site was detected between the D-33 and P-34, residues leading to a mature protein of 845 aa with a theoretical molecular mass of 92393 Da.

Analysis of the deduced amino acid sequence of Cel9X revealed a multidomain architecture consisting of a family 4 carbohydrate-binding module (CBM), an immunoglobulin domain and a glycosyl hydrolase family 9 catalytic domain (GH9) followed by a duplicated sequence corresponding to a type I dockerin module (Fig. 33). The CBM4 domain displays 50, 46 and 30% of identity with the CBM4 of respectively *C. cellulovorans* EngK, *C. cellulolyticum* Cel9E and *C. thermocellum* CbhA cellulases. Structure-based sequence alignment shown in Fig. 34, revealed that the structure of the CBM4 from the *C. acetobutylicum* Cel9X cellulase seems to be very similar to the previously elucidated family 4 CBM, including a nine-stranded jelly-roll topology which exhibits distinctive structural element consistent with family 3 CBM . The CBM4 is followed by an immunoglobulin domain.

The Cel9X catalytic domain showed a high degree of identity with those of the *C. cellulolyticum* Cel9E (67%), and *C. cellulovorans* EngK (87%) cellulases (Fig. 35). Moreover, the residues involved in the catalysis of family 9 glycosyl hydrolases were conserved in Cel9X (Lee *et al*., 1985). We can speculate that the conserved Glu734 residue is the catalytic proton donor and the conserved Asp376 or Asp379 are the bases. The presence of a dockerin module is a clear indication that Cel9X can dock to the cohesin modules of the CipA scaffolding protein, and explain why Cel9X was found to be one of the component of the cellulosome complex (Sabathé *et al*., 2002).


Putative signal peptide
Asp376
AAA73869
AAF06107
AAK78540
aaa73869

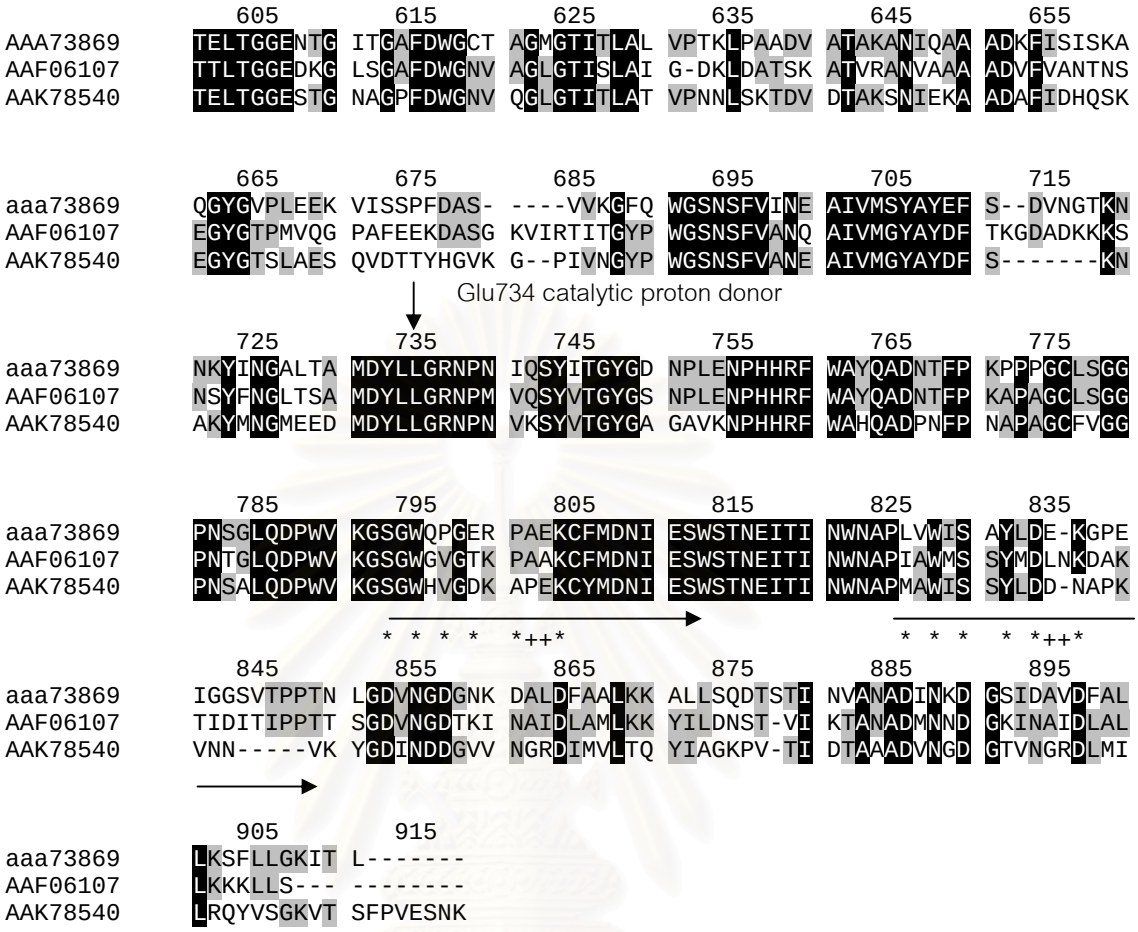


Figure 35. Multiple alignments of selected cellulases with Cel9X from *C. acetobutylicum***.** Arterisks indicate putative catalytic residues. (*) Residues involves in calcium binding and (+) residues involves in species specific of cohesion/dockerin recognition. Horizontal arrows indicated duplicated segments. Secondary-structural elements are indicated and enumerated.

The sequences shown in the figure are from the following proteins: Cel9E from *C. cellulolyticum* and EngK from *C. cellulovorans*. Residues which are identical are on a solid background, whereas residues which are similar or identical in at least 50% of the proteins are dark shaded background.

4.1.3 Production and purification of Cel9C and Cel9X protein

Recombinant Cel9C and Cel9X were purified from the periplasmic fraction of *E. coli* BL21(DE3) harboring pET-C or pET-X. Purification was performed in a one-step process by using binding properties of the His-tag. Initial attempts to express Cel9C or Cel9X in *E. coli*, encountered problem of inclusion bodies when induction was performed at 37°C. By optimizing the induction conditions, it was possible to produce soluble proteins for both cellulases. After purification of Cel9C on a Ni-nitrilotriacetic acid column, a single band with an estimated molecular weight of 75 kDa was detected (Fig. 35) by SDS-PAGE, in good agreement with the 76225 Da theoretical value. Using the same purification protocol, Cel9X gave a single band in SDS-PAGE with an estimated molecular weight of 94 kDa. in good agreement with the 95902 Da theoretical value. CipA expression/purification was performed from the cytoplasmic fraction of *E. coli* BL21(DE3) harboring pET-CipA (Sabathé and Soucaille, 2002) as previously described. Respectively, 1.5, 1 and 4 mg of pure Cel9C, Cel9X and CipA were produced.

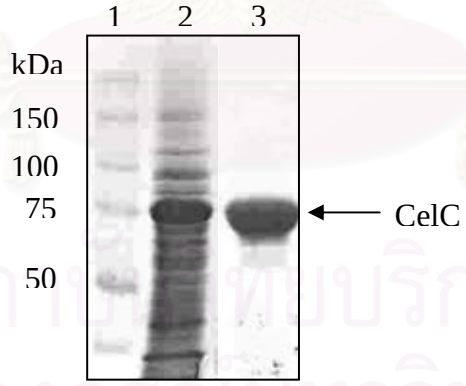


Figure 36. SDS-PAGE analysis of induced *E. coli* BL21(DE3)[pET22-9C-1]. Lane 1: marker protein with molecular masses indicated on the left. Lane 2: Soluble crude fraction of strain BL21(DE3)[pET22-9C-1]. Lane 3: purified Cel9C. The arrow indicates the position of Ce9lC at 75 kDa

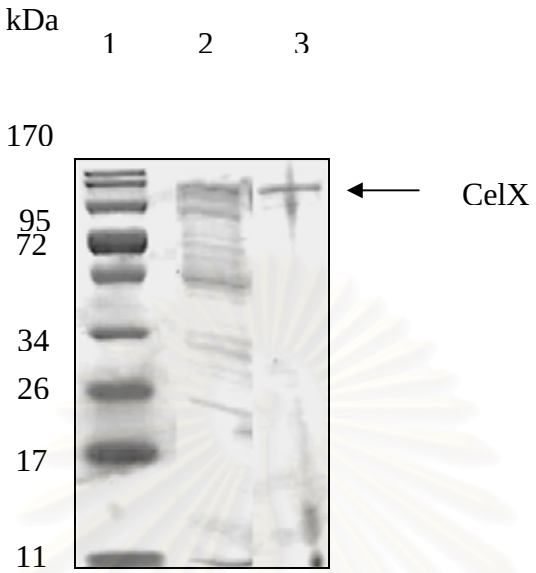


Figure 37. SDS-PAGE analysis of induced *E. coli* BL21(DE3)[pET22-9X-C1]. Lane 1: marker protein with molecular masses indicated on the left. Lane 2: Soluble crude fraction of strain BL21(DE3)[pET22-9X-C1]. Lane 3: purified Cel9X. The arrow indicates the position of Ce9lX at 95 kDa

4.1.4 Catalytic properties of Cel9C and Cel9X

The specific activities of Cel9C and Cel9X were studied, using CMC, PASC and Avicel as substrates. The degradation activities were followed for at least two hours. The results are summarized in Table 1 and compared with the activities of homologous Cel9G and Cel9E from *C. cellulolyticum*, which are known to act efficiently on crystalline cellulose (Bélaïch *et al*., 2002; Gal *et al*., 1997). Cel9C and Cel9X exhibited similar level of activities than respectively Cel9G and Cel9E from *C. cellulolyticum* on CMC and PASC but also on crystalline cellulose.

Table 9. Activity of Cel9C and Cel9X from *C. acetobutylicum* on various substrates and comparison with activities of previously characterized family 9 cellulases from *C. cellulolyticum*.

Substrate	Activity (IU/µM)			
	C. acetobutylicum		*C. cellulolyticum*	
	Cel9C	Cel9X	Cel9G	Cel9E
CMC	1217	30.2	1170	13.5
PASC	25.8	36.7	38	42.5
Avicel	2.6	2.1	5	5.8

4.1.5 Biomolecular interaction between Cel9C and Cel9X with the CipA scaffolding protein

The interaction between Cel9C and CipA was studied first using surface plasmon resonance (SPR). CipA was covalently immobilized on a dextrane sensor chip and Cel9C was used as the ligand. When the experiments were performed in a $CaCl_2$ free buffer, no binding was detected suggesting that $CaCl_2$ was required for the interaction, as it was demonstrated for *C. thermocellum* (Yaron *et al*., 1995) and *C. cellulolyticum* (Pagès *et al*., 1997). In addition to this experiment, sequence comparisons strongly suggest that the dockerin domain of Cel9C contains two EF-Hand calcium binding sites, located at the beginning of each conserved segment (Chauvaux *et al*., 1990) (Fig. 34). To check the hypothesis of a $CaCl_2$ dependent Cel9C/CipA interaction, a calcium dose response experiment was done. Results shown in Fig. 37 indicate an increase of the resonance units (RU) directly linked to a higher level of bound Cel9C when the $CaCl_2$ concentration was shifted from 1 µM to 100 µM. This confirms that the cohesin/dockerin interaction between CipA and Cel9C is strictly calcium dependent. A stepwise increase in $CaCl_2$ concentration to 1 mM had no effect on the RU response. Surprisingly, at higher $CaCl_2$ concentration (10 mM) a decrease of the RU response was observed (data not shown).

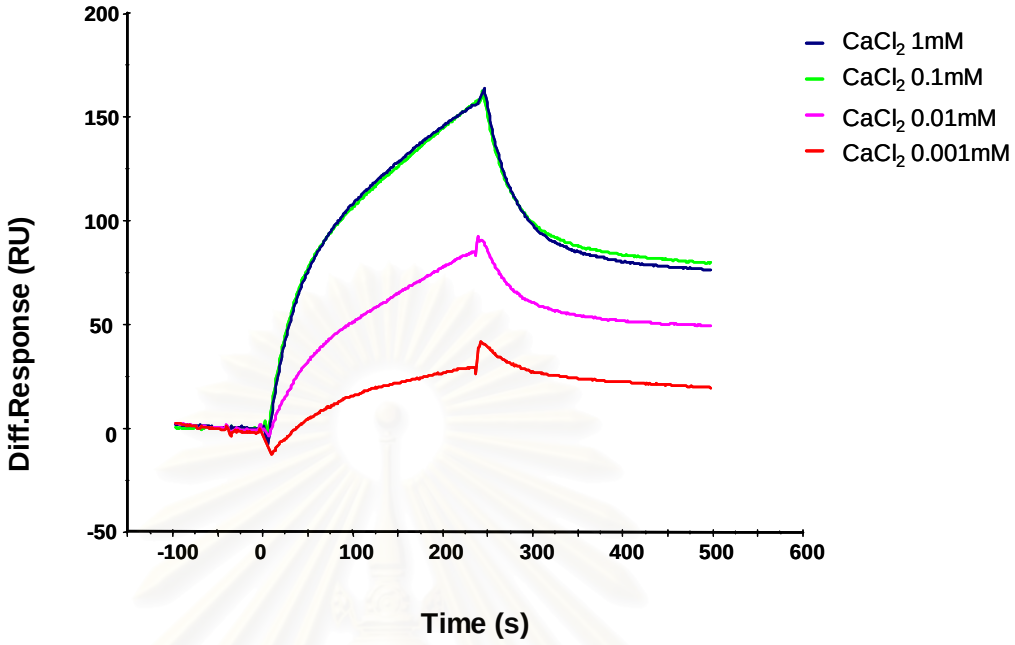


Figure 38. Effects of calcium on the binding of CelC fusion protein to immobilized CipA protein (Sabathé *et al*., unpublished data)

Cel9C fusion protein, at a concentration of 5 µM, was diluted in increasing concentrations of $CaCl_2$ ranging from 1 µM to 1 mM before injection on immobilized CipA protein. CelC mixed to $CaCl_2$ was injected at a flow rate of 20 µl/min, exposed to the surface for 240 s (association phase) followed by a 180-s flow running during which the dissociation occurred. Amounts of immobilized CipA protein were 2250 RU on flow cell 2. Sensorgrams are representative of specific interactions (differential response) where non-specific binding that occured on flow cell 1 was deduced from binding that occurred on flow cell 2. Results are expressed as resonance units (RU) as a function of time in seconds (figure 38).

In the second set of experiments, all the interactions between Cel9C and Cel9X with CipA were performed in the presence of 100 µM $CaCl_2$. Five sensorgrams were obtained for five different Cel9C and Cel9X concentrations. The BIAevaluation software was used to analyze each sensorgram and to calculate the association and dissociation constants (k_{on} and k_{off}). The results reported in Table 10 show that Cel9C and Cel9X bound to the CipA cohesin domains with a K_D of respectively $1.24\ 10^{-7}$ M and $0.97\ 10^{-7}$ M.

Table 10. Biomolecular interaction between Cel9C and Cel9X with the CipA scaffolding protein (Sabathé *et al*., unpublished data).

	Cel9C	**Cel9X**
k_{on}	$3.85+0.3 \times 10^{4}$	$4.62+0.4 \times 10^{4}$
k_{off}	$4.75+0.5 \times 10^{-3}$	$4.5+0.4 \times 10^{-3}$
K_D	1.24×10^{-7}	0.97×10^{-7}

4.1.6 Discussion

Based on sequence comparisons, the Cel9C and Cel9X cellulases from *C. acetobutylicum*, belongs to the family 9 of glycosyl hydrolases. This is the most taxonomically diverse family, containing cellulases from plants, bacteria, slime molds and termites, but not from fungi. In the cellulosomal cluster from *C. acetobutylicum*, five genes, *cel9C, cel9E, cel9F, cel9H* and *cel9X* code for GH9 enzymes (Sabathé *et al*., 2002). The Cel9C, Cel9E, Cel9H and Cel9X corresponding proteins are multimodular. Cel9C and Cel9E contain a family 3c CBM and share 46% identity, whereas Cel9H and Cel9X contain respectively a family 3b and 4 CBM. The cellulosomal gene clusters from *C. cellulolyticum* and *C. cellulovorans* contain respectively five and four genes coding for GH9 enzymes (Bélaïch *et al*., 1997, Tamaru *et al*., 2000).

Cel9G is a multidomain endoglucanase of the active cellulosome of *C. cellulolyticum* (Gal *et al*., 1997) that is very homologous to Cel9C. The most highly conserved region is in the catalytic core with 69% of identity. Furthermore, sequence alignment revealed that the residues involved in catalysis are perfectly conserved (Fig. 33). Concerning the CBM3a of Cel9C, the general fold of two antiparallel β-sheets with the topology of a jelly role β-sandwich seems to be respected. The two Ca^{2+} binding sites that are present in all members of the family 3 (Tormo *et al*., 1996) are also conserved in Cel9C. In agreement with this high degree of homology, Cel9C from *C. acetobutylicum* and Cel9G from *C. cellulolyticum* (Gal *et al*., 1997), had also very similar activities on both CMC (1217 UI/µM against 1170 UI/µM), PASC (25.8

UI/µM against 38 UI/µM) and avicel (2.6 UI/µM against 5 UI/µM). The good activity of Cel9C on crystalline cellulose clearly indicates that if the cellulosome of *C. acetobutylicum* is inactive on this substrate (Sabathé *et al*., 2002), it is not due to a problem with Cel9C.

In the same way, Cel9E is a multidomain processive endoglucanase of the active cellulosome of *C. cellulolyticum* that is homologous to Cel9X. Sequences alignment are appeared to be involved in catalysis are totally conserved which has 67% homology in conserved region (Fig. 33). CBM4 of Cel9X are shown antiparallel β-sheets with the topology of a jelly role β-sandwich like in CBM3 of Cel9C. In comparation with high degree of homology, Cel9X from *C. acetobutylicum* and Cel9E from *C. cellulolyticum* (table 9), had also higher activities on CMC (30.2 UI/µM against 13.5 UI/µM), lower activities on PASC (36.7 UI/µM against 42.5 UI/µM) and avicel lower activities (2.1 UI/µM against 5.8 UI/µM). Even the activity over crystalline cellulose does not present the same. This also indicated that Cel9X still active on crystalline cellulose and does not related to malfunctional of *C. acetobutylicum* cellulosome.

It is now clearly established that the cohesin/dockerin interaction plays a key role in the assembly of the cellulosome (Pagès *et al*., 1997, Yaron *et al*., 1995). Cel9C and Cel9X contain a C-terminal dockerin domain. The presence of a dockerin suggests that Cel9C and Cel9X can dock to the CipA five cohesin domains. In order to understand the formation of the cellulosome of *C. acetobutylicum,* it was necessary to determine the affinity of its components for the scaffolding protein. The entire CipA was purified and the apparent equilibrium dissociation constants of the Cel9C-CipA and Cel9X-CipA complexes were measured. No binding of Cel9C to CipA was detected in the absence of $CaCl_2$, suggesting a strictly calcium dependent cohesin/dockerin interaction as it was observed for the cellulosomes of *C. thermocellum* and *C. cellulolyticum* (Bélaïch *et al*., 2002, Yaron *et al*., 1995). Most of the experiments reported so far in the literature were done in the presence of 10 mM $CaCl_2$ (Fierobe *et al*., 1999). In the case of *C. acetobutylicum*, the association between Cel9C and CipA increased with the $CaCl_2$ concentration but at concentration higher than 1 mM a negative effect on binding started to be observed. At a 100 µM $CaCl_2$

concentration, the apparent dissociation constants, K_d, for the Cel9C-CipA and Cel9X-CipA complexes were 1.24 10^{-7} M and 10^{-7} M respectively. This value corresponds to an average global affinity of the whole CipA five cohesin domains. It differs strikingly from the one obtained by Pagès et al., for the *C. cellulolyticum* Cel5A-miniCipC complexes (Pagès *et al*., 1997): K_d for the CelA-miniCipC1 and CelA-miniCipC8 complexes were respectively 2.1 x 10^{-10} M and 2.6 10^{-10} M. Furthermore, the K_d between the wild-type endoglucanase D from *C. thermocellum* and the seventh cohesin domain of CipA from *C. thermocellum* was studied by isothermal titration calorimetry (Miras *et al*., 2002), and was found to be 5 x 10^{-10} M.

In comparison with these data, we can conclude that Cel9C and Cel9X from *C. acetobutylicum* bind to their scaffolding protein with lower affinity than the cellulases from *C. cellulolyticum* and *C. thermocellum*. These results show that the cellulosome produced by *C. acetobutylicum* should be less stable than those of *C. thermocellum* and *C. cellulolyticum*. Nevertheless, as Cel9C and Cel9X have been demonstrated (Sabathé *et al*., 2002) to be two of the main components of the *C. acetobutylicum* cellulosome, it cannot be the explanation for its non functionality.

It is known that during evolution of bacterial genomes, when selection is not strong enough, genes are partially lost by large deletions or inactivated by point mutations or insertion. As *C. acetobutylicum* has been maintained under laboratory conditions for many years, it is possible that this strain has evolved to a non cellulolytic one by point mutations in a gene encoding a key component of the cellulosome. Further experiments with Cel48A, the third major cellulase of the *C. acetobutylicum* cellulosome should allow a better understanding of the non-functionality of the cellulosome.

4.2 Cel48A, suspect and evidences for malfunction cellulosome

4.2.1 Modified a major cellulase of *Clostridium acetobutylicum* ATCC824 cellulosome to exhibit enhanced activity on crystalline cellulose

The structure polysaccharides cellulose and hemicellulose account for greater than 50% of plant biomass and are consequently the most abundant organic materials on earth. Considerable potential exists for increasing the effective utilization of plant biomass in biological process that involve the enzymatic hydrolysis of cellulose and hemicellulose, and the subsequent fermentation of liberated sugars to yield products ranging from the fine chemical to bulk organics and biofuels (Chen *et al*., 1993). The ability to partially or completely degrade cellulose or hemicellulose has been attributed to several clostridia, by extracellular cellulases subunit called *cellulosome*. *Clostridium* sp. has several advantages for fuel fermentation from biomass due to several factors. The cellulolytic and ethanogenic nature are allowing saccharification and fermentation in a single step. It has amenability to cocultire with other ethanol-producing and pentose-fermenting organisms. Moreover, the anaerobic nature negating the need for expensve oxygen transfer lets the process and fermentor tank simple without ventilator.

The Clostridia cellulosome is an extracellular supramolecular machine that can efficiently degrade crystalline cellulosic substrates and associated plant cell wall polysaccharide (Shoham *et al*., 1999). This complexes produced by cellulolytic anaerobic bacteria. To date, seven genes coding for the components of *Clostridium cellulolyticum* cellulosome have been identified. *celA* (Faure *et al*., 1989) and *celD* (Shima *et al*., 1991) were described as monocistronic transcription units. More recently, a cluster of *cel* genes was described (Bagnara-Tardif *et al*., 1992) including ORF1, *celC, celG* and *celE*, with *celC* and *celG* associated in a polycistronic transcription unit. It has recently been reported that celF of cellulase gene of *Clostridium cellulolyticum* is active on degraded swallen avicel more efficiently than substitued soluble CM-cellulose or crystalline avicel (Reverbel-Leroy *et al*., 1996). These models of mesophilic clostridial cellulosomes have a large cluster which is very similar to that of *Clostridium acetobutylicum* ATCC824. In contrast, *C. acetobutylicum* is unable to grow on cellulosic substrate even it can produce high molecular mass cellulosomal complex (Sabathe *et al*., 2002). They conclude that *C.*

acetobutylicum can produce cellulosome but inactive. This may relates to the major catalytic subunit of the cellulosome which classified as a member of family 48 glycosyl hydrolases. Although it is the most abundant catalytic subunit of the cellulosome, its low or complete lack of activities on CMC and other synthetic substrates explains why it had been difficult to clone it gene. A recent study on *Ruminococcus albus* mutants that are defective in cellulose degradation found that all mutants failed to produce family 48 and family 9 enzymes. It can be indicate that this fimily 48 enzyme are crucial for bacterial cellulose system (Demain *et al*, 2005).

The absolute goal of this study is modify *cel*48A by direct mutagenesis method and engineering to restoring the activity protein. In previous work (Fierobe, H.-P., Mingardon, F., Mechaly, A., Bélaïch, A., Rincon, M., Pagès, S., Lamed, R., Tardif, C., Bélaïch, J.-P., Bayer, E. A. (2005), they design and produce active cellulosome chimeras complexes between *C. cellulolyticum* and *C. thermocellum* by construct high-affinity species specific cohesion-dockerin molecules. Using this strategy, appropriate dockerin-containing enzymes could be assembled precisely and by design into a desired complex. Compared with the mixture of free cellulasees, the resultant cellulosome chimeras exhibited enhanced synergistic action on crystalline cellulose (Fierobe et al., 2001). Nowsaday, we try to construct new hybrid molecules of *cel*48A dockerin domain and *cel*48F catalytic domain. Chimeric protein molecules were designed by use gene for express Cel48F catalytic domain from plasmid pET_{Fc} (Reverbel-leroy *et al*., 1996) with modification for correct protein expression. We modify this plasmid by replace the Cel48F dockerin domain with Cel48A dockerin domain. This dockerin domain is appropriate for *C. acetobutylicum* ATCC824 cellulosome assembly. The replacement is very important because the construct of cellulosome in *Clostridium* sp. needs species-specific dockerin domain. Finally, this hybrid cellulose enzyme was detected activity on various cellulosic substrates to examine their capacity to degrade microcrystalline cellulose.

4.2.2 Amino acid sequence analysis of the recombinant protein

The recombinant protein sequence of Cel48A protein was analysed. Three mutation points were identified in amino acid sequence. For this reason, we hypothesize that this point mutation is the cause of defective cellulosome of *C.*

acetobutylicum. Previous report showed that there is a high level of similarity between the cellulosome gene clusters of *C. acetobutylicum* and *C. cellulolyticum*, suggesting a close taxonomic relatedness (Sabathé *et al*., 2002). They presented interesting data that the highest amino acid sequence homologies between the two species are obtained with Cel48F and Cel9E, the two major cellulases of the *C. cellulolyticum* cellulosome, and respectively Cel48A and Cel9X from *C. acetobutylicum*. Because of this point, we try to observe the homology between sufficient Cel48F of *C. cellulolyticum* and defective Cel48A of *C. acetobutylicum*. Cel48A shows 58% amino acid sequence similarity when compaired with Cel48F (Figure 38). This data is important for understand the function of those two cellulases in family 48 because of different activity on cellulolytic substrate.

```
Query: 4   ISKNFKKIMAVALTSTVIFGSLSGLLTNKVAAATTTDSSLKVDNAYTQRFETMYNKMHDA 63
           +SKNFK++ AVA+ + +   SLS + T  + AA++       +  Y  RFE+MY+K+ D
Sbjct: 1   MSKNFKRVGAVAVAAAM---SLSIMATTSINAASSP-----ANKVYQDRFESMYSKIKDP 52

Query: 64  NNGYFSKDGVPYHSVETFMVEAPDYGHETTSEAFSYYMWLEAMQGKITGNFSGVNTAWDT 123
            NGYFS+ G+PYHS+ET MVEAPDYGH TTSEA SYYMWLEAM G+ +G+F+G + +W
Sbjct: 53  ANGYFSEQGIPYHSIETLMVEAPDYGHVTTSEAMSYYMWLEAMHGRFSGDFTGFDKSWSV 112

Query: 124 AEKYMIPSHQDQPG--MDRYNASSPATYSPEWEDPSKYPSRMDQGAAKGQDPISDELKSA 181
            E+Y+IP+ +DQP   M RY+A+ PATY+PE++DPSKYPS +D     G+DPI+ +L SA
Sbjct: 113 TEQYLIPTEKDQPNTSMSRYDANKPATYAPEFQDPSKYPSPLDTSQPVGRDPINSQLTSA 172

Query: 182 YGTSDMYGMHWLMDVDNWYGFGNHEDGTSKNVYINTYQRGEQESVFETVPQPCWDAFKYG 241
           YGTS +YGMHW++DVDNWYGFG   DGTSK  YINT+QRGEQES +ET+PQPCWD  K+G
Sbjct: 173 YGTSMLYGMHWILDVDNWYGFGARADGTSKPSYINTFQRGEQESTWETIPQPCWDEHKFG 232

Query: 242 GKNGYLDLFTGDNSY-AKQAKYTDAPDADXXXXXXXXXXXXXXKEDGVDLSSIVGKASKM 300
           G+ G+LDLFT D    AKQ KYT+APDAD              KE G  +S+ VGKA+KM
Sbjct: 233 GQYGFLDLFTKDTGTPAKQFKYTNAPDADARAVQATYWADQWAKEQGKSVSTSVGKATKM 292

Query: 301 GDYLRYAMFDKYFRKIGNSTQAGNGKDSMHYLLSWYYAWGGSQNNDWSWKIGCSHSHFGY 360
           GDYLRY+ FDKYFRKIG  +QAG G D+ HYLLSWYYAWGG  ++ WSW IG SH+HFGY
Sbjct: 293 GDYLRYSFFDKYFRKIGQPSQAGTGYDAAHYLLSWYYAWGGGIDSTWSWIIGSSHNHFGY 352

Query: 361 QNPLTAWVLSTDSQFKPKSATGATDWAKSLTTQVDFYQWLQSSEGAIAGGASNSNHGRYE 420
           QNP  AWVLSTD+ FKPKS+ GA+DWAKSL  Q++FYQWLQS+EGAIAGGA+NS +GRYE
Sbjct: 353 QNPFAAWVLSTDANFKPKSSNGASDWAKSLDRQLEFYQWLQSAEGAIAGGATNSWNGRYE 412

Query: 421 AWPEGTATFDGMGYQEEPVYHDPGSNTWFGMQSWSMQRMAQYYYQSKDPKAKALLDXXXX 480
           A P GT+TF GMGY E PVY DPGSNTWFGMQ WSMQR+A+ YY++ D +AK LLD
Sbjct: 413 AVPSGTSTFYGMGYVENPVYADPGSNTWFGMQVWSMQRVAELYYKTGDARAKKLLDKWAK 472

Query: 481 XXXXXXXXNPNGAGTFEVPSKLSWTGQPDTW--TGSYTGNPNLHVNVDSYTTDIGXXXXX 538
                N +  GTF++PS + W GQPDTW  T  YTGN NLHV V +Y TD+G
Sbjct: 473 WINGEIKFNAD--GTFQIPSTIDWEGQPDTWNPTQGYTGNANLHVKVVNYGTDLGCASSL 530

Query: 539 XXXXXXXXXXXGDKDSQALSKTILDDIWKNYQDAKGVSAPEQM-DYSRVFNQEVYIPQGW 597
                   GD+ S+  ++ +LD +W NY D+KG+S  EQ  DY R  +QEV++P GW
Sbjct: 531 ANTLTYYAAKSGDETSRQNAQKLLDAMWNNYSDSKGISTVEQRGDYHRFLDQEVFVPAGW 590

Query: 598 TGTMPNGDVIKSGNKFIDIRSQYKNDPDYARVKSDVEAGKSSTFNYHRFWAESEYAIANA 657
           TG MPNGDVIKSG KFIDIRS+YK DP++  + + ++AG+  T   HRFWA+SE+A+AN
Sbjct: 591 TGKMPNGDVIKSGVKFIDIRSKYKQDPEWQTMVAALQAGQVPTQRLHRFWAQSEFAVANG 650

Query: 658 NYGTLFANTATPXXXXXXXXXXRDLME---LRQYLAGKLDASKINLAAADVNNDGVVNG 714
            Y  LF +             D ++  L++YL     ++ IN A AD+N+D  ++
Sbjct: 651 VYAILFPDQGPEKLLGDVNGDETVDAIDLAILKKYLLN--SSTTINTANADMNSDNAIDA 708
```

Figure 39. Results from amino acid sequence alignment between Cel48A of *C. acetobutylicum* and Cel48F of *C. cellulolyticum.* BLAST result shows identities percentage about 58%.

From cellulases amino acid sequence, we can devide the sequence into two parts. The first part is responsible for catalytic activity of enzyme, we known it as "*catalytic domain*". The second parts always found at the C-terminus of each cellulase. This "*dockerin*" part is very important to the assembly of the cellulosome because it always different and depend on the species. The significant of this part is nessesary for cellulosome association in each cellulosome producing clostridia because of species specificity.

4.2.3 Cloning and sequence of modified Cel48A protein

We generate two recombinant plasmids for express modified Cel48A protein by construct two plasmid pET-Cel48A and pET-FH3. The plasmid constructions are shown below. pET-Cel48A was constructed by insert corrected Cel48A gene into appropriate vector that design for expression in *E. coli.* Based on Cel48A sequence, we try to eliminate point mutation by replace with the right sequence for translate the desired amino acid sequence for correct expression.

Regularly, degradation of cellulose by cellulosomes occurred by this dominant family-48 processive endoglucanase by synergistic activity. By this construct, we expect that direct reverse mutagenesis that induced to the synthetic Cel48A gene is effect to the enzyme activity. At least, retain the stability of enzyme or promote the activity on cellulosic substrate. The role and mechanism to retain the enzyme activity is still unknown. However, usual cellulosomes, such as those produced by *C. cellulolyticum* observation give some convenient answer. Catalytic domain is quite conserved except some points that mutation occurred. Hybrid protein was designed to give more valuable information to gain further insight into the function of natural Cel48A.

Plasmid pET-Cel48A was previously constructed by our team. In parallel, plasmid pET-FH3 was constructed in this study. From previous study, pET-Fc and pET-Ft use to be constructed for express recombinant protein between catalytic domain of *C. cellulolyticum* CelF cellulose gene and dockerin domain of *C. thermocellum* CelF dockerinn domain (Fierobe et al., 2001). Using this strategy, appropriate dockerin-containing enzymes could be assembled precisely and by design into a desired complex. Cellulolytic activities of this recombinant protein were

measured and found that *Chimeras*—Cellulosome chimeras were generally found to be more active than simple mixtures of the free enzyme pairs. In this work, plasmid contain gene encode hybrid functional protein was constructed and expressed in *E. coli*.

Table 11. Bacterial strain, plasmid and primer

Strain, plasmid or primer	Relevant characteristics	Source or reference
Strain		
DH5α	F-Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *deo*R *rec*A1 *ton*A	Invitrogen
BL21 CodonPlus(DE3)-RIL strain*	*E. coli* BF *omp*T *hsd*S (r_B^- m_B^-) *dcm*$^+$ Tet^r *gal*λ (DE3) *end*A The [*argU ileY leuW* Cam^r]	Stratagene
C. acetobutylicum	Wild type	ATCC
Plasmid		
pET-Fc	F1 ORI, Amp^r, *lac* promoter, T7 promotor	Reverbel-Leroy, 1996
pET-Cel48A	F1 origin, Carb^r, , *lac* promoter, T7 promotor	This study
pET-FH3	F1 origin, Amp^r, *lac* promoter, T7 promotor	This study
Primer		
Ccf1	5'GACCTAGGTTGTGCTTCTTCA3'	Reverbel-Leroy, 1996
Ccf2	5'TACTTTATATGTCATGCTCGGGAAGAGTATTG CATAAACTC3'	
Dock1	5'TTCCCGAGCATGACATATAAAGTATTCGCTAA TACAGCTACACCAGG3'	This study
DockII	5'TTTCTCGAGTTTTGCAATTAATTTAGTAAGTTC CATAATATC3'	

* This strains, derivative of *E. coli* B, are general protein expression strains that lack both the L_{on} protease and the OmpT protease, which can degrade protein during purification. Dcm methylase, naturally lacking in *E. coli* B, is inserted into the genomes (Stratagene).

In order to design hybrid protein, amino acid sequence of *C. acetobutylicum* Cel48A compare with Cel48F of *C. cellulolyticum* was analysed. Starting point of dockerin domain of Cel48A was identified by amino acid sequence alignment. Normally, *C. acetobutylicum* dockerins contain a duplicated sequence of about 22 amino acid residues, the first 12 of which are homologous to known structure, the calcium-binding loop in the EF-hand motif (Sabathé *et al*., 2002). The end point of catalytic domain of Cel48F was determined by estimate from the primer that previously used by Fierobe *et al*., 2001. This make it more clearly that where the catalytic domain start and end. By the way, the end of catalytic is closed to start point of dockerin domain.

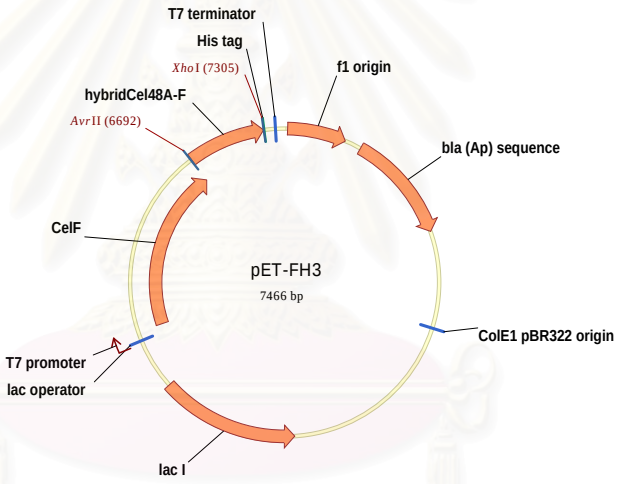


Figure 40. Shematic diagram of pET-FH3 hybrid Cel48F catalytic domain and Cel48A dockerin domain

The hybrid molecule between Cel48F catalytic domain and Cel48A dockerin domain coding sequence has length of 624 bp. This fragment was cloned into pET-Fc (~7 kb) plasmid digested with *Xho*I and *Avr*II. After that, plasmid was transformed into *E. coli* DH5α. Ampicillin resistant clones were selected to analysis restriction pattern. After select the demanded clone by restriction analysis, they were sent to sequencing. From the examination, sequence results between T7 terminator universal primer (New England Biolabs Inc.) is appear correctly as prediction. This recombinant clone was selected for express hybrid protein to investigate the purified hybrid

enzyme activities on several cellulose substrates.

Figure 41 shows strategy to construct hybrid Cel48F-Cel48A molecules. At first, we designed suitable primers for synthesis recombinant Cel48 molecules. First primer pair was used to make the terminal part of Cel48F catalytic domain and some part of linker. The second pair was designed for synthesis all dockerin domain that following natural Cel48A dockerin. After that, appropriate restriction enzymes were used to cut this synthetic part in order to insert the replaced Cel48F-A dockerin domain into pET-Fc (Reverbel-Leroy *et al*., 1996). The final plasmid was the recombinant plasmid that selected for express new molecules.

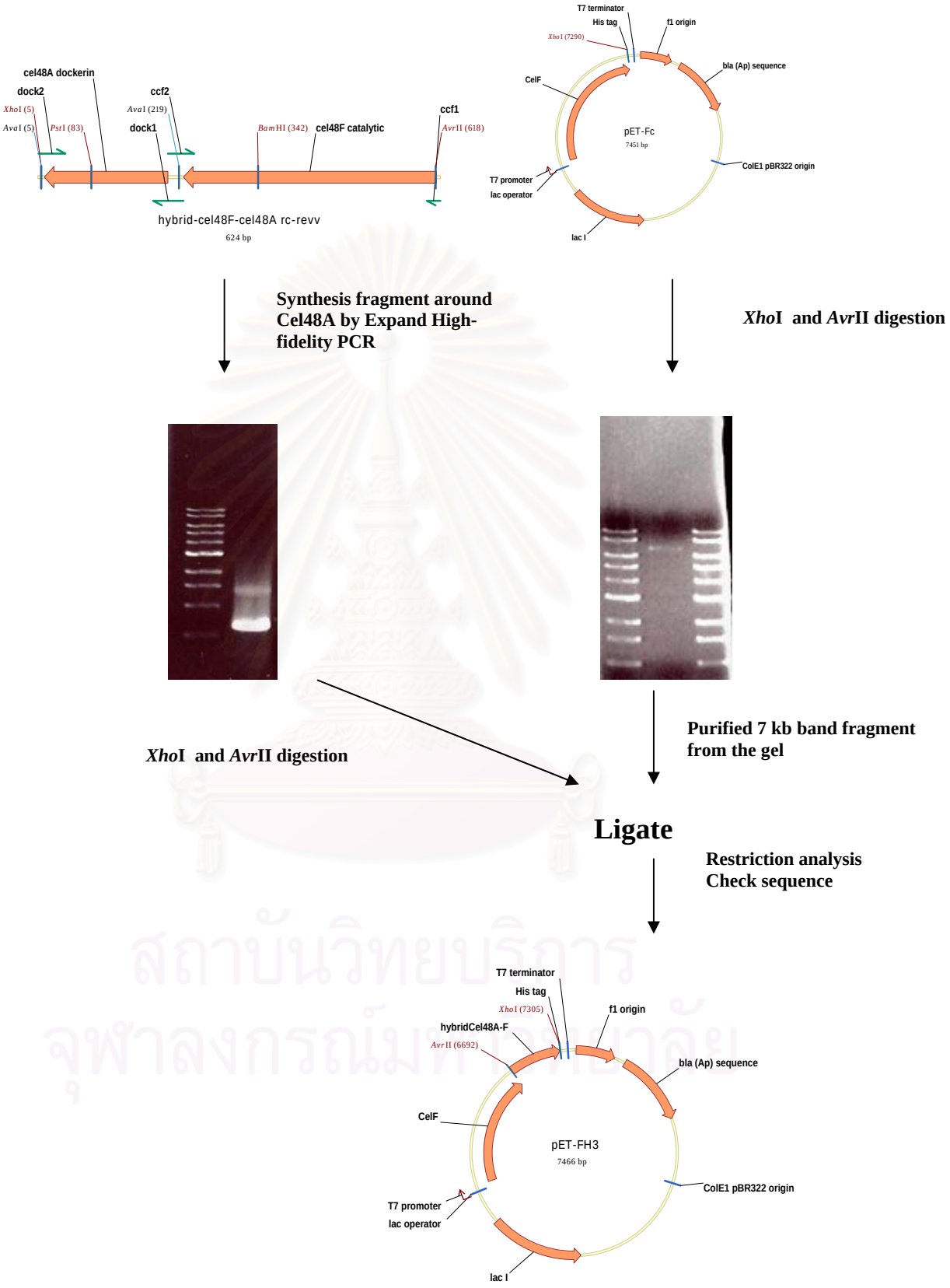


Figure 41. Schematic representation of plasmid construction of the recombinant proteins used in this study.

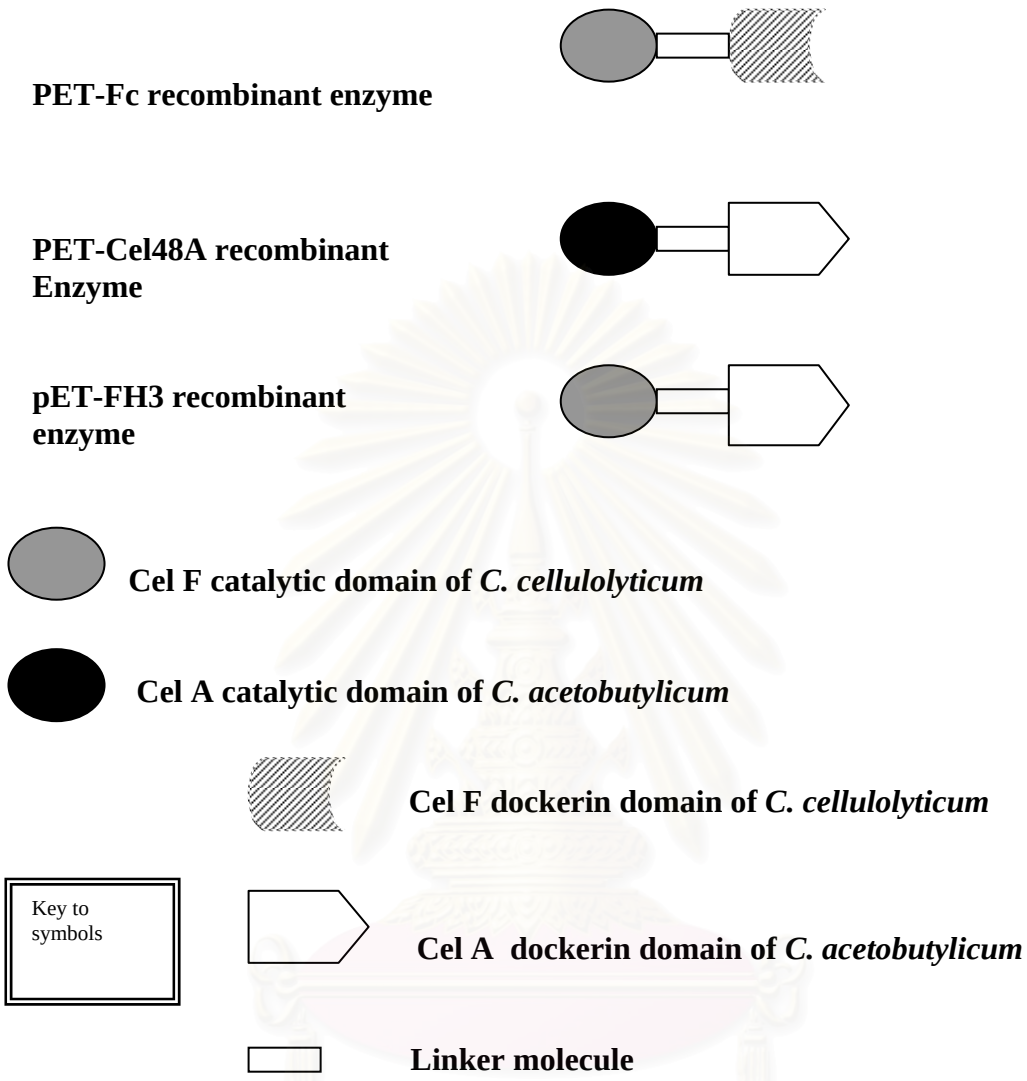


Figure 42. Schematic representation of the recombinant proteins used in this study. *Black* (*C. acetobutylicum*) and *gray* (*C. cellulolyticum*) *symbols* denote the source of the respective domains (see "*Key to symbols*").

***Avr*II site**

5’GACCTAGGTTGTGCTTCTTCACTTGCAAACACATTGACTTACTATGCTGC
TAAATCAGGAGATGAAACTTCAAGGCAGAATGCACAGAAATTACTTGACG
CTATGTGGAATAACTATAGCGATTCAAAGGGTATATCAACTGTTGAACAG
CGTGGTGATTACCATAGATTCCTTGATCAGGAAGTTTTTGTACCAGCTGGT
TGGACTGGAAAAATGCCTAACGGCGACGTAATCAAATCTGGTGTCAAGTT
CATAGACATTCGTTCCAAGTACAAGCAGGATCCTGAATGGCAGACAATGG
TTGCTGCATTACAGGCAGGACAGGTTCCAACTCAGAGATTACACCGTTTCT
GGGCTCAGAGTGAATTTGCAGTTGCAAATG**GAGTTTATGCAATACTCTT**
CCCGAGCATGACATATAAAGTATTCGCTAATACAGCTACACCAGGAGAT
GTAAATGGAGACGGCGTAGTTAATGGAAGAGATCTTATGGAGCTTAGACA
GTATTTAGCAGGTAAGTTAGATGCTAGTAAAATAAATTTAGCTGCAGCAG
ATGTTAATAATGATGGTGTAGTTAATGGAAGAGATATTATGGAACTTACTA
AATTAATTGCAAAACTCGAGAAA 3’

***Xho*I site**

Figure 43. Nucleotide sequence of hybrid Cel48F catalytic and Cel48A dockerin domain gene fragment. Bold underlined refer to linker sequence.

4.2.4 Activity of the purified hybrid protein on various cellulolytic substrates

Overproduction of Cel48A protein in *E. coli*

The modified Cel48A protein expression was induced in *E. coli* (DE3) CodonPlus. This strain contains plasmid enhance efficiency of heterologous protein production in *E. coli.* In order to express, cells were grown until OD_{600} approximate 0.8. Then induction was performed by IPTG. The expression was performed at 16°c to avoid accumulation of foreign protein in form of inclusion body. However, the nature of this protein is non-soluble protein. So, the SDS-PAGE and Western blott results show high amount of protein in the pellet fraction.

To confirm the correct of protein that expressed again, anti-HIS immunodetection method was performed again. This results bellowed show the appearance of band after detection with anti-HIS approximatelt 80 kDA. Some protein degradation occurred also as the band exhibit in figure 43.

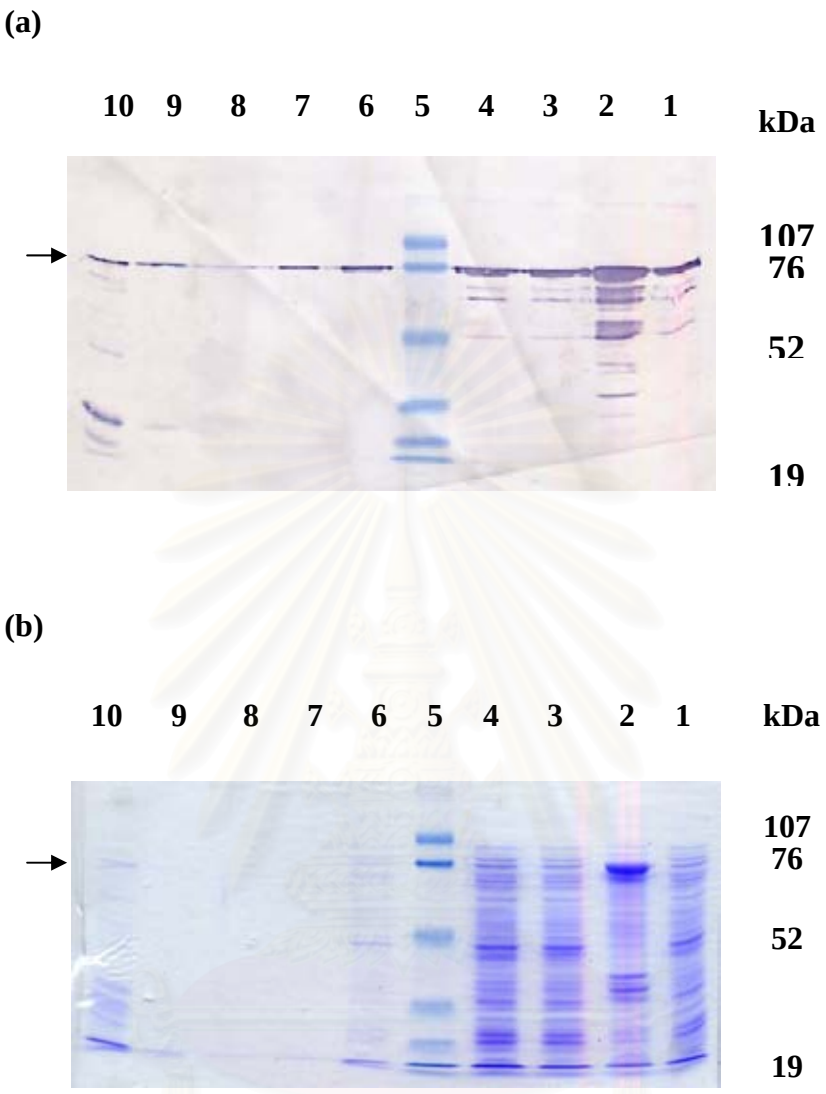
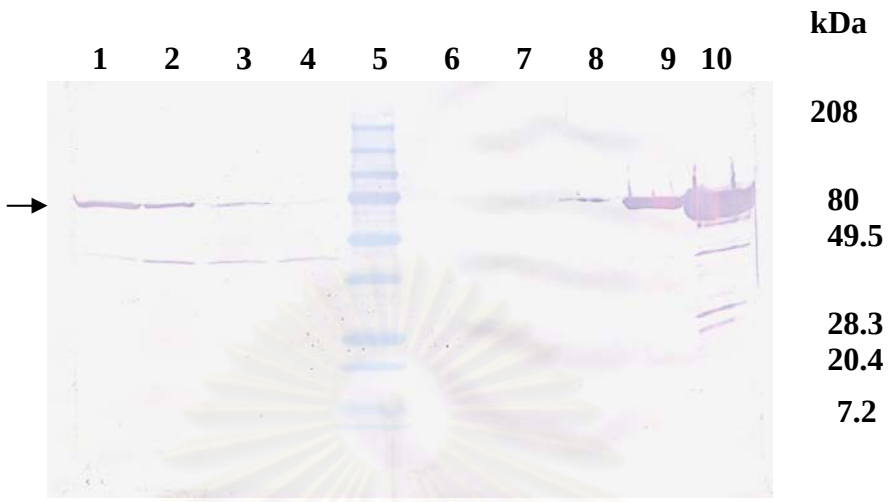


Figure 44. (a) Western-blott analysis of induced *E. coli* (DE3) CodonPlus (pET-Cel48A). Lanes: 1, supernatant of crude extract; 2, pellet after crude extract centrifugation; 3, first fraction after passing through the column; 4, second after passing through the column; 5 low molecular weight protein marker; 6, first washing fraction; 7, second washing fraction; 8, third washing fraction; 9, elution fraction; 10, concentrated protein after elution as arrow indicated.

(b) SDS-PAGE gel of 40µl protein fraction after purification. Purified protein has size approximately 80 KD.

(a)



(b)

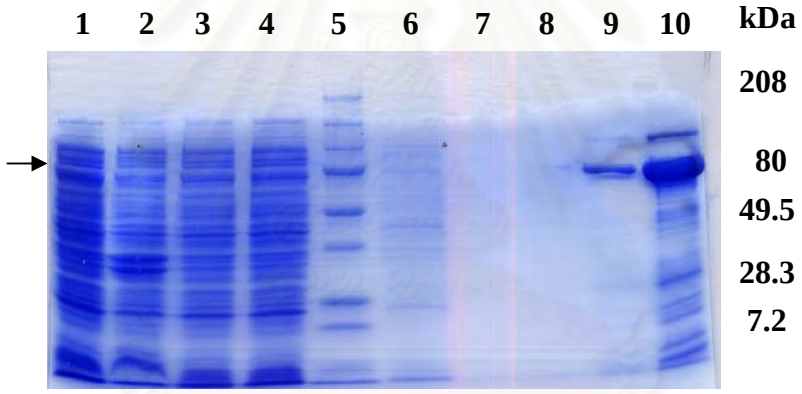


Figure 45. (a) Western-blott analysis of induced *E. coli* (DE3) CodonPlus (pET-FH3). Lanes: 1, supernatant of crude extract; 2, pellet after crude extract centrifugation; 3, first fraction after passing through the column; 4, second after passing through the column; 5 all molecular weight protein marker standard; 6, first washing fraction; 7, second washing fraction; 8, third washing fraction; 9, elution fraction; 10, concentrated protein after elution as arrow indicated.

(b) SDS-PAGE gel of 40µl protein fraction after purification. Purified protein has size 84 KD.

The solubility and the localization of the recombinant proteins, crude cell extracts, and soluble and insolubles fractions were examined by SDS-PAGE. Both Cel48A and hybridCel48 were produced in soluble and insoluble form (Figure 44 lane1 and lane2). Lane 1 shows soluble form of recombinant protein at average molecular mass of about 78 kDa. Expression of recombinant proteins were performed in *E. coli* (DE3) CodonPlus. Both recombinant proteins were purified by Ni^{2+}-Histidine affinity chromatography, as intimate at C-terminus end of recombinant proteins. Insoluble proteins are indicated as inclusion bodies because of the presence of aggregate heterologous proteins. Interestingly, modified Cel48A by direct mutagenesis showed more accumulated recombinant proteins in form of inclusion bodies than in hybrid Cel48A. As described in previous report, CelA soluble/insoluble form was estimated about 0.7 (Reverbel-Leroy *et al*., 1997). We proposed that when express in high level, Cel48A always form higher insoluble form than that in soluble form. In contrast, hybrid Cel48A-Cel48F protein expression displayed soluble form than insoluble form. As always occurred, Cel48F proteins normally exhibit soluble form/insoluble form ratio reach to 19. We assume that catalytic domain of cellulases has effect to the structure, solubility and proximity of enzyme which is related to enzyme activity. Clearly that the stimulation of activity was mainly a function of two factors: (i) enzyme proximity within the complex and (ii) the substrate-targeting effect of CBM (Fierobe *et al*., 2005). To indicate enzyme activities, we purified and concentrated soluble form of recombinant enzyme. This soluble form still active and can be used to determine chimeric enzyme activities on cellulose efficiently.

4.2.5 Activity assay

Both on carboxymethyl cellulose (CM-cellulose) and phosphoric acid swallen cellulose (PASC), native Cel48A showed very low activities (Table. 12). No enzyme activity observed on crystalline cellulose (Avicel-cellulose). In the same way, site-directed modified Cel48A still has no activity on crystalline cellulose and the activity on the other cellulosic substrates still unclear. In contrast, the purified hybrid Cel48A and Cel48F has proficient activity on all cellulosic substrates that we inspected. Interestingly, the hybrid protein shows activity on crystalline cellulose (Avicel-cellulose) as good as effective native Cel48F (11.882 and 13.4 IU/μmol respectively).

Table 12. Enzyme activities (IU/μM) recombinant Cel48A protein compare between wild type and recombinant.

substrate	**Specific activity (IU/μM)**		
	C. acetobutylicum		***C. cellulolyticum***
	pET-FH3 hybrid	**Cel48A**	**Cel48F**
CMC	0.339	0.011	ND
PASC	5.542	0.0020	42.5
Avicel	11.882	-	13.4
BMCC	ND	-	ND

4.2.6 Plasmid designed and transformation to *C. acetobutylicum* ATCC824

Plasmid p952-SA-FA was designed in order to transform hybrid Cel48A-48F protein to *C. acetobutylicum* ATCC824. Plasmid contained P-thiolase-lac operator for express in *Clostridium* sp., ampicillin resistant gene and Erythromycin resistant gene as marker for selection.

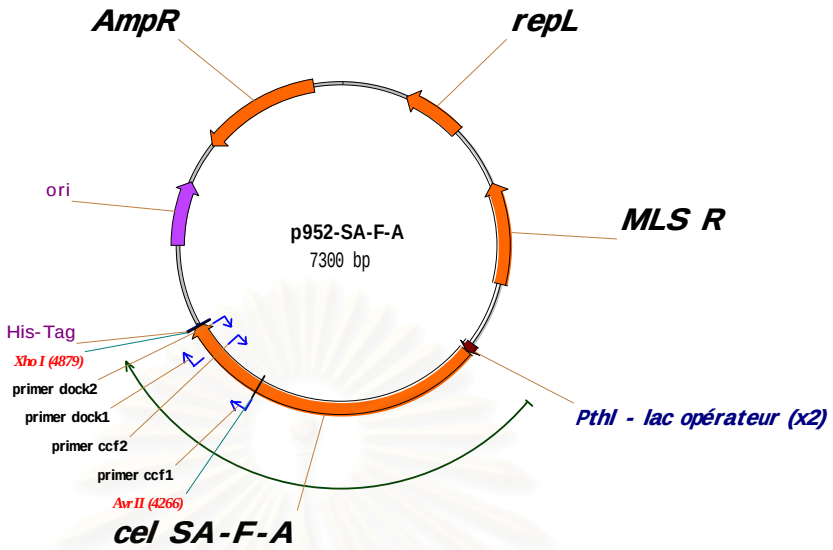


Figure 46. Shematic diagram of p952-SA-F-A plasmid

After transformation, plasmid availability, correction and stability were checked by PCR amplification. *C. acetobutylicum* ATCC 824 transformed with p952-SA-FA that express ampicillin and erythromycin resistant phenotype were extracted for genomic DNA. Then, PCR amplification by primer pREPD and Dock I results PCR product approximately 800 bp.

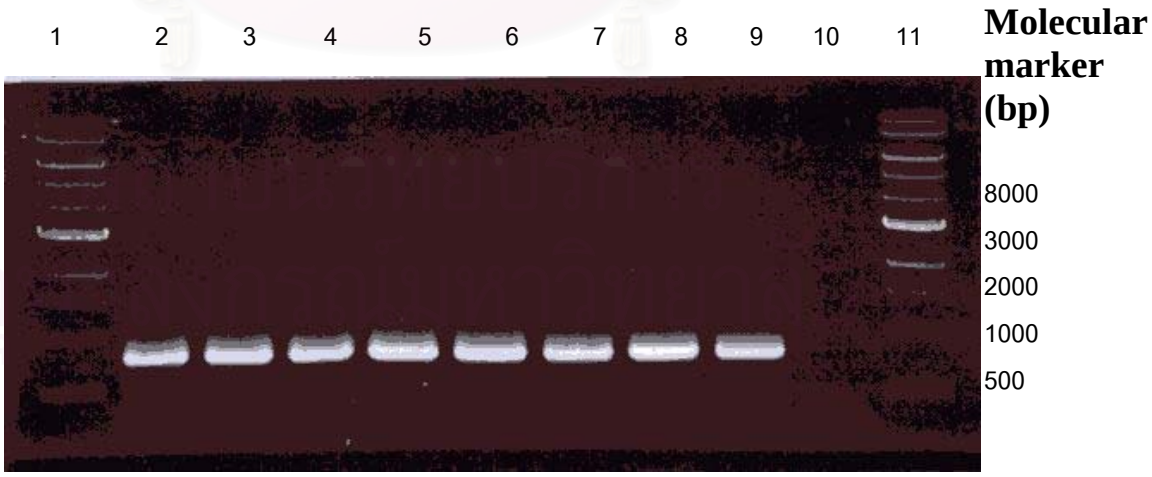


Figure 47. Amplification checking 7 clones of p952-SA-FA plasmid in transformed *C. acetobutylicum* ATCC824. Lane 9 refers to positive control and lane 10 refers to negative control.

4.2.7 Cellulosome expression and purification

Cellulosome of *C. acetobutylicum* both wild type strain ATCC824 and hybrid strain were expressed and purified. Cellulosomal fractions were detected by combination of anti-CelF, anti-CelC from *C. cellulolyticum* as 1st antibody. Wild type shows a single band protein cellulosome production (56 kDa) noticeably. In conflict, hybrid *C. acetobutylicum* produce detected various cellulosome subunits between 72-130 kDa and the main subunit 56 kDa. At least, this can be elucidated difference between wild type and hybrid cellulosome (Figure 47). This also can be expected that it is truncated hybrid Cel48A protein. Since the molecular weight of hybrid Cel48A is 84 kDa.

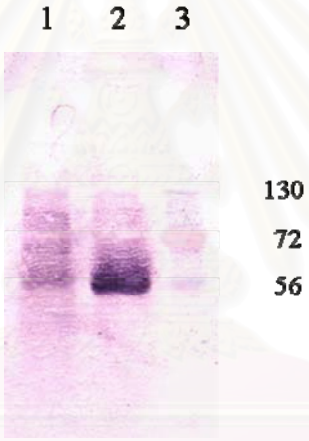


Figure 48. Cellulosome purified fraction of *C. acetobutylicum.* Lanes: 1, Cellulosome fraction of *C. acetobutylicum* ATCC824 Wildtype; 2, Cellulosome purified fraction of *C. acetobutylicum* hybrid p952-SA-FA.

4.2.8 Discussion

The recombinant hybrid Cel48A dockerin-Cel48F catalytic domain can express improved activity on cellulosic substrate (avicel, CMC and PASC) in *E. coli* BL21CodonPlus. This hybrid can be transformed successfully to *C. acetobutylicum* ATCC824. The recombinant cellulosome can be expressed in *C. acetobutylicum*. Cellulosome fraction was purified to compare different between recombinant and wild type *C. acetobutylicum* ATCC824. This different also might due to the expression of truncated hybrid Cel48A protein. Since cellulosome detection requires various

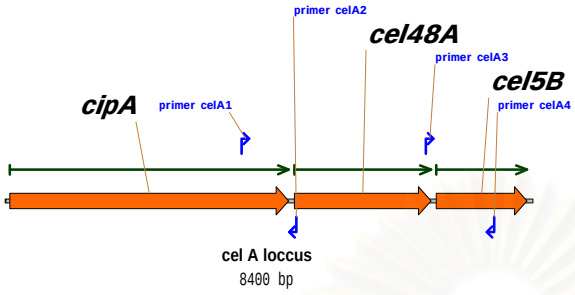
antibody for detection, this difference may be according to many factors such as the sensibility different between hybrid protein and wild type. In this study, antibody for detect equivalent protein of Cel9C, Cel48A and scaffolding protein were used as cocktail mixture. Actually, These antibody were more specific to *C. cellulolyticum* cellulosome. Clearly that it may needs more specific enzyme subunit antibody for efficiently detection. For further application, confirmation is necessary to determine cellulolytic activity in recombinant *C. acetobutylicum* ATCC824 with p952-SA-FA.

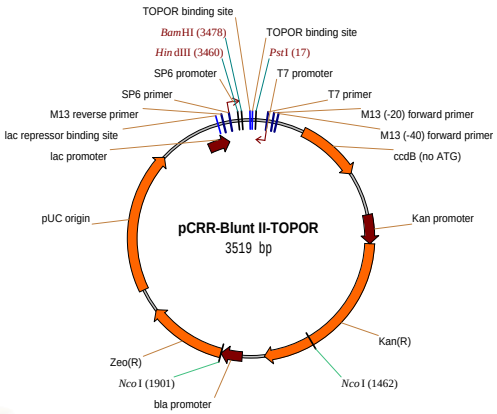
4.3 Ineffective *cel*48A gene deletion

4.3.1 Gene deletion

It is a clue in metabolic engineering strategies that use for recombinant strains construction. It is useful for the over-expression of homologous or heterologous genes by inactivation/deletion another branch in metabolic pathway.

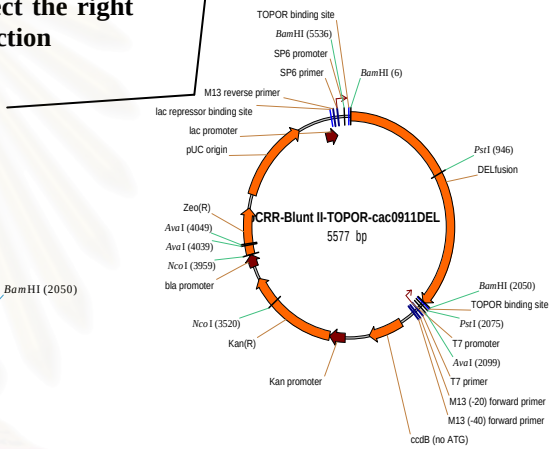
The briefly strategy was shown in figure 49. Firstly, PCR amplify FRT-flanked antibiotic resistant gene was prepared. These flanking regions were designed to recognize by FLP recombinase enzyme. Antibiotic resistant gene was provided for selective marker after transformation. In this case, Primer pair was designed for amplified fragment around *cel48A* gene (CAC0911). Then, PCR-fusion amplification between two gene fragments around *cel48A* gene was performed. This fragment was inserted into commercial vector (pCRBluntII-TOPO or another plasmid without *Stu*I site). FRT-MLSR-FRT cassette was inserted in between two gene fragment around *cel48A* gene by replace on *Stu*I site of insert. After that, pCRBluntII-TOPO vector was replaced by replicative plasmid of *C. acetobutylicum* ATCC824.This was transform into *C. acetobutylicum*WT or ΔCAC1501 DNA-methyltransferase strain. Corrected phenotypic strains were selected for genomic DNA extraction. Transformants were checked for double-crossing-over integrants. In this step, target gene was deleted. The replaced antibiotics gene was eliminated by FLP recombinase enzyme system.

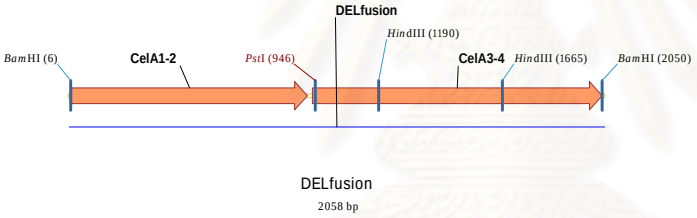


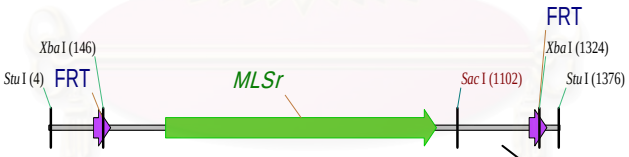


Synthesis fragment around Cel48A by Expand High-fidelity PCR

Ligate and transform into *E. coli*, restriction analysis was done to select the right plasmid construction



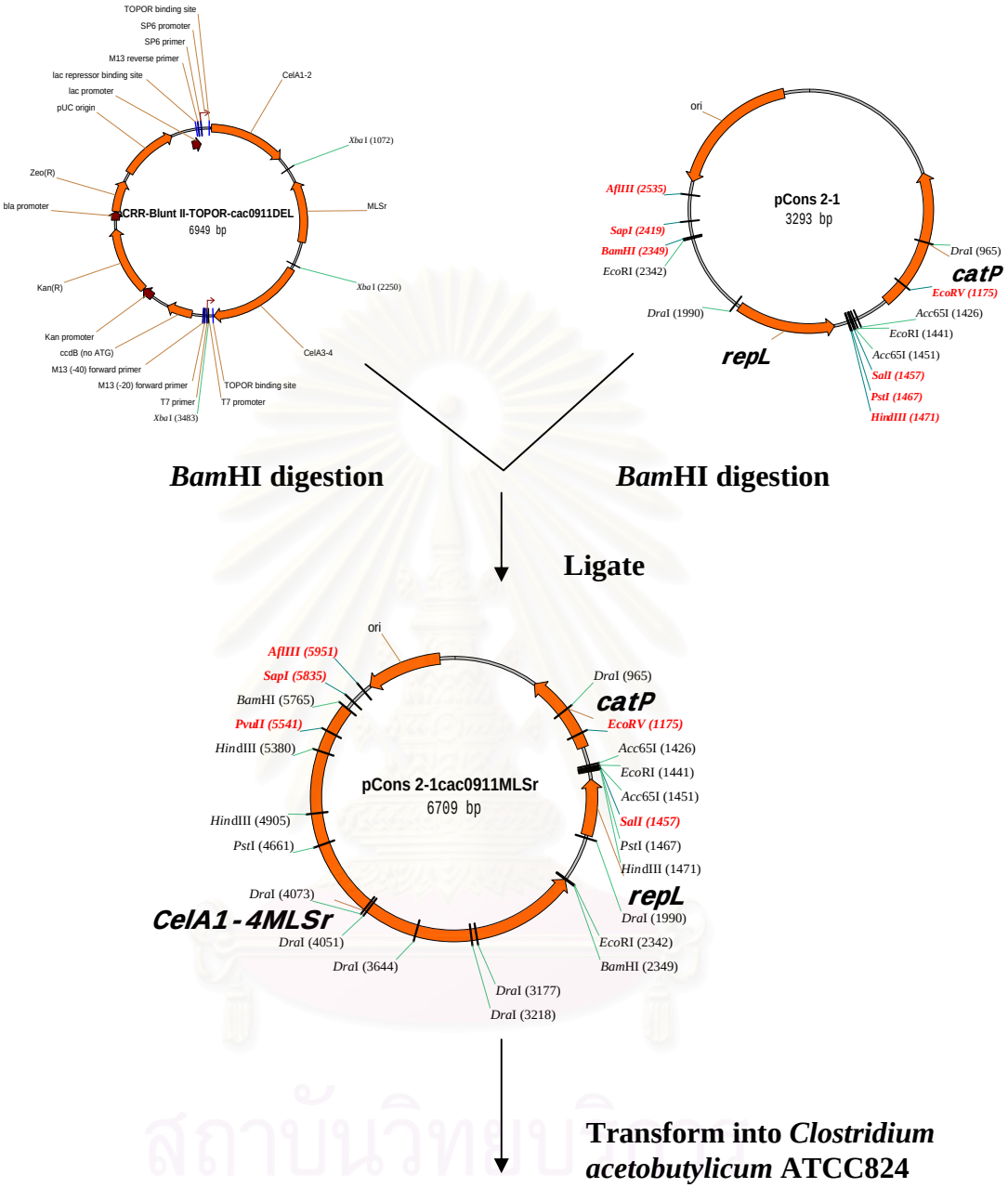




fragment FRT-MLSr-FRT
1378 bp

***Stu*I digestion** ***StuI* digestion**

Ligate and do the restriction analysis to select the clone that the MLS^{r} contained in the plasmid



Select for *Clostridium acetobutylicum* that contain the erythromycin resistance ability and contain no Cel48A gene

Figure 49. Construction of plasmid pCRBlunt-TOPOcac0911 and pCons2-1CelA1-4MLS^{r}

For each plasmid, the location and directions of transcription of relevant gene are indicated (arrows). Relevant restriction sites are shown. Abbreviations: cipA, Probably cellulosomal scaffolding protein; cel48A, Possible processive endoglucanase family 48; cel5B, Possible non-processive endoglucanase family5; Kan(R), Kanamycin resistance gene; MLS^r, macrolide-lincosamide-streptogramin resistance gene; repL, gene required for replication in gram-positive microorganisms.

4.3.2 Step-By-Step *in vitro* Construction results

Our standard protocol is illustrated in Figure 48. PCR products were generated by using two pairs of 40- to 50-nt-long primers (Table 15).

4.3.2.1 Design primer for amplified fragment around *cel48A* gene (CAC0911)

CelA1-A4 amplification by using *Clostridium acetobutylicum* ATCC824 chromosomal DNA as template for amplification.

Table 13. Bacterial strain and plasmid

strain or plasmid	Relevant characteristics	Source or reference
Strain		Invitrogen
DH5α	F-Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F)U169 *deo*R *rec*A1 *ton*A	
C. acetobutylicum	Wild type	ATCC
Plasmid		
pCR®II-Blunt-TOPO®	pUC ORI, Kan^r, *lac* promoter, T7 promotor	Invitrogen
pCR® 4Blunt-TOPO®	pUC ORI, Kan^r, Amp^r, *lac*Z promoter, T7 promotor	Invitrogen
pCons 2-1CelA1-4MLS^r	Cam^r, RepL, MLS^r	This study

Table 14. Oligonucleotides used for PCR and sequencing[a]

Primer name	Oligonucleotide sequence (5'-3')	Use
CelA1	AAAAGGGATCCCAGATGGAAAAATAACAATCACAGGTTCAGCACCAG	PCR CelA1-2
CelA2	**GGGG***AGGCCT***AAAAAGGGGG**AAATATAACTGTAGATGTAAGAGCTACAGCC	PCR CelA1-2
CelA3	**CCCCCTTTTT***AGGCCT***CCCC**TTAGCTGCAGCAGATGTTAATAATGATGGTG	PCR CelA3-4
CelA4	AAAAGGATCCTCTTGTATATAGAACCAGCTTGGACGCC	PCR CelA3-4

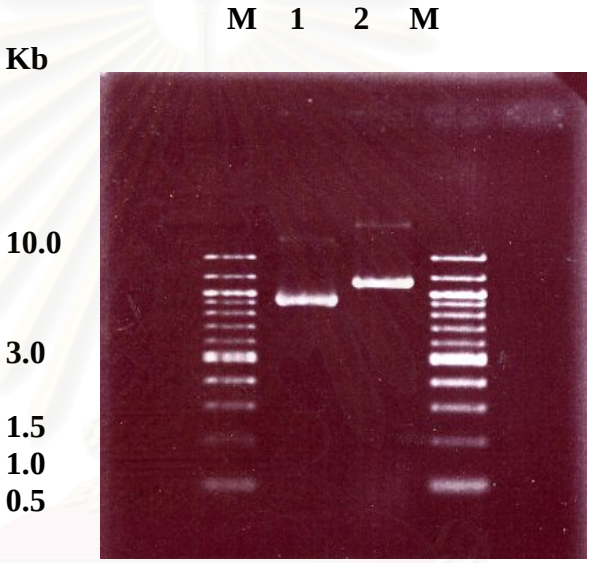


Figure 50. CelA1-2 and CelA3-4 amplification. Lane M, 1Kb DNA ladder; 1, Amplified product from CelA1-2 primer results 935 bp DNA fragment; 2, Amplified product from CelA3-4 primer results 1142 bp DNA fragment.

4.3.2.2 PCR-fusion amplification between two gene fragments around *cel48A* gene

Fusion fragment around *cel48A* gene was amplified. Results DNA fragment size 2058 bp (fusion between 1142+935 bp) as shown in figure 51.

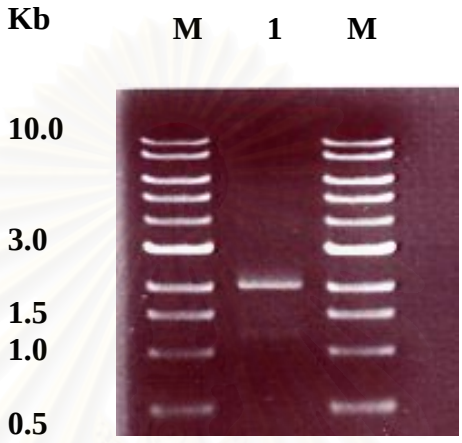


Figure 51. CelA1-2 and CelA3-4 fusion amplification. Lane M, 1Kb DNA ladder; 1, Amplified fusion product between CelA1-2 primer and CelA3-4 primer results 2058 bp DNA fragment.

4.3.2.3 Insert fragment into commercial vector (pCRBluntII-TOPO or another plasmid without *Stu*I site)

After got this purified product for Cel48A inactivation, cel48A deletion vector was constructed by ligation with pCR®II-Blunt-TOPO® vector (Invitrogen). This plasmid is supplied linearized with Vaccinia virus DNA topoisomerase which can bind to duplex DNA and form covalent bond between the 3' phosphate of the cleaved strand and a tyrosyl residue of topoisomerase I (Invitrogen). This vector plasmid contains TOPO cloning site and Kanaycin resistance gene which use as the selective pressure for the transformant.

The restriction analysis was performed to check for the correct location and direction of gene in the vector. In this case, *Eco*RI restriction site were checked plasmid direction. From the restriction profile, *Eco*RI provides 2

DNA fragments in case the digestion are occurred completely. This generated 3501 and 2076 bp respectively.

4.3.2.4 Insertion of FRT-MLSR-FRT cassette by replace on *Stu*I site

In order to insert the erythromycin resistant gene into the vector, *Stu*I site is designed in the amplification fragment of CelA1-A4. So, we use this restriction site to insert MLS^r fragment into the plasmid pCons2-1 cac0911MLSr. After that, use *Nco*I and *Pst*I were used as the restriction site for analyse correct plasmid from 10 transformants. Figure 52-53 showed correct restriction pattern in clone number 7, 8 and 9. These clones are refered as clone 4A1E7, 4A1E8 and 4A1E9.

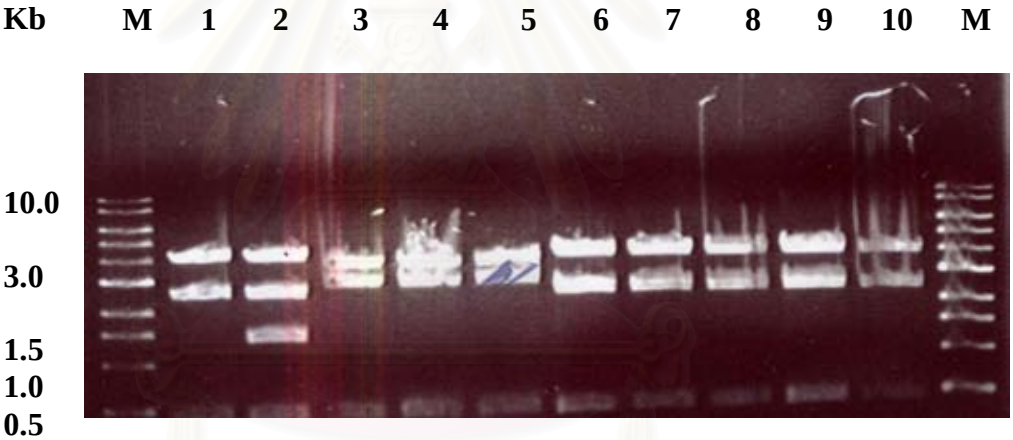


Figure 52. Transformants restriction pattern were generated by *Nco*I enzyme. Restriction profile from the gel shows the expected band (3871, 2639 and 439 bp) in lane1, 6-10.

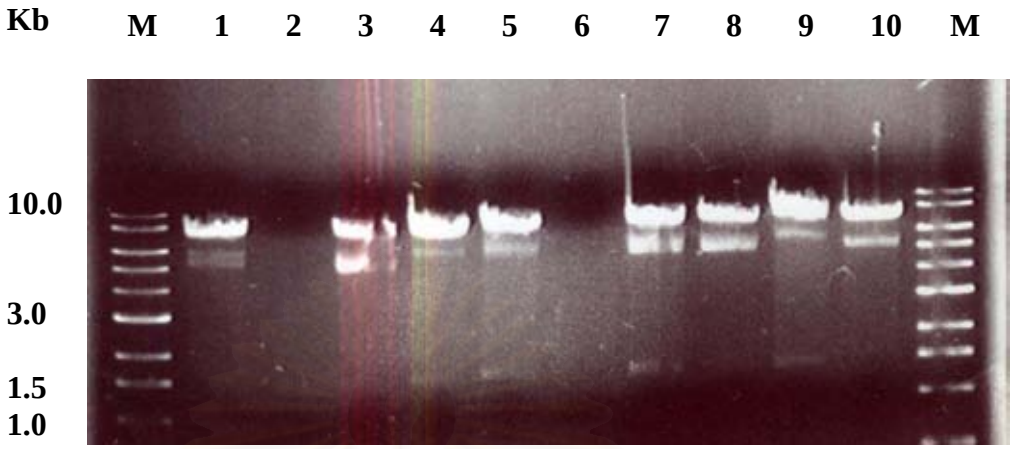


Figure 53. Restriction analysis pattern of transformant by using *Stu*I enzyme. Restriction profile from the gel shows the expected band (5577 and 1372 bp) in lane 4. 5. 7. 8. 9.

4.3.2.5 Replace pCRBluntII-TOPO vector by replicative plasmid of *C. acetobutylicum* ATCC824

From the results, 2 clones (4A1E8 and 4A1E9) were selected for extract plasmid then cut plasmid with *Bam*HI for moving the deletion cassette to insert to pConS2-1 vector for *C. acetobutylicum* ATCC824.

However, we have the problem about *Bam*HI restriction site in the selected clone. So, we try to analyse and find the reason why this restriction site is malfunction.

Protein translation checking for 4A1E8 and 4A1E9

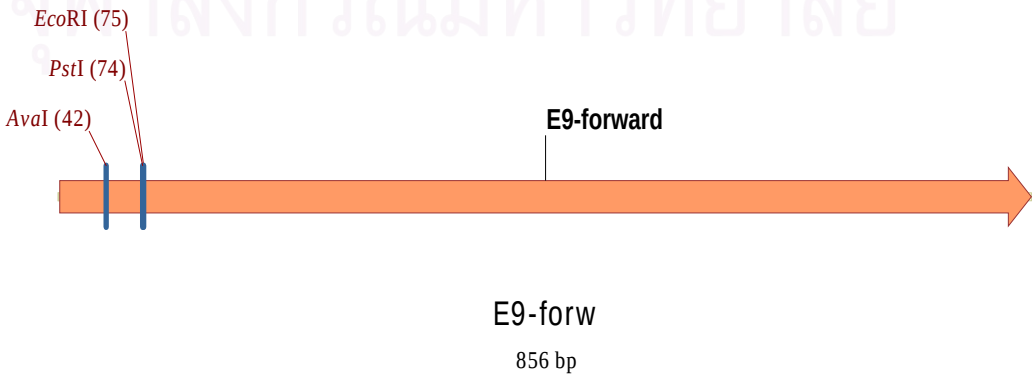


Figure 54. Restriction profile from forward sequence of E9 clone.

This sequence can be identified clone into 2 parts.

1. First part: first 71 base pair (from BLAST sequence: first part match to gi/52850698/emb/aj5859.1): *Thermococcales archaeon* partial 16SrDNA gene: 71/71=100%)

2. Second part: (from BLAST sequence: first part match to gi/25168256/gAE001437): *C. acetobutylicum* ATCC824 Probably cellulosomal scaffolding protein

However, forward sequence of clone is not correct since *Bam*HI site is missing.

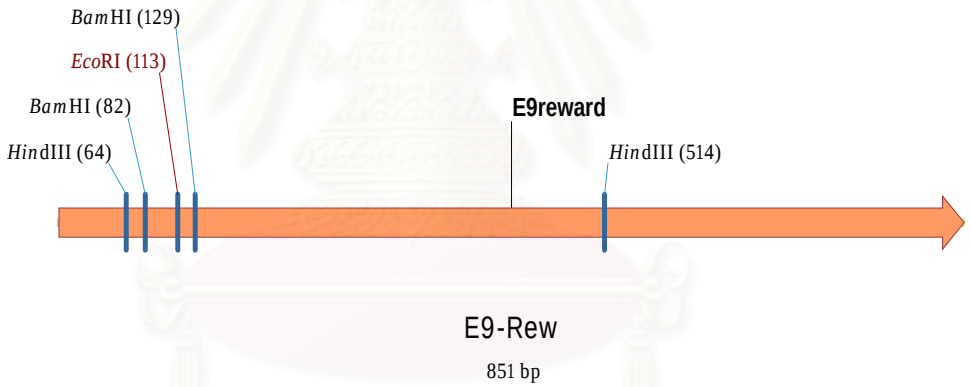


Figure 55. Restriction profile from reverse sequence of E9 clone

Reverse sequence of E9 clone showed complement between 762 bp from 849 bp in model (87 mismatchs from model). The sequence was checked and found *Bam*HI site inside sequence.

New primer was designed for compensate CelA4 to solve that problem because one of *Bam*HI site was lost in CelA4 amplified direction. This is important since *Bam*HI site is necessary for transfer cel48A deletion fragment to pCon-S2-1 plasmid which is a shuttle vector appropriate for *E. coli* and *C. acetobutylicum* ATCC824.

Table 15. Oligonucleotides of primer CelAP

Primer name	Oligonucleotide sequence (5'-3')	Use
CelAP	GTTCTTAAGGATCCAAGTGATTACACAGTAACAAGTGATGG	PCR CelA1-AP

3258 bp-.PCR fusion product was amplified again by CelA1-CelAP primer. DNA template was received from 4A1E8 clone.

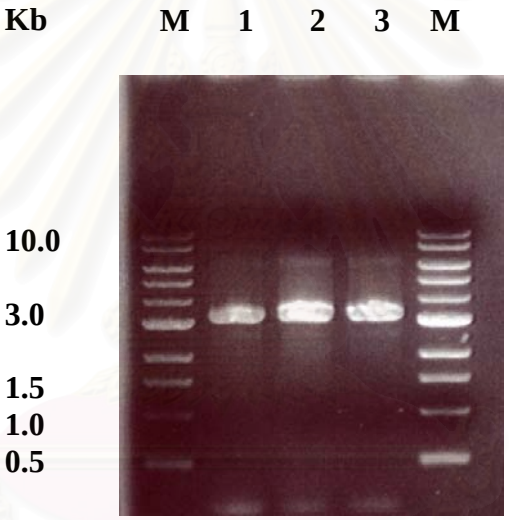


Figure 56. 3258 bp-PCR fusion product was amplified again by CelA1-CelAP

This fragment was ligated with pCR-Blunt II-TOPO® vector again. After ligation and transformation into *E. coli* DH5α, selected clones (Bl1-Bl10) were extracted for plasmid. Restriction patterns were analysed by 4 restriction enzymes *Eco*RI, *Pst*I, *Bam*HI and *Nco*I. Clones that expressed corrected restriction pattern were kept for the next step. The results revealed that all clones got the plasmid that insert in the inverse direction.

***Eco*RI:** expected band 3501 + 3276

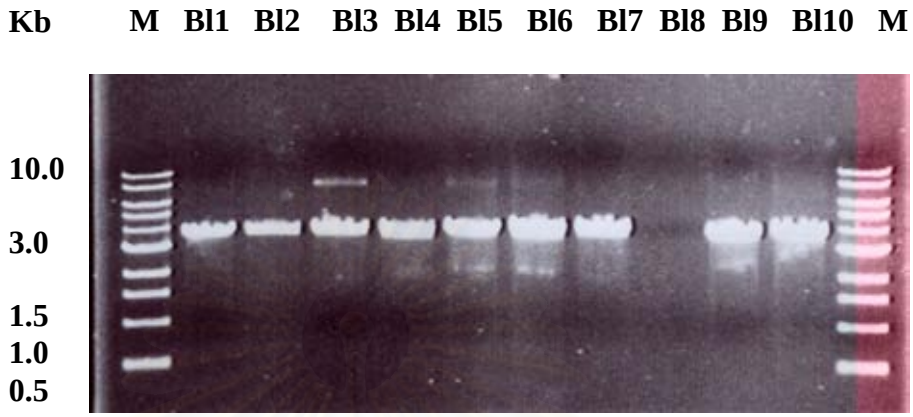


Figure 57. Restriction analysis results from clone Bl1-Bl10 by *Eco*RI. Expected and is 3501 + 3276 for insert in forward direction.

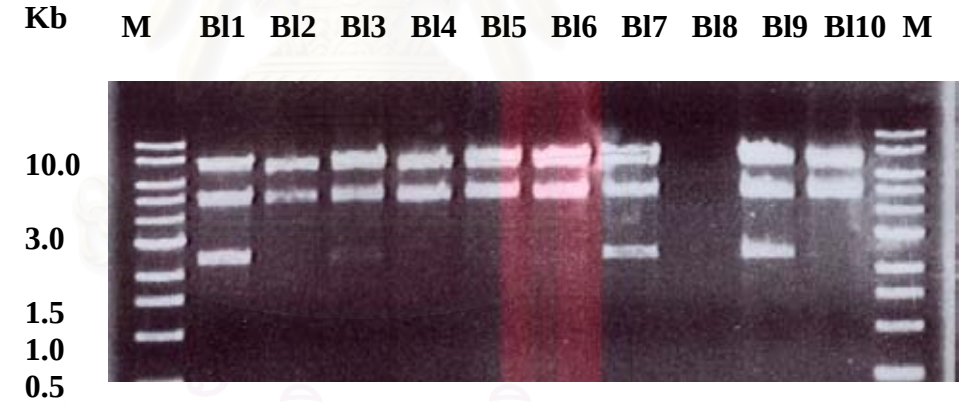


Figure 58. Restriction analysis results from clone Bl1-Bl10 by *Pst*I.

Expected band is 5648 + 1129 bp for insert in forward direction (Figure 58). The results revealed that clone Bl4, Bl5, Bl6 and Bl10 got the plasmid that insert in the inverse direction. Expected band is 4579+2190 bp. Other clones showed incorrected restricyion profiles on *Pst*I. *Bam*HI restriction analysis was done to confirm corrected clone again. From figure 60, the results revealed that only clone Bl4 and Bl6 got the corrected profile. Clone Bl4 and Bl6 showed the inverse direction of insertion. Expected band is

3527+3240 bp. This *Bam*HI restriction site checking is important since *Bam*HI site is necessary for switch deletion fragment into vector appropriate for *Clostridium* sp.

***Bam*HI:** inverse insertion showed fragment 3527+3240 bp

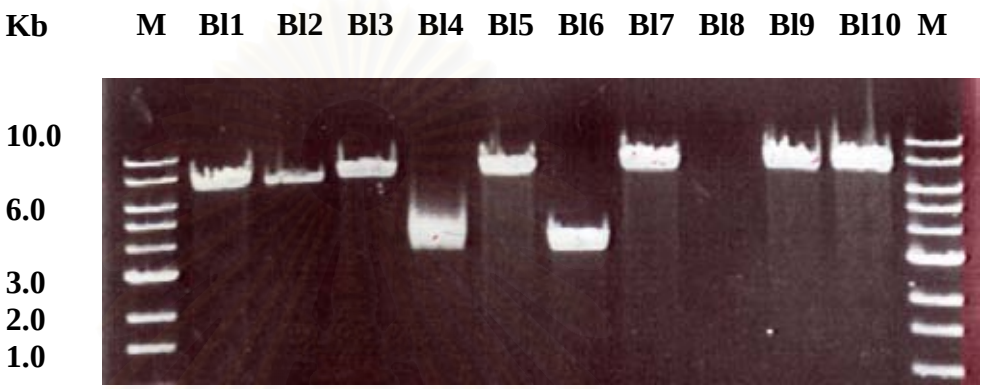


Figure 59. Restriction analysis results from clone Bl1-Bl10 by *Bam*HI.

This can be concluded that all transformants plasmid are reverse insertion as shown in figure 60 belowed.

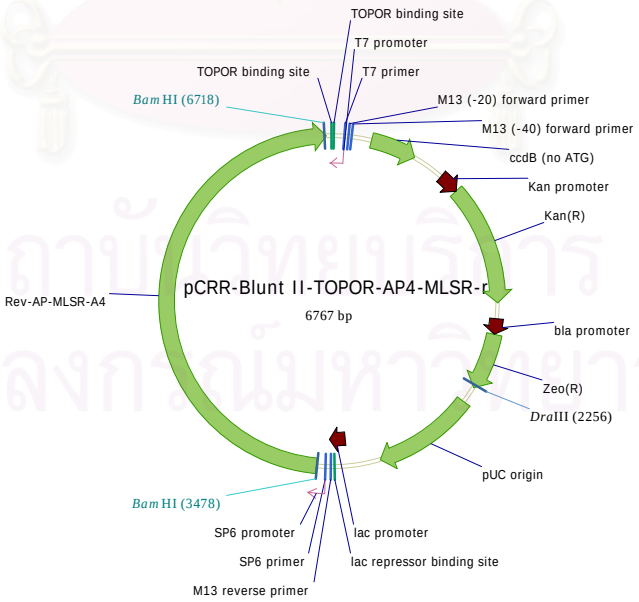


Figure 60. Restriction map of expected reverse insertion in all clones got.

In the same time, Bl4 plasmid was sent to sequencing in order to check the correction of sequence. BLAST analysis was performed to check the similarity between the designed model in vector NTI and the plasmid selected from clone. Results revealed that the identity between two sequences is 100% identity. So, we decided to use this clone for insert into pREPcac plasmid which is derivative plasmid of pCon-S2-1 plasmid.

Bl4 and Bl6 were selected to digest with *Bam*HI. 3240 bp-DNA fragment was ligated to pREPcac plasmid. Final plasmid before transform into *C. acetobutylicum* was shown as figure 62.

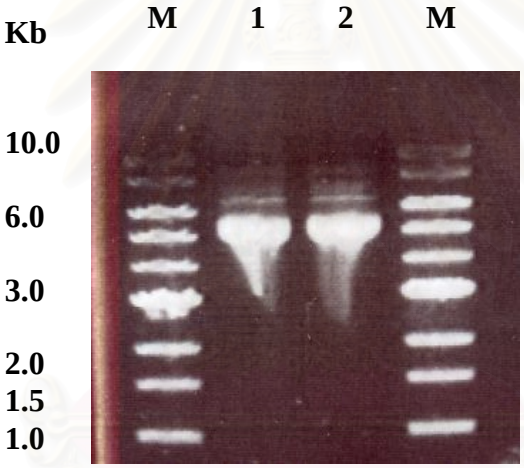
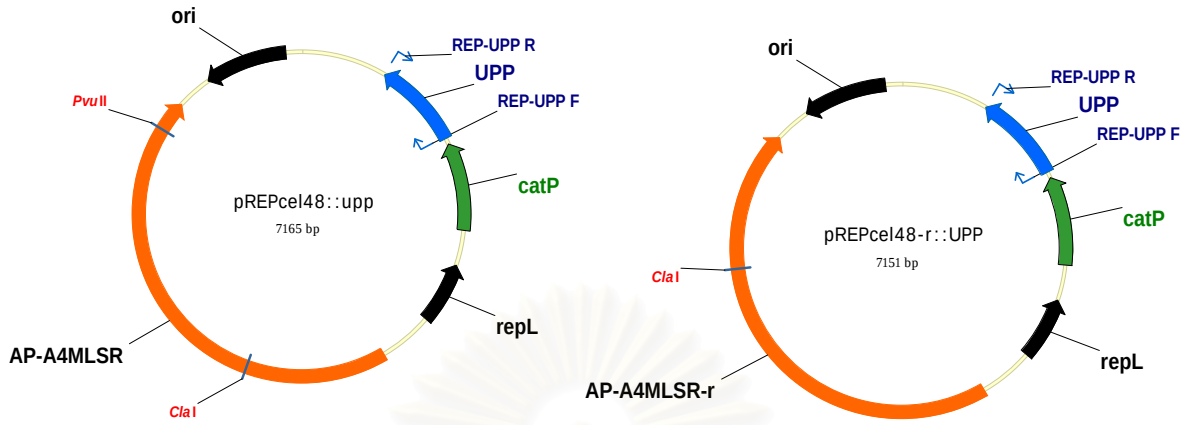


Figure 61. Plasmid Bl4 and Bl6. Lane1: plasmid Bl4 from clone extraction. Lane2: plasmid Bl6. The expected size is 6767 bp.



Restriction fragment:

***Sma*I:** 5290+1875

***Pst*I:** 4133+3032

***Pvu*II:** 7561

Restriction fragment:

***Sma*I:** 3966+3185

***Pst*I:** 5157+1994

***Pvu*II:** 7151

Figure 62. Expected plasmid pREPcel48A:: upp: forward form and reverse form.

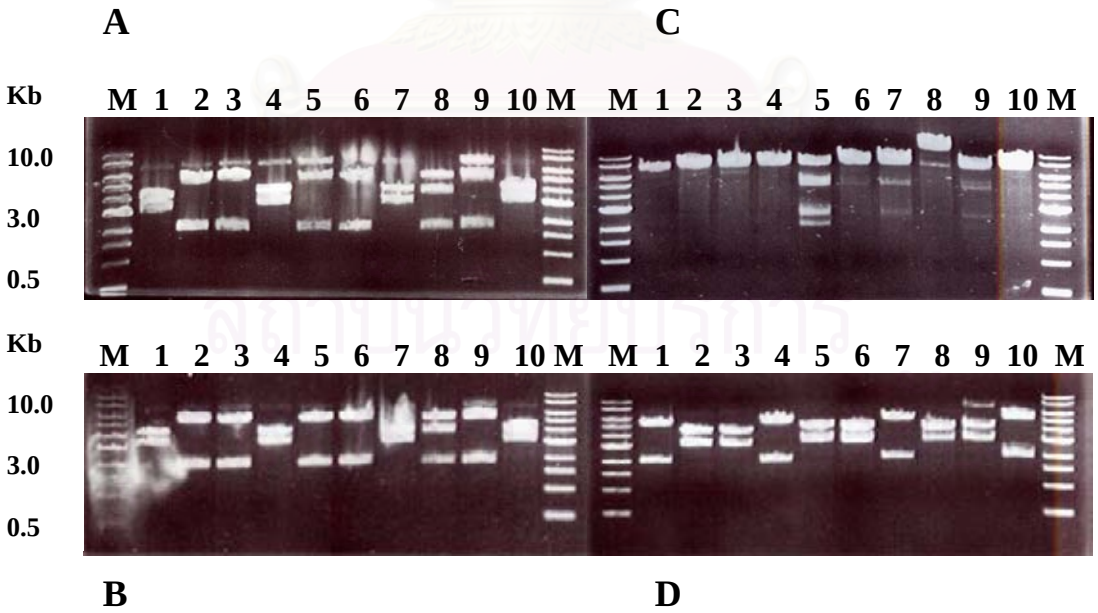


Figure 63. Restriction analysis of clone 27-1 to 27-10.

Figure 63, Those clones obtained from pREP*cel48*::upp transformation to *E. coli* DH5α then check restriction pattern. Results clones 27-1 to 27-10. *Sma*I, *Pst*I and *Pvu*II were select for distinguish different clone pattern. A) and B) Restriction profile of clone 27-1 to 27-10 analyzed with *Sma*I. C) Restriction profile of *Pvu*II analysis. D) Restriction profile of *Pst*I analysis. Clones 27-2, 3, 5, 6, 8, 9 show pattern match to expected digestion of forward form. Clones 27-1, 4, 7 show pattern of reverse form.

4.3.2.6 Transform into *C. acetobutylicum* into WT or ΔCAC1501 DNA-methyltransferase strain

Transformations were done in the anaerobic chamber. This process is necessary to culture *C. acetobutylicum* from spore directly before use in process. Time constants for electroporation were observed around 7.14-7.48 millisecond. Clones were appeared within 4 days cultivation after transformation process.

4.3.2.7 Screening for corrected phenotypic strains ($Thiam^S$ and Ery^R)

C. acetobutylicum transformants were screening for double crossing-over integrants. Colonies were picked and inoculate directly to synthetic medium supplemented with 40 μg/ml erythromycin. Incubate in 37 ºC overnight in anaerobic condition. Regenerated clones were replicated on reinforce clostridial agar (RCA) supplemented with 40 μg/ml erythromycin. Replica plates were made on both RCA and RCA supplemented with 50 μg/ml Thiamphenical. Correct phenotype clones were selected (Ery^R, $Thiam^S$) and check for spore-producing ability as shown in figure 64-65.

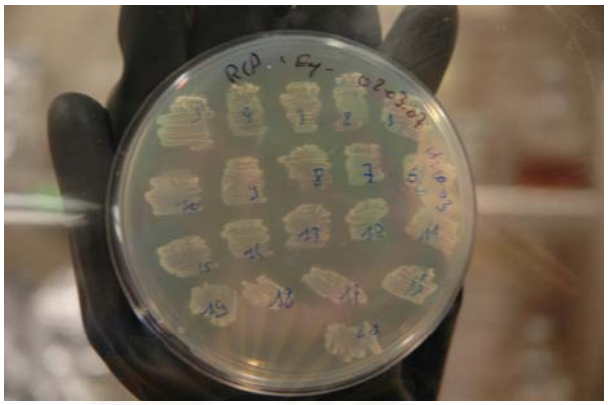


Figure 64. Erythromycin resistance ability of selected transformants

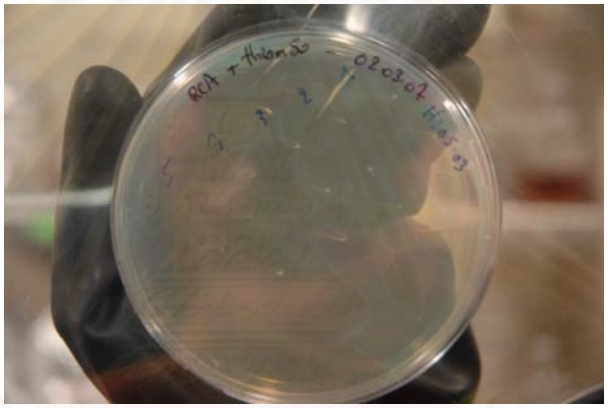


Figure 65. Thiamphenical sensitive ability of selected transformants

4.3.2.8 *C. acetobutylicum* genomic DNA extraction and amplification checking

After screening for Ery^R, $Thiam^S$ transformants, colonies were picked and cultivated in anaerobic condition. Genomic DNA of each clone was prepared. PCR amplification were performed to check availability of pREPcel48A:: upp plasmid. CelA1 and CelAP primers were used in reaction. Expected PCR amplification results from 13 transformants (+1 control) are indicated by 1.6 kbp PCR product as shown in figure 66.

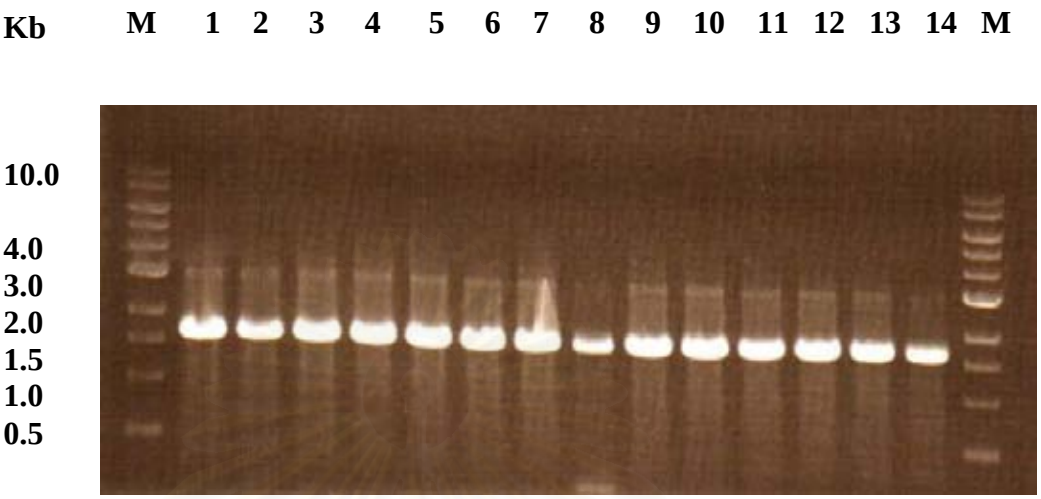


Figure 66. Amplification checking for pREPcel48A:: upp plasmid availability in *C. acetobutylicum* transformants

4.3.2.9 Screening for double-crossing-over integrants

Double-crossing-over process is the normal process by naturally between *C. acetobutylicum* chromosomal DNA and plasmid designed for gene deletion. These integrants were screened by detection of antibiotic resistant gene (erythromycin). Firstly, Ery^R, $Thiam^S$ transformants were kept and checking for pREPcel48A:: upp plasmid availability. This transformants were selected replicately on erythromycin supplemented media (3 times in 2YTG broth supplemented with 40 μg/ml erythromycin and then 3 times on clostridia growth agar supplemented with 40 μg/ml erythromycin). Results indicated all 14 clones are Ery^R.

4.3.2.10 MLS^R gene elimination by FLP recombinase enzyme system

In theory, erythromycin resistant gene was eliminated from clones that contained double-crossing-over integrants by FLP recombinase enzyme system. This FLP recombinase encoded gene was contained in plasmid pCLF1-1 (4911 bp, contained FLP recombinase gene as shown in figure 67)

and transformed to *C. acetobutylicum* directly. After MLS^R gene elimination, phenotype of transformants were changed from Ery^R, $Thiam^S$ to Ery^s, $Thiam^R$ since pCLF1-1 plasmid contained thiamphenical resistant gene as shown in figure 67. DNA amplification of gene around deletion cassette was reduced in size around 1.3 kb. In the experiment, we could not obtain this erythromycin gene eliminated clones since we had problems about pCLF1-1 transformation. This may be related to the FLP expression by thiolase promoter. This promoter is strong promoter for *C. acetobutylicum*. It may produce too much amount of enzyme that toxic to cells.

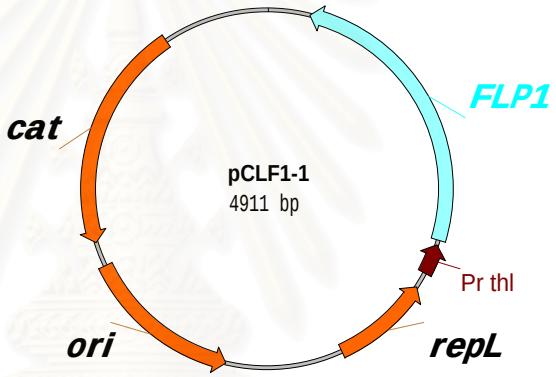


Figure 67. Plasmid pCLF1-1

4.3.2 Discussion

A clue in Metabolic Engineering strategies is the construction of recombinant strains. The inactivation/deletion of genes is one way to regulate metabolic pathway or abnormal gene in organism. In this article, gene deletion technique was applied for one enzyme subunit of *C. acetobutylicum* ATCC 824 cellulosome. *C. acetobutylicum* ATCC824 is one of the best-studied solventogenic Clostridia. Eventhough large cellulosomal gene cluster was found, this organism doesn't consume cellulose. Cellulosome improvement by genetic engineering was challenged for new recombinant strain in consolidate bioprocess without biomass pre-treatment. Sequence comparative analysis and activity assay were done in each main enzyme subunit of *C. acetobutylicum* and *C. cellulolyticum*. Enzyme activity lacking and mutation points were found in *C. acetobutylicum cel*48A gene which affected to overall cellulosome

activity. Purpose of work is to develop strategy for *cel*48A gene elimination. The approach is to improve cellulosome action by knock-out inactive gene in cellulosome cluster. Finally, pREPcel48::upp plasmid was constructed and transformed into *C. acetobutylicum*. Plasmid availability and phenotypic expression were observed in recombinant strain. In conclusion, this designed vector can be applied for gene crossing-over inactivation in *C. acetobutylicum* ATCC824. This designed vector is the first vector for knock off *cel*48A gene that successfully transformed into *C. acetobutylicum* ATCC824.

System to delete *cel*48A gene was desiged *in silico* and *in vitro* successfully. The recombinant *cel*48A was constructed notably and recombinant p952-SA-FA plasmid is ready-to-use in *C. acetobutylicum*. Available *cel*48A deletion strain will be use as p952-SA-FA recipient in further study.

4.4 Screening: alternative ways for improve sustainable strains in process

4.4.1 Solventogenic-cellulolytic Clostridia from 4-step-screening process in agricultural waste and cow intestinal tract

"Consolidated bioprocessing" (CBP) to produce solvent from cellulose and hemicellulose agricultural substrates is one of the most single process steps for an ideal goal to generate solvent by large-scale fermentation process. This means that cellulose production, cellulose hydrolysis, and production of a desired product are accomplished in a single process step (Ozkan *et al.*, 2001). In addition, the cost of fermentation process can be reduced and save the time for production at once time. Although many strains have been screened and successfully used in the industrial fermentation, only a limited number of them have survived in strain collections (Montoya *et al.*, 2001). Moreover, the lack of cellulolytic activity in the organism appears to be one of the main important problems for the efficient production of fuel from alternative cheaper substrate. Screening a variety of strains may be one of the useful strategies for developing new powerful bioprocess system.

Clostridium are the obligate anaerobes which are relatively large rods forming terminal or subterminal spores. They are able to ferment several organic compounds. Their end products include butyric acid, acetic acid, butanol, acetone, and large amounts of carbon dioxide and hydrogen gas. Most cells are Gram-positive which has variable fermentative metabolism. *Clostridium* can produce wide varieties of extracellular enzymes. They are usually found in soil and in the gastrointestinal tracts of animals and humans. As known that plant-cell-wall polysaccharides only become available as a sources of energy to the host animal through the activity of the rumen microflora (Ekinci *et al.*, 2001). For this reason, many previous studies focused on screening for cellulolytic and hemicellulolytic enzymes from anaerobic rumen bacterium. Previously described species of anaerobic, cellulolytic, moderately thermophillic and mesophilic bacteria include *Clostridium thermocellum, C. stercorarium, C. thermolacticum, C. thermocapriae, C. aerotolerans, C. acetobutylicum, C. cellobioparum* and *C. thermopapyrolyticum* (Ozkan *et al.*, 2001; Hazelwood *et al.*, 1993).

To identify the genus, biochemical tests are reliable process that accepted for a long times. However, it can not be used to identify species. More physiological and morphological tests are needed. Gelatin hydrolysis, glucose fermentation, nitrate reduction, spore shape and location, lecithinase, lipase, urease, lactose, maltose and etc. (Finegold *et al.*, 1986) are well-known tests for accuracy identification. Nowadays, some commercial kits are provided for that purposes. However, errors could occur from color observation by raw eyes (Clarridge, 2004). An alternative system for rapid strain selection and identification was then suggested in this article. The isolates of cellulolytic bacteria in this study came from various sources. Soil, cellulolytic dry decompose, and feaces from animal gastrointestinal tracts were observed. In this research, we developed methods to investigate and isolate the cellulolytic, solvent producing clostridia. Samples were collected from natural sources in Thailand. After isolated, the bacterial strains were characterized by a designed system from cellulolytic-enzyme-activity and fermentation-production points of view.

The objective of this research is to describe an alternative simple four-step screening process for cellulolytic Clostridia. An endospore formation, sulfite reducing ability and metabolic products were used for genus identification. Consequently, 16S rDNA sequences were analyzed for species identification. The results indicated that all selected isolates of chosen cellulolytic anaerobes can be identified and belonged to *Clostridium*.

4.4.2 Isolation of cellulolytic bacteria

All of the samples were inoculated into the cellulolytic selective media as described above. This media contained basal medium added with avicel as the carbon source. Butanol was added until final concentration in each bottle reached to 5 g/l to select butanol-tolerant anarobic bacteria. After subculture process was done for three times, Thirty isolates were found and tested to identify bacterial genus. Only 15 isolates (Fea-PA, Fea-P1, Sc-THB, Sc-THC, Sc-THD, Sc-THE, Deco-TH5, Deco-TH6, Decom-PA, Decom-PB, Decom-PE, Fea-PB, Fea-PC, Mix-P2, Sc-THA) were found belonging to the genus *Clostridium* after used four criterias for selection (see Table 1). Sources of the 15 cellulolytic stains were listed in Table 2 with most of them originating from decompose, cow dung and bagasses. All of the isolates were spore

forming, Gram-positive, rod-shaped bacteria. Morphological properties showed diverse differences as described in Table 16.

Table 16. Sources of cellulolytic clostridia.

Bacterial strain	Source	Colony morphology	Colony size (mm)	Total solvent production (g/l)
Fea-PA	Cow dung	Small, transparent, slime	2	3.221
Fea-P1	Cow dung	Very small	ND	1.998
Sc-THB	Bagasses	Small, pale, yellow shiny	1	0.263
Sc-THC	Bagasses	Medium, pink, round, shiny	2	4.00
Sc-THD	Bagasses	Very small	ND	0.897
Sc-THE	Bagasses	Pulvinate, white, undulate	2.5	0.719
Deco-TH5	Decompose	White with pink pulvinate, shiny, undulate	1.5	0.487
Deco-TH6	Decompose	Yellow, pulvinate, shiny	1	1.371
Decom-PA	Decompose	Circular, convex	2.0	2.724
Decom-PB	Decompose	Circular, convex	2.5	2.449
Decom-PE	Decompose	Circular, convex	1.0	2.74
Fea-PB	Cow dung	Very small	ND	3.633
Fea-PC	Cow dung	Circular, convex	2.5	0.073
Mix-P2	Cow decompose	Circular, convex	1.0	4.134
Sc-THA	Bagasses	Circular, convex	2.0	2.745

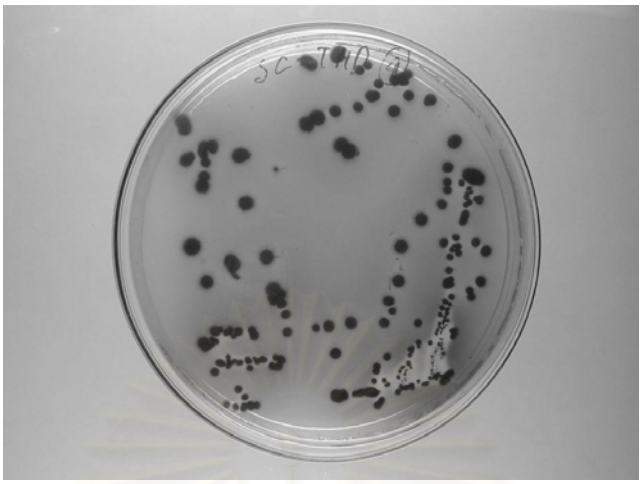


Figure 68. Black precipitation of iron and sulfide produced by bacterial isolates. Precipitation occurred inside colony of Sc-THD isolate

In order to identify genus of the bacteria, Sulfite reducing ability were tested on all of the isolates by using Differential Clostridial agar (see Figure 68). This medium contains rich base medium, appropriate starch concentration which induces spore germination (Figure 69). Sulfite and iron sources play the role as indicator. Most of sulfite will be changed to be sulfide by sulfite producing clostridia. We can observe this reaction by black precipitate with the iron in the medium (Gibbs and Freame, 1965). Results from this study revealed that some of cellulolytic butanol-tolerant clostridia show the ability to reduce sulfite. Twelve isolates (Fea-PA, Fea-P1, Sc-THB, Sc-THC, Sc-THD, Sc-THE, Deco-TH5, Deco-TH6, Decom-PE, Mix-P2, Fea-PB and Fea-PC) exhibit the positive reaction on the Differential Clostridial media as shown in figure1. On the other hand, three isolates (Decom-PA, Decom-PB and Sc-THA) showed negative results.

Figure 69. Terminal endospores presented in Sc-THE isolate appeared on 1000x microscope

4.4.3 Nutritional characteristics

Two types of media were used to determine the ability of solvent formation. Basal media were prepared and supplemented with avicel or glucose as carbon sources until final concentration equal to 20 g/l. After inoculation with 7-day-old culture in reinforced clostridial media and then incubated at 37°C for 96 hrs, from observation, all isolates produced ethanol and acetate. As Table 17, Sc-THC was able to produce ethanol more than other isolates in this research (3.939 g/l in avicel supplemented media). The same results also found in the case of acetate production by glucose supplemented media (6.23 g/l). Three isolates (Fea-PA, Sc-THD, Sc-THE) had an ability to produce butanol when using glucose as the source of carbon. This may indicate that butanol-producing bacteria may produce butanol when concentration of glucose in media is enough for generate energy for all activity in metabolism and solvent formation pathway.

Table 17. Anaerobic metabolism of each isolate in two different media after inoculation for 96 hours

Pure culture	Anaerobic metabolism in 2 different media at 96 hours (g/l)									
	acetone		butanol		ethanol		acetate		butyrate	
	avicel	glucose	avicel	glucose	avicel	glucose	avicel	glucose	avicel	glucose
Fea-P1	0.044	0.075	-	-	1.954	1.113	4.429	-	-	-
Fea-PA	0.032	2.960	-	0.153	3.189	1.634	7.385	-	-	-
Sc-THC	0.061	0.959	-	-	3.939	-	4.829	6.230	-	-
Deco-TH5	0.012	-	-	-	0.475	1.659	1.563	4.196	-	1.124
Sc-THB	-	-	-	-	0.263	1.252	0.0147	4.323	-	-
Sc-THD	-	0.126	-	0.225	0.897	2.257	2.123	7.425	-	-
Deco-TH6	-	0.047	-	-	1.371	1.836	4.313	-	-	-
Sc-THE	-	-	-	0.073	0.719	1.169	2.950	5.461	-	2.818
Decom-PA	-	0.564	-	1.964	0.027	0.524	2.697	2.907	-	1.789
Decom-PB	-	0.152	-	0.560	0.026	0.445	2.423	2.605	-	3.631
Decom-PE	-	0.242	-	0.723	0.012	0.181	2.728	2.589	-	2.008
Fea-PB	0.011	0.042	0.039	0.156	0.029	0.105	3.554	1.965	-	3.602
Fea-PC	-	0.019	0.031	0.068	0.042	0.168	-	4.161	-	4.663
Mix-P2	-	0.435	0.039	1.124	0.055	0.833	4.040	3.484	-	1.692
Sc-THA	-	0.005	0.028	0.198	0.032	0.582	2.685	2.092	-	3.659

Table 17 shows results of anaerobic metabolism in each isolates from screening. Samples were collected after cultures were incubated at 37ºC for 96 hours. Metabolisms of the isolates can be characterized into six groups depend on the studied conditions. Normally, types of metabolites produced alone can not group bacteria because cells can be effected by physiological and external conditions. It should be combined with other appropriate criterias. In this study, grouping of bacteria by metabolite products is helpful to decide which strains were suitable for further study especially solvent production ability and cellulose consumption. The first group was the bacteria producing acetone, ethanol and acetate as their metabolic final products. Bacteria belonged to this group was the isolates Fea-P1, Sc-THC and Deco-TH6. In the second group, acetone, butanol, ethanol and acetate were always the final products of this group. Bacteria in this group were Fea-PA and Sc-THD. The third group was identified by acetone, ethanol, acetate, and butyrate production. Deco-TH5 express metabolism in this way. Group 4 was the bacteria that can produce ethanol and acetate. Sc-THB belonged to this group. Group 5 was the bacteria producing butanol, ethanol, acetate and butyrate. Sc-THE was in this group. The last group was the largest in this screening. All of them produced acetone, butanol, acetate and butyrate.

Decom-PA, Decom-PB, Decom-PE, Fea-PB, Fea-PC, Mix-P2 and Sc-THA were included in this group. They can produce most various products in metabolism when compared with the other groups from screening. From the results analyzed, the highest acetone producer from all screened groups was Fea-PA (2.96 g/l). Best butanol producer was Decom-PA (1.964 g/l). Best ethanol producer was Sc-THC (3.939 g/l). This strain was interesting since it could produce ethanol from crystalline substrate directly at high level. Best acetate producer was Fea-PA (7.385g/l). Previous results showed that this strain also gave highest acetone production among all groups. The last one, Decom-PB, was the best butyrate producer (3.631 g/l). These higher-solvent production strains were selected to test cellulose utilization ability whereas Sc-THC, Fea-PA, Decom-PB and Decom-PA were selected to test cellulolytic activity. Firstly, colony observation was done on CMCase plate. This step was important to do according to cellulosome properties. Naturally, cellulosome is produced and activates cellulose degradation by its enzyme attached to cell surface. This leads to the extracellular cellulase activity. So, all strains produce halo-zone indicate were supposed to produce cellulosome hypothetically. All other 11 strains from screening process were also checked.

4.4.4 PCR amplification and sequencing

The 16S rDNA fragments were successfully amplified from genomic DNA templates of the bacterial samples. We prepared the DNA templates using rapidly heating technique by picking single colony of bacteria diameter around 0.5 cm from Clostridial differential media directly after cultured for seven days. The bacterial cells were disrupted by heat and then centrifuged to remove cell membranes. DNA supernatants were used directly as the templates for amplification. The 1.6 kb PCR fragments (Figure 70) were harvested and sequenced. Taxonomic status of each bacterial isolate was effectively identified using several bioinformatic tools below.

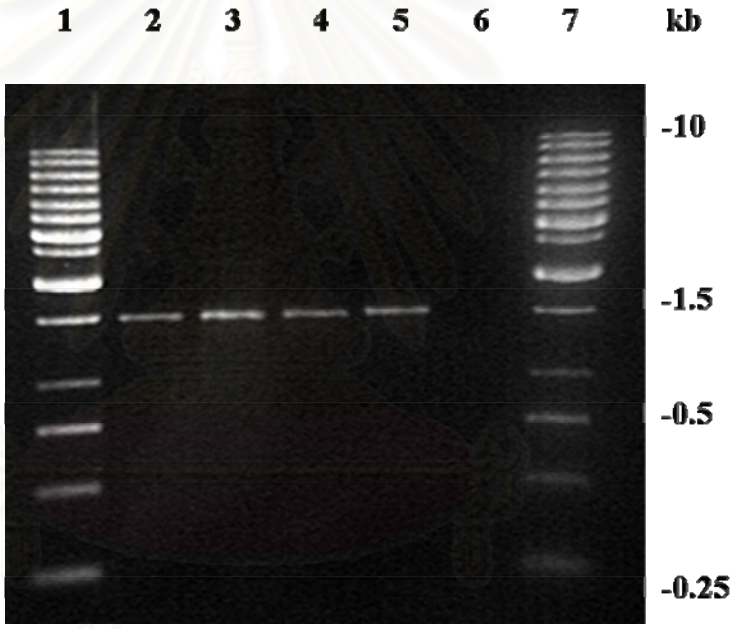


Figure 70. Analysis of 1.6 kb PCR products amplified from the isolates using 27F and 1492R primers

The gel was loaded with PCR products. Lane 1 and 7: loaded with 1-kb DNA standard (Fermentas); Lane 2-5: 16S rDNA PCR fragments of Decom-PE, Fea-P1, Fea-PA, and Mix-P2, respectively; Lane6: Amplification in the absence of template as the negative control.

4.4.5 Bioinformatic tools for sequence analyses

The 16S rDNA sequences were analyzed and compared with Ribosomal Database Project II. Identity percentages were obtained from nucleotide similarity between the sample sequences and the sequence database. Family classification from RDP II showed in the Table 3 with confidential threshold at 80%. The sequences were identified as in the family Clostridiaceae with sequence similarity ranging between 83-100%. The isolates were identified to the genus *Clostridium* by being compared to NCBI database before phylogenetic analysis. Some suspected species were suggested from RDP II with estimated similarity score = 1.000 (Table 18).

Table 18. Obtained sequences were compared with Ribosomal Database Project II (RDP II) database

Isolate	Family classification (RDP II)/ percentage by confidential threshold 80%
Fea-P1	Clostridiaceae / 100%
Fea-PA	Clostridiaceae / 100%
DecoTH5	Unclassified-Fermicutes / 83%
DecoTH6	Unclassified bacteria
Sc-THB	Unclassified bacteria
Sc-THC	Clostridiaceae / 100 %
Sc-THD	Clostridiaceae / 100%
Sc-THE	Clostridiaceae / 100 %
Decom-PA	Clostridiaceae / 100 %
Decom-PB	Clostridiaceae / 100 %
Decom-PE	Clostridiaceae / 100 %
Fea-PB	Unclassified
Fea-PC	Clostridiaceae / 100 %
Mix-P2	Clostridiaceae / 100%
Sc-THA	Clostridiaceae / 100 %

Phylogenetic analyses using NJ method revealed possible groupings of some of the bacterial isolates with known *Clostridium* species, though with less bootstrap supporting values, as shown in the Figure 71. Notably, Sc-THC was grouped with the pair of *C. aerotolerans* and *C. xylanolyticum* (99% bootstrap). The pair of *C. ramosum* and *C. spiroforme* was grouped (47% bootstrap) with both cluster of Deco-TH5, Sc-THB and Fea-PB, and cluster of Sc-THD, Fea-P1 and Fea-PA. Decom-PB, Sc-THA and Sc-THE were clustered together and close to *C. beijerinckii* while *C. diolis* was grouped with Decom-PA and Decom-PB couple. Last, Deco-TH6 was strongly paired with *C. nitrophenolicum* (92% bootstrap).

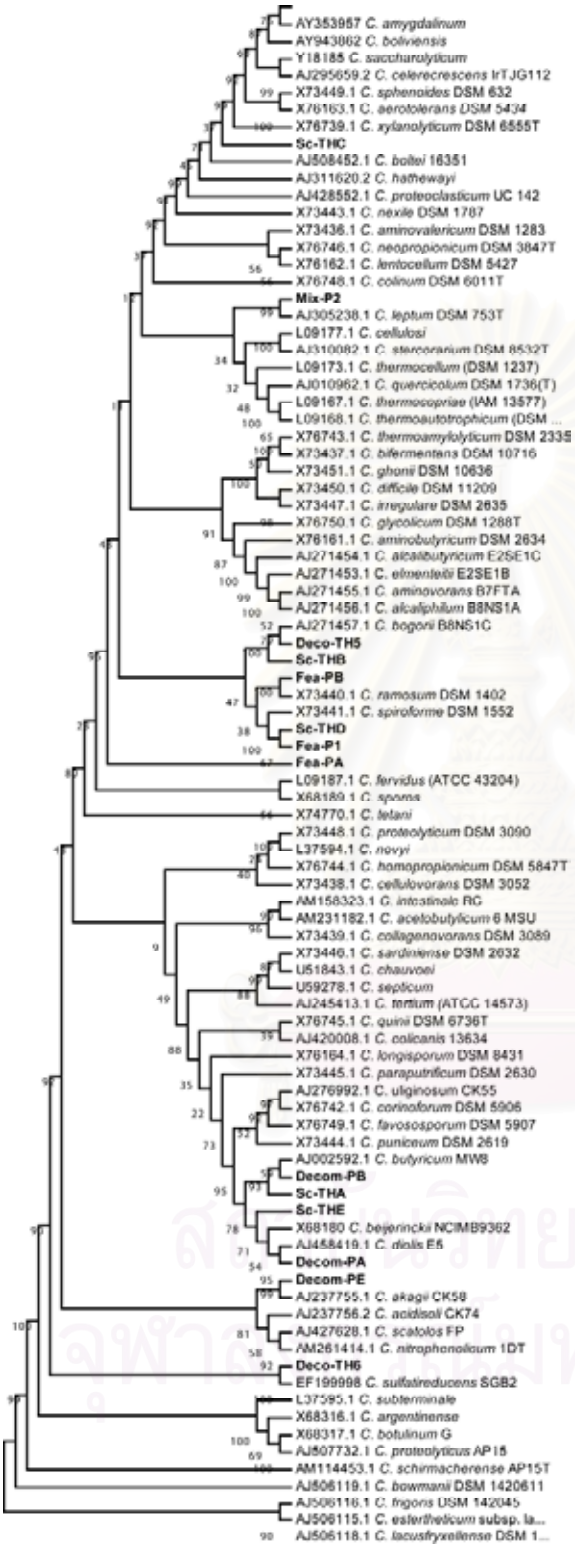


Figure 71. Phylogenetic tree analysis using Neighbour-Joining method of the 14 bacterial isolates compared with 74 *Clostridium* species. The percentages of bootstrap supporting values (1000 replicates) are shown along the branches

4.4.6 Enzyme assay

CMCase activity assays were tested on each identified isolate. Firstly, CMCase activity was estimated by clear halo-zone appearance. Halo-zone size refers to efficiency of enzyme secretion. Fea-PA and Decom-PB generated clear results of the halo zone. Colors of the halo zones were various depending on where the colonies formed. A large zone appeared when the cell grew on the surface while a diameter of the halo zone was hard to estimate because of diffused halo appearances (Figure 72). An embedded colony produced a more sharply clear zone with a visible colony. Fea-PA can express halo that can be measured an exact end of the zone while another halo appearance was diffused by distance from the colony. However, Decom-PA which was the best butanol producer in this study did not produce a halo-zone on the CMC plate. Interestingly, Decom-PE strain also produced a clear halo zone around colony. All of the isolates that expressed halo on CMC were tested for prior non-cellulosomal cellulase activity (exoglucanase and endoglucanase activities).

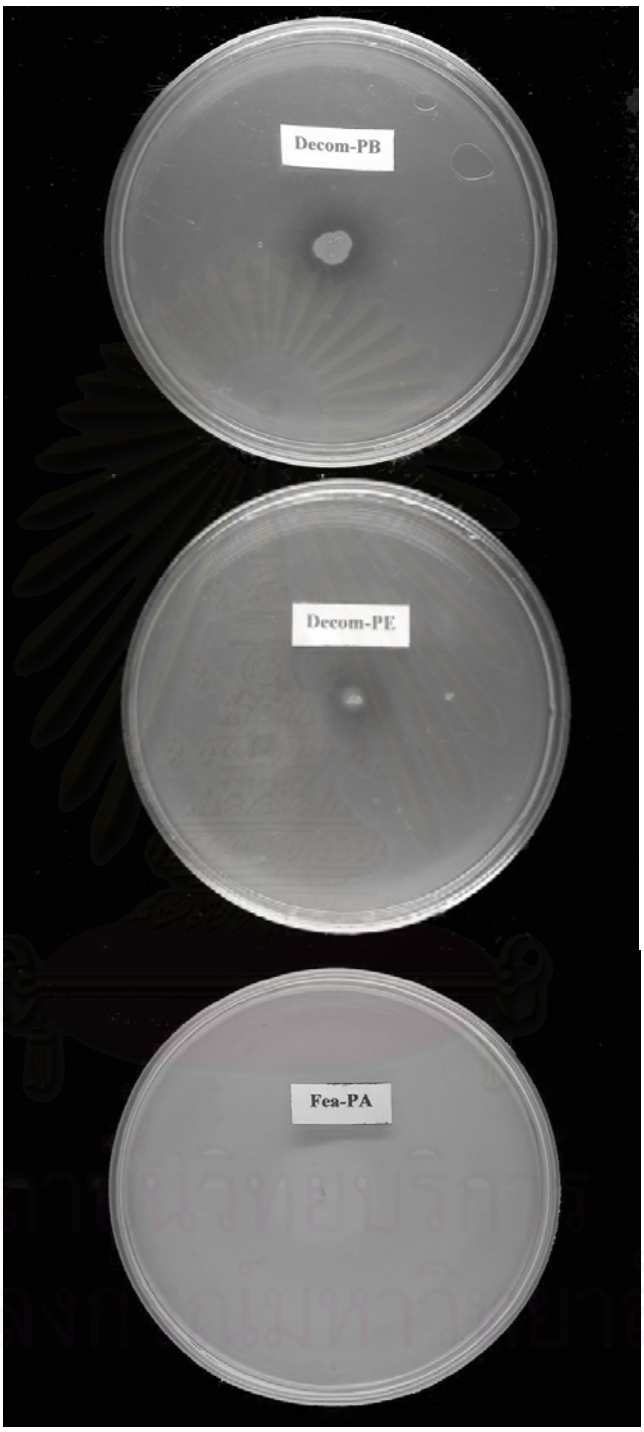
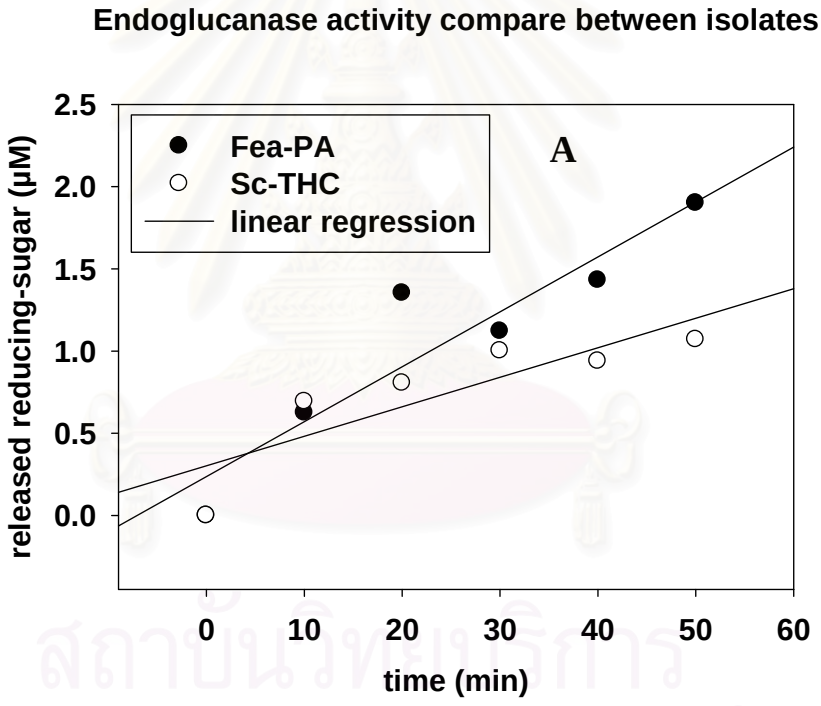


Figure 72. Orange halo on a red background indicated CMC activity on CMCase test after incubated for 72 hrs. Smaller clear zone appeared on the tested media by embedded Decom-PB isolate and Decom-PE isolate. Fea-PA isolate showed a larger clearing zone without colony

Fea-PA and Sc-THC were selected to test cellulolytic activity as their ability to produce solvents. From the results analyzed, the highest acetone producer from the screened group was Fea-PA (2.96 g/l) while the ethanol producer was Sc-THC (3.939 g/l). Even though these two isolates were identified as the acetone-ethanol production group by metabolic products, Fea-PA can produce a clear halo zone when detected on PC/agar plates containing 0.1% CM cellulose while this halo can not be detected in Sc-THC (Figure 72). In order to understand mechanism of the clear zone production, exoglucanase and endoglucanase activity assays were tested on both isolates (Figure 73).



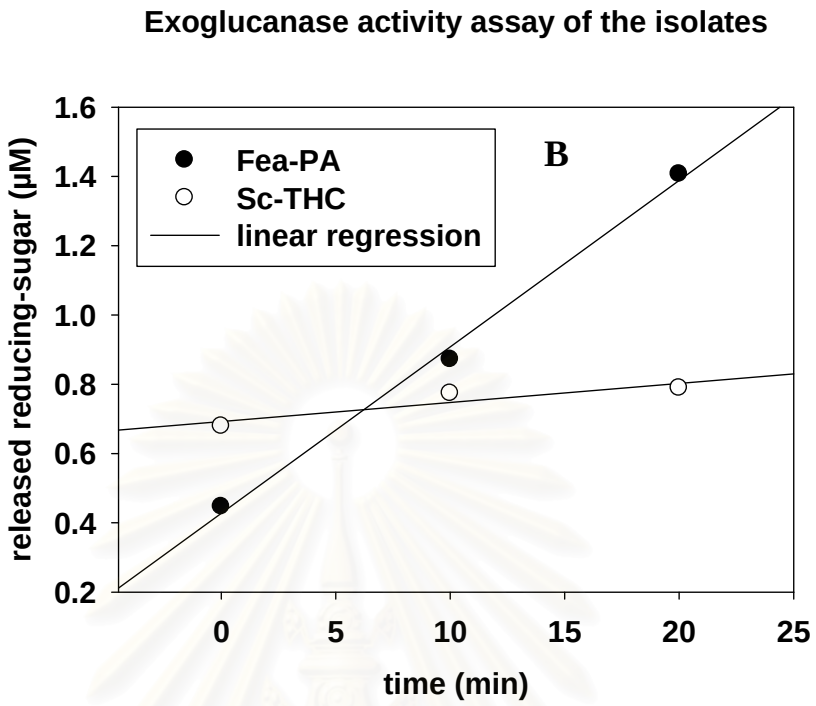


Figure 73. A) Endoglucanase activity tested by crude enzyme from Fea-PA and Sc-THC isolates. B) Exoglucanase activity tested by crude enzyme from Fea-PA and Sc-THC isolates. Reducing sugar was determined by 2, 2-BCA assay every 10 minutes of incubation

4.4.7 Cellulosome purification

All strains producing clear CMCase halo-zone were tested for cellulosome purification. Fea-PA, Decom-PB, Decom-PE isolates and *C. acetobutylicum* ATCC824 wild type stain were selected for cellulosome extraction. However, non-halo producers such as Sc-THC and Decom-PA were also studied. Expression media supplemented with glucose and cellobiose as inducers (MS+glucose+20 g/l cellobiose + 10 g/l avicel) showed no expression of cellulosome even antibody were changed (1^{st} antibody: anti-CelF from *C. cellulolyticum* and combination of anti-CelF, anti-CelC from *C. cellulolyticum*). This may be related to glucose concentration which can suppress cellulosome gene expression. However, non-cellulosome cellulase activity can be detected in the isolates producing a halo-zone on CMC both endo- and exoglucanase activities (table 19). Specific enzyme activities were calculated in International Unit/mg enzymes (IU/mg). Molecular weights of these cellulases will be identified further.

Table 19. Non-cellulosome cellulase productivity of the isolates determined by 2, 2-BCA assay

Isolate	Protein (µg/ml)	Endoglucanase activity (CMCase activity) on PC/agar plates containing 0.1% CM cellulose	Endoglucanase specific activity (IU/mg protein)	Exoglucanase specific activity (IU/mg protein)
Fea-PA	32.43	+	$5.1x10^{-3}$	$2.4x10^{-3}$
ScTHC	36.65	-	$2.4x10^{-3}$	$0.3x10^{-3}$
recombinant	34.9	+	ND	$4.03x10^{-3}$

4.4.8 Discussion

Based on their nature, anaerobic cellulolytic species are limited in their carbohydrate ranges. They are unable to use lipids or proteins as energy sources for growth. Media for selective screening is one of the important criteria to select interested strain. In this study, selective media was designed to grow and select only cellulolytic and solvent-producing bacteria. However, cellulolytic carbon sources are diverse. A cellulolytic bacteria like *C. thermocellum* does not grow easily or well on monosaccharides. It uses cellobiose in preference to glucose when both substrates present (Lynd *et al.*, 2002). Oligosaccharides and polysaccharides are more appropriate carbon sources for them. Avicel which is crystalline cellulose was selected to be the main carbon source to generate energy for cells. The cultures were grown and attempted to select until maintained pure culture. A single colony was picked and streaked at least three times. Metabolic product results showed all isolates in this study can use both crystalline cellulose and monosaccharide like glucose as carbon sources.

Since a fermentable carbohydrate was required for animal growth (Varel *et al.*, 1992), many researches have been reported on screening and characteristics of cellulolytic isolates from animal intestinal tracts, especially of pigs (Varel *et al.*, 1992; Dehority and Grubb, 1976; Robinson *et al.*, 1984; Russel, 1979). Biochemical reactions like the growth test on substrates (cellulose, cellobiose, maltose, starch or glycogen) are still necessary for determining substrate utilization of bacterial strains. Other biochemical tests for characterizing species of *Clostridium* are needed

absolutely (nitrate reduction, catalase, oxidase and urease production). In this study, three main criterias of biochemical tests were selected for rapid species identification. The first was sulfite-reducing, second was spore-forming ability and the last was secondary metabolic production. Almost all isolates of Clostridia species produce black precipitation of sulfide on differential clostridial media. However, some of *Clostridium* may lack of this character. Sc-THE isolate showed no precipitation by DCA. It is clear that not all of *Clostridium* species can express positive results on sulfite-reducing test. Spore-forming ability is one of the important characteristics to support the decision. Two groups can be divided from spore formation. First group was the *Clostridium* species producing round-shaped spores. This kind of spores were always found in numerous. Fea-P1, Fea-PA, Mix-P2, Sc-THB, Sc-THC, Sc-THD, Fea-PC and Sc-THA belonged to this first group. Another group presented oval-shaped spores at terminus and were rarely found (1-2 spores per observation area under microscope). Sc-THE, Deco-TH5, Deco-TH6, Decom-PA, Decom-PE and Fea-PB were belonged to the second group.

Enzymatic cellulose hydrolysis is generally a slow and incomplete process. However, in relatively short time (up to 48 hrs) the microbial consortium in the bovine rumen may hydrolyze cellulose to 60-65% (Schwarz, 2001). Although crystalline cellulose is homogeneous chemical, no single enzyme is able to hydrolyze it. In anaerobic bacteria, cellulose will be hydrolyzed by a multi-enzyme system. A unique extracellular supramolecular molecule was called “cellulosome”. It has potential activity on cellulolytic substrates.

After finish identification process, many questions about cellulase production of the isolates were set. Much knowledge has been obtained about the cellulosome and non-cellulosomal cellulases from other thermophilic cellulolytic species such as *C. thermocellum* but less is known about the anaerobic, mesophilic Clostridia species and their cellulose activities. However, the general picture emerging about the Clostridia is that of the cellulosome (Doi *et al.*, 1998, Lamed and Bayer, 1988). Experiments were designed in order to determine both cellulosomal cellulase and non-cellulosomal cellulases in selected isolates. Two non-cellulosomal cellulases have been identified: EngD and EngF (Doi *et al.*, 1998). EngD is very active on endoglucanase and

xylanase while EngF has lower activity (Foong and Doi, 1992). Fea-PA isolate showed endoglucanase activity both on the assay and carboxymethyl cellulose (CMC) but can not detect cellulosome (anti-CelF, anti-CelC from *C. cellulolyticum*). It can be noticed that this result exhibited the expression of EngD cellulase gene which is non-cellulosomal cellulases. Sc-THC isolate also showed endoglucanase activity, even though its halo zone can not be observed on CMC plates.

However, it can not be concluded that there was no cellulosome production among the isolates. It may be related to other factors such as specificity of antibody in used. Anti-CelF and CelC of *C. cellulolyticum* which used for immunoreaction may not be specific to the isolates. Since based on structural features, these enzymes have been divided into 14 different families of glycosidic hydrolases. Moreover, it is well-known that cellulosome can not be produced, or observed very few, in medium supplemented with glucose or cellobiose. However, it can be produced under condition with cellobiose in *C. acetobutylicum* without glucose (Sabathé and Soucaille, 2003). In this research cellulosome expression experiments were done with synthetic medium supplemented glucose as an initiate carbon source. Thus, cellobiose acts like inducer for cellulosome gene expression. It may be concluded that glucose exhibits as cellulosome gene suppressor. Carle-Urioste *et al*. (1997) reported the production of cellulases was repressed by CRE1 in the presence of glucose, but a basal level of production occured in the absence of glucose. However, the mechanism of repression of cellulosome expression by glucose or cellobiose is unknown.

CHAPTER V

CONCLUSION

5.1 Modification of *C. acetobutylicum* cellulosome by using genetic engineering technique

Since *C. acetobutylicum* cellulosome doesn't function on crystalline cellulose. This strain can not utilize crystalline cellulose as carbon source. Cellulolytic activity of each main cellulosome subunit was tested on on crystalline cellulose. Two hypotheses were set. Firstly, it related to cellulolytic activity of three main subunit in cellulosome Cel9X Cel9C and Cel48A. Secondly, binding ability between three main suunit to their scaffolding protein is loosening. So, our experiments were designed to proof that possibility and answer that question. Since cellulolytic activity of Cel9C and Cel9X subunit are clearly expressed as well as in equivalent Cel9E and Cel9G of *C. cellulolyticum*. Cel9C is a subunit has high ability to convert cellulolytic substrate to reducing sugar (Table 9). It can be concluded that Cel9C and Cel9X from *C. acetobutylicum* bind to their scaffolding protein with lower affinity than the cellulases from *C. cellulolyticum* and *C. thermocellum*. These results show that the cellulosome produced by *C. acetobutylicum* should be less stable than those of *C. thermocellum* and *C. cellulolyticum*. Thus, inactive cellulosome may be affected from both cellulosome assembly ability and point mutation occurred in Cel48A subunit.

Hybrid Cel48A-48F protein was designed to solve that problem. Overall results revealed that the chimera protein exhibited enhanced action on cellulose than in native *C. acetobutylicum* catalytic domain of cellulose. The action of hybrid enzyme over carboxymethylcellulose (CMC) developed for 30.8 folds from native (native Cel48A: 0.011 IU/μM, hybrid: 0.339 IU/μM). Action over Phosphoric acid swollen cellulose (PASC) is improved for 2771 folds (native Cel48A: 0.002 IU/μM, hybrid: 5.542 IU/μM). Then, p952-SA-FA plasmid was constructed in order to transform and promote hybrid gene expression in *C. acetobutylicum* ATCC824. Phenotypic expression and stability odf plasmid were checked before express protein. Cellulosome fraction was purified and identified in hybrid strain *C. acetobutylicum*.

5.2 Isolation and characterization of cellulolytic, solvent producing clostridia from soil samples in Thailand

Thirty isolates of cellulolytic and butanol tolerant anaerobic bacteria were obtained from screening in this medium. Fifteen isolates were rapidly classified as in the class Clostridia by three selected criterias (endospore formation, sulfite reducing ability and metabolic products). Secondary metabolites of the bacteria such as acetone, butanol and ethanol were varied depending on the process. Clostridial differential medium was used as a genus identification tool. Finally, PCR-Amplified gene fragments coding for 16S rDNA were analyzed as a key to identify bacteria species. This process can be used to screen and identify *Clostridium* species in short period. Cellulosome and non-cellulosome cellulases productivity were analyzed. The results revealed that the selected cellulolytic strains (such as Fea-PA) exhibited EngD non-cellulosome cellulase activity especially endoglucanase activity on carboxymethyl cellulose. The selective system in this research was appropriate for the screening of *Clostridiaceae* in a similarity range between 83-100%.

5.3 Cellulolytic activity and solvent producing ability comparative analysis between naturally selected and recombinant Clostridia

Developed strain of *C. acetobutylicum* ATCC824 with recombinant plasmid p952-SA-FA expressed enhanced *in vitro* cellulolytic activity in *E. coli* while strain type ATCC824 does not. Cellulosome fraction of recombinant strain exhibit different form of protein when compare with ATCC824. However, this strain needs to be tested further for cellulolytic activity and binding ability to scaffolding protein *in vivo* directly.

Comparative analysis between engineered strain p952-SA-FA and naturally selected clostridia shows improved exoglucanase activity of recombinant strain as shown in table 19. Phenotypic expression of p952-SA-FA strain exhibits growth in synthetic media supplemented with cellobiose and avicel (crystalline cellulose) after three day cultivation. Cellulose aggregation and gas production were observed in

three-seventh day culture. This may related to the exoenzyme action over crystalline cellulose substrate. It is a good sign that indicate improved cellulose activity. However, this needs to be proven and repeately tested before use in consolidate ferment application in the future.

REFERENCES

Ahsan, M. M., Kimura,T., Karita, S., Sakka, K., and Ohmiya, K. 1996. Cloning, DNA Sequencing, and Expression of the Gene Encoding *Clostridium thermocellum* Cellulase CelJ, the Largest Catalytic Component of the Cellulosome. J. Bacteriol. 178: 5732-5740.

Arai, T., Ohara, H., Karita, T., Kimura, T., Sakka, K., and Ohmiya, K. 2001. Sequence of *celQ* and properties of CelQ, a component of the *Clostridium thermocellum* cellulosome. Appl. Microbiol. Biotechnol. 57:660-666.

Attala, R. H., Hackney, J. M., Uhlin, I., and Thompson, N. S. 1993. Hemicelluloses as structure regulators in the aggregation of native cellulose. Int. J. Biol. Macromol. 15: 109-112.

Bagnara-Tardif , C., Gaudin, C., Belaich, A., Hoest, P., Citard, T., Belaich, J. P. 1992. Sequence analysis of a gene cluster encoding cellulases from *Clostridium cellulolyticum*. Gene. 119: 17-18.

Bayer E.A., Shoham, Y., and Lamed, R. 2000. Cellulose-decomposing prokaryotes and their enzyme systems. In Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K. H., and Stackebrandt, E. (eds.), The Prokaryotes: An evolving electronic resource for the microbiological community. pp. 135-144. New York: Springer-Verlag.

Bayer, E. A., Chanzy, H., Lamed, R., and Shoham, Y. 1998. Cellulose, cellulases and cellulosomes. Curr. Opin. Struct. Biol. 8: 548–557.

Bayer, E. A., Setter, E., and Lamed, R. 1985. Organization and distribution of the cellulosome in *Clostridium thermocellum*. J. Bacteriol. 163: 552-559.

Bayer, E. A., Kenig, R., and Lamed, R. 1983. Adherence of *Clostridium thermocellum* to cellulose. J. Bacteriol. 156: 818-827.

Beguin, P., Cornet, P., and Aubert, J. P. 1985. Sequence of a Cellulase Gene of the Thermophilic Bacterium *Clostridium thermocellum*. J. Bacteriol. 162: 102-105.

Bélaïch, A., Parsiegla, G., Gal, L., Villard, C., Haser, R., and Belaich, J-P. 2002. Cel9M, a new family 9 cellulase of the *Clostridium cellulolyticum* cellulosome. J. Bacteriol. 184: 1378-1384.

Bélaïch, J. P., Tardif, C., Bélaïch, A., and Gaudin, C. 1997. The cellulolytic system of *Clostridium cellulolyticum*. J. Biotechnol. 57: 3-14.

Bray, M. R., Johnson, P. E., Gilkes, N. R., Mcintosh, L. P., Kilburn, D. G., and Warren, R. A. J. 1996. Probing the role of tryptophan residues in a cellulose-binding domain by chemical modification. Protein Science. 5: 2311-2318.

Brown, R. M. J., and Saxena, I. M. 2000. Cellulose biosynthesis: a model for understanding the assembly of biopolymers. Plant. Physiol. Biochem. 38: 57-67.

Bronenmeier, K., and Staudenbauer, W. L. 1988. Resolution of *Clostridium stercorarium* cellulase by fast protein liquid cgromatography (FPLC). Appl. Microbiol. Biotechnol. 27: 432-436.

Carle-Urioste, J. C., Escobar-Vera, J., El-Gogary, S., Henrique-Silva, F., Torigoi, E., Crivellaro, O., Herrera-Estrella, A., El-Dorry, H. 1997. Cellulase induction in *Trichoderma reesei* by cellulose requires its own basal expression. J. Biol. Chem. 272:10169–10174.

Cavalho, A. L., Dias, F. M. V., Prates, J. A. M., Nagy, T., Gilbert, H. J., Davies, G., Ferreila, J., Romão, L. M. A., and Fontes, C. M. G. A. 2003. Cellulosome assembly revealed by the crystal structure of the cohesin-dockerin complex. PNAS. 100: 13809-13814.

Chauvaux, S., Béguin, P., Aubert, J. P., Bhat, K. M., Gow, L. A., Wood, T. M., and Bairoch, A.. 1990. Calcium-binding affinity and calcium-enhanced activity of *Clostridium thermocellum* endoglucanases D. Biochem. J. 265: 261-265.

Chen, J-S. 1993. Properties of acid- and solvent-forming enzymes of clostridia. In Woods, D. R. (eds.), The clostridia and biotechnology. pp. 51–76. Stoneham, Mass: Butterworth-Heinemann.

Clarridge, J. E. 2004. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. Clin. Microbiol. Rev. 17: 840-862.

Compere, A. L., and Griffith, W. L. 1979. Evaluation of substrates for butanol production. Dev. Ind. Microbiol. 20: 509-517.

Cole, B., and Fort, R. 2007. The composition and structure of wood [Online]. Available from: www. chemistry.umeche.maine.edu/Fort/Cole-Fort.html [2007, November 12].

Complex Carbohydrate Research Center, The University of Georgia. 2007. Plant cell walls [Online]. Available from: www.ccrc.uga.edu/~mao/intro/ouline.htm [2007 , September 10].

Coutinho, J. B., Moser, B., Kilburn, D. G., Warren, R. A. J., and Miller, R. C. Jr. 1991. Nucleotide sequence of endoglucanase C (CenC) of *Cellulomonas fimi*, its high-level expression in *Escherichia coli*, and characterization of its products. Mol. Microbiol. 5: 1221–1223.

Darvill, N. M., and Albersheim, P. 1980. The primary cell walls of flowering plant. In Tolbert, N. E. (eds.), The biochemistry of plants vol. 1. pp. 91–162. New York, Academic Press.

Davies, G., and Henrissat, B. 1995. Structures and mechanisms of glycosyl hydrolases. Structure 3: 853–859.

Dehority, B. A., and Grubb, J. A. 1977. Basal medium for the selective enumeration of rumen bacteria utilizing specific energy sources. Appl. Eviron. Microbiol. 32: 703-710.

Demain, A. L., Newcomb, M., and David W. J. H. 2005. Cellulase, Clostridia, and Ethanol. Microbiology and Molecular Biology Reviews. 69: 124-154.

Desvaux, M. 2005. The cellulosome of *Clostridium cellulolyticum*. Enz. Microb. Tech. 37: 373-385.

Dien, B. S., Cotta, M. A., Jeffries, T. W. 2003. Bacteria engineered for fuel ethanol production : current status. Appl. Microbiol. Biotechnol. 63: 258-266.

Doi, R. H., Park, J. S., Liu, C. C., Malburg, L. M., Tamaru, Y., Ichiishi, A., Ibrahim A. 1998. Cellulosome and noncellulosomal cellulases of *Clostridium cellulovorans*. Extremophiles. 2: 53-60.

Doi, R. H., Kosugi, A., Murashima, K., Tamaru, Y., and Han, S. O. 2003. Cellulosomes from mesophilic bacteria. J. Bacteriol. 185: 5907-5914.

Doi, R. H., and Kosugi, A. 2004. Cellulosome: plant-cell-wall degrading enzyme complexes. Nature Rev. Microb. 2: 541-551.

Doi, R. H., Golstein, M., Hashida, S., Park, J. S., and Takagi, M. 1994. The *Clostridium cellulovorans* cellulosome. Crit. Rev. Microbiol. 20: 87-93.

Doner, L. W., Bécard, G. 1991. Solubilization of gellan gels by chelation of cations. Biotechnol. Tech. 5: 25–29.

Doner, L. W., Irwin, P. L. 1992. Assay of reducing End-Groups in Oligosaccharide Homologues with 2, 2'-Bicinchoninate. Anal. Biochem. 202: 50-53.

Doner. L. W. 1998. An Improved Process for Isolation of Corn Fiber Gum," Cereal Chemistry. 75: 408-411.

Dunning, J. W., and Lathrop, E. C. 1945. The saccharification of agricultural residues. Ind. Eng. Chem. 28: 350-356.

Ekinci, M. S., Flint, H. J. 2001. Expression of bifunctional gene encoding xylanase and ß (1, 3-1, 4)-glucanase in gram-positive bacteria. Turk. J. Vet. Sci. 25: 771-775.

Enzyme Nomenclature [Online]. 2007. Available from : www.chem.qmul.ac.uk /iubmb/enzyme/EC3/2/1/15 [2007, August 9]

Faure, E., Belaich, A., Bagnara, C., Gaudin, C. and Belaich, J. P. 1989. Sequence analysis of the *Clostridium cellulolyticum* endoglucanase-A-encoding gene, *celCCA*. Gene. 84: 39–46.

Felix, C. R., and Ljungdahl, L. G. 1993. The cellulosome: the exocellularorganelle of *Clostridium*. Annu. Rev. Microbiol. 47: 791–819.

Fierobe, H. P., Pagès, S., Bélaïch, A., Champ, S., Lexa, D., and Bélaïch, J. P. 1999. Cellulosome of *Clostridium cellulolyticum*: Molecular study of the dockerin/cohesin interaction. Biochemistry. 38: 12822-12832.

Fierobe, H. P., Mechaly, A., Tardif, C., Belaich, A., Lamed, R., Shoham, Y., Belaich, J-P., and Bayer, E. A. 2001. Design and Production of Active Cellulosome Chimeras Selective Incorporation of Dockerin-containing Enzymes into Definited Functional Complexes. J. Biol. Chem. 276: 21257-21261.

Fierobe, H.-P., Bayer, E. A., Tardif, C., Czjzek, M., Mechaly, A., Belaich, A., Lamed, R., Shoham, Y., Belaich, J.-P. 2002. Degradation of Cellulose Substrates by Cellulosome Chimeras Substrate Targeting versus Proximity of enzyme Components. J. Biol. Chem. 277: 49621-49630.

Fierobe, H. P., Bagnara-Tardif, C., Gaudin, C., Guerlesquin, F., Sauve, P., Bélaïch, A., and Bélaïch, J. P. 1993. Purification and characterization of Endoglucanase C from *Clostridium cellulolyticum* Catalytic comparison with Endoglucanase A. Eur. J. Biochem. 217: 557-565.

Fierobe, H. P., Gaudin, C., Bélaïch, A., Loutfi, M., Faure, E., Bagnara, C., Baty, D., and Bélaïch, J. P. 1991. Characterization of Endoglucanase A from *Clostridium cellulolyticum*. J. Bacteriol. 173: 7956-7962.

Fierobe, H. P., Mechaly, A., Tardif, C., Belaich, A., Lamed, R., Shoham, Y., Belaich, J-P., and Bayer, E. A. 2001. Design and production of active cellulosome chimeras. Selective incorporation of dockerin-containing enzymes into defined functional complexes. J. Biol. Chem. 276: 21257–21261.

Fierobe, H-P., Mingardon, F., Mechaly, A., Belaich, A., Rincon, M. T., Pages, S., Lamed, R., Tardif, C., Belaich, J-P., Bayer, E. A. 2005. Action of Designer Cellulosomes on Homogeneous Versus Complex Substrates: Controlled Incorporation of Three distinct enzymes into a Defined Trifunctional Scaffodin. J. Biol. Chem. 280: 16325-16334.

Finegold, S. M., Barron, E. J. (1986). Bailey and Scott's Diagnostic Microbiology. pp. 590–633. Toronto: The CV Mosby Co.

Fischer, R. J., Helms, J., and Dürre, P. 1993. Cloning, sequencing, and molecular study of the dockerin/cohesion interaction. Biochemistry. 38: 12822-12832.

Fontes, C. M. G. A., Hazlewood, G. P., Morag, E., Hall, J., Hirst, B. H., and Gilbert, H. J. 1995. Evidence for a general role for non-catalytic thermostabilizing domains in xylanases from thermophilic bacteria. Biochem. J. 307: 151-158.

Foong, F. C-F., Doi, R. H. 1992. Characterization and Comparison of *Clostridium cellulovorans* Endoglucanase-Xylanase EngB and EngD Hyperexpressed in *Escherichia coli.* J. Bacteriol. 174: 1403-1409.

Gal, L., Gaudin, C., Bélaïch, A., Pagès, S., Tardif, C., and Bélaïch, J.P. 1997. CelG from *Clostridium cellulolyticum*: a multidomain endoglucanases acting efficiently on crystalline cellulose. J. Bacteriol. 179: 6595-6601.

Gal, L., Pages, S., Gaudin, C., Belaich, A., Reverbel-Leroy, C., Tardif, C., and Belaich, J-P. 1997. Characterization of the Cellulolytic Complex (Cellulosome) Produced by *Clostridium cellulolyticum*. Appl. Environ. Microbiol. 63: 903-909.

Gal, L., Pagès, S., Gaudin, C., Bélaïch, A., Reverbel-Leroy, C., and Tardif, C. 1997. Characterization of the cellulolytic complex (cellulosome) produced by *Clostridium cellulolyticum*. Appl. Environ. Microbiol. 63: 903–909.

Gaudin, C., Bélaïch, A., Champ, S., and Bélaïch, J. P. 2000. CelE, a multidomain cellulase from *Clostridium cellulolyticum*: a key enzyme in the cellulosome?. J. Bacteriol. 182: 1910–1915.

Gibbs, B. M., Freame, B. 1965. Methods for the recovery of clostridia from foods. J. Appl. Bact. 28: 95-111.

Gilad, R., Rabinovich, L., Yaron, S., Bayer, E. A., Lamed, R., Gilbert, H. J., and Shoham, Y. 2003. CelI, a Noncellulosomal Family 9 Enzyme from *Clostridium thermocellum*, Is a Processive Endoglucanase That degrades Crystalline Cellulose. J. Bacteriol. 185: 391-398.

Gillan, D. C., Speksnijder, A. G. C. L., Zwart, G., and Ridder, C. D. 1998. Genetic Diversity of the Biofilm *Montacuta ferruginosa* (Mollusca, Bivalva) as Evaluated by Denaturing Gradient Gel Electrophoresis Analysis and Clonong of PCR-Amplified Gene Fragments Coding for 16S rRNA. Appl. Environ. Microbiol. 64: 3463-3472.

Girbal, L., and Soucaille, P. 1994. Regulation of *Clostridium acetobutylicum* Metabolism as Revealed by Mixed-Substrate Steady-State Continuous Cultures: Role of NADH/NAD Ratio and ATP Pool. J. Bacteriol. 176: 6433-6438.

Gottschalk, G. (1986). Bacterial Metabolism. pp. 208-282. Springer-Verlag: New York.

Green, E. M., Boynton, Z. L., Harris, L. M., Rudolph, F. B., Papoutsakis, E. T., and Bennett, G. N. 1996. Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. Microbiology. 142: 2079–2086.

Grepinet, O., Chebrou, M. C., and Beguin, P. 1988. Purification of *Clostridium thermocellum* xylanase Z expressed in *Escherichia coli* and identification of the corresponding product in the culture medium of *C. thermocellum*." J. Bacteriol. 170: 4576-81.

Halstead, J. R., Vercoe, P. E., Gilbert, H. J., Davidson, K., Hazlewood, G. P. 1999. A family 26 mannanase produced by *Clostridium thermocellum* as a component of the cellulosome contains a domain which is conserved in mannanases from anaerobic fungi. Microbiology. 145: 3101-3108.

Hamamoto, T., Foong, F., Shoseyov, S., and Doi, R. H. 1992. Analysis of functional domains of endoglucanases from *Clostridium cellulovorans* by gene cloning, nucleotide sequencing, and chimeric protein construction. Mol. Gen. Genet. 231: 472–479.

Hammel, H. B., Pater, I. D., Gibbard, S., Lockwood, G. W. and Rages, K. 2005. Uranus in 2003: Zonal winds, banded structure, and discrete features. Icarus 175: 534–545.

Handelsman, T., Barak, Y., Nakar, D., Mechaly, A., Lamed, R., Shoham, Y., and Bayer, E. A. 2004. Cohesin–dockerin interaction in cellulosome assembly: a single Asp-to-Asn mutation disrupts high-affinity cohesin–dockerin binding. FEBS Lett. 572: 195-200.

Harris, L. M., Welker, N. E., and Papoutsakis, E. T. 2002. Northern, morphological, and fermentation analysis of *spo0A* inactivation and overexpression in *Clostridium acetobutylicum* ATCC 824. J. Bacteriol. 184: 3586–3597.

Hayashi, H., Takehara, M., Hattori, T., Kimura, T., Karita, S., Sakka, K., and Ohmiya, K. 1999. Nucleotide sequences of two contiguous and highly homologous xylanase genes *xynA* and *xynB* and characterization of XynA from *Clostridium thermocellum.* Appl. Microbiol. Biotechnol. 51: 348-357.

Hayashi, H., Takagi, K-I., Fukumara, M., Kimura, T., Karita, S., Sakka, K., and Ohmiya, K. 1997. Sequence of *xynC* and Properties of XynC, a Major Component of the *Clostridium thermocellum* Cellulosome. J. Bacteriol. 179: 4246-4253.

Hazelwood, G. P., Davidson, K., Laurie, J. I., Huskison, N. S., and Gilbert, H. J. 1993. Gene sequence and properties of CelI, a family E endoglucanase from *Clostridium thermocellum*. J. Gen. Microbiol. 139: 307-316.

Hazlewood, G. P., Davidson, K., Clarke, J. H., Durrant, A. J., Hall, J. and Gilbert, H. J. 1990. Cloning, DNA Sequencing, and Expression of the Gene Encoding *Clostridium thermocellum* Cellulase CelJ, the Largest Catalytic Component of the Cellulosome. J. Bacteriol. 178: 5732-5740.

Henrissat, B., and Davies, G. J. 2000. Glycoside Hydrolases and Glycosyltransferases. Families, Modules, and Implications for Genomics. Plant Physiology. 124: 1515–1519.

Hogsett, D. A., Ahn, H-J., Bernadez, T. D., South, C. R., and Lynd, L. R. 1992. Direct microbial conversion. Prospects, progress, and obstacles. Appl. Biochem. Biotechnol. 35: 527-541.

Ingram, L. O., Conway, D. P., Sewell, G. W., Preston, J. F. 1987. Genetic Engineering of Ethanol Production in *Escherichia coli*. Appl. Environ. Microbiol. 53(10): 2420-2425.

Irwin D., Shin, D. H., Zhang, S., Barr, B. K., Sakon, J., Karplus, P. A., and Wilson, D. B. 1998. Roles of the catalytic domain and two cellulose binding domains of *Thermonospora fusca* E4 on cellulose hydrolysis. J. Bacteriol. 180: 1709-1714.

Irwin, D. C., Zhang, S., and Wilson, D. B. 2000. Cloning, expression and characterization of a Family 48 exocellulase, Cel48A, from *Thermobifida fusca*. Eur. J. Biochem. 267: 4988-4997.

Johnson, E. A., Madia, A., and Demain, A. L. 1981. Chemically defined medium for growth of the anaerobic cellulolytic thermophile, *Clostridium thermocellum*. Appl. Environ. Microbiol. 41: 1060-1062.

Jolif, F. G., Béguin, P., and Aubert, J. P. 1986. Nucleotide sequence of the cellulase gene *celD* encoding endoglucanase D of *Clostridium thermocellum.* Nuc. Acids. Res. 14: 8605-8613.

Kakiuchi, M., Isui, A., Suzuki, K., Fijino, T., Fujino, E., Kimura, T., Karita, S., Sakka, K., and Ohmiya, K. 1998. Cloning and DNA sequencing of the genes encoding

Clostridium josui scaffolding protein CipA and cellulase CelD and identification of their gene products as major components of the cellulosome. J. Bacteriol. 180: 4303-4308.

Kataewa, I. A., Seidel, R. D., Li, X. L., and Ljungdahl, L. G. 2001. Property and Mutation Analysis of the CelK Cellulose-Binding Domain from the *Clostridium thermocellum* Cellulosome. J. Bacteriol. 183: 1552-1559.

Kataewa, I., Guglielmi, G., and Beguin, P. 1997. Interaction between *Clostridium thermocellum* endoglucanase CelD and polypeptides derived from the cellulosome-integrating protein CipA: stoicheometry and cellulolytic activity of the complexes. J. Bacteriol. 326: 617-624.

Khademi S., Guarino, L. A., Watanabe, H., Tokuda, G., and Meyer, E. F. 2002. Structure of an endoglucanases from termite, *Nasutitermes takasagoensis*. Acta. Cryst. D. 58: 563-659.

Kim, A. Y., Attwood, G. T., Holt, S. M., White, B. A., and Blaschek, H. P. 1994. Heterologous Expression of Endo-β-1,4-D- Glucanase from *Clostridium cellulovorans* in *Clostridium acetobutylicum* ATCC824 following Transformation of the *engB* Gene. Appl. Environ. Microbiol. 60: 337-340.

Kim, A. Y., Vertes, A. A., and Blaschek, H. P. 1990. Isolation of a Single-Stranded Plasmid from *Clostridium acetobutylicum* NCIB6444. Appl. Environ. Microbiol. 56: 1725-1728.

Kim, J., and Yun, S. 2006. Discovery of Cellulose as a Smart Material. Macromolecules. 39: 4202-4206.

Koukiekolo, R., Cho, H-Y., Kosugi, A., Inui, M., Yukawa, H., and Doi, R. H. 2005. Degradation of Corn Fiber by *Clostridium cellulovorans* Cellulases and Hemicellulases and Contribution of Scaffolding Protein CbpA. Appl. Environ. Microbiol. 71: 3504-3511.

Kraulis, J., Clore, G. M., Nilges, M., Jones, T. A., Pettersson, G., Knowles, J., Gronenborn, A. M. 1989. Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from Trichoderma reesei. A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. Biochemistry. 18: 7241-57.

Kruus, K., Luaa, A. C., Demain, A. S., and Wu, J. H. D. 1995. The anchorage function of CipA (CelL), a scaffolding protein of the *Clostridium thermocellum* cellulosome. Proc. Natl. Acad. Sci. 92: 9254-9258.

Kruus, K., Wang, W. K., Ching, J., and Wu, J. H. D. 1995. Exoglucanase Activities of the Recombinant *Clostridium thermocellum* CelS, a Major Cellulosome Component. J. Bacteriol. 177: 1641-1644.

Kurokawa, J., Hemjinda, E., Arai, T., Kimura, T., Sakka, K., and Ohmiya, K. 2002. *Clostridium thermocellum* cellulase CelT, a family 9 endoglucanase without an Ig-like domain or family 3c carbohydrate-binding module. Appl. Microbiol. Biotechnol. 59: 455-461.

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

Lamed, R., Bayer, E. A. 1988. The cellulosome of *Clostridium thermocellum*. Adv. Appl. Microbiol. 33: 1-46.

Lamed, R., Setter, E., and Bayer, E. A. 1983. Characterization of a cellulose-binding, cellulase-containing complex in *Clostridium thermocellum*. J. Bacteriol. 156: 828-836.

Lamed, R., Kenig, R., Setter, E., and Bayer, E. A. 1985. Major characteristics of the cellulolytic system of *Clostridium thermocellum* coincide with those of the purified cellulosome. Enzyme Microb. Technol. 7: 37-41.

Lee, S. F., Forsberg, C. W., and Gibbins, L. N. 1985. Cellulolytic activity of *Clostridium acetobutylicum*. Appl. Environ. Microbiol. 50: 220–228.

Lemaire, M., and Béguin, P. 1993. Nucleotide sequence of the *celG* gene of *Clostridium thermocellum* and characterization of its product, endoglucanase CelG. J. Bacteriol. 175: 3353-3360.

Leschine, S. B. 2004. Microbial Physiology and Diversity: Cellulose and Chitin Decomposition, Biofilms on Natural Polymers, Fuels from Biomass [Online]. Available from: www.bio.umass.edu/micro/faculty/leschine.html [2007, September 12]

López-Contreras, A. M., Martens, A. A., Szijarto, N., Mooibroek, H., Classen, P. A. M., Oost, J. V. D., and Vos, W. M. D. 2003. Production by *Clostridium*

acetobutylicum ATCC824 of CelG, a Cellulosomal Glycoside Hydrolase Belonging to Family 9. Appl. Environ. Microbiol. 69: 869-877.

Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. 193: 265-275.

Lynd, L. R., Weimer, P. J., Zyl, W. H. V., and Pretorius, I. S. 2002. Microbial Cellulose Utilization: Fundamentals and Biotechnology. Microbiol. Molec. Biol. Rev. 66: 506-577.

Margolles-Clark, E. 1996. Molecular characterization of side-chain cleaving hemicellulases of *Tricoderma reesei*. Academic dissertation. Department of Biosciences. Division of Genetics. U. Helsinki, Finland.

Matsunaka, T. 2001. Plant nutritional biochemistry [Online]. Available from: www. cse.naro.affrc.go.jp/.../interest/interest_e.htm [2007, December 10]

Mayer, F., Coughlan, M. P., Mori, Y., and Ljungdahl, L. G. 1987. Macromolecular organization of the cellulolytic enzyme complex of *Clostridium thermocellum* as revealed by electron microscopy. Appl. Environ. Microbiol. 53: 2785-2792.

Mechaly, A., Fierobe, H-P., Belaich, A., Belaich, J-P., Lamed, R., Shoham, Y., and Bayer, E. A. 2001. Cohesin-Dockerin Interactioni in Cellulosome assembly. A single hydroxyl group of a dockerin domain distinguished between nonrecognition and high affinity recognition. J. Bacteriol. Chem. 276: 9883-9888.

Mermelstein, L. D., Papoutsakis, E. T., Petersen, D. J., and Bennett, G. N. 1993. Metabolic engineering of *Clostridium acetobutylicum* for increased solvent production by enhancement of acetone formation enzyme activities using a synthetic acetone operon. Biotechnol. Bioeng. 42: 1053-1060.

Mingardon, F., Perret, S., Belaich, A., Tardif, C., Belaich, J.-P., Fierobe, H.-P. 2005. Heterologous Production, Assembly, and Secretion of a Minicellulosome by *Clostridium acetobutylicum* ATCC 824. Appl. Environ. Microbiol. 71: 1215-1222.

Miras, I., Schaeffer, F., Béguin, P., and Alzari, P. M. 2002. Mapping by site-directed mutagenesis of region responsible for cohesin-dockerin interaction on the

surface of the seventh cohesin domain of *Clostridium thermocellum* CipA. Biochem. 41: 2115-2119.

Montoya, D., Arevalo, C., Gonzales, S., Aristizabal, F., and Schwarz, W. H. 2001. New solvent producing Clostridium sp. Strains, hydrolyzing a wide range of polysaccharides, are closely related to *Clostridium butyricum*. J. Ind. Microbiol. Biotechnol. 27: 329-335.

Morag, E., Bayer, E. A., Hazlewood, G. P., Gilbert, H. J., and Lamed, R. 1993. Cellulase S_s (CelS) is synonymous with the major cellobiohydrolase (subunit S8) from the cellulosome of *Clostridium thermocellum*. Appl. Biochem. Biotechnol. 43: 147-151.

Naganuma, T., Kato, C., Hirayama, H., Moriyama, N., Hashimoto, J., and Horikoshi, K. 1997. Intracellular Occurance of ε-Proteobacterial 16S rDNA Sequences in the Vestimantiferan Trophosome. J. Oceanography. 53: 193-197.

Nakhamanovich, B. M., and Sticheblykina, N. A. 1959. Fermentation of pentoses of corn cob hydrolyzates by *Clostridium acetobutylicum*. Mikrobiologiya. 28: 91-96.

Navarro, A., Chebrou, M-C., Béguin, P., and Aubert, J.-P. 1991. Nucleotide sequence of the cellulase gene *celF* of *Clostridium thermocellum*. Res. Microbiol. 142: 927-936.

Nölling, J., Breton, G., Omelchenko, M. V., Makarova, K. S., Zeng, Q., Gibson, R., Lee, H. M., Dubois, J., Qiu, D., Hitti, J., GTC Sequencing Center Production, Finishing and Bioinformatics Teams, Wolf, y. L., Tatusov, R. L., Sabathé, F., Doucette-Sam, L., Soucaille, P., Daly, M. J., Bennett, G. N., Koonin, E. V., and Smith, D. G. 2001. Genome Sequence and Comparative Analysis of the Solvent-producing Bacterium *Clostridium acetobutylicum*. J. Bacteriol. 183: 4823-4838.

Ozkan, M., Desai, S. G., Zhang, Y., Stevenson, D. M., Beane, J., White, E. A., Guerinot, M. L., and Lynd, L. R. 2001. Caracterization of 13 newly isolated strains of anaerobic, cellulolytic, thermophilic bacteria. J. Ind. Microbiol. Biotchnol. 27: 275-280.

Pagès, S., Bélaich, A., Tardif, C., Reverbel-Leroy, C., Gaudin, C., and Bélaich, J. P. 1996. Interaction between the endoglucanase CelA and the scaffolding protein CipC of the *Clostridium cellulolyticum* cellulosome. J. Bacteriol. 178: 2279-2286.

Pagès, S., Bélaïch, A., Bélaïch, J. P., Morag, E., Lamed, R., Shoham, Y., and Bayer, E.A. 1997. Species-specificity of the cohesin-dockerin interaction between *Clostridium thermocellum* and *Clostridium cellulolyticum*: Prediction of specificity determinants of the dockerin domain. Proteins. 29: 517-527.

Parsiegla, G., Juy, M., Reverbel-Leroy, C., Tardif, C., Bélaïch, J. P., and Driguez, H. 1998.The crystal structure of the processive endocellulase CelF of *Clostridium cellulolyticum* in complex with a thiooligosaccharide inhibitor at 2.0 A resolution, EMBO J. 17: 5551–5562.

Rapp, P. and Beerman, A. 1991. Bacterial cellulases. In Haigler, C. H., and Weimer, P. J. (eds.), Biosynthesis and biodegradation of cellulose. pp.535-595. New York: Marcel Dekker Inc.

Reverbel-Leroy, C., Belaich, A., Bernadac, A., Gaudin, C., Belaich, J. P., and Tardif, C. 1996. Molecular study and overexpression of the *Clostridium cellulolyticum celF* cellulase gene in *Escherichia coli*. Microbiology. 142: 1013–1023.

Reverbel-Leroy, C., Pages, S., Bélaich, A., Bélaich, J-P., and Tardif, C. 1997. The processive endocellulase CelF, a major component of the *Clostridium cellulolyticum* cellulosome: purification and characterization of the recombinant form. J. Bacteriol. 179: 46-52

Robinson, I. M., Whipp, S. C., Bucklin, J. A., and Alison, M. J. 1984. Characterization of predominant bacteria from the colons of normal and dysenteric pig. Appl. Environ. Microbiol. 48: 964-969.

Roger, P., and Gottschalk, G. 1993. Biochemistry and Regulation of acid and Solvent Production in Clostridia. In Woods, D. R. (ed.), The Clostridia and Biotechnology. pp. 42. United States of America: Butterworth-Heinemann Press.

Russel, E. G. 1979. Type and distribution of anaerobic bacteria in the large intestine of pigs. Appl. Environ. Microbiol. 37: 187-193.

Sabathé F, and Soucaille P. Characterization of the CipA scaffolding protein and *in vivo* production of a minicellulosome in *Clostridium acetobutylicum*. 2003. J Bacteriol. 185: 1092-1096.

Sabathé, F. 2002. Le métabolisme des polyosides chez *Clostridium acetobutylicum*: Etude Fonctionnelle du Cellulosome. En vue d'obtenir le grade de Docteur de l'Université de Provence. Laboratoire de Biotechnologie-Bioprocédés, Département de Génie Biochimique et Alimentaire de l'INSA.

Sabathé, F., Bélaïch, A., and Soucaille, P. 2002. Characterization of the cellulolytic complex (cellulosome) of *Clostridium acetobutylicum*. FEMS. Microbiol. Lett. 217: 15-22.

Sabathé, F., and P. Soucaille. 2002. Characterization of the scaffolding protein CipA from the cellulosome of *Clostridium acetobutylicum*. J. Bacteriol. 185:1092-1096.

Sabathé, F., Lopez, F., Girbal, L., and Soucaille, P. 2000. Charaterization of the Cel9C cellulases from *Clostridium acetobutylicum*, and its interactions with the CipA scaffolding protein. FEMS. Microbiol. Lett. 217: 15-22.

Saha, B. C. 2003. Hemicellulose bioconversion. J. Ind. Microbiol. Biotechnol. 30: 279-291.

Sakon, J., Irwin, D., Wilson, D. B., and Karplus, P. A. 1997. Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca*. Nature Struct. Biol. 4: 810-818.

Salamitou, S., Raynaud, O., Lemaire, M., Coughlan, M., Béguin, P., and Aubert, J. P. 1994. Recognition specificity of the duplicated segments present in *Clostridium thermocellum* endoglucanase CelD and in the cellulosome-integrating protein CipA. J. Bacteriol. 176: 2822-2827.

Schülein, M. 2000. Protein engineering of cellulases. Biochim. Biophys. Acta. 1543: 239–252.

Schwarz, W. H., Zverlov, V. V., and Bahl, H. 2004. Extracellular glycosyl hydrolases from clostridia. Adv. Appl. Microbiol. 56: 215-261.

Schwarz, W. H. 2001. The cellulosome and cellulose degradation by anaerobic bacteria. Appl. Microbiol. Biotech. 56: 634-649.

Shima, S., Igararashi, Y., and Kodama, T. 1991. Nucleotide sequence analysis of the endoglucanase-encoding gene, *celCCD*, of *Clostridium cellulolyticum*. Gene, 104: 33–38.

Shimon, L. J. W., Bayer, E. A., Morag, E., Lamed, R., Yaron, S., Shoham, Y., and Frolow, F. 1997. A Cohesin Domain from *Clostridium thermocellum*: the crystal structure provides new insights into cellulosome assembly. Structure. 5: 381–390.

Shimon, L. J. W., Pagès, S., Bélaïch, A., Bélaïch, J. P., Bayer, E. A., Lamed, R., Shoham, Y., and Frolow, F. 2001. Structure of family IIIa scaffoldin CBD from the cellulosome of *Clostridium cellulolyticum* at 2.2 Å resolution. Acta. Cryst. D. 56: 1560-1568.

Shoham, Y., Lamed, R., Bayer, E. A. 1999. The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides. Trends. Microbiol. 7: 275–281.

Shoseyov, O., Warren, R. A. J. 1997. Cellulose binding domains-A novel fusion technology for efficient, low cost purification and immobilization of recombinant proteins. Innovation Newsletter of Novagen, Inc.

Somerville, C. 2004. Toward a system approach to understanding plant cell walls. Science. 306: 2206-2211.

South, C. R., Hogsett, D. A. L., and Lynd, L. R. 1995. Modeling simultaneous saccharification and fermentation of lignocellulose to ethanol in batch and continuous reactors. Enzyme Microb. Technol. 17: 797–803.

South, C. R., Hogsett, D. A. L., and Lynd, L. R. 1993. Continuous fermentation of cellulosic biomass to ethanol. Appl. Biochem. Biotechnol. 39/40: 587-600.

Spinelli, S., Fierobe, H. P., Belaich, A., Belaich, J. P., Henrissat, B. and Cambillau, C. 2000. Crystal structure of a cohesin module from *Clostridium cellulolyticum*: implications for dockerin recognition. J. Mol. Biol. 304:189–200.

Taiz, L., and Zeiger, E. (1991). Plant physiology. Benjamin Cummings Publishing, CA.

Tali, W. D., Morag, E., Rolider, A., Bayer, E. A., Lamed, R., and Shoham, Y. 2003. Regulation of the Cellulosomal *celS* (*cel48A*) Gene of *Clostridium thermocellum* Is Growth Rate Dependent. J. Bacteriol. 185: 3042–3048.

Tamaru, Y., Karita, S., Ibrahim, A., Chan, H., and Doi, R. H. 2000. A Large Gene Cluster for the *Clostridium cellulovorans* Cellulosome. J. Bacteriol. 182: 5906-5940.

Tamura, K., Dudley, J., Nei, M., Kumar, S. 2000. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0 [Computer Program]. Mol Biol Evol., No. 10.1093/molbev/msm092.

Tokatlidis, K., Salamitou, S., Béguin, P., Dhurjati, P., and Aubert, J. P. 1991. Interaction of the duplicated segment carried by *Clostridium thermocellum* cellulases with cellulosome components. FEBS. Lett. 291: 185–188.

Tokatlidis, K., Dhurjati, P., and Béguin, P. 1993. Properties conferred on *Clostridium thermocellum* endoglucanase CelC by grafting the duplicated segment of endoglucanase CelD. Protein Eng. 6: 947–952.

Tomme, P., Driver, D. P., Amandoron, E. A., Miller, R. C., Antony, Jr. R., Warren, J., and Kilburn, D. G. 1995. Comparison of a Fugal (Family I) and Bacterial (Family II) Cellulose-Binding Domain. J. Bacteriol. 177: 4356-4363.

Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y., and Steitz, T. A. 1996. Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose. EMBO. 15: 5739-5751.

U. S. Department of Energy Office of Science. 2007. Genomic: GTL Systems Biology for Energy and Environment [Online]. Available from: www.genomics. energy.gov [2007, November, 11].

Varel, V. H., Tanner, R. S., and Woese, C. R. 1995. *Clostridium herbivorans* sp. nov., a cellulolytic anaerobe from the pig intestine. Int. J. Systematic. Bacteriol. 45: 490–494.

Vasconcelos, I., Girbal, L., and Soucaille, P. 1994. Regulation of carbon and electron flow in *Clostridium acetobutylicum* grown in chemostat culture at neutral pH on mixtures of glucose and glycerol. J. Bacteriol. 176: 1443-1450.

Walseth, C. S. 1952. Occurrence of cellulase in enzyme preparations from microorganisms. TAPPI. 35: 228-233.

Wang, W. K., Kruus, K., and David, W. J. H. 1993. Cloning and DNA Sequence of the Gene Coding for *Clostridium thermocellum* Cellulase Ss (CelS), a Major Cellulosome Component. J. Bacteriol. 175: 1293-1302.

Weenk, G. Fitzmaurice, E., and Mossel, D. A. A. 1991. Selective enumeration of spores of *Clostridium* species in dried foods. J. Appl. Bact. 70: 135-143.

Wilson, D. B., and Irwin, D. C. 1999. Genetics and properties of cellulases. Adv. Biochem. Eng. Biotech. 65: 50-71.

Wood, T. M., and Bhat, K. M. 1988. Methods for measuring cellulase activities. Methods Enzymol. 160: 19-25.

Woodward, J., Hayes, M. K., and Lee, N. E. 1988. Hydrolysis of Cellulose by Saturating and Non-Saturating Concentrations of Cellulase: Implications for Synergism. Bio/Technology. 6: 301-304.

Yaguee, E., Beguin, P., and Aubert J-P. 1990. Nucleotide sequence and deletion analysis of the cellulase-encoding gene celH of *Clostridium thermocellum.* Gene. 89: 61-67.

Yaron, S., Morag, E., Bayer, E. A., Lamed, R., and Shoham, Y. 1995. Expression, purification and subunit-binding properties of cohesin 2 and 3 of the *Clostridium thermocellum* cellulosome. FEBS. Lett. 360: 121-124.

Zhang, M., Eddy, C., Deanda, K., Finkelstein, M., and Picataggio, S. 1995. Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis*. Science. 267: 240–243.

Zhang, Y-H.P., Lynd, L. R. 2004. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: Noncomplexed cellulase systems. Biotechnol. Bioeng. 88: 797–824.

Zverlov, V. V., and Schwarz, W. H. 2003. The *Clostridium thermocellum* Cellulosome-the Paradigm of a Multienzyme Complex [Online]. Available from http://www.wzw.tum.de/ mbiotec/ cellmo.htm[2007, August 10]

Zverlov, V. V., Velikodvorskaya, G. A., and Schwarz , W. H. 2003. Two new cellulosome components encoded downstream of *celI* in the genome of *Clostridium thermocellum*: the non-processive endoglucanase CelN and the possibly structural protein CseP. Microbiol. 149: 515-524.

Zverlov, V. V., Velikodvorskaya, G. A., 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.

Zverlov, V. V., Velikodvorskaya, G. A., Schwarz, W. H., Kellermann, J. and Staudenbauer, W. L. 1999. Duplicated *Clostridium thermocellum* cellobiohydrolase gene encoding cellulosomal subunits S3 and S5. Appl. Microbiol. Biotechnol. 51: 852-859.

Zverlov, V. V., Velikodvorskaya, G. V., Schwarz, W. H., Bronnenmeier, K., Kellermann, J., Staudenbauer, W. L. 1998. Multidomain Structure and Cellulosomal Localization of the *Clostridium thermocellum* Cellobiohydrolase CbhA. J. Bacteriol. 180: 3091-3099.

Zverlov, V. V., Fuchs, K. P., Schwarz , W. H., and Velikodvorskaya , G. A. 1994. Purification and cellulosomal localization of*Clostridium thermocellum* mixed linkage β-glucanase LicB (1,3–1,4-β-D-glucanase). Biotechnol. Lett. 16: 29-34.

APPENDICES

APPENDIX A:
Sequence alignment between pCR-Blunt II TOPO-AP4 MLSR-r and Bl4 plasmid sequence by using M13 primer

Sequence 1: lcl|1_seq_1 = Bl4 plasmid sequence by using M13 primer
Length = 909 (1 .. 909)

Sequence 2: lcl|2_seq_2 = pCR-Blunt II TOPO-AP4 MLSR-r
Length = 3248 (1 .. 3248)

Score = 1583 bits (823), Expect = 0.0
Identities = 837/837 (100%), Gaps = 0/837 (0%)
Strand=Plus/Minus

```
Query  73    GTTCTTAAGGATCCAAGTGATTACACAGTAACAAGTGATGGAATTACACTAAGCCAAAGC
132
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3248  GTTCTTAAGGATCCAAGTGATTACACAGTAACAAGTGATGGAATTACACTAAGCCAAAGC
3189

Query  133   TATCTTGCTACTCTAGCAGCAGGAACTTACACATATACAGTTGATTTTAGTGCAGGAAAT
192
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3188  TATCTTGCTACTCTAGCAGCAGGAACTTACACATATACAGTTGATTTTAGTGCAGGAAAT
3129

Query  193   GCAGGTACATTTACTGTTGTTGTTAAGGGAAAAGCAGTAGTAAATAAAACTACTTTAGCA
252
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3128  GCAGGTACATTTACTGTTGTTGTTAAGGGAAAAGCAGTAGTAAATAAAACTACTTTAGCA
3069

Query  253   GTAGGAGCTGCATCAGGAAAAGCAGGAGATACTGTTAAGGTGCCTGTAACTATAAGTAAA
312
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3068  GTAGGAGCTGCATCAGGAAAAGCAGGAGATACTGTTAAGGTGCCTGTAACTATAAGTAAA
3009

Query  313   GTAACAACACCAGTAGGTTTAATATGCATGGAAATAGATTATGATGCAAGTAAGTTTACT
372
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3008  GTAACAACACCAGTAGGTTTAATATGCATGGAAATAGATTATGATGCAAGTAAGTTTACT
2949

Query  373   GTTAAGGATGTACTTCCTAATACAGATCTTGTAAAAGATACTGATAACTACAGCTTTATT
432
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2948  GTTAAGGATGTACTTCCTAATACAGATCTTGTAAAAGATACTGATAACTACAGCTTTATT
2889

Query  433   GCTAATACGACATCAGCAGGAAAAATCAGTATTACATTTACAGATCCAACACTTGAGAAA
492
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2888  GCTAATACGACATCAGCAGGAAAAATCAGTATTACATTTACAGATCCAACACTTGAGAAA
2829

Query  493   TTCCCAATAAGTGCAGATGGAGTTATAGCAAATATAGATTTTGTTGTAAATTCAGGTGCA
552
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2828  TTCCCAATAAGTGCAGATGGAGTTATAGCAAATATAGATTTTGTTGTAAATTCAGGTGCA
2769
```

```
Query  553   GCAACTGGTGATAGCGATTTAACAGTAAATTCATCAGGTTTCATTGTTGCAGATGAAAGT
612

Sbjct  2768  GCAACTGGTGATAGCGATTTAACAGTAAATTCATCAGGTTTCATTGTTGCAGATGAAAGT
2709

Query  613   GATACAGATATAGATCATGTATCAACAAATGGAAAAATAACTGTTGAATAATCAATGACA
672

Sbjct  2708  GATACAGATATAGATCATGTATCAACAAATGGAAAAATAACTGTTGAATAATCAATGACA
2649

Query  673   TAATACTGCCGCCATTATAGGATAGCGGCAGTATACTATAAAATTTTAATTAATTATTTT
732

Sbjct  2648  TAATACTGCCGCCATTATAGGATAGCGGCAGTATACTATAAAATTTTAATTAATTATTTT
2589

Query  733   TAAAGGAGATAGAAAAATATGTTAAAGATAAGTAAGAATTTTnnnnnnnTAATGGCTGTA
792

Sbjct  2588  TAAAGGAGATAGAAAAATATGTTAAAGATAAGTAAGAATTTTAAAAAAATAATGGCTGTA
2529

Query  793   GCTCTTACATCTACAGTTATATTTCCCCCTTTTTAGGCCTGGGATGTAACGCACTGAGAA
852

Sbjct  2528  GCTCTTACATCTACAGTTATATTTCCCCCTTTTTAGGCCTGGGATGTAACGCACTGAGAA
2469

Query  853   GCCCTTAGAGCCTCTCAAAGCAATTTTGAGTGACACAGGAACACTTAACGGCTGACA  909

Sbjct  2468  GCCCTTAGAGCCTCTCAAAGCAATTTTGAGTGACACAGGAACACTTAACGGCTGACA  2412
```

APPENDIX B: Media preparation

Luria-Bertani medium (LB)

Tryptone	10 g/l
Yeast extract	5 g/l
NaCl	10 g/l
Agar (for agar plate)	12 g/l

Autoclave

Clostridial Basal medium (CGM)

Yeast extract	5 g/l
K_2HPO_4	0.75 g/l
KH_2PO_4	0.75 g/l
$MgSO_4.7H_2O$	0.4 g/l
NaCl	1 g/l
Asparagine	2 g/l
$(NH_4)_2SO_4$	2 g/l
$MnSO_4.H_2O$	10 mg/l
FeSO4.7H2O	10 mg/l

Adjust pH 6.6 and supplemented by 0.4 mg/l resazurin

Remove dissolved O_2 in medium by boiling and N_2 flow through

Autoclave

Add Add 1ml 30X cystein and 1ml of 125 g/l glucose per each bottle of media (total volume/each bottle 30 ml)

Media Synthetic (MS)

Glucose	60 g/l
KH_2PO_4	0.5 g/l
$FeSO_4$	10 g/l
K_2HPO_4	0.25 g/l
$MgSO_4.7H_2O$	0.2 g/l
Acetic acid (conc.)	2.2 ml/l

Adjust pH 6.0 by NH_3

Supplemented by 8 mg/l Para-aminobenzoic acid and 80 µg/l Biotin

Add 0.4 mg/l rezasurin

Remove dissolved O_2 in medium by boiling and N_2 flow through

Autoclave

Add 1ml of 30X cystein freshly before use media

2YTG

Bactotryptone	16 g/l
Yeast extract	10 g/l
NaCl	4 g/l
Glucose	2.5 g/l

Adjust pH 5.2 by HCl (conc.)

Supplemented by 0.4 mg/l rezasurin

Remove dissolved O_2 in medium by boiling and N_2 flow through

Autoclave

Add 1ml of 30X cystein freshly before use media

Reinforced Clostridial Media (RCM)

Meat extract	10.0 g/l
Peptone	5.0 g/l
Yeast extract	3.0 g/l
D (+) glucose	5.0 g/l
Starch	1.0 g/l
Sodium chloride	5.0 g/l
Sodium acetate	3.0 g/l
L-cystenium chloride	0.5 g/l
Agar	0.5 g/l

Adjust pH 6.8 (±0.2) by NH_3

Autoclave

Differential Clostridia Media (DCM)

Peptone from casein	5.0 g/l

Peptone from meat 5.0 g/l (or Universal peptone)

Meat extract 8.0 g/l

Yeast extract 1.0 g/l

Starch 1.0 g/l

D (+) glucose 1.0 g/l

L-cystenium chloride 0.5 g/l

Resazurin 0.002 g/l

Adjust pH 7.6 (±0.2) by NH3

Agar 20 g

Prepare separated solution:

Ferric (III) ammonium citrate 2.0 g/10 ml

Sodium sulfite x $7H_2O$ 1.5 g/ 15 ml

Before plating: supplemeted by (per 1000 ml media)

5.0 ml Ferric (III) ammonium citrate

7.5 ml Sodium sulfite x $7H_2O$

APPENDIX C: Solution preparation

PBS buffer pH 7.3

Potassium phosphate 50 mM

NaCl 150 mM

Adjust pH 7.3 with 1 N NaOH

PBS Complete

PBS buffer pH 7.3 50 ml

Tween 20 500 µl

Add distilled water until final volume reach to 500 ml

Electroporation buffer

Saccharose 270 mM 10 ml + 200 mM Phosphate buffer pH 7.4 150 µl

CIRRICULUM VITAE

Miss Chompunuch Virunanon was born on the September 12th, 1980 at Bangkok, Thailand. In 2002, she graduated in the Bachelor of Science, Major of Genetics from faculty of Sciences, Chulalongkorn University, Thailand (2nd class honor). She applied and started for Ph. D. in Biological Science Program, Faculty of Sciences, Chulalongkorn University in 2002.

Publication

Virunanon, C., Chantaroopamai, S., Dendoungbaripant, J., and Chulalaksananukul, W. 2007. Solventogenic-cellulolytic Clostridia from 4-step-screening process in Agricultural Waste and Cow Intestinal Tract. Anaerobe. (In press).

Scholarships and award

1988-2001	Scholarship for accelerating production of Graduate in Sciences and Mathematics from Chulalongkorn University
2002	Scholarship from Graduate School, Chulalongkorn University (72 PANSA)
2002-2007	University Development Scholarship (UDC), Chulalongkorn University
2002-2007	Royal Golden Jubilee Ph.D. Project, Thailand Research Fund
2005	French embassy's support Second prize in oral communication entitled "*In silico* and *In vitro* Cel48A engineer in *Clostridium acetobutylicum* ATCC824" in 10th Biological Science Graduate Congress, National University of Singapore
2006	The Commission of Higher education support for research in France