BIOCHEMICAL CHARACTERIZATION OF A PECTATE LYASE FROM CALDICELLULOSIRUPTOR BESCII

by

PUJA CHANDRAYAN

(Under the Direction of Michael W.W Adams)

ABSTRACT

Caldicellulosiruptor bescii is a gram-positive thermophilic bacterium with the ability to grow on untreated plant biomass. Upon growth on switchgrass at 78°C, a set of genes encoding 94 proteins was up-regulated transcriptionally by more than 4-fold, in comparison to growth on glucose. Enzymes with activity towards pectin were among those that were most highly up-regulated. Of a total of four predicted pectinolytic enzymes, Cbes_1854, a multidomain protein of polysaccharide lyase family 3, was chosen for overexpression and purification in *E. coli*. The full-length protein (X-PL3) and just the catalytic domain (PL3) were purified with yields of approximately 10 mg/liter *E. coli* culture. Both purified recombinant proteins display alkaline pH optima and have a half-life of approximately one hour at 90°C. Under *in vitro* assay conditions, domain X does not play a significant role in catalysis. The catalytic domain, PL3, also shows a synergistic effect on sugar release from untreated switchgrass biomass during *in vitro* assays with a cellulase (Cel A) from *C. bescii*. In a collaborative study, the crystal structure of the PL3 domain was obtained. This provided insight into the catalytic mechanism of the enzyme.

INDEX WORDS: *Caldicellulosiruptor bescii*, Cbes_1854; Pectinolytic enzymes, Polysaccharide lyase family 3; Catalytic domain (PL3), Full-length protein (X-PL3); *E. coli*, Cellulase (Cel A)

BIOCHEMICAL CHARACTERIZATION OF A PECTATE LYASE FROM CALDICELLULOSIRUPTOR BESCII

by

PUJA CHANDRAYAN

BS, CCS University, India, 2007

MS, Brakatullah University, India, 2008

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment

of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2013

© 2013

PUJA CHANDRAYAN

All Rights Reserved

BIOCHEMICAL CHARACTERIZATION OF A PECTATE LYASE FROM CALDICELLULOSIRUPTOR BESCII

by

PUJA CHANDRAYAN

Major Professor:

Michael W.W Adams

Committee:

Robert J Maier William N Lanzilotta

Electronic Version Approved:

Maureen Grasso Dean of the Graduate School The University of Georgia December 2013

DEDICATION

This work dedicated to my parents:

Shri Laliteshwar Kumar Singh and Veena Kumari

For their blessings

ACKNOWLEDGEMENTS

I express my sincere gratitude to Dr. Adams for giving me an opportunity to work under his guidance. I would like to thank Dr. Lanzilotta and Dr. Maier for their advice on my research project. I am grateful to Dr. Irina Kataeva, who helped immensely in design and completion of the proposed research work. I am also thankful to Dr. Mirko Basen for giving me useful advice for the execution of my experiments. I would also like to thank my BESC team members Amanda Rhaesa and Israel Scott for reading drafts of my thesis and for many useful comments.

I am also thankful to the lab manager, Farris Poole, for his continuous effort to make the work place more efficient and productive. I also offer my sincere thanks to all other Adams lab members who helped me whenever it was needed.

At last but not the least, I would like to thank my family members for their love, moral support, patience, encouragement and affection during the entire program. This thesis would not have been possible without their presence in my life.

TABLE OF CONTENTS

Page				
ACKNOWLEDGEMENTSV				
LIST OF TABLES				
LIST OF FIGURES IX				
CHAPTER				
1 INTRODUCTIONS AND LITERATURE REVIEW				
Recalcitrance of plant biomass and role of pectin1				
General features of the pectate lyases				
Caldicellulosiruptor bescii: the most thermophilic lignocellulolytic bacterium5				
Plant biomass and pectin deconstructing enzymes in C. bescii				
Goals of study14				
2 CLONING, EXPRESSION AND PURIFICATION OF A PECTATE LYASE IN				
E. COLI16				
Introduction16				
Materials and Methods17				
Results and Discussion21				
3 BIOCHEMICAL CHARACTERIZATION OF TWO DOMAINS OF CBES_1854,				
A FAMILY 3 PECTATE LYASE				
Introduction				
Materials and Methods28				

	Results and Discussion	32
4	SUMMARY	41
REFE	RENCES	44

LIST OF TABLES

Page

Table 1.1: Up-regulated genes of <i>C. bescii</i> after growth on switchgrass	12
Table 2.1: List of primers used in this study	20
Table 3.1: Comparison of different family 3 pectate lyases	29
Table 3.2: Predicted biophysical properties of the catalytic domain (PL3) and of the	holoenzyme
form (X-PL3) of Cbes_1854 from C. bescii	31
Table 3.3: Substrates used for pectate lyase assay	33
Table 3.4: Activity of X-PL3 with different substrates	

LIST OF FIGURES

Page

Figure 1.1: Growth temperature and morphology of <i>C. bescii</i>
Figure 1.2: Growth of <i>C. bescii</i> on untreated switchgrass
Figure 1.3: Growth of <i>C. bescii</i> on high concentrations of crystalline cellulose and switchgrass 10
Figure 1.4: Gene and domain organization of four predicted polysaccharide lyases of <i>C. bescii</i> .13
Figure 2.1: Plasmid map for overexpression constructs
Figure 2.2: The protein (upper) and DNA sequence (lower) of Cbes_1854 (X-PL3)23
Figure 2.3: SDS PAGE gel of expressed proteins
Figure 2.4: Diagram of the three-dimensional structure of the <i>C. bescii</i> PL3 domain25
Figure 3.1: Schematic diagram of the typical β -elimination cleavage of polygalacturonic acid34
Figure 3.2: Temperature, pH and calcium (Ca ⁺⁺) dependence of pectate lyase activity of PL3 and
X-PL3
Figure 3.3: Thermostability of PL3 and X-PL3
Figure 3.4: Diagram of the catalytic domain PL3 with the reaction product trigalacturonic acid 40

CHAPTER 1

INTRODUCTIONS AND LITERATURE REVIEW

Recalcitrance of plant biomass and role of pectin

The fossil fuel crisis has accelerated the search for renewable sources of energy [1-3]. Amongst the various available alternatives to fossil fuels, ethanol from cellulose has become a viable option [4-7]. As a feedstock, cellulose is the most abundant organic biopolymer, but exists in a complex form of lignocellulosic biomass [8,9]. Of the total lignocellulosic biomass, 70% is represented by plant cell walls but due to cell wall recalcitrance only 2% can be currently used for the production of biofuels [9,10]. It is a major challenge to develop a cost-effective bioconversion process [3,11,12]. The reported poor yield of this process is largely due to the natural resistance of the complex plant cell wall to deconstruction [8,13,14]. Cell wall resistance to enzymatic degradation is primarily due to the crystalline nature of cellulose, which is a major component of the plant cell wall, and its intimate association with cell wall polymers such as hemicellulose, pectin and lignin [9,15]. Ultrastructure studies show that the cell wall has three layers: (1) the middle lamella, which is outermost layer and has a significant amount of pectic polysaccharide and proteins; (2) the primary cell wall, which is the next layer after the middle lamella and largely consists of cellulose microfibrils cross-linked by hemicellulose and (3) the innermost layer of the plant cell wall termed the secondary cell wall, and this is mainly composed of cellulose microfibrils and lignin [9,15,16]. In general, there are primarily two types of cell walls depending on the type of crosslinking. Type-I plant cell walls are abundant in dicot plants and non-commelinoid monocot and contains xyloglucans (XyGs) as cross-linking polysaccharides [9]. Type-II plant cell walls are prevalent in commenlinoid grasses such as switchgrass and contain glucuronoarabinoxylans (GAXs) for crosslinking of the cell wall. Switchgrass is a widespread grass in eastern United States and has been proposed as an ideal plant for lignocellulose conversion [9,17,18]. Besides differences in hemicellulose content, cell walls also differ in the amounts of pectic polysaccharide and soluble protein that they contain [19].

Among the different polysaccharides within plant cell walls, pectin represents a very small amount, typically less than 5% of the total polysaccharide [19]. Nevertheless, pectin has an essential role in maintaining the structural integrity, mechanical strength and shape of the cell wall. Pectic polysaccharides together with a variety of proteins are present in the middle lamella and primary cell wall. There are three major types of pectic polysaccharide that have been found in plant cell walls and these are: (1) homogalacturonan, (HG) (2) rhamnogalacturonan I (RG-I) and (3) rhamnogalacturonan II (RG-II). HG and RG-I have a polygalacturonic acid backbone joined together by α -1-4 linkage while RG-II has a disaccharide backbone of rhamnose and polygalacturonic acid. Although RG-I and HG have a similar backbone, RG-I is heavily modified at the O-2 and O-3 positions by multiple hexose and pentose sugars while HG are mostly esterified by methyl group or by a simple xylose sugar [19].

By proportion in the primary cell wall, HG constitutes 65% of the total pectic polysaccharide, 20-25% is made up of RG-I and rest consists of RG-II and XGA. It is still not very clear how pectin and other polysaccharides are organized into the primary cell wall. To understand the organization in the primary cell wall, an *in vitro* attempt to extract pectin was carried out by mild salt and base that resulted into an insoluble pectin fraction with cellulose and hemicellulose. The insoluble pectin fraction with hemicellulose constitutes almost 10-50% of the

total pectin [20]. This observation demonstrates that the pectin lies in the insoluble network of polysaccharides within the plant cell wall rather than soluble phase. Recently, it has been shown that mutation of the homogalacturonan synthesis gene, GAUT1, of the model plant *Arabidiopsis thaliana*, gives rise to a phenotype in which the cell wall contains less pectin [21]. *Caldicellulosiruptor saccharolyticus*, which lacks pectin-degrading capability, has been shown to grow better on transgenic plants with less pectin [21]. In a separate study, reduction of demethyl-esterified homogalacturonan (HG) in *A. thaliana* increases the efficiency of enzymatic sugar release in bioconversion processes [21,22]. All of these observations suggest an important role for pectin in cell wall recalcitrance.

General features of the pectate lyases

According to enzyme nomenclature, pectate lyase fits into enzyme class EC 4.2.2.2 and is a carbon-oxygen lyase. It uses uronic acid containing polysaccharide as a substrate. Uronic acid is a sugar acid with a carbonyl or hydroxyl group at position C1 or C6 of a monosaccharide that has undergone oxidation to form a carboxylic acid. In D- polygalacturonic acid the hydroxyl group at C6 position of D-galactose has been oxidized into a carboxylic acid [19,23,24]. Pectin is one of the complex polysaccharides made up from uronic acid carrying polysaccharides joined together by α 1-4 glycosidic linkages. The substrate utilized by pectate lyase is a type of pectin that contains homogalacturonans and forms an integral part of plant cell wall. On the basis of the mechanism of catalysis and nature of substrate, the class of enzyme is categorized into the polysaccharide lyase family [24].

Pectate lyases from fungi and bacteria have been extensively characterized and have provided valuable insight regarding the general properties of this type of enzyme. It cleaves the carbon–oxygen bond of α 1-4 glycosidic linkage between uronic acids, leaving an unsaturated

C4-C5 bond at the newly formed non-reducing end. Absorption at 235nm of the unsaturated product has been utilized for *in vitro* assays of pectate lyase. Pectate lyase activities are typically maximal in the alkaline pH range (8.5-10.5) and a metal ion is essential for catalysis. Ca^{2+} is used as a metal cofactor in most enzymes, except for the few that use Co^{2+} , Mn^{2+} , and Ni^{2+} . The divalent cations participate like a cofactor in enzyme-substrate interaction by binding both the enzyme and the charged sugar substrate [23,24].

A systematic study of distribution of polysaccharide lyases among different domains of life has been made based on their protein sequence similarity and known catalytic features. On this basis, they have been organized into 22 families of polysaccharide lyase (PL) [25]. Among these only five families, PL1, PL2, PL3, PL9 and PL10, have been shown experimentally to be pectate lyases. Several structures have been determined for representatives of all five polysaccharide lyase families. By analyzing all available three-dimensional structures, it has been observed that there are three different kinds of topologies for the catalytic module: (1) a right-handed β -helix fold which is common with the families PL-1, PL-3, and PL-9, (2) an (α/α) 7 toroid in family PL2 and (3) an (α/α) 3 toroid in family PL-10. β -helix fold is more abundant among different pectate lyase and is structurally very unique [24].

Despite differences in substrate specificities and variation in structural topology all 22 families still have a common conserved mode of catalysis [26]. Catalysis involves a β -elimination reaction for the cleavage of the glycosidic bond. In contrast to glycosyl hydrolyase, polysaccharide lyase breaks the bond without the involvement of a water molecule and results in the formation of an unsaturated bond at the non-reducing end. The catalytic mechanism can be summarized into three steps: (1) abstraction of the C-5 proton on the sugar ring of a uronic acid or ester by a basic catalytic amino acid, (2) stabilization of the resulting anion by charge

delocalization into the C-6 carbonyl group and (3) lytic cleavage of the O-4: C-4 bond, facilitated by proton donation from a catalytic acid to yield a hexenuronic acid and hexuronic acid. The unsaturated product is one of the crucial pieces of evidence for the β -elimination [27].

Caldicellulosiruptor bescii: the most thermophilic lignocellulolytic bacterium

The crystalline nature of cellulose and presence of other heterogeneous polysaccharides has been reported as a major obstacle in development of a cost-effective process for making biofuels [12]. A strategy to develop an optimized and cost-effective pathway for production of biofuel or chemicals from insoluble complex cellulose is the subject of extensive research efforts and poses a major scientific and technical challenge. Current processes involve two major strategies. The first strategy is physio-chemical pretreatment of lignocellulosic biomass followed by enzymatic conversion into fermentable sugars, which are subsequently fermented to ethanol. The second strategy has been proposed on the basis of pioneering work done on *Clostridium* thermocellum. This process is known as consolidated bioprocessing (CBP), in which C. thermocellum carries out both saccharification and fermentation [4,10,28-30]. C. thermocellum has optimal growth temperature of 60°C and it is thought that high temperature increases the efficiency of plant biomass deconstruction [31]. High temperature also reduces the chances of microbial contamination, enhances the rate of catalysis, facilitates the simple purification of liquid biofuels from the growth medium and decreases the viscosity. Reduced viscosity helps in the fermentation of high loads of insoluble biomass [32]. Therefore, thermophilic cellulolytic microbes have been sought for efficient production of biofuels [31,33,34].

Among known thermophilic bacteria, species of the *Caldicellulosiruptor* genus have potential to be used in CBP due to their cellulolytic capability in the temperature range of 75-80°C [35]. All members of this genus are gram-positive, anaerobe, thermophilic and non-spore

forming bacteria. At present nine different species of Caldicellulosiruptor have been isolated and growth studies have been carried out [35]. Among them, C. bescii has the highest growth temperature and can grow up to 90°C with an optimum growth temperature of 78°C [36,37]. This strain was originally deposited in 1990 in the German Collection of Microorganisms and Cell Cultures (DSMZ) as DSM 6725 [38]. It was classified as a member of a new genus Anaerocellum and named as Anaerocellum thermophilum strain Z-1320. Subsequent growth studies on strain DSM 6725 with cellulose, hemicellulose and pectin showed marked differences compared to what was originally reported with strain Z-1320. For example, DSM 6725 was capable of growing on xylose and pectin while earlier it was shown that the strain Z-1320 could not utilize xylose and pectin [38]. The other notable difference was in the growth temperature of DSM 6725 (Figure 1.1 A) [38]. DSM 6725 strain could grow up to 90°C while the Z-1320 strain can grow up to 83°C only (Figure 1.1 A). These differences, together with 16S rRNA sequence analyses, showed that DSM 6725 was highly similar to Caldicellulosiruptor saccharolyticus suggesting that it is a new *Caldicellulosiruptor* genus. Subsequently, DSM 6725 was proposed as type strain of a new species named as *Caldicellulosiruptor bescii*. The rod-shaped morphology of C. bescii is shown in Figure 1.1 B [38].

C. bescii grows well on cellulose, xylan and pectin and, remarkably, was shown to grow well on unpretreated plant biomass, including poplar and switchgrass [37]. Its capability of growing on untreated switchgrass considerably enhances the utility of *C. bescii* as a potential CBP microbe. A growth curve of *C. bescii* on wild type switchgrass is shown in Figure 1.2. A cell density of 1×10^8 cells/ml was reached on all complex substrates, which is similar to the cell density reported after growth on the simple sugar cellobiose (Figure 1.1) [37]. Quantitatively, after three successive growths on spent switchgrass, *C. bescii* can degrade up to 85% (w/w) of



В

A



Figure 1.1. Growth temperature and morphology of *C. bescii.* (A) Comparison of growth temperature of *C. bescii* with *A. thermophilum* Z-1320 [38]. (B) Scanning electron micrograph of *C. bescii* grown on 0.5% cellobiose. Taken from reference [38].

the intial switchgrass biomass (Figure 1.2) [39].Considering its potential future application on an industrial scale, *C. bescii* exceeds all other cellulolytic microbes by showing uninhibited growth using up to 200 g/liter switchgrass as well as on similar concentration of crystalline cellulose (Figure 1.3) [40]. Interestingly, acid pretreated switchgrass (50 g/liter) inhibits growth of the *C. bescii*, showing the problems associated with the harsh pretreatment conditions [40].

Another crucial property of *C. bescii* is its ability to solubilize lignin, which could be very advantageous in enhancing the efficacy of CBP [39]. *C. bescii* can solubilize the cellulose microfibrils as well as the lignin components of cell walls at high temperature, thus making it a very promising candidate for CBP [36,37]. Currently, the bottleneck in using *C. bescii* on an industrial scale is its inability to produce ethanol, which is in contrast to members of the genus *Clostridium*, which generate ethanol when grown on cellulose [31]. In pursuit of a better technology and considering the recent development of genetic tools for *C. bescii*, it is possible to envision *C. bescii* as an efficient CBP-microbe in the not too distant future [41,42].

Plant biomass and pectin deconstructing enzymes in C. bescii

The *C. bescii* genome encodes 2666 proteins among which 259 are categorized as carbohydrate active enzymes (CAZy) [25,43]. These CAZy genes are predicted to make 180 transcriptional units, 111 of which are multigene operons while 69 genes exist as single genes. CAZy enzymes often contain at least one cellulose-binding module (CBM), which factor greatly in their capacity to degrade cellulose and plant biomass [43,44]. For example, CBM3-containing enzymes attach to crystalline cellulose and CBM22-containing enzymes bind to xylan. *C. bescii* has enzymes carrying both of these CBM modules and moreover they also have other classes of CBM-containing enzymes. Importantly, some of the CBMs are specific to *C. bescii*. For example, CBM4_9 is present in all three extracellular pectinolytic enzymes [43,44].



В

Figure 1.2. Growth of *C. bescii* on untreated switchgrass. (A) Growth curve with hot water washed biomass (wSG) and successive growth curves on spent switchgrass SG1 and SG2, which is the insoluble biomass remaining at the end of growth that is used for next culture (taken from reference [37]). (B) Switchgrass degradation in successive growth experiments is plotted by comparing the residual switchgrass (%, w/w) in each step. The control involves incubating switchgrass without *C. bescii*. Taken from reference [39].



Figure 1.3. Growth of *C. bescii* on high concentrations of crystalline cellulose and switchgrass. (A) Green triangles, 5 g/liter; orange circles, 200 g/liter; and blue diamonds, 200 g/liter of crystalline cellulose or (B) unpretreated switchgrass. Taken from reference [40].

The potential role of CAZy enzymes in plant biomass degradation has been studied by transcriptional analyses and glycan immunoanalyses. Transcriptional analyses of the growth of *C. bescii* on unpretreated switchgrass versus growth on glucose revealed that there are 94 enzymes whose genes are up-regulated more than 4-fold when grown on untreated switchgrass and among these 18 have a functional role in carbohydrate metabolism. Of these 18 enzymes, 12 have the potential to bind and hydrolyze the cellulose, xylan, xyloglucans, pectin and mannan [39]. Up-regulated genes are shown in Table 1.1 [39].

Enzymes that hydrolyze cellulose, hemicellulose, xylan and xyloglucan are all involved in degradation of plant biomass [35,45]. However with the *C. bescii*, the presence of pectindegrading enzymes among those that are up-regulated upon growth on unpretreated switchgrass is a novel and very interesting finding. Remarkably, two of the predicted pectin-degrading enzymes are up-regulated by more than 20-fold [39]. The importance of these pectin-degrading enzymes is also confirmed by immunoanalyases where pectin-related epitopes showed increased extractability during switchgrass degradation [39]. This suggests that hydrolysis of specific pectin components facilitates biomass degradation by *C. bescii*.

Pectin-degrading enzymes fall into the category of polysaccharide lyase (PL) enzymes. The *C. bescii* genome contains 4 PLs. These are PL3 (Cbes_1854), PL9 (Cbes_1855), PL11 (Cbes_1853) and an unclassified PL (Cbes_2353), putatively active with polygalacturonate and other pectins, including rhamnogalacturonans. Gene clusters showing all the classified polysaccharide lyases has been shown in Figure 1.4. The three PL genes are located in a large CAZy cluster encoding several multidomain, multi-functional enzymes active against all major components of plant biomass. PL3 and PL9 have an N-terminal X domain of unknown function and and PL11 has a C-terminal CBM3 domain that is proposed to specifically bind with

Gene	SP	CAZy modules	Catalytic activity	Binding	Fold			
Pectin-degrading enzymes								
Cbes_1853	Y	PL11-CBM3	rhamnogalacturonan lyase	cellulose	9.6			
Cbes_1854	Y	CBM_4_9-PL3	pectate lyase	pectin	51.0			
Cbes_1855	Y	CBM_4_9-PL9	pectate lyase, exopolygalacturonate lyase	pectin	27.9			
Cbes_2353	N	PL (?)	unclassified pectate lyase	pectin	4.8			
Xylan- and cellulose degrading enzymes								
Cbes_0618	Y	CBM22-CBM22-GH10	endo-1,4-β-xylanase, endo-1,3-β-xylanase	xylan	4.5			
Cbes_0183	Y	CBM22-CBM22-GH10	endo-1,4-β-xylanase, endo-1,3-β-xylanase	xylan	13.7			
Cbes_0184	N	GH39	β-xylosidase	xylan	17.3			
Cbes_0187	N	GH39	β-xylosidase	xylan	5.8			
Cbes_1857	Y	GH10-CBM3-CBM3-GH48	 10: endo-1,4-β-xylanase, endo-1,3-β-xylanase 48: cellobiohydrolase, endo-processive cellulase 	cellulose	3.9			
Cbes_1865	Y	GH9-CBM3-CBM3-CBM3-GH5	 9: cellobiohydrolase, endoglucanase 5: β-mannosidase, cellulase, endo-β-1,4-xylanase 	cellulose	22.8			
Cbes_1866	Y	GH5-CBM3-CBM3-CBM3-GH5	5: β-mannosidase, cellulase, endo-β-1,4-xylanase	cellulose	6.6			
Cbes_1867	Y	GH9-CBM3-CBM3-CBM3-GH48	9: cellobiohydrolase, endoglucanase 48: cellobiohydrolase, endo-processive cellulase	cellulose	2.3			
Cbes_0234	N	GH5	5: β -mannosidase, cellulase, endo- β -1,4-xylanase	cellulose	17.3			

SP: signal peptide. Taken from reference [39].



Figure 1.4. Gene and domain organization of four predicted polysaccharide lyases of *C. bescii*.(A) Domain organization for each gene is shown. The abbreviations are: SS, signal sequence;PL, polysaccharide lyase; CBM, cellulose binding modules and X, domain of unknown function.(B) Gene organization of three extracellular pectate lyase genes as a cluster in the genome.

crystalline cellulose. The roles of these three PLs are to degrade untreated plant biomass as the three polysaccharide lyases are not highly up-regulated upon growth of *C. bescii* on cellulose or cellobiose. Two of them, PL3 and PL9, were part of the secretome at the beginning of growth on cellulose and cellobiose growth but they are downregulated as growth continues [46]. This result is not surprising because growth on cellulose does not release the same hydrolysis products as grow on switchgrass. Presumably, the initial hydrolysis of switchgrass biomass leads to exposure of pectin epitopes and release of soluble pectin-derived products that cause the up-regulation of genes encoding pectin-degrading enzymes.

Goals of study

Considering the importance of pectin-degrading enzymes in plant biomass degradation, the primary goal of this study was to carry out a biochemical study of three extracellular PL-type enzymes from *C. bescii* in order to provide insight their roles to plant biomass deconstruction. Ultimately, this information would determine which enzymes should be up-regulated to increase the efficiency of plant biomass degradation or be used to better design of enzyme mixtures for in *vitro* plant biomass degradation.

Earlier efforts to make multidomain cellulase A (Cbes_1867) from *C. bescii* in *E. coli* were unsuccessful due to proteolysis in linker region that is present between two domains of cellulase A [47]. In this study, pectin-degrading enzymes were chosen to demonstrate overexpression in *E. coli* due to its proposed role in untreated switchgrass degradation. Pectin-degrading enzymes also have similar linker regions like Cbes_1867. Therefore, overexpression of pectin-degrading enzyme in *E. coli* is a challenge due to potential proteolytic susceptibility in the domain linker region of the protein [48].

Here the goal was the heterologous expression of pectin-degrading enzymes and the determination of their biochemical properties, including pH and temperature optima and temperature tolerance.

CHAPTER 2

CLONING, EXPRESSION AND PURIFICATION OF A PECTATE LYASE IN E. COLI Introduction

The genes encoding four predicted pectin-degrading enzymes are up-regulated during growth of *C. bescii* on switchgrass in comparison with growth on glucose [39]. The corresponding genes are Cbes_1853, Cbes_1854, Cbes_1855 and Cbes_2353. Cbes_2353 is predicted to be an intracellular protein and belongs to an unclassified pectate lyase family [43]. The CAZY classification is unknown for this protein. The other three enzymes are predicted to be extracellular. All three extracellular enzymes share similar domain architecture with signal sequence at the N-terminus. Following the signal sequence, there are catalytic and substrate-binding domains. The domain arrangement is shown in Figure 1.4 (chapter 1). In all three proteins the catalytic and binding domains are separated by linkers, composed of repetitive amino acid sequences rich in proline, threonine and serine. Linkers promote proper interaction of the catalytic domains with insoluble substrate [49-52]. Presumably, proline residues provide conformational rigidity to binding and catalytic domains while serine and threonine provides sites for O-glycosylation that plays roles in enhancing the protein stability in the extracellular environment [53].

Cbes_1853 has a catalytic module of the PL11 family followed by a CBM3 domain at the C-terminus. The PL11 family has predicted rhammogalacturonan lyase activity and the CBM3 domain proposed to bind crystalline cellulose. In comparison to Cbes_1853, the arrangement of catalytic and substrate binding domains is reversed in Cbes_1854 and Cbes_1855. Both proteins

contain catalytic domains, PL3 and PL9, respectively. The PL3 and PL9 families have been characterized as pectate lyases. The substrate-binding domain of both proteins shares homology with CBM4 (subfamily 9). The family 4 CBM module is composed of approximately 150 residues and is mostly found in mostly in bacterial enzymes not in archaeal enzymes. The binding of this module to xylan, glucan and amorphous cellulose has been confirmed [54].

The goal here was to overexpress multidomain Cbes_1854 in *E. coli*. Besides the fulllength protein, its two domains, 'X' and PL3, were also cloned and expressed separately in *E. coli* for biochemical and structural studies.

Materials and Methods

Polymerase chain reaction (PCR): PCR amplification of X, X-PL3 and PL3 was carried out with 20 ng *C. bescii* genomic DNA as template, in a final reaction volume of 25 μ l, using AccuPrimeTM Pfx DNA Polymerase from Invitrogen (Carlsbad, CA, USA) according to the manufacturer's instructions. Primers are shown in Table 2.1.

Restriction endonuclease digestion of DNA: DNA samples were digested with restriction endonucleases, KpnI and XhoI, in their specific reaction buffers. Both restriction enzymes and the buffers were purchased from New England Biolabs (NEB, Ipswich, MA, USA). Digestions were carried out overnight (using 1U of enzyme/µg DNA) at 37°C in the presence of 0.1 mg/ml bovine serum albumin (BSA) as recommended by the supplier.

Agarose gel electrophoresis: DNA purification was performed using the StrataPrep gel extraction kit (Agilent Technologies, Santa Clara, CA, USA) and DNA was finally eluted in either elution buffer (10 mM Tris.Cl, pH 8.5) or in autoclaved distilled water.

Ligation and transformation: Digested and purified plasmids and inserts were set up for cohesive-end ligation using T4 DNA Ligase (NEB) in 1X buffer supplied. An insert: vector ratio

of 3:1 was used. Ligation reactions were incubated at 14-16°C for 18-20 h with 0.1 mg/ml BSA. For heat shock transformation, the ligation mix (10 μ l containing ~ 20 ng DNA) was mixed with *E coli* XL1 Blue-MRF' (Agilent Technologies, Santa Clara, CA) and after heat shock at 42°C for 1 min, 1 ml Luria Broth medium was added into the transformed cells. This was followed by incubation at 37°C for 1 hr. Aliquots of the 50-100 μ l cells was then plated on LB agar supplemented with an antibiotic ampicillin (100 μ g/ml).

Plasmid preparation for DNA sequencing: Plasmid DNA for sequencing was purified using the StrataPrep Plasmid Miniprep Kit (Agilent). Purified DNA was sent for sequencing to the Macrogen sequencing facility (Macrogen, MD, USA) using T7 promoter primers.

Design of the overexpression vector: Three overexpression vectors for expression of full-length Cbes_1854(X-PL3), the catalytic domain only (PL3) and domain X was designed under the control of the T7 promoter using Novagen expression vector pET-45b. A plasmid map of the two constructs those were successfully expressed in *E. coli* is shown in Figure 2.1. Expression of recombinant proteins in E. coli: For expression of the proteins from the cloned genes in pET-45b-PL3 and pET-45b-X-PL3, the constructs were transformed into the E. coli strain BL21-Codon plus(DE3)-RIL strain. Cells were inoculated to a final concentration of 1 % in a fresh LB medium containing ampicillin (100 ug/ml) and were grown overnight to an OD600 of 0.6-0.7 at 37°C with shaking at 220 rpm. Production of recombinant proteins was induced with 1 mM IPTG.

Heat treatment of cell free extract: Cells expressing PL3 and X-PL3 domains were lysed by treatment with 1 mg/ml lysozyme and heated at 80°C for 30 minutes. Extracts were clarified by centrifuging at 14000 rpm for 30 min.

Rapid screening of small expression cultures: Expression of proteins PL3 and X-L3 in BL21-



Figure 2.1. Plasmid map for overexpression constructs. (A) Overexpression vector for full-length protein (X-PL3) and (B) overexpression vector for only catalytic domain (PL3).

Primer name	Sequence (5' – 3')
X (Forward)	CGGGAATTCCGTAATGAAGTTGCAAAGGCAGCG
X (Reverse)	CCGCTCGAGTTATGCTGTACCGCTCACACTTGAATTTAC
PL3 (Forward)	CGGGAATTCAATACGGGTGGTGTTTTAGTTATTACAGAT
PL3 (Reverse)	CCGCTCGAGTTAGTATTGATGTATCTGTGATTGGGACGG

Table 2.1: List of primers used in this study

Codon plus(DE3)-RIL cells (10 ml culture) were confirmed by affinity chromatography of heattreated cell free extracts using the N-terminal 6x-His tag. Cells were induced with 1 mM IPTG for 16 h at 15°C for overnight. The cell pellet was collected and suspended in 20 mM sodium phosphate pH 7.0 containing 0.3 M NaCl (buffer A) and the cells were disrupted by the addition of 1 mg/ml lysozyme. Cell debris was removed by centrifugation at 10,000 rpm for 30 min. The Supernatant was heated at 80°C for 20 min, and, after removing denatured protein by centrifugation at 10,000 rpm for 30 min, the supernatant was applied onto a minispin HisTrap FF affinity column (GE Healthcare, USA) equilibrated with buffer A. After washing the column, the protein was eluted using imidazole (0.5 M) in buffer A.

Protein expression and purification: BL21-Codon plus(DE3)-RIL cells from a 1 liter culture were induced with 1 mM IPTG for 16 h at 15°C. The cell pellet was collected and suspended in buffer A and the cells were disrupted using a French press (800 psi). Cell debris was removed by centrifugation at 10000 rpm for 30 min. The supernatant was heated at 80°C for 20 min, and after removing denatured protein by centrifugation at 10 000g for 30 min, the supernatant was applied onto a His Trap FF affinity column (5 ml, GE Healthcare) equilibrated with buffer A. After washing the column, the PL3 module was eluted using an imidazole gradient (0–0.5 M) in buffer A. Based on SDS–PAGE analysis, fractions containing PL3 were combined, concentrated using ultrafiltration (Amicon, 10 kDa cutoff) and dialyzed against 20 mM phosphate of pH 7.0.

Results and Discussion

Characterization of PL3 and X-PL3: The full-length protein contains a PL3 family catalytic domain, which has pectate lyase activity, while the X domain shows some similarity with carbohydrate-binding domains present in several CAZy proteins [25]. Sequence alignment analysis of the X domain (39-222) predicts it to be a member of concanavalin A-like lectin superfamily, which binds with sugars, glycoproteins and glycolipids. Association of this X

domain with pectate lyase suggests that this is a putative pectin-binding domain. The X domain has also been proposed to be a member of cellulose binding domain family 4 CBM4_9 [39].

Purification of PL3 and X-PL3: X-PL3 was cloned from a thermophilic bacterium, therefore, it was expected that the X and PL3 domains, as well as full-size X-PL3, would tolerate high temperature. We found that the X domain was insoluble upon induction in *E. coli* at either 37°C or 18°C. We were unable to recover the X domain from the insoluble fraction. In contrast, full-length X-PL3 was soluble. This indicates that the presence of the PL3 domain, which is separated from the X domain by a long linker, mediates proper folding of the X domain occurs due to specific domain interactions with the PL3 domain as shown for other CAZy enzymes [55]. The presence of a 44 amino acid long flexible linker (Figure 2.2) might play an important role in mediating domain interactions. In many cases, the presence of a linker sequence separating domains in a recombinant protein expressed in *E. coli* results in proteolysis of the linker region [47,56]. This can arise because of lack of glycosylation or phosphorylation of the linkers region by host *E. coli*. Consequently, in recombinant proteins linkers are very susceptible to intracellular proteolysis. It is important to note that we did not observe proteolysis of X-PL3.

The PL3 and X-PL3 proteins were not degraded after incubation at 70°C for 30 min as the intensities of their bands on the SDS gel were similar to those seen in unheated cell extracts (Figure 2.3 A). The proteins also retained 100% of their activity after heating of the cell extract, again confirmed that they were unaffected by heat treatment. These conditions were therefore used as a first heat-treatment step during the large-scale purification of each enzyme. Both enzymes contain an N-terminal 6xHis tag, which was used for one-step purification by affinity chromatography using a 5ml GE-HisFF column. Each protein bound to the column and eluted MIKSKNKKEEVWVMSNRKILAIVVSLIMVVSLFTGIGLRNEVAKAATLLTDDFEDGNRDG WSTSNGSWSVVVDGSKVLKQASTGSEARAYTGSSDWSDYTVEAKVKVLNVKDSSS GAGVIVRYKNSGNFYALVLRGSKIEIGKKLNSNWSTLAFKSFTLDQDTWYNVKLEV NGSKLVGYVNGSQVLSASDLSITTGKAGLIADRCVAEFDDVVVNSSVSGTAPTPTPT TSSVTPTPTSTPTPTKTPTPTSTPVPTQTPAVTPTPTNTGGVLVITDTIIVKSGQTYDGKGI KIIAQGMGDGSQSENQKPIFKLEKGANLKNVIIGAPGCDGIHCYGDNVVENVVWED VGEDALTVKSEGVVEVIGGSAKEAADKVFQLNAPCTFKVKNFTATNIGKLVRQNG NTTFKVVIYLEDVTLNNVKSCVAKSDSPVSELWYHNLNVNNCKTLFEFPSQSQIHQ Y

Figure 2.2. The protein (upper) and DNA sequence (lower) of Cbes_1854 (X-PL3). The indicated residues are: residues 1-38, signal sequence (italics); Residues 39-222, X domain (bold); Residues 223-267, linker sequence (italics) and Residues 268-460, PL3 domain (bold). The primer sequences are underlined in the DNA sequence.

23



Figure 2.3. SDS PAGE gel of expressed proteins. (A) 1, protein marker; 2, heated cell-free extract of PL3; 3, unheated cell-free extract of PL3; Lane 4, unheated cell free extract of X-PL3 and Lane 5, heated cell-free extract of X-PL3. Lanes 1-4 depict the purification of PL3 domain and lanes 5-8 depict the purification of X-PL3. The lanes are (B) 1, flowthrough; 2, wash; 3-4, elution fractions; 5, flowthrough; 6, wash; 7-8, elution fractions and 9, protein marker.



Figure 2.4. Diagram of the three-dimensional structure of the *C. bescii* PL3 domain. Diagram was made using PyMol software using structure of PL3 domain (PDB: 3T9G) [59]. The calcium ion is in magenta.

with increasing concentration of imidazole. This indicates that the tags were readily accessible. After two steps of purification, X-PL3 and PL3 were homogeneous, as indicated by SDS gel electrophoresis (Figure 2.3). The molecular masses estimated from the SDS gel were similar to those calculated based on the deduced amino acid sequences (46 and 21 kDa for X-PL3 and PL3, respectively) suggesting that no proteolysis occurred. After purification from two liters of E. coli cultures, approximately 20 mg of each protein (PL3 and X-PL3) was obtained. The purified proteins were used for crystallization.

Crystallization of PL3 domain: The purified proteins were supplied to Dr. M. Himmel of National Renewable Energy Laboratory, CO, for structure studies. A 1.5-Å resolution X-ray structure of the catalytic module of C. bescii family 3 pectate lyase was obtained (PDB: 3T9G) [57]. The structure is similar to the previously solved structure of a family 3 pectate lyase from Bacillus sp. strain KSM-P15 (Pel15; PDB: 1EE6) [58]. The overall structure has β -helix fold, which is common for pectate lyases from families 1 and 3. The structural model for the *C. bescii* PL3 domain is shown in Figure 2.4. One calcium ion is bound with each chain in the crystal. Structural comparison with Pel15 suggests a strong structural similarity with a root-mean-square deviation of 0.93 Å. This structural similarity is significant considering that *C. bescii* is a thermophile and *Bacillus* sp. is a mesophile and that the sequence identity between PL3 and Pel 15 is only 53%. Notably, Pel15 from *Bacillus sp.* has an optimum temperature of 30°C while PL3 of *C. bescii* is expected to be optimally active and stable at the growth temperature of *C. bescii* (78°C), see chapter 3.

However our collaborators were unable to crystallize X-PL3 under any of the conditions used. This was not surprising, as the majority of multi-domain CAZy proteins with domains separated by even short linkers have not been crystallized. The only approach at present to determine the structure of multi-domain proteins is to use protein modeling using structures of the individual domains [60].

CHAPTER 3

BIOCHEMICAL CHARACTERIZATION OF TWO DOMAINS OF CBES_1854, A FAMILY 3 PECATE LYASE

Introduction

Cbes_1854 of the thermophilic bacterium C. bescii is a family 3 pectate lyase and belongs to subfamily 1, as predicted by amino acid sequence analysis. In addition to the C. bescii enzyme, there are three other pectate lyases in subfamily 1, from the mesophilic bacteria *Bacillus* sp. KSM 15, Bacillus subtlis subsp. and Penibacillus barcinonensis BP 23 (Table 3.1). Among these, the structures of the enzyme from Bacillus sp. KSM P15 and Bacillus subtilis subsp. are available. It is assumed that family 3 pectate lyase have the same catalytic mechanism as that of family 1 pectate lyase but this has not been confirmed structurally [58]. This assumption is based on the fact that both families share the same β -helix fold. The goal of this work was the biochemical characterization of C. bescii X-PL3 and PL3 including determining their pH optima. It was expected that they would be similar to those of family 1 pectate lyases, which prefer alkaline pH. The other goals were to (A) determine the temperature optimum of both enzymes; (B) evaluate the substrate specificity of X-PL3; (C) reveal the catalytic mechanism through a collaborative effort by structure determination with bound substrate and (D) determine if this enzyme works synergistically with another cellulolytic enzyme of C. bescii, cellulase A. This is the most abundant cellulolytic enzyme in extracellular fraction of C. bescii [46].

Materials and Methods

Protein sample preparation: The PL3 and X-PL3 proteins were obtained by expression and one

Protein Name	Organism	Temp. opt.(°C)	pH opt.	Subfamil y	References
Pectin lyase (Ply; PlyAI4)	Bacillus sp. 14	50	10.5	3	[61]
Pectate lyase (Pel15)	Bacillus sp. KSM15	50-55	10.5	1	[62]
Pectate lyase C (PelC)	Bacillus subtilis subsp.	65	10	1	[63]
Pectate lyase I (PelI)	Dickeya dadantii 3937	37	9	5	[64]
Pectate lyase A (PelA)	Penibacillus barcinonensis BP- 23	50	10	1	[65]
Pectate lyase B (PelB)	Pectobacterium carotovorum SCC3193	37	9.5	5	[66]
Pectate lyase (Pel-3)	Pectobacterium carotovorum subsp.cart .71	37	ND	5	[67]
Endo Pectate lyase I (PL I)	Streptomyces thermocarboxydus B-1	50	9	4	[68]
Pectate lyase (X-PL3)	C. bescii	85	9	1	This work

Table 3.1: Comparison of different family 3 pectate lyases

ND: not determined

Pel15 and PelC have three-dimensional structure.

step purification as described in chapter 2. Samples were prepared by dialysis of protein from the Ni-NTA elution fractions against 20mM Tris buffer pH 8.0 containing 100 mM KCl. The protein was quantified by absorption at 280 nm, using the calculated extinction coefficient of the PL3 and X-PL3 shown in Table 3.2. Some other properties based on amino acid sequence was predicted by Vector NTI software are also shown in Table 3.2 [69].

Pectate lyase assay: A substrate stock of 0.25% polygalacturonic acid (PGA) was prepared in 100 ml water and the pH adjusted 9.0 by the addition of NaOH pellets. For the assay, 100 mM N-cyclohexyl-3-aminopropanesulfonic acid (CAPS) buffer containing1mM CaCl₂ at 80°C was used, unless otherwise indicated. The amount of product was measured by the absorption at 235 nm, using a molar extinction coefficient of 4600 $M^{-1}cm^{-1}$. One unit of enzyme activity is defined as the amount of enzyme that produces 1 µmol of product /min [70].

To determine the pH dependence of the enzyme the same amount were added to different buffers, ranging from pH 4.0 to 11.0, and were incubated with 0.1% PGA for 3 min at 80°C. Data was collected at one-minute intervals to calculate the slope, which was used to calculate the specific activity. To determine the temperature profile, the same assay protocol was used at pH 9.0 but the temperature range used was between 40 and 90°C. For the comparative studies the results are expressed as relative activity (in %).

The optimum concentration of calcium for activity was determined using the same protocol. The substrate specificity was determined using 0.1% of all of the substrates listed in Table 3.3. The thermal stability of the protein was determined by incubating it at different temperatures for up to 48 hours. The thermal stability was assessed at pH 9.0 and pH 7.0. At different time intervals samples were removed and assayed at 80°C.

Table: 3.2. Predicted biophysical properties of the catalytic domain (PL3) and of the

Properties	X-PL3	PL3
Molecular weight (kDa)	46,178	20,023
Isoelectric point (pI)	6.32	6.29
Molar extinction coefficient (A ₂₈₀)	55,265	18,500

holoenzyme form (X-PL3) of Cbes_1854 from C. bescii

Results and Discussion

The pH, temperature and calcium dependence of the pectate lyase activity of PL3 and X-PL3 were determined using polygalacturonic acid as the substrate. The assay is based on measurement of the unsaturated reaction product, hexenuronic acid, formed after cleavage of the α 1-4 glycosidic linkage of PGA. Cleavage follows β elimination reaction and results in the unsaturated product, hexenuronic acid (Figure 3.1).

As shown in Figure 3.2 calcium ions were indispensable for the activity of both X-PL3 and PL3 with an optimum concentration of 1mM Ca^{++.} Both enzymes also display pH optima of 9.0. At pH 9.0, the optimum temperature for the activity is between 85 and 90°C. As shown in Figure 3.3, the half-life for both, PL3 and X-PL3, at 90°C is approximately 1 hr but at pH 7.5 this increases to 5 hr and 2 hr for PL3 and X-PL3, respectively. These data suggests that X-PL3 is less thermostable than PL3. To determine the specificity of X- PL3, different types of esterified forms of citrus pectin were used as substrates (Table 3.3). It was found that X-PL3 shows the highest activity with 90% esterified citrus pectin, as shown in Table 3.4. No significant activity was observed when rhamnogalacturonan was used as substrate, although X-PL3 still shows comparable amount of activity using non-esterified polygalacturonan as a substrate.

The biochemical properties of enzymes from the PL3 family (Table 3.1) suggest that these enzymes have the same catalytic mechanism as that of family 1 pectate lyases (PL1). The catalytic mechanism of *Bacillus subtilis* pectate lyase (BsPel), the closest homolog to X-PL3 and a member of PL1, has been studied in detail. BsPel has two basic residues, lysine and arginine, as part of the active site where arginine acts as a catalytic base and abstracts a proton. The role of the arginine residue has been established by mutation [27]. As shown in Figure 3.1, the proposed

Substrate	Description
Apple pectin	75% methylation
Citrus pectin (mixed)	55-70% methylation
Rhamnogalactomanan	Galacturonic acid and rhamnose backbone
Pectic galactan	Gal : Ara : Rha : GalUA = 82 : 6 : 3 : 9
Citrus pectin (30%)	30% esterified
Citrus pectin (60%)	60% esterified
Citrus pectin (90%)	60% esterified
Polygalacturonan	Pectin backbone

Table 3.3: Substrates used for pectate lyase assay

The detail descriptions of the substrates are present in following references [70-72].



Figure 3.1. Schematic diagram of the typical β -elimination cleavage of polygalacturonic acid. P+, Ca⁺²; B, catalytic base that abstracts the proton from C-5; and A-H, does protonation to the glycosidic oxygen. The unsaturated product is Hexenuronic acid and the diagram is taken from reference [26].



Figure 3.2. Temperature, pH and calcium (Ca⁺⁺) dependence of pectate lyase activity of PL3 and X-PL3. (A) Specific activity profile at pH 9.0; (B) Calcium requirement for pectate lyase activity; (C) pH profile and (D) Temperature profile. All assays were carried out using 0.1% polygalaturonic as the substrate.



Figure 3.3. Thermostability of PL3 and X-PL3.These plots depict the activity and thermal stability relationship between X-PL3 (A, C) and PL3 (B, D) at pH 9.0 (A, B) and 7.5 (C, D). The temperature range for measurement of thermal stability is between 60-90°C, as shown in D.

catalytic mechanism of BsPel is an anti- β elimination reaction that requires a basic residue to abstract the proton from the C5 atom of the substrate that has been charge stabilized by the two catalytically important Ca²⁺ ions, together with the concerted action of an acidic residue that donates a proton to the carbonyl group, creating an enol intermediate. This is followed by elimination of the leaving group together with protonation of the O4 atom by another acidic residue or a water molecule acidified by the two catalytically important Ca²⁺ ions, while the Ca⁺² atoms stabilize the enol intermediate [27,73].

In our collaborative effort, the three-dimensional structure of the catalytic domain (PL3) of *C. bescii* pectate lyase was obtained with the substrate trigalacturonic acid bound to the enzyme [74]. A comparison of catalytic residues in PL3 and BsPel, from family PL1, revealed that there is Lys 130 in the catalytic domain of PL3 in place of the catalytic residue Arg 279 of BsPel. The evidence for this comes through the substrate-bound structure of the PL3 domain (PDB: 4EW9) at 1.6-Angstrom resolution [74]. Figure 3.4 shows a diagram of the catalytic domain of PL3 with trigalacturonic acid bound. In this three-dimensional structure, Lys 130 is not correctly positioned to abstract the proton from C-5, rather Lys 108 is better positioned for proton abstraction in the active site (Figure 3.4). Therefore, it is likely that C. *bescii* PL3 uses this Lys 108 as a catalytic residue rather than Lys 130. This observation supports a different catalytic mechanism for the family 3 pectate lyase. Based on this, our collaborators have proposed a novel reaction mechanism based on Lys 108 as a catalytic base for the family 3 pectate lyase [74].

Based on the biochemical results presented here, the unknown domain 'X' does not provide any advantage to the catalytic PL3 domain in either degrading polygalacturonic acid or providing protein stability. From the NCBI-BLAST of just the 'X' domain against the non-

Substrate	Sp. Activity (U/mg)	Activity (%)
Polygalacturonic acid	225 ± 23	45 ± 10
Apple pectin	394 ±13	78 ± 3
Citrus pectin (mixed)	385 ±16	76 ± 4
Rhamnogalactomannan	6 ± 1	1 ± 0.5
Pectic galactan	10 ± 2	2 ± 20
Citrus pectin (30%)	274 ± 34	54 ± 12
Citrus pectin (60%)	362 ± 14	72 ± 4
Citrus pectin (90%)	506 ± 15	100 ± 3

Table 3.4: Activity of X-PL3 with different substrates

One unit is 1 μmol of unsaturated galacturonic acid product /min.

redundant protein database, this domain was not found in archaea and is mainly found associated with the catalytic domain of the pectate lyases in different *Caldicellulosiruptor* species. In some cases it was found with a domain of glycosyl hydrolase family 28. GH28 is also proposed to function as a pectin-degrading enzyme. The association of domain X with pectin-degrading enzymes suggests that domain 'X' might functions as a pectin-binding domain.

During these studies it was found that the recombinant form X-PL3 and PL3 tended to aggregate at higher temperature (85° C) if the protein concentrations were above ~0.2mg/ml. This was not expected for a protein of thermophilic origin and suggests that glycosylation of the linker region of X-PL3 may have a positive role in bridging the interaction the domain 'X' and its catalytic domain. The linker region provides sites for O-glycosylation but this does not occur in recombinant proteins expressed in *E. coli*. These data suggest that glycosylation plays an important role in the protein stability at higher temperature.

Another objective of the present study was to provide purified pectate lyase in order to study of enzymatic degradation of biomass. The essential role of cellulases in biomass degradation has been well established but it is not known if other extracellular enzymes, such as pectate lyases, could enhance the biomass degradation. In collaboration studies with Dr. M. Himmel of National Renewable Energy Laboratory, CO, it was shown that the presence of *C. bescii* PL3 increases the amount of sugar released from untreated switchgrass biomass in comparison to using *C. bescii* cellulase A alone. In fact, the presence of PL3 addition reduces the load of cellulase A by \sim 33% [74]. Therefore, adding *C. bescii* pectate lyase to cellulases appears to be a promising approach to increase the efficiency of enzymatic biomass degradation.



Figure 3.4. Diagram of the catalytic domain PL3 with the reaction product trigalacturonic acid. Reaction products are marked and labeled (PDB: 4EW9) [74]. The catalytic residue, Lys 108, is in bold and labeled. The diagram was prepared using PyMol software [59].

SUMMARY

Caldicellulosiruptor bescii is the most thermophilic bacterium known with potential for plant biomass degradation. Remarkably, it grows on switchgrass biomass without any chemical treatment at 78°C [37,43]. *C. bescii* possesses a variety of enzymes with carbohydrate active domains (CAZy enzymes) that are potentially involved in the conversion of insoluble switchgrass into soluble fermentable sugars [43]. Transcriptional analyses after growth of *C. bescii* on glucose and switchgrass revealed that the genes encoding 94 enzymes were upregulated and therefore apparently involved in switchgrass degradation. Among these 94 enzymes, 18 have a proposed functional role in carbohydrate metabolism. These include four potential pectinolytic enzymes, three of which are predicted to be extracellular (Cbes_1853,Cbes_1854 and Cbes_1855) [39]. Interestingly, Cbes_1854 is the most highly upregulated enzyme during growth on switchgrass versus glucose. This was unexpected because pectin represents less than 5% of the total polysaccharides in plant biomass. Therefore, Cbes_1854 was chosen as a model multidomain enzyme for overexpression, purification and biochemical studies.

Cbes_1854 belongs to family 3 pectate lyases, of which there is no structure available for the catalytic domain with a bound substrate. In addition to biochemical studies, we also sought to provide the pure enzyme for the structural analysis so that information could be obtained on the mechanism of family 3 pectate lyases. Cbes_1854 encodes two domains. The X domain is of unknown function while PL3 is the catalytic domain. The role of the X domain is probably to bind the insoluble pectin-rich component of the complex biomass. Three constructs were cloned and expressed separately in *E. coli:* domain X, catalytic domain (PL3) and full-length protein having both domains (X-PL3). Domain X was found in the insoluble fraction after expression in *E. coli* and attempts to recover the protein were not successful. In contrast, the other two proteins remained in the soluble fraction and could be purified using the 6X-His tag present at the N-terminus by affinity chromatography. The yields of each of these two proteins were approximately 10 mg/liter of *E. coli* culture.

Biochemical studies of PL3 and X-PL3 using polygalacturonic acid (PGA) as a substrate showed that both enzymes have similar alkaline pH optima (pH 8-9) and temperature optima (85-90°C). This observation also highlights the fact that under *in vitro* assay conditions domain X does not influence the catalytic properties of the PL3 domain. Interestingly, it was observed that the catalytic domain is more active against 90% esterified citrus pectin than it is with unesterified PGA. PL3 and X-PL3 are thermostable and each has a half-life of 1 h at 90°C at pH 9.0. Although the protein stability increases at physiological pH (pH 7.0), they retain only 20-30% of total activity at this pH value.

The three-dimensional structure of catalytic domain PL3 was obtained in a collaborative study and this has been deposited (PDB: 1EE6). It consists of a parallel β -helix fold. The structure with product of trigalacturonic acid, 4, 5-unsaturated digalacturonic acid and D-galacturonic acid, was also obtained and deposited (PDB: 4EW9). Surprisingly, unlike the previously described structure of a pectate lyase, that of BsPel, a homolog with same structural fold belonging to PL1 family, the PL3 catalytic module has only one basic residue, lysine, rather than two conserved basic residues lysine and arginine, present in active site. This suggests that PL3 has different catalytic mechanism from what was observed with BsPel.

To probe the role of Cbes_1854 in biomass degradation, untreated switchgrass biomass

was incubated with the most abundant cellulolytic enzyme, Cel A, from *C. bescii* supplemented with recombinant PL3. The presence of PL3 reduced the load of Cel A by ~33% while maintaining the same amount of sugar released from biomass. This demonstrates the synergistic effect of pectinolytic enzymes in biomass degradation. In the future, recombinant PL3 and X-PL3 have the potential to be used in enzyme mixtures to increase the efficiency of biomass degradation.

REFERENCES

- 1. Turner JA (1999) A realizable renewable energy future. Science 285: 687-689.
- 2. Sheehan JJ (2009) Biofuels and the conundrum of sustainability. Curr Opin Biotechnol 20: 318-324.
- 3. Jenkins R, Alles C (2011) Field to fuel: developing sustainable biorefineries. Ecol Appl 21: 1096-1104.
- 4. Lynd LR, Cushman JH, Nichols RJ, Wyman CE (1991) Fuel ethanol from cellulosic biomass. Science 251: 1318-1323.
- 5. Louime C, Uckelmann H (2008) Cellulosic ethanol: securing the planet future energy needs. Int J Mol Sci 9: 838-841.
- 6. Carroll A, Somerville C (2009) Cellulosic biofuels. Annu Rev Plant Biol 60: 165-182.
- 7. Dale BE (2011) Cellulosic biofuels and the road to energy security. Environ Sci Technol 45: 9823.
- Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, et al. (2007) Biomass recalcitrance: Engineering plants and enzymes for biofuels production. Science 315: 804-807.
- 9. Pauly M, Keegstra K (2008) Cell-wall carbohydrates and their modification as a resource for biofuels. Plant J 54: 559-568.
- 10. Schubert C (2006) Can biofuels finally take center stage? Nat Biotechnol 24: 777-784.
- 11. Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, et al. (2006) The path forward for biofuels and biomaterials. Science 311: 484-489.
- 12. Davison BH, Keller M, Fowler VS (2009) The goals and research of the bioenergy sciences center (BESC): developing cost-effective and sustainable means of producing biofuels by overcoming biomass recalcitrance. Bioenergy Research 2: 177-178.
- 13. Zhao XB, Zhang LH, Liu DH (2012) Biomass recalcitrance. Part I: the chemical compositions and physical structures affecting the enzymatic hydrolysis of lignocellulose. Biofuels Bioproducts & Biorefining-Biofpr 6: 465-482.
- Zhao XB, Zhang LH, Liu DH (2012) Biomass recalcitrance. Part II: Fundamentals of different pre-treatments to increase the enzymatic digestibility of lignocellulose. Biofuels Bioproducts & Biorefining-Biofpr 6: 561-579.
- 15. Pauly M, Keegstra K (2008) Physiology and metabolism 'Tear down this wall'. Curr Opin Plant Biol 11: 233-235.
- 16. Carpita NC, Gibeaut DM (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. Plant J 3: 1-30.
- 17. McLaughlin SB, Kszos LA (2005) Development of switchgrass (*Panicum virgatum*) as a bioenergy feedstock in the United States. Biomass & Bioenergy 28: 515-535.
- 18. Sanderson MA, Adler PR, Boateng AA, Casler MD, Sarath G (2006) Switchgrass as a biofuels feedstock in the USA. Canadian Journal of Plant Science 86: 1315-1325.
- 19. Mohnen D (2008) Pectin structure and biosynthesis. Curr Opin Plant Biol 11: 266-277.

- 20. Atmodjo MA, Hao Z, Mohnen D (2013) Evolving views of pectin biosynthesis. Annu Rev Plant Biol 64: 747-779.
- 21. Mohnen D, Biswal A, Hao Z, Hunt K, Gelineo-Albersheim I, et al. (2011) Plants with altered cell wall biosynthesis and methods of use. WO Patent 2,011,130,666.
- 22. Lionetti V, Francocci F, Ferrari S, Volpi C, Bellincampi D, et al. (2010) Engineering the cell wall by reducing de-methyl-esterified homogalacturonan improves saccharification of plant tissues for bioconversion. Proc Natl Acad Sci U S A 107: 616-621.
- 23. Marin-Rodriguez MC, Orchard J, Seymour GB (2002) Pectate lyases, cell wall degradation and fruit softening. J Exp Bot 53: 2115-2119.
- 24. Abbott DW, Boraston AB (2008) Structural biology of pectin degradation by Enterobacteriaceae. Microbiol Mol Biol Rev 72: 301-316, table of contents.
- 25. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, et al. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res 37: D233-238.
- 26. Charnock SJ, Brown IE, Turkenburg JP, Black GW, Davies GJ (2002) Convergent evolution sheds light on the anti-β-elimination mechanism common to family 1 and 10 polysaccharide lyases. Proc Natl Acad Sci U S A 99: 12067-12072.
- 27. Seyedarabi A, To TT, Ali S, Hussain S, Fries M, et al. (2010) Structural insights into substrate specificity and the anti beta-elimination mechanism of pectate lyase. Biochemistry 49: 539-546.
- 28. Yee KL, Rodriguez M, Jr., Tschaplinski TJ, Engle NL, Martin MZ, et al. (2012) Evaluation of the bioconversion of genetically modified switchgrass using simultaneous saccharification and fermentation and a consolidated bioprocessing approach. Biotechnol Biofuels 5: 81.
- 29. Olson DG, McBride JE, Shaw AJ, Lynd LR (2012) Recent progress in consolidated bioprocessing. Curr Opin Biotechnol 23: 396-405.
- 30. Li H, Cann AF, Liao JC (2010) Biofuels: biomolecular engineering fundamentals and advances. Annu Rev Chem Biomol Eng 1: 19-36.
- 31. Blumer-Schuette SE, Brown SD, Sander KB, Bayer EA, Kataeva I, et al. (2013) Thermophilic lignocellulose deconstruction. FEMS Microbiol Rev.(*in press*)
- 32. Frock AD, Kelly RM (2012) Extreme thermophiles: Moving beyond single-enzyme biocatalysis. Curr Opin Chem Eng 1: 363-372.
- 33. Blumer-Schuette SE, Kataeva I, Westpheling J, Adams MW, Kelly RM (2008) Extremely thermophilic microorganisms for biomass conversion: status and prospects. Curr Opin Biotechnol 19: 210-217.
- 34. Chang T, Yao S (2011) Thermophilic, lignocellulolytic bacteria for ethanol production: current state and perspectives. Appl Microbiol Biotechnol 92: 13-27.
- 35. Blumer-Schuette SE, Giannone RJ, Zurawski JV, Ozdemir I, Ma Q, et al. (2012) *Caldicellulosiruptor* core and pangenomes reveal determinants for noncellulosomal thermophilic deconstruction of plant biomass. J Bacteriol 194: 4015-4028.
- 36. Blumer-Schuette SE, Lewis DL, Kelly RM (2010) Phylogenetic, microbiological, and glycoside hydrolase diversities within the extremely thermophilic, plant biomass-degrading genus *Caldicellulosiruptor*. Appl Environ Microbiol 76: 8084-8092.
- 37. Yang SJ, Kataeva I, Hamilton-Brehm SD, Engle NL, Tschaplinski TJ, et al. (2009) Efficient degradation of lignocellulosic plant biomass, without pretreatment, by the thermophilic

anaerobe "Anaerocellum thermophilum" DSM 6725. Appl Environ Microbiol 75: 4762-4769.

- 38. Yang SJ, Kataeva I, Wiegel J, Yin Y, Dam P, et al. (2010) Classification of 'Anaerocellum thermophilum' strain DSM 6725 as Caldicellulosiruptor bescii sp. nov. Int J Syst Evol Microbiol 60: 2011-2015.
- 39. Kataeva I, Foston MB, Yang SJ, Pattathil S, Biswal AK, et al. (2013) Carbohydrate and lignin are simultaneously solubilized from unpretreated switchgrass by microbial action at high temperature. Energy & Environmental Science 6: 2186-2195.
- 40. Basen M, Rhaesa AM, Kataeva I, Prybol CJ, Scott IM, et al. (2013) Degradation of high loads of crystalline cellulose and of unpretreated plant biomass by the thermophilic bacterium *Caldicellulosiruptor bescii*. Bioresource Technology.(*in press*)
- 41. Chung D, Farkas J, Westpheling J (2013) Overcoming restriction as a barrier to DNA transformation in *Caldicellulosiruptor* species results in efficient marker replacement. Biotechnol Biofuels 6: 82.
- 42. Cha M, Chung D, Elkins JG, Guss AM, Westpheling J (2013) Metabolic engineering of *Caldicellulosiruptor bescii* yields increased hydrogen production from lignocellulosic biomass. Biotechnol Biofuels 6: 85.
- 43. Dam P, Kataeva I, Yang SJ, Zhou F, Yin Y, et al. (2011) Insights into plant biomass conversion from the genome of the anaerobic thermophilic bacterium *Caldicellulosiruptor bescii DSM 6725*. Nucleic Acids Res 39: 3240-3254.
- 44. Kataeva IA, Yang SJ, Dam P, Poole FL, 2nd, Yin Y, et al. (2009) Genome sequence of the anaerobic, thermophilic, and cellulolytic bacterium "*Anaerocellum thermophilum*" DSM 6725. J Bacteriol 191: 3760-3761.
- 45. VanFossen AL, Ozdemir I, Zelin SL, Kelly RM (2011) Glycoside hydrolase inventory drives plant polysaccharide deconstruction by the extremely thermophilic bacterium *Caldicellulosiruptor saccharolyticus*. Biotechnol Bioeng 108: 1559-1569.
- 46. Lochner A, Giannone RJ, Rodriguez M, Jr., Shah MB, Mielenz JR, et al. (2011) Use of labelfree quantitative proteomics to distinguish the secreted cellulolytic systems of *Caldicellulosiruptor bescii* and *Caldicellulosiruptor obsidiansis*. Appl Environ Microbiol 77: 4042-4054.
- 47. Zverlov V, Mahr S, Riedel K, Bronnenmeier K (1998) Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) from the extreme thermophile 'Anaerocellum thermophilum' with separate glycosyl hydrolase family 9 and 48 catalytic domains. Microbiology 144 (Pt 2): 457-465.
- 48. Jung SK, Parisutham V, Jeong SH, Lee SK (2012) Heterologous expression of plant cell wall degrading enzymes for effective production of cellulosic biofuels. J Biomed Biotechnol 2012: 405842.
- 49. Beguin P, Aubert JP (1994) The biological degradation of cellulose. FEMS Microbiology Reviews 13: 25-58.
- 50. Kataeva IA, Uversky VN, Brewer JM, Schubot F, Rose JP, et al. (2004) Interactions between immunoglobulin-like and catalytic modules in *Clostridium thermocellum* cellulosomal cellobiohydrolase CbhA. Protein Eng Des Sel 17: 759-769.
- 51. Kataeva IA, Brewer JM, Uversky VN, Ljungdahl LG (2005) Domain coupling in a multimodular cellobiohydrolase CbhA from *Clostridium thermocellum*. FEBS Lett 579: 4367-4373.

- 52. Sammond DW, Payne CM, Brunecky R, Himmel ME, Crowley MF, et al. (2012) Cellulase Linkers Are Optimized Based on Domain Type and Function: Insights from Sequence Analysis, Biophysical Measurements, and Molecular Simulation. Plos One 7.e48615.
- 53. Jeoh T, Michener W, Himmel ME, Decker SR, Adney WS (2008) Implications of cellobiohydrolase glycosylation for use in biomass conversion. Biotechnology for Biofuels 1.10.
- 54. Gunnarsson LC, Montanier C, Tunnicliffe RB, Williamson MR, Gilbert HJ, et al. (2007) Novel xylan-binding properties of an engineered family 4 carbohydrate-binding module. Biochemical Journal 406: 209-214.
- 55. Kataeva IA, Uversky VN, Ljungdahl LG (2003) Calcium and domain interactions contribute to the thermostability of domains of the multimodular cellobiohydrolase, CbhA, a subunit of the *Clostridium thermocellum* cellulosome. Biochem J 372: 151-161.
- 56. Kataeva I, Chang J, Xu H, Luan CH, Zhou J, et al. (2005) Improving solubility of *Shewanella oneidensis MR-1* and *Clostridium thermocellum JW-20* proteins expressed into Esherichia coli. J Proteome Res 4: 1942-1951.
- 57. Alahuhta M, Chandrayan P, Kataeva I, Adams MW, Himmel ME, et al. (2011) A 1.5 A resolution X-ray structure of the catalytic module of *Caldicellulosiruptor bescii* family 3 pectate lyase. Acta Crystallogr Sect F Struct Biol Cryst Commun 67: 1498-1500.
- 58. Akita M, Suzuki A, Kobayashi T, Ito S, Yamane T (2001) The first structure of pectate lyase belonging to polysaccharide lyase family 3. Acta Crystallography D Biol Crystallography 57: 1786-1792.
- 59. DeLano W The PyMOL Molecular Graphics System. 2002 Version 1.2 r3pre. Schrödinger, LLC.
- 60. Rappsilber J (2011) The beginning of a beautiful friendship: cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. J Struct Biol 173: 530-540.
- 61. Zhou JP, Dong YY, Gao YJ, Tang XH, Li JJ, et al. (2012) Characterization of a family 3 polysaccharide lyase with broad temperature adaptability, thermo-alkali stability, and ethanol tolerance. Biotechnology and Bioprocess Engineering 17: 729-738.
- 62. Hatada Y, Saito K, Koike K, Yoshimatsu T, Ozawa T, et al. (2000) Deduced amino-acid sequence and possible catalytic residues of a novel pectate lyase from an alkaliphilic strain of *Bacillus*. European Journal of Biochemistry 267: 2268-2275.
- 63. Soriano M, Diaz P, Pastor FIJ (2006) Pectate lyase C from *Bacillus subtilis*: a novel endocleaving enzyme with activity on highly methylated pectin. Microbiology-Sgm 152: 617-625.
- 64. Creze C, Castang S, Derivery E, Haser R, Hugouvieux-Cotte-Pattat N, et al. (2008) The crystal structure of pectate lyase PelI from soft rot pathogen *Erwinia chrysanthemi* in complex with its substrate. Journal of Biological Chemistry 283: 18260-18268.
- 65. Soriano M, Blanco A, Diaz P, Pastor FIJ (2000) An unusual pectate lyase from a *Bacillus sp* with high activity on pectin: cloning and characterization. Microbiology-Uk 146: 89-95.
- 66. Heikinheimo R, Flego D, Pirhonen M, Karlsson MB, Eriksson A, et al. (1995) Characterization of a Novel Pectate Lyase from *Erwinia-Carotovora Subsp Carotovora*. Molecular Plant-Microbe Interactions 8: 207-217.
- 67. Liu Y, Chatterjee A, Chatterjee AK (1994) Nucleotide-sequence and expression of a novel pectate lyase gene (*pel-3*) and a closely linked endopolygalacturonase gene (*peh-1*) of *Erwinia-carotovora subsp carotovora-71*. Appl Environ Microbiol 60: 2545-2552.

- 68. Tonouchi A, Hara Y, Umehara R, Sanuki T, Fukusawa T, et al. (2010) Cloning of the Gene Encoding an Endo-Acting Pectate Lyase from *Streptomyces thermocarboxydus*. Bioscience Biotechnology and Biochemistry 74: 433-436.
- 69. Lu G, Moriyama EN (2004) Vector NTI, a balanced all-in-one sequence analysis suite. Briefings in bioinformatics 5: 378-388.
- 70. Jayani RS, Saxena S, Gupta R (2005) Microbial pectinolytic enzymes: a review. Process Biochemistry 40: 2931-2944.
- 71. van Alebeek GJ, van Scherpenzeel K, Beldman G, Schols HA, Voragen AG (2003) Partially esterified oligogalacturonides are the preferred substrates for pectin methylesterase of *Aspergillus niger*. Biochem J 372: 211-218.
- 72. Dongowski G, Lorenz A, Anger H (2000) Degradation of pectins with different degrees of esterification by *Bacteroides thetaiotaomicron* isolated from human gut flora. Appl Environ Microbiol 66: 1321-1327.
- 73. Scavetta RD, Herron SR, Hotchkiss AT, Kita N, Keen NT, et al. (1999) Structure of a plant cell wall fragment complexed to pectate lyase C. Plant Cell 11: 1081-1092.
- 74. Alahuhta M, Brunecky R, Chandrayan P, Kataeva I, Adams MW, et al. (2013) The structure and mode of action of *Caldicellulosiruptor bescii* family 3 pectate lyase in biomass deconstruction. Acta Crystallography Sect D Biol Crystallography 69: 534-539.