EXPLORATION OF THE DIVERSE CAZOME OF *PAENIBACILLUS AMYLOLYTICUS* 27C64 REVEALS A NOVEL PECTIN DECONSTRUCTION SYSTEM

by

CHRISTIAN KEGGI

(Under the Direction of Joy Doran-Peterson)

ABSTRACT

Paenibacillus amylolyticus 27C64 is a Gram-positive bacterium that was isolated from the hindgut of an aquatic crane fly larva, an insect which relies on a microbial community in its gut to break down the plant material it eats. This bacterium in particular showed promise as a source of novel enzymes that could be used to improve deconstruction of plant cell wall polysaccharides. Whole genome sequencing revealed a diverse set of more than 300 CAZymes including pectinases, xylanases, and cellulases. This organism was able to use pectins and xylan as sole carbon sources for growth and appears to specialize in the deconstruction of non-cellulosic polysaccharides. Dynamic regulation of putative pectinases in response to the specific substrate supplied for growth revealed a complex system with potentially novel elements. Characterizing key enzymes within the homogalacturonan deconstruction system revealed that pectin methylesterase activity is not required for efficient deconstruction. This stands in stark contrast to systems that have already been described. Three of the homogalacturonan degrading enzymes also have significant potential to improve numerous industrial applications. The

genomic analysis also revealed a novel bifunctional enzyme involved in rhamnogalacturonan I deconstruction. This enzyme, RglA, has a complex multidomain structure that includes both rhamnogalacturonan acetylesterase and rhamnogalacturonan lyase domains along with eight non-catalytic fibronectin type III like domains. Five of these non-catalytic domains are likely carbohydrate binding domains. When characterized *in vitro*, RglA displayed both lyase and esterase activities as predicted, as well as a preference for acetylated pectins. This ability to function on acetylated substrates has not been previously described; acetylation is typically inhibitory to rhamnogalacturonan lyases. Collectively, this work to understand the pectinolytic system of *P. amylolyticus* 27C64 has contributed to overall understanding of microbial pectin deconstruction and identified novel enzymes that may have industrial uses.

INDEX WORDS: *Paenibacillus amylolyticus*, 27C64, pectinase, pectate lyase, pectin lyase, rhamnogalacturonan lyase, rhamnogalacturonan acetylesterase

EXPLORATION OF THE DIVERSE CAZOME OF *PAENIBACILLUS AMYLOLYTICUS* 27C64 REVEALS A NOVEL PECTIN DECONSTRUCTION SYSTEM

by

CHRISTIAN KEGGI

BS, Wake Forest University, 2014

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2019

EXPLORATION OF THE DIVERSE CAZOME OF *PAENIBACILLUS AMYLOLYTICUS* 27C64 REVEALS A NOVEL PECTIN DECONSTRUCTION SYSTEM

by

CHRISTIAN KEGGI

Major Professor:	Joy Doran-Peterson
Committee:	David Blum
	Timothy Hoover
	Elizabeth Ottesen
	Diana Downs

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2019

ACKNOWLEDGEMENTS

I would like to thank the countless people who have furthered my education and provided the mentoring and help I needed to finish my doctoral degree. My major advisor, Dr. Joy Doran-Peterson, has guided my scientific journey for five long years and enabled all of the work within this dissertation. I would also like to thank my committee members, Dr. David Blum, Dr. Tim Hoover, Dr. Elizabeth Ottesen, and Dr. Diana Downs for the time they have taken to provide important feedback as I worked towards my degree. My classmates, especially Olivia Thompson and Joshua Chu, were always helpful when I needed someone to provide feedback on experimental design, papers, posters, and presentations throughout my time at UGA. The Bioexpression and Fermentation Facility was also a tremendously helpful resource, and my thanks go out to all of the staff who took time to help me. Rebecca Gardner helped with parts of this work as an undergraduate, and I appreciate her effort and dedication. I would be remiss if I did not also thank my parents who ensured I received the education I needed to pursue my own path. Finally, I owe my wife Leah the most gratitude. Her unwavering support made the past five years so much easier than it would have been otherwise.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 PURPOSE

When Robert Hooke described the first biological cell after viewing thin sections of cork under a microscope [1], he could not have known that understanding the structure of the plant cell walls he had just drawn would remain an area of keen scientific interest more than three and a half centuries later. Today we recognize that plant cell walls are built from rigid cellulose fibers interwoven with matrix polysaccharides (hemicelluloses and pectins) that are cemented in place with lignin, an aromatic heteropolymer. These cell wall structures store the majority of the 55 billion tons of carbon that are fixed by plants in terrestrial ecosystems every year [2] and are the most abundant carbon source available to heterotrophic herbivores at the base of the food web. Even in the human intestinal tract where cellulose is not digested, competition for the carbon available in matrix polysaccharides drives selection of the gut microbiome [3, 4]. Plant material can also be converted to thermal energy by burning and subsequently to electrical energy, or to intermediate energy-rich fuels. Although biomass provides an alternative to complete reliance on long-sequestered hydrocarbon fuels, only about two percent of this renewable carbon source is utilized by humans today [5].

Recognition that conversion of renewable biomass to biofuels could lower carbon emissions, reduce the United States' dependence on imported oil, and minimize the impact

of oil price fluctuations on the domestic economy led to passage of the Energy Independence and Security of Act of 2007, a law that defined targets for domestic renewable fuel production. In pursuit of these targets, the US Department of Energy commissioned a series of studies to evaluate the feasibility of large scale biofuel production and estimated that more than one billion dry tons of biomass per year could be made available by 2040, enough to displace more than 30% of annual US petroleum consumption [6]. By using a variety of agricultural and industrial waste residues along with energy crops, these production levels can be achieved without disrupting current food, feed, and fiber demands [6]. The sources evaluated in this assessment include first-generation feedstocks, such as corn and sugar, which presently account for virtually all US ethanol production, as well as second-generation lignocellulosic feedstocks. Second-generation or advanced biomass sources are non-food crops that are low in starch and free sugar. Instead, the majority of the carbon in second-generation feedstocks is stored in cell wall polysaccharides. A wide variety of sources are considered second-generation, including agricultural wastes, such as corn stover, and purpose-grown energy crops including poplar and switchgrass [6]. Future increases in biofuel production must rely on second-generation feedstocks since corn ethanol is capped at current production levels of 15 billion gallons per year. Advanced biofuel production is also critical to the goal of reducing carbon emissions because corn ethanol can only be carbon neutral in existing cropland; converting forest or prairie lands for corn production results in substantial soil carbon losses that are difficult to offset [7]. In contrast, lignocellulosic feedstocks such as grasses that require minimal fertilizer or irrigation incur little to no carbon debt [7–9], while agricultural waste resources carry no land use change consequences and are already available [6, 10].

Current conversion technology for lignocellulosic feedstocks includes two divergent processes: thermochemical conversion and biochemical conversion. Thermochemical conversion involves heating biomass to break it down into intermediate products, such as syngas (hydrogen and carbon monoxide), or liquid intermediates that can be chemically converted into fuels and other chemicals [11]. In contrast, biochemical conversion leverages plant cell wall deconstructing enzymes to release sugars from biomass that are in turn fermented to fuels or other chemicals [11]. This conversion process typically involves four key steps: mechanical processing, pre-treatment, enzymatic saccharification, and fermentation. Mechanical processing reduces the particle size and increases the surface area to volume ratio prior to pretreatment. A wide range of pretreatment technologies exist and may rely on chemicals, such as dilute acids or alkali, physical treatments including heating or steam explosion (pressurization and rapid depressurization), or a combination of methods. Every method has the advantage of improving enzymatic digestibility by expanding the accessible cell wall area, and some may degrade or remove hemicelluloses, pectins, or lignin [12]. Unfortunately, many of these pretreatment methods also release aromatic compounds or produce sugar-degradation products that are toxic to fermenting organisms, with severe pretreatments often producing high concentrations of these fermentation inhibitors [12, 13]. Pre-treated biomass is next subjected to enzymatic saccharification to release soluble sugars that are fermented to ethanol or other chemicals that can be recovered from the fermentation broth [11]. To reduce the cost of biochemical conversion and make the process competitive with gasoline production even when the price of oil is low, improvements in converting the mechanically processed biomass to fermentable sugars must be made [11]. The enzymatic

saccharification step is the most important point of leverage to achieve this goal because its improvement may reduce enzyme cost, reduce the severity of pretreatment or even eliminate it entirely, and improve overall process viability. However, a more complete understanding of diverse plant cell wall deconstruction systems is first required.

The majority of work on lignocellulosic biomass deconstruction has focused on the production of biofuels largely because ethanol production from sugar is an established technology and government funding aimed at achieving energy independence has focused on this goal. However, unlocking the energy available in plant cell walls could allow for the low-cost production of countless renewable chemicals. Interest in the use of lignocellulosic feedstocks for lactic acid production is growing [14], and there is a long-established history of producing acetone and butanol from carbohydrates, including some agricultural waste sources [15, 16]. Improved understanding of plant cell wall polysaccharide deconstruction also has the potential to improve other manufacturing sectors that work with plant material. For example, the pulp and paper, cotton and textiles, and food industries all routinely process plant material in ways that may be improved by the discovery of new enzymes that modify cell wall structure or extracted polysaccharides.

1.2 PLANT CELL WALL STRUCTURE

1.2.1 Cell Wall Anatomy and Organization

Plant cell walls are built in single layers called lamellae that are deposited as they are synthesized. These layers rely primarily on cellulose for their structural integrity, have matrix polysaccharides (hemicelluloses and pectins) that surround and embed the cellulose, and may contain lignin, an aromatic heteropolymer that crosslinks cell wall elements and

rigidifies the wall. The lamellae vary in composition and eventually build distinct regions of the cell wall, each with distinguishing features and a different function [17]. The outermost portion of the cell wall, the middle lamella, is composed primarily of pectin and is responsible for cell-cell adhesion. The primary cell wall, directly beneath the middle lamella, is the first true wall layer and contains cellulose embedded in matrix polysaccharides that are easily remodeled as the cell grows [17]. All plant cells have a primary cell wall, and in herbaceous plants most cells have only a primary wall. Some cells, especially in woody plants, also have very thick secondary cell walls. These secondary walls have significantly more hemicellulose than pectin, and lignin deposited in these walls increases rigidity and largely excludes water [17]. In both primary and secondary layers of the cell wall, a complex set of interactions between polymers stabilize the wall superstructure and interfere with enzymatic deconstruction. This resistance to saccharification is referred to as biomass recalcitrance and is the primary challenge to using second-generation lignocellulosic feedstocks as a carbon source for biofuel or chemical production.

1.2.2 Cellulose

Cellulose is simple in its primary structure; it is a β-1,4 linked glucose polymer in which the repeating glucose disaccharide is termed cellobiose (Figure 1.1a) [18]. The complexity of cellulose lies in the superstructures it is able to form. Nascent chains are synthesized by rosettes of cellulose synthase subunits that move through the plasma membrane, extruding 36 parallel glucose molecules that crystallize into a microfibril through extensive chain-to-chain hydrogen bonding [17]. These microfibrils can also form larger macrofibrils that are synthesized by arrayed cellulose synthase rosettes and can

assume different crystalline forms [19]. The crystalline superstructure of cellulose inherently excludes water and enzymes and is a major contributor to biomass recalcitrance. For this reason, many pretreatment techniques focus on disrupting the crystallinity of cellulose microfibrils to improve enzyme access [12].

1.2.3 Hemicelluloses

The term hemicellulose actually refers to a family of structurally distinct polysaccharides. Members of this group are generally named for the primary carbohydrate in their backbone, and many hemicelluloses have heterogenous sidechains that vary between cell wall layers, tissues, and species [18]. The most important hemicellulosic polysaccharides are the xylans. They are composed of a β-1,4 linked D-xylose backbone that can be acetylated and carry various short carbohydrate sidechains (Figure 1.1b). Xylans are the most abundant hemicellulose in the secondary cell walls of most plants [18]. Softwoods are the exception to this rule, where xylans are present, but less common than mannans [18]. Xylans also play a smaller role in the primary cell walls of many plants. Mannans have a randomly ordered backbone of β-1,4 linked D-mannose and D-glucose monomers, and as with xylans, can be acetylated, carry sidechains, and are important to the secondary cell wall structure of most plants (Figure 1.1b). Many primary cell walls contain glucans, a less abundant type of hemicellulose that has a β-1,4 or β-1,3 lined Glu backbone (Figure 1.1b) [18].

Hemicelluloses contribute to biomass recalcitrance by tightly associating with the surface of spatially separated cellulose microfibrils and tethering them together [18]. Some sidechains can also carry hydroxycinnamate esters, such as ferulic acid, which form covalent crosslinks between accessible tether regions [18]. Although hemicellulose can be

enzymatically deconstructed, many pretreatment methods attempt to degrade or disrupt these polysaccharides to allow enzymes direct access to cellulose without delay [12].

1.2.4 Pectic Polysaccharides

Pectin, unlike cellulose and hemicellulose, plays a critical role in the primary cell wall but has only a minimal role in the secondary cell wall. It is also the main component of the middle lamella. The term pectin actually refers to a family of five galacturonic acid (GalA) -rich pectic polysaccharides: homogalacturonan (HG), rhamnogalacturonan I (RG-I), rhamnogalacturonan II (RG-II), xylogalacturonan (XGA), and apiogalacturonan (AG). HG has a linear backbone of 1,4-linked α-D-GalA residues that can be methylated at the C-6 carboxyl group or acetylated at the O-2 or O-3 positions, but which do not have sidechains (Figure 1.1c) [20]. Methylation can be extensive, even in excess of 90%, because pectin is synthesized in its methylated form in the Golgi before being deposited in the cell wall. Once in the wall, stretches of HG are demethylated to convert the C-6 carboxyl to its negatively charged state, which can form calcium bridges with adjacent demethylated regions of separate HG chains [21]. Acetylation is typically far less common; many pectins have virtually no acetylation and even the most highly acetylated pectins rarely surpass 30-35% [22]. The functional role of acetylation or the cause of variation between different plant species is not understood. HG is the most abundant pectic polysaccharide, comprising about 65% of pectin by weight in most plants [20]. The next most abundant pectic polysaccharide, RG-I, typically represents 20-35% of pectin and has an alternating GalA-rhamnose (Rha) backbone comprised of the repeating [-α-D-GalA-1,2-α-L-Rha-1,4-] disaccharide (Figure 1.1c) [20]. Twenty to eighty percent of the Rha residues in the backbone are decorated at the C-4 position with single, linear, or branched

arabinans, galactans, or arabinogalactans [20]. The fact that these sidechains vary between plants and even between different tissues within a single plant suggests that the sidechains of RG-I have an important functional role. GalA residues on the RG-I backbone are commonly acetylated at the O-2 and O-3 positions as in HG, but it is unclear whether methylation of these residues can occur [20]. The third major pectic polysaccharide, RG-II, is a substituted homogalacturonan and typically accounts for 10% of pectin by weight [20]. Four distinct side chains comprised of 12 different monosaccharides and 20 different linkages decorate a HG backbone, making RG-II the most complex plant cell wall polysaccharide [23]. The structures of these sidechains are highly conserved across tissues and species because they form covalent borate cross-links that allow these pectins to dimerize, a function that is critical to wall integrity as evidenced by dwarfism in RG-II sidechain mutants and borate-deficient plants [20, 23]. The two other types of pectin, XGA and AG, are simply homogalacturonan substituted with xylose and apiose, respectively. However, the former is only found in reproductive tissues, while the latter is restricted to the walls of specific aquatic monocots, and neither is important to the wall structure or integrity of most higher plants [21].

While the structures of individual pectic polysaccharides have been clearly resolved within the last two decades, the superstructure of the polymer remains an area of active research. Until fairly recently, it was assumed that pectin was one long chain with short RG-I and RG-II regions interspersed along the HG backbone. However, when an alternate model was proposed that suggested RG-I was a sidechain of HG [24], this assumption was revisited. There is clear evidence that both RG-I [25] and RG-II [26] are covalently linked to HG, but it seems that there are only short (4-10 residue long) stretches of HG separating

RG-I regions [27]. Whether the long stretches of HG that can be isolated (>300 residues) are also attached to RG-I [28] remains unknown, as does whether RG-I and RG-II exist on the same linear chain. Perhaps the most important recent development is the recognition that RG-I is also covalently linked to an arabinogalactan protein and arabinoxylan in *Arabidopsis thaliana*, which suggests that other proteins may also be involved in interconnection of pectic polysaccharides [29].

The chemistry of linkages between pectin and other cell wall components is another area of active research. Traditional models of the cell wall frequently describe pectin as a separate but co-extensive network in which pectin is not tightly bound to or covalently associated with the cellulose-hemicellulose network [18]. However, this thinking has changed. The strongest evidence for tight associations between these cell wall elements is the previously mentioned arabinogalactan protein in *A. thaliana* that covalently links RG-I and arabinoxylan [29]. It has also recently been demonstrated that pectin can covalently attach to cellulose in carrot cell walls low in xyloglucan [30]. These newer findings agree with older evidence that pectin can covalently attach to xyloglucans [31], has ferulic acid on the sidechains of RG-I that can cross-link with hemicelluloses [32], and is likely attached to lignin [33].

Pectin's contribution to biomass recalcitrance in pectin-rich feedstocks that lack secondary cell walls is well established [10], and efficient deconstruction of this type of biomass is hindered without pectinases [34]. However, a series of recent publications have demonstrated that pectin contributes to biomass recalcitrance in lignocellulosic feedstocks even though its concentration in this type of material is generally so low that its role has been ignored. Expression of pectinase enzymes in *A. thaliana* [35, 36] and aspen (*Populus*

tremula L. x *tremuloides* Michx.) [37] has successfully improved sugar release from these plants. Downregulation of native genes involved in HG synthesis in *Populus deltoides* has also improved saccharification yields [38, 39]. Analysis of natural variation in *A. thaliana* is consistent with these results, showing that a reduction in the amount of demethylated HG (which is capable of calcium cross-linking) improves saccharification [40]. However, the most convincing evidence that pectin plays a critical role in biomass recalcitrance comes not from a plant mutant, but from a deletion in the biocatalyst *Caldicellulosiruptor bescii*. *C. bescii* is able to grow on unpretreated biomass, but deletion of a pectinase gene cluster resulted in a significant growth defect on whole plant material [41]. Put together, this recent evidence has clarified that pectin is in fact intimately associated with other plant cell wall elements and contributes to the recalcitrance of all biomass types, even when it is a minor component of the cell wall.

1.2.5 Lignin

Lignin is an aromatic heteropolymer composed of three phenylpropanoid compounds — *p*-coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol — collectively referred to as monolignols [42]. These compounds are synthesized intracellularly then secreted to secondary cell walls where they are assembled into random structures through radical polymerization reactions [21]. Lignin is only deposited in the secondary walls when cells have stopped growing because it rigidifies the wall and prevents further expansion.

In secondary cell walls, lignin is perhaps the largest single factor contributing to biomass recalcitrance because it is covalently crosslinked to cellulose, hemicellulose, and pectin, and fills space in the cell wall to reduce its porosity and exclude water and enzymes [18, 33]. Lignin also poses a unique challenge because its efficient enzymatic

deconstruction has yet to be achieved. However, a number of pretreatments are able to partially remove lignin or modify its structure and allow for high recovery of the carbohydrates from secondary cell walls [12].

1.3 ENZYMATIC SACCHARIFICATION OF PLANT CELL WALL POLYSACCHARIDES

1.3.1 Classification of Plant Cell Wall Polysaccharide Deconstruction Enzymes

Given the complexity of plant cell wall structure and superstructure, it is unsurprising that the enzymes that deconstruct cell wall polysaccharides are remarkably diverse and are classified into scores of different families. The need to clearly organize these families led to the creation of the Carbohydrate Active Enzymes database (CAZy database; www.cazy.org), which today contains 340 distinct families including glycoside hydrolases (GH), polysaccharide lyases (PL), carbohydrate esterases (CE), auxiliary activities (AA), and glycosyl transferases (GT) [43]. The database includes not just those families known to break down plant cell wall polysaccharides, but rather all enzymes that synthesize, deconstruct, or modify carbohydrates and polysaccharides in all domains of life. The enzymes included in this broad definition are frequently referred to as Carbohydrate Active Enzymes (CAZymes), and the sum of all CAZymes in a genome is referred to as a CAZome. Importantly, the database also describes 85 different non-catalytic carbohydrate binding motifs (CBM; also carbohydrate binding domains) that mediate carbohydrate binding and are critical to the function of many CAZymes [44].

1.3.2 Cellulose Deconstruction

The conventional model of cellulose deconstruction is derived from the filamentous fungus *Hypocrea jecorina* (*Trichoderma reesei*) and includes three synergistic groups of glycoside hydrolases distributed across many CAZy families: endo-cellulases, exo-cellulases, and β-glucosidases. Endo-cellulases (EC 3.2.1.4) hydrolyze internal sites of non-crystalline amorphous cellulose, while exo-cellulases (cellobiohydrolases; EC 3.2.1.91 and EC 3.2.1.176) work from either the reducing or non-reducing ends of cellulose chains to release cellobiose. The β-glucosidases (EC 3.2.1.21) then split cellobiose into two glucose molecules [45, 46]. Collectively, these enzymes work slowly inward from the outer ends of microfibrils and produce inefficient saccharification of highly crystalline celluloses. Subsequent work has recognized a number of mechanisms used by diverse organisms to directly attack crystalline cellulose and rapidly increase the number of sites available to exo-cellulases. This group of enzymes includes some traditional glycoside hydrolases with unique biochemical mechanisms, such as a *H. jecorina* cellulase that pries a single cellulose chain away from a crystalline bundle [46, 47]; an endo-processive cellulase from *Clostridium phytofermentans* that binds directly to crystalline cellulose and is essential for growth of the bacterium on this substrate [48]; and a combined GH48-GH9 cellulase from *C. bescii* that can burrow into crystalline cellulose to create cavities that serve as access points for other enzymes [49, 50]. Lytic polysaccharide monooxygenases (LPMOs), a group of enzymes first described as acting on chitin and other substrates [51], also contribute to cellulose deconstruction [52, 53]. These enzymes can bind to the surface of microfibrils and nick the cellulose chains, working synergistically with glycoside hydrolases by increasing the number of sites available to traditional cellulases. These more

recent findings make it clear efficient cellulose deconstruction requires a means of attacking the crystalline structures of cellulose directly, after which the more common endo-cellulases, cellobiohydrolases, and β-glucosidases can finish saccharifying the substrate.

1.3.3 Hemicellulose Deconstruction

Each hemicellulosic polysaccharide has a different suite of enzymes responsible for its deconstruction, but the overall process for each polymer is similar. In general, exo-hydrolases remove the branched sidechains from xylans, mannans, and glucans to strip them down to their backbones [45]. Enzymes capable of removing esterified acetyl or ferulic acid groups also exist for those hemicelluloses that are known to carry these decorations [43]. Glucan backbones are broken down primarily by endo-hydrolases across a wide range of CAZy families including GH5, GH7, GH12, GH16, GH44, and GH74 [43, 45]. Mannans are deconstructed with a combination of GH5 and GH26 endo-hydrolases, as well as GH2 β-mannosidases, while xylan deconstruction relies primarily on GH10 and GH11 xylanases [43, 45]. In general, these groups of enzymes excel at degrading soluble hemicellulose, but deconstruction can be hindered by the tight interactions between hemicellulose and cellulose, as well as the limited porosity of the cell wall. Both of these limitations are typically addressed through pretreatment.

1.3.4 Pectin Deconstruction

Microbial pectinases are distributed across all three domains of life and are remarkably diverse in their mechanisms and substrate specificities. In general, the term ‘pectinase’ refers to enzymes that deconstruct the HG or RG-I backbones underlying the different pectic polysaccharides. These two types of pectin make up the majority of the

polymer [20], and degradation of the backbones alone is sufficient for pectin removal in many industrial processes [54–56]. However, RG-I sidechains can be removed by endo- and exo- arabinases and galactanases [57], while deconstruction of the RG-II sidechains is a carefully coordinated process requiring numerous glycoside hydrolases [58].

Deconstruction of the pectic backbones that remain after sidechain removal is a multistep process that varies by organism. Although hundreds of pectinases are listed in the CAZy database, many have been characterized at only a very basic level [43], and the vast majority of what we know about microbial pectin deconstruction comes from just three phylogenetic groups: filamentous fungi in the genus *Aspergillus,* which have been used as a source of commercially available pectinases; members of the Gram-negative *Enterobacteriaceae* family, which includes two genera of important plant pathogens (*Dickeya* and *Pectobacterium,* formerly contained with *Erwinia*); and the Gram-negative human gut symbiont *Bacteroides thetaiotaomicron,* which competes for non-cellulosic polysaccharides available in the human digestive tract. With very few exceptions, the only pectinases that have been studied in the context of their native biological system come from these groups of organisms. The RG-I deconstruction capability of *Bacillus subtilis* has also received some attention, but much less is known about its mechanism of HG deconstruction.

The major classes of pectinases that have been identified thus far can be broadly categorized into two groups: esterases and depolymerases. Esterases remove the esterified methyl and acetyl groups decorating the pectic backbones and are traditionally subdivided into pectin methylesterases (PME; EC 3.1.1.11), which are CE8 family enzymes, CE12 pectin acetylesterases (PAE), and CE12 rhamnogalacturonan acetylesterases (RGAE) [56,

59–61]. Plant PAEs and RGAEs are CE13 family enzymes. Each of these are named for the type of decoration they remove and the backbone on which they are active. In the context of pectinase nomenclature, the term pectin refers to highly methylated or acetylated HG, while pectate refers to de-esterified HG. Rhamnogalacturonan always refers to the RG-I type alternating backbone in this context. The second group, the depolymerases, deconstruct the pectic backbones and are likewise subdivided based on the substrate they act upon, but are also split into two catalytic groups: hydrolases that break glycosidic bonds with a hydrolytic mechanism and lyases that rely on β-elimination [56, 59–61]. Each depolymerase subgroup contains examples of endo-acting enzymes that cut at random internal sites, and exo-acting enzymes that bind the end of a polysaccharide molecule and work processively inward.

Endo- and exo-hydrolases active on HG are referred to as polygalacturonases (EC 3.2.1.15) and exo-polygalacturonases (EC 3.2.1.67, EC 3.2.1.82), respectively, and are assigned to GH28 [43, 59]. All of these enzymes cut demethylated sites in the HG backbone; polygalacturonases that cut methylated sites have not been described. In contrast, pectin lyases (EC 4.2.2.10) cut only at methylated sites in HG and belong to PL1. Pectate lyases (EC 4.2.2.2), active on demethylated HG, are spread across PL1, PL2, PL3, PL9, and PL10 with exo-pectate lyases (EC 4.2.2.9) present only in PL1, PL2, and PL9 [43]. Only one exo-pectin lyase has been described, but it has not yet been assigned to a CAZy family [62].

Unlike HG-specific enzymes, RG depolymerases are not subdivided by methylation preference because RG-I is not methylated. Acetylation, however, is common on this type of backbone, but all known RG depolymerases are inhibited by acetylation and

cut only at deacetylated sites [57]. This includes rhamnogalacturonases (EC 3.2.1.171), which cleave internal Rha-1,4-α-D-GalA bonds and belong to GH28 along with the polygalacturonases; RG lyases (EC 4.2.2.23), which cleave the same bond with a different mechanism and fall into PL4 and PL11; and exo-RG lyases (EC 4.2.2.24), which occupy PL11 and PL26 [43, 57]. In general, these major classes of pectinases are shared between phylogenetic domains, although most pectinolytic organisms rely only on some of these classes and supplement these activities with additional enzymes to degrade each pectic polysaccharide to monomeric sugars. A summary of pectinase classification, including the supplemental enzymes used by disparate organisms to saccharify substrates completely, is provided (Table 1.1).

In every pectinolytic organism, HG deconstruction begins extracellularly with pectate lyases, pectin lyases, and polygalacturonases that solubilize pectic fragments and make them available for further deconstruction. Since HG is often methylated extensively, removing the methyl groups is critical to initial size reduction. *Dickeya dadantii* (formerly *Erwinia chrysanthemi*) relies on one secreted PME [63] and one outer membrane associated PME [64] to achieve this. Likewise, three *Apsergillus* species [65–67] and *Bacillus licheniformis* [68] have similar extracellular enzymes, although none have been described in *B. subtilis* or *B. thetaiotaomicron* [43]. Removal of acetyl groups by extracellular PAEs at this stage has been described in *D. dadantii* [69, 70], *B. subtilis*[71], *B. licheniformis* [72], *A. niger* [73], and *A. aculeatus* [74]. Once inhibitory acetyl groups are gone, depolymerases can begin degrading long HG chains to smaller fragments. In general, fungi rely primarily or exclusively on polygalacturonases [59], while bacteria rely mostly on pectate lyases [75, 76]. This is likely because hydrolases function most

effectively in the mildly acidic pH ranges preferred by fungi, whereas lyases function best across the neutral to alkaline pH range preferred by these bacteria. Pectin lyases are also involved in initial size reduction of HG chains but are much less common. Only one pectin lyase has been described within the *Enterobacteriaceae* [77], and many bacteria lack pectin lyases entirely, while fungal genomes contain very few [43]. Instead, these systems rely on the combined action of PMEs and methylation-sensitive depolymerases. After de-esterification and initial depolymerization, filamentous fungi use exo-polygalacturonases, and in one case an exo-pectin lyase, to degrade remaining fragments to monomeric sugars that can be imported and catabolized [59, 62, 78]. In contrast, both *Enterobacteriaceae* and *B. thetaiotaomicron* initially import oligogalacturonides into the periplasm where an additional set of pectate lyases and polygalacturonases further degrade the oligosaccharides [75, 76]. *Enterobacteriaceae* use exo-polygalacturonases and PL2 exo-pectate lyases to produce a mixture of saturated and Δ4,5-unsaturated digalacturonates that are split into monomers by a cytoplasmic PL22 oligogalacturonide lyase (EC 4.2.2.6) [75, 79]. *B. thetaiotaomicron,* however, uses an endo-acting PL1 pectate lyase that produces very short oligosaccharides with Δ4,5-unsaturated GalA residues at their non-reducing ends. The terminal unsaturated residue is removed by a GH105 Δ4,5-unsaturated α-galacturonidase, and the remaining saturated fragments are degraded to monomeric GalA by an exo-polygalacturonase [76]. Only monomers are transported to the cytoplasm [76]. The complete mechanisms by which *B. subtilis* breaks down HG is not understood even though a few extracellular enzymes [71, 80, 81] and an unusual NAD+ dependent GH4 α-D-galacturonidase [82] have been described. Unfortunately, this means that there is no good model of Gram-positive HG deconstruction.

RG-I deconstruction proceeds similarly in that deacetylation of the backbone begins the process of deconstructing solubilized fragments. RG acetylesterases have been described in *Aspergillus* [73, 83–85]*, Bacillus* [86, 87]*,* and *B. thetaiotaomicron* [76]. In both *B. subtilis* and *B. thetaiotaomicron*, initial extracellular size reduction is accomplished exclusively with RG lyases, but these organisms take different approaches to saccharifying the released oligosaccharides [76, 88, 89]. *B. subtilis* has an extracellular exo-RG lyase that further degrades RG fragments to a disaccharide (Δ4,5-unsaturated D-GalA-1,2-α-L-Rha) that is imported to the cytoplasm and cleaved by a GH105 unsaturated RG hydrolase [88, 89]. In contrast, *B. thetaiotaomicron* transports oligosaccharides to the periplasm and uses additional endo-RG lyases to produce a mixture of oligomeric products [76, 88, 89]. The terminal Δ4,5-unsaturated GalA residues are removed from these mixed-length oligosaccharides by a GH105 hydrolase similar to the one in *B. subtilis*, and the remaining saturated products are saccharified by a GH106 RG α-L-rhamnosidase (EC 3.2.1.174) and GH28 α-D-galacturonidases that alternate removal of terminal Rha and GalA residues [76]. The released monosaccharides are then transported to the cytoplasm [76]. Although some *Aspergillus* species do have RG lyases [43, 90], their systems rely primarily or exclusively on RG hydrolases for initial size reduction. Exo-acting GH28 RG galacturonohydrolases (EC 3.2.1.173) and GH78 RG rhamnohydrolases (EC. 3.2.1.174) then alternate removal of terminal residues until oligosaccharides are degraded to monomeric GalA and Rha [57]. Removal of unsaturated terminal GalA residues from RG fragments generated by those strains that carry RG lyases has not been directly demonstrated, but all the major species of *Aspergillus* contain uncharacterized GH105 sequences that likely fill this role [43]. Since RG-I deconstruction has been studied primarily after work on the pectinolytic systems of

Enterobacteriaceae slowed, little is known about how this family deconstructs the RG backbone, and only one RG lyase has been described within this phylogenetic group [91].

Although much is known about microbial pectin deconstruction, there are substantial gaps in our understanding of this process. The majority of well-studied pectinases come from *Aspergillus* and Gram-negative bacteria, while less work has been done in Gram-positives. There has also been a lot of focus on key CAZy families like GH28 and PL1, which contain 219 and 110 characterized sequences, respectively, while nine other CAZy families have fewer than ten characterized sequences [43]. Even within GH28, the largest and best characterized pectinase family, most studies have focused on the structure and function of enzymes from *Aspergillus* [92]. While characterization of representative proteins within each family provides important information about the family as a whole, an adequate assessment of the diversity within each family has not been pursued in tandem. Likewise, a system-level understanding of pectin deconstruction in phylogenetically diverse organisms needs to be explored further. The fact that recent work in understudied groups has consistently led to the description of new classes of pectinases is a testament to the value of exploring diverse systems.

1.4 SOURCES OF NOVEL CAZYMES

Exploration of microbial communities responsible for breaking down plant material will likely continue to yield new and useful enzymes for overcoming biomass recalcitrance and help make the use of lignocellulose-derived carbohydrates viable for the production of low-cost fuels and chemicals. One such microbial community exists in the hindgut of *Tipula abdominalis* (aquatic crane fly) larvae. These insects live in riparian ecosystems and

consume primarily conditioned leaf litter, processing it in a manner that parallels a biorefinery [93]. Ingested material is mechanically processed by chewing before passing through the alkaline (pH 11) midgut, which in effect chemically pretreats the cell walls. Enzymatic saccharification is mediated by microbial enzymes that ferment released carbohydrates to organic acids the insect can absorb [94]. Screening organisms isolated from this hindgut community using differential plate media resulted in the identification of numerous bacteria capable of breaking down plant cell wall polysaccharides or fluorescent carbohydrate substrates, but only the isolates identified as *Paenibacillus amylolyticus* 27C64 were able to break down all of the substrates evaluated [95]. Since this initial identification, other evidence has supported the idea that *P. amylolyticus* is a potentially valuable source of CAZymes. Adding the organism to yeast fermentations of pretreated pine wood improved ethanol yields [96]. In addition, two pectate lyases (PelA and PelB) with unusually high activity on methylated substrates were identified by screening a genomic library [97]. PelB was able to eliminate the need for a commercial pectinase mixture in fermentations of cull peaches despite the fact that the commercial product contained a wide variety of pectinases, including esterases and depolymerases [34]. This broad substrate specificity has also been noted in some of the very few pectinases characterized in other species of *Paenibacillus,* which suggests that this capability is conserved, although no published work has addressed how it is functionally important to the pectinolytic systems of this genus [98–103]. Collectively, these data point to *P. amylolyticus* 27C64 as a promising source of potentially novel pectinases and other CAZymes.

1.5 OBJECTIVE

The broad objective of this work is to describe the breadth of *P. amylolyticus* 27C64's plant cell wall degradation capabilities and study its novel aspects in detail, especially those enzymes that have potential to improve biomass deconstruction. *P. amylolyticus* is well-suited to expand the diversity of known CAZymes since it is phylogenetically distinct from the thermophilic anaerobes and filamentous fungi that are the source of much of our knowledge about cellulases and hemicellulases, as well as from the fungal and Gram-negative models of pectin deconstruction.

In pursuit of this objective, the genome of *P. amylolyticus* was sequenced to identify potential CAZymes, including cellulases, hemicellulases, and pectinases (Chapter 2). Dynamic regulation of the putative pectinolytic system in response to different pectic polysaccharides was also evaluated. Key depolymerases were purified and characterized *in vitro*, shedding light into this organism's novel approach to HG deconstruction (Chapter 3). A bifunctional RG lyase and RG acetylesterase, initially identified in the genome sequence, was also studied *in vitro* (Chapter 4). Collectively, these findings have contributed to a better understanding of CAZyme diversity.

Table 1.1 Carbohydrate Active Enzyme (CAZy) families involved in pectin deconstruction

Family	Number Characterized	Activity	Sub-families	EC Number
Pectate lyases				
PL1	110	Pectate lyase	1-3, 5-7	4.2.2.2
		Exo-pectate lyase	5	4.2.2.9
		Pectin lyase	4, 8	4.2.2.10
PL2	6	Pectate lyase	1	4.2.2.2
		Exo-pectate lyase	2	4.2.2.9
PL3	24	Pectate lyase	1, 2, 3	4.2.2.2
PL4	6	Rhamnogalacturonan lyase	1-4	4.2.2.23
PL9	12	Pectate lyase	1	4.2.2.2
		Exo-pectate lyase	1	4.2.2.9
		Rhamnogalacturonan lyase	1	4.2.2.23
PL10	7	Pectate lyase	1	4.2.2.2
PL11	6	Rhamnogalacturonan lyase	1	4.2.2.23
		Exo-rhamnogalacturonan lyase	1	4.2.2.24
PL22	2	Oligogalacturonate lyase	--	4.2.2.6
PL26	1	Exo-rhamnogalacturonan lyase	--	4.2.2.24
Glycoside hydrolases				
GH4	27	α-D-galacturonidase	--	3.2.1.67
GH28	219	Polygalacturonase	--	3.2.1.15
		Exo-polygalacturonase	--	3.2.1.67
		Exo-polygalacturonosidase	--	3.2.1.82
		α-L-rhamnosidase	--	3.2.1.40
		Rhamnogalacturonase	--	3.2.1.171
		RG galacturonohydrolase	--	3.2.1.173
		Endo-xylogalacturonan hydrolase	--	3.2.1.-
GH78	30	α-L-rhamnosidase	--	3.2.1.40
		RG rhamnohydrolase	--	3.2.1.174
GH105	9	Unsaturated RG hydrolase	--	3.2.1.172
GH106	6	α-L-rhamnosidase	--	3.2.1.40
GH138	2	α-D-galacturonidase	--	3.2.1.-
Carbohydrate esterases				
CE8	48	Pectin methylesterase	--	3.2.1.11
CE12	11	Pectin acetylesterase	--	3.1.1.-
		RG acetylesterase	--	3.1.1.86
CE13	2	Pectin acetylesterase	--	3.1.1.-

A. Cellulose

→4)- β-D-Glu*p*-(1→4)-β-D-Glu*p*-(1→4)-β-D-Glu*p*-(1→4)-β-D-Glu*p*-(1→

B. Hemicelluloses

Xylan backbone

→4)- β-D-Xyl*p*-(1→4)-β-D-Xyl*p*-(1→4)-β-D-Xyl*p*-(1→4)-β-D-Xyl*p*-(1→

Mannan backbone

→4)- β-D-Man*p*-(1→4)-β-D-Man*p*-(1→4)-β-D-Glu*p*-(1→4)-β-D-Man*p*-(1→

Glucan backbone

→4)- β-D-Glu*p*-(1→3)-β-D-Glu*p*-(1→4)-β-D-Glu*p*-(1→3)-β-D-Glu*p*-(1→

C. Pectins

Homogalacturonan

Me ↓ 6 Me ↓ 6

→4)-α-D-Gal*p*A-(1→4)-α-D-Gal*p*A-(1→4)-α-D-Gal*p*A-(1→4)-α-D-Gal*p*A-(1→

2 ↑ Ac 3 ↑ Ac

Variable sidechains including linear and branched arabinans (1,5-α-L-Ara*f*), galactans (1,4-β-D-Gal*p*), or arabinogalactan sidechains ↓ 4

Rhamnogalacturonan I

→4)-α-D-Gal*p*A-(1→2)-α-L-Rha*p*-(1→4)-α-D-Gal*p*A-(1→2)-α-L-Rha*p*-(1→

2 ↑ Ac 3 ↑ Ac

Figure 1.1 Structures of selected plant cell wall polysaccharides. Xylans, mannans, and glucans have various sidechains not depicted.

1.6 REFERENCES

1. Hooke R (1665) Micrographia, or some physiological descriptions of minute bodies made by magnifying glasses, with observations and inquiries thereupon. The Royal Society, London

2. Running SW, Nemani RR, Heinsch FA, Zhao M, Reeves M, Hashimoto H (2004) A continuous satellite-derived measure of global terrestrial primary production. Bioscience 54:547–560

3. Martens EC, Lowe EC, Chiang H, Pudlo NA, Wu M, McNulty NP, Abbott DW, Henrissat B, Gilbert HJ, Bolam DN, Gordon JI (2011) Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. PLOS Biol 9:1–16. doi:10.1371/journal.pbio.1001221

4. Koropatkin NM, Cameron EA, Martens EC (2012) How glycan metabolism shapes the human gut microbiota. Nat Rev Microbiol 10:323–335. doi:10.1038/nrmicro2746

5. Pauly M, Keegstra K (2008) Cell-wall carbohydrates and their modification as a resource for biofuels. Plant J 54:559–568. doi:10.1111/j.1365-313X.2008.03463.x

6. US Department of Energy (2016) 2016 Billion-ton report: advancing domestic resources for a thriving bioeconomy, volume 1: economic availability of feedstocks. Oak Ridge National Laboratory, Oak Ridge, TN

7. Fargione J, Hill J, Tilman D, Polasky S, Hawthorne P (2008) Land clearing and the biofuel carbon debt. Science (80-) 319:1235–1238. doi:10.1016/0024-4937(90)90045-3

8. Tilman D, Hill J, Lehman C (2006) Carbon-negative biofuels from low-input high-diversity grassland biomass. Science (80-) 314:1598–1600. doi:10.1126/science.1133306

9. US Department of Energy (2017) 2016 Billion-ton report: advancing domestic resources for a thriving bioeconomy, volume 2: environmental sustainability effects of select scenarios from volume 1. Oak Ridge National Laboratory, Oak Ridge, TN

10. Edwards MC, Doran-Peterson J (2012) Pectin-rich biomass as feedstock for fuel ethanol production. Appl Microbiol Biotechnol 95:565–75. doi:10.1007/s00253-012-4173-2

11. Foust TD, Ibsen KN, Dayton DC, Hess JR, Kenney KE (2008) The biorefinery. In: Himmel ME (ed) Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy. Blackwell, pp 7–37

12. Johnson DK, Elander RT (2008) Pretreatments for enhanced digestibility of feedstocks. In: Himmel ME (ed) Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy. Blackwell, pp 436–453

13. Wang S, Sun X, Yuan Q (2018) Strategies for enhancing microbial tolerance to inhibitors for biofuel production: a review. Bioresour Technol 258:302–309. doi:https://doi.org/10.1016/j.biortech.2018.03.064

14. Abdel-Rahman MA, Tashiro Y, Sonomoto K (2011) Lactic acid production from lignocellulose-derived sugars using lactic acid bacteria: overview and limits. J Biotechnol 156:286–301. doi:https://doi.org/10.1016/j.jbiotec.2011.06.017

15. Jones DT, Woods DR (1986) Acetone-butanol fermentation revisited. Microbiol Rev 50:484–524

16. Zverlov V V, Berezina O, Velikodvorskaya G a, Schwarz WH (2006) Bacterial acetone and butanol production by industrial fermentation in the Soviet Union: use of hydrolyzed agricultural waste for biorefinery. Appl Microbiol Biotechnol 71:587–97. doi:10.1007/s00253-006-0445-z

17. Ding S, Himmel ME (2008) Anatomy and ultrastructure of maize cell walls: an example of energy plants. In: Himmel ME (ed) Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy. Blackwell, pp 38–60

18. Harris PJ, Stone BA (2008) Chemistry and molecular organization of plant cell walls. In: Himmel ME (ed) Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy. Blackwell, pp 61–93

19. Atalla RH, Brady JW, Matthews JF, Ding S, Himmel ME (2008) Structures of plant cell wall celluloses. In: Himmel ME (ed) Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy. Blackwell, pp 188–212

20. Mohnen D (2008) Pectin structure and biosynthesis. Curr Opin Plant Biol 11:266–277. doi:10.1016/j.pbi.2008.03.006

21. Mohnen D, Bar-peled M, Somerville C (2008) Cell wall polysaccharide synthesis. In: Himmel ME (ed) Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy. Blackwell, pp 94–187

22. Sun R, Hughes S (1998) Extraction and physico-chemical characterization of pectins from sugar beet pulp. Polym J 30:671–677. doi:10.1295/polymj.30.671

23. O'Neill MA, Ishii T, Albersheim P, Darvill AG (2004) Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. Annu Rev Plant Biol 55:109–139. doi:10.1146/annurev.arplant.55.031903.141750

24. Vincken J-P, Schols HA, Oomen RJFJ, McCann MC, Ulvskov P, Voragen AGJ, Visser RGF (2003) If homogalacturonan were a side chain of rhamnogalacturonan I. Implications for cell wall architecture. Plant Physiol 132:1781–1789. doi:10.1104/pp.103.022350

25. Coenen GJ, Bakx EJ, Verhoef RP, Schols HA, Voragen AGJ (2007) Identification of the connecting linkage between homo- or xylogalacturonan and rhamnogalacturonan type I. Carbohydr Polym 70:224–235. doi:10.1016/j.carbpol.2007.04.007

26. Ishii T, Matsunaga T (2001) Pectic polysaccharide rhamnogalacturonan II is covalently linked to homogalacturonan. Phytochemistry 57:969–974. doi:https://doi.org/10.1016/S0031-9422(01)00047-4

27. NAKAMURA A, FURUTA H, MAEDA H, TAKAO T, NAGAMATSU Y (2003) Structural studies by stepwise enzymatic degradation of the main backbone of soybean soluble polysaccharides consisting of galacturonan and rhamnogalacturonan. Biosci Biotechnol Biochem 66:1301–1313. doi:10.1271/bbb.66.1301

28. Atmodjo MA, Hao Z, Mohnen D (2013) Evolving views of pectin biosynthesis. Annu Rev Plant Biol 64:747–79. doi:10.1146/annurev-arplant-042811-105534

29. Tan L, Eberhard S, Pattathil S, Warder C, Glushka J, Yuan C, Hao Z, Zhu X, Avci U, Miller JS, Baldwin D, Pham C, Orlando R, Darvill A, Hahn MG, Kieliszewski MJ, Mohnen D (2013) An Arabidopsis cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. Plant Cell 25:270–287. doi:10.1105/tpc.112.107334

30. Broxterman SE, Schols HA (2018) Interactions between pectin and cellulose in primary plant cell walls. Carbohydr Polym 192:263–272. doi:https://doi.org/10.1016/j.carbpol.2018.03.070

31. Popper ZA, Fry SC (2005) Widespread occurrence of a covalent linkage between xyloglucan and acidic polysaccharides in suspension-cultured angiosperm cells. Ann Bot 96:91–99. doi:10.1093/aob/mci153

32. Bunzel M, Ralph J, Steinhart H (2005) Association of non-starch polysaccharides and ferulic acid in grain amaranth (Amaranthus caudatus L.) dietary fiber. Mol Nutr Food Res 49:551–559. doi:10.1002/mnfr.200500030

33. Meshitsuka G, Lee ZZ, Nakano J, Eda S (1982) Studies on the nature of lignin - carbohydrate bonding. J Wood Chem Technol 2:251–267. doi:10.1080/02773818208085134

34. Edwards MC, Williams T, Pattathil S, Hahn MG, Doran-Peterson J (2014) Replacing a suite of commercial pectinases with a single enzyme, pectate lyase B, in Saccharomyces cerevisiae fermentations of cull peaches. J Ind Microbiol Biotechnol 41:679–686. doi:10.1007/s10295-013-1394-z

35. Lionetti V, Francocci F, Ferrari S, Volpi C, Bellincampi D, Galletti R, D'Ovidio R, De Lorenzo G, Cervone F (2010) Engineering the cell wall by reducing de-methyl-esterified homogalacturonan improves saccharification of plant tissues for bioconversion. Proc Natl Acad Sci 107:616–621. doi:10.1073/pnas.0907549107

36. Tomassetti S, Pontiggia D, Verrascina I, Reca IB, Francocci F, Salvi G, Cervone F, Ferrari S (2015) Controlled expression of pectic enzymes in Arabidopsis thaliana enhances biomass conversion without adverse effects on growth. Phytochemistry 112:221–230. doi:10.1016/j.phytochem.2014.08.026

37. Biswal AK, Soeno K, Gandla M, Immerzeel P, Pattathil S, Lucenius J, Serimaa R, Hahn MG, Moritz T, Jönsson LJ, Israelsson-Nordström M, Mellerowicz EJ (2014) Aspen pectate lyase PtxtPL1-27 mobilizes matrix polysaccharides from woody tissues and improves saccharification yield. Biotechnol Biofuels 7:11. doi:10.1186/1754-6834-7-11

38. Biswal AK, Hao Z, Pattathil S, Yang X, Winkeler K, Collins C, Mohanty SS, Richardson EA, Gelineo-Albersheim I, Hunt K, Ryno D, Sykes RW, Turner GB, Ziebell A, Gjersing E, Lukowitz W, Davis MF, Decker SR, Hahn MG, Mohnen D (2015) Downregulation of GAUT12 in Populus deltoides by RNA silencing results in reduced recalcitrance, increased growth and reduced xylan and pectin in a woody biofuel feedstock. Biotechnol Biofuels 8:41. doi:10.1186/s13068-015-0218-y

39. Biswal AK, Atmodjo MA, Li M, Baxter HL, Yoo CG, Pu Y, Lee Y-C, Mazarei M, Black IM, Zhang J-Y, Ramanna H, Bray AL, King ZR, LaFayette PR, Pattathil S, Donohoe BS, Mohanty SS, Ryno D, Yee K, Thompson OA, Rodriguez M, Dumitrache A, Natzke J, Winkeler K, Collins C, Yang X, Tan L, Sykes RW, Gjersing EL, Ziebell A, Turner GB, Decker SR, Hahn MG, Davison BH, Udvardi MK, Mielenz JR, Davis MF, Nelson RS, Parrott WA, Ragauskas AJ, Neal Stewart C, Mohnen D (2018) Sugar release and growth of biofuel crops are improved by downregulation of pectin biosynthesis. Nat Biotechnol. doi: 10.1038/nbt.4067

40. Francocci F, Bastianelli E, Lionetti V, Ferrari S, De Lorenzo G, Bellincampi D, Cervone F (2013) Analysis of pectin mutants and natural accessions of Arabidopsis highlights the impact of de-methyl-esterified homogalacturonan on tissue saccharification. Biotechnol Biofuels 6:163. doi:10.1186/1754-6834-6-163

41. Chung D, Pattathil S, Biswal AK, Hahn MG, Mohnen D, Westpheling J (2014) Deletion of a gene cluster encoding pectin degrading enzymes in Caldicellulosiruptor bescii reveals an important role for pectin in plant biomass recalcitrance. Biotechnol Biofuels 7:147. doi:10.1186/s13068-014-0147-1

42. Davin LB, Patten AM, Lewis NG (2008) Lignins: a twenty-first century challenge. In: Himmel ME (ed) Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy. Blackwell, pp 213–305

43. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:490–495. doi:10.1093/nar/gkt1178

44. Boraston AB, Bolam DN, Gilbert HJ, Davies GJ (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem J 382:769–781. doi:10.1042/BJ20040892

45. Gilbert HJ (2010) The biochemistry and structural biology of plant cell wall deconstruction. Plant Physiol 153:444–455. doi:10.1104/pp.110.156646

46. Chundawat SPS, Beckham GT, Himmel ME, Dale BE (2011) Deconstruction of lignocellulosic biomass to fuels and chemicals. Annu Rev Chem Biomol Eng 2:121–145. doi:10.1146/annurev-chembioeng-061010-114205

47. Koivula A, Kinnari T, Harjunpaa V, Ruohonen L, Teleman A, Drakenberg T, Rouvinen J, Jones TA, Teeri TT (1998) Tryptophan 272: an essential determinant of crystalline cellulose degradation by Trichoderma reesei cellobiohydrolase Cel6A. FEBS Lett 429:341–346

48. Tolonen AC, Chilaka AC, Church GM (2009) Targeted gene inactivation in Clostridium phytofermentans shows that cellulose degradation requires the family 9 hydrolase Cphy3367. Mol Microbiol 74:1300–1313. doi:10.1111/j.1365-2958.2009.06890.x

49. Young J, Chung D, Bomble YJ, Himmel ME, Westpheling J (2014) Deletion of Caldicellulosiruptor bescii CelA reveals its crucial role in the deconstruction of lignocellulosic biomass. Biotechnol Biofuels 7:142. doi:10.1186/s13068-014-0142-6

50. Brunecky R, Donohoe BS, Yarbrough JM, Mittal A, Scott BR, Ding H, Taylor Ii LE, Russell JF, Chung D, Westpheling J, Teter SA, Himmel ME, Bomble YJ (2017) The multi domain Caldicellulosiruptor bescii CelA cellulase excels at the hydrolysis of crystalline cellulose. Sci Rep 7:9622. doi:10.1038/s41598-017-08985-w

51. Gilbert HJ (2014) Developing novel enzyme repertoires for the efficient deconstruction of plant biomass tailored for the bioenergy industry. In: McCann MC, Buckeridge MS, Carpita NC (eds) Plants and BioEnergy. Springer, New York, NY, pp 197–209

52. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. Biotechnol Biofuels 6:41. doi:10.1186/1754-6834-6-41

53. Forsberg Z, Vaaje-Kolstad G, Westereng B, Bunaes AC, Stenstrom Y, MacKenzie A, Sorlie M, Horn SJ, Eijsink VGH (2011) Cleavage of cellulose by a CBM33 protein. Protein Sci 20:1479–1483. doi:10.1002/pro.689

54. Kashyap DR, Vohra PK, Chopra S, Tewari R (2001) Applications of pectinases in the commercial sector: a review. Bioresour Technol 77:215–227. doi:10.1016/S0960-8524(00)00118-8

55. Hoondal G, Tiwari R, Tewari R, Dahiya N, Beg Q (2002) Microbial alkaline pectinases and their industrial applications: a review. Appl Microbiol Biotechnol 59:409–418. doi:10.1007/s00253-002-1061-1

56. Pedrolli DB, Monteiro AC, Gomes E, Carmona EC (2009) Pectin and pectinases: production, characterization and industrial application of microbial pectinolytic enzymes. Open Biotechnol J 3:9–18. doi:10.2174/1874070700903010009

57. Silva IR, Jers C, Meyer AS, Mikkelsen JD (2015) Rhamnogalacturonan I modifying enzymes: an update. N Biotechnol 33:41–54. doi:10.1016/j.nbt.2015.07.008

58. Ndeh D, Rogowski A, Cartmell A, Luis AS, Baslé A, Gray J, Venditto I, Briggs J, Zhang X, Labourel A, Terrapon N, Buffetto F, Nepogodiev S, Xiao Y, Field RA, Zhu Y, O'Neill MA, Urbanowicz BR, York WS, Davies GJ, Abbott DW, Ralet M-C, Martens EC, Henrissat B, Gilbert HJ (2017) Complex pectin metabolism by gut bacteria reveals novel catalytic functions. Nature 544:65–70. doi:10.1038/nature21725

59. Jayani RS, Saxena S, Gupta R (2005) Microbial pectinolytic enzymes: a review. Process Biochem 40:2931–2944. doi:10.1016/j.procbio.2005.03.026

60. Kohli P, Gupta R (2015) Alkaline pectinases: a review. Biocatal Agric Biotechnol 4:279–285. doi:10.1016/j.bcab.2015.07.001

61. Bonnin E, Garnier C, Ralet MC (2014) Pectin-modifying enzymes and pectin-derived materials: applications and impacts. Appl Microbiol Biotechnol 98:519–532. doi:10.1007/s00253-013-5388-6

62. Pedrolli DB, Carmona EC (2014) Purification and characterization of a unique pectin lyase from Aspergillus giganteus able to release unsaturated monogalacturonate during pectin degradation. Enzyme Res 2014:353915. doi:10.1155/2014/353915

63. Laurent F, Kotoujansky A, Labesse G, Bertheau Y (1993) Characterization and overexpression of the pem gene encoding pectin methylesterase of Erwinia chrysanthemi strain 3937. Gene 131:17–25. doi:https://doi.org/10.1016/0378-1119(93)90664-O

64. Shevchik VE, Condemine G, Hugouvieux-Cotte-Pattat N, Robert-Baudouy J (1996) Characterization of pectin methylesterase B, an outer membrane lipoprotein of Erwinia chrysanthemi 3937. Mol Microbiol 19:455–466. doi:10.1046/j.1365-2958.1996.389922.x

65. Khanh NQ, Ruttkowski E, Leidinger K, Albrecht H, Gottschalk M (1991) Characterization and expression of a genomic pectin methyl esterase-encoding gene in Aspergillus niger. Gene 106:71–77. doi:https://doi.org/10.1016/0378-1119(91)90567-U

66. Kitamoto N, Okada H, Yoshino S, Ohmiya K, Tsukagoshi N (1999) Pectin methylesterase gene (pmeA) from Aspergillus oryzae KBN616: its sequence analysis and overexpression, and characterization of the gene product. Biosci Biotechnol Biochem 63:120–124

67. Christgau S, Kofod L V, Halkier T, Andersen LN, Hockauf M, Dorreich K, Dalboge H, Kauppinen S (1996) Pectin methyl esterase from Aspergillus aculeatus: expression cloning in yeast and characterization of the recombinant enzyme. Biochem J 319 (Pt 3:705–712. doi:10.1042/bj3190705

68. Remoroza C, Wagenknecht M, Buchholt HC, Moerschbacher BM, Gruppen H, Schols HA (2015) Mode of action of Bacillus licheniformis pectin methylesterase on highly methylesterified and acetylated pectins. Carbohydr Polym 115:540–550. doi:10.1016/j.carbpol.2014.09.016

69. Shevchik VE, Hugouvieux-Cotte-Pattat N (1997) Identification of a bacterial pectin acetyl esterase in Erwinia chrysanthemi 3937. Mol Microbiol 24:1285–1301

70. Shevchik VE, Hugouvieux-Cotte-Pattat N (2003) PaeX, a second pectin acetylesterase of Erwinia chrysanthemi 3937. J Bacteriol 185:3091–100

71. Bolvig PU, Pauly M, Orfila C, Scheller HV, Schnorr K (2003) Sequence analysis and characterisation of a novel pectin acetyl esterase from Bacillus subtilis. In: Voragen F, Schols H, Visser R (eds) Advances in Pectin and Pectinase Research. Springer Netherlands, Dordrecht, pp 315–330

72. Remoroza C, Wagenknecht M, Gu F, Buchholt HC, Moerschbacher BM, Schols HA, Gruppen H (2014) A Bacillus licheniformis pectin acetylesterase is specific for homogalacturonans acetylated at O-3. Carbohydr Polym 107:85–93. doi:10.1016/j.carbpol.2014.02.006

73. Searle-van Leeuwen MJF, Vincken J-P, Schipper D, Voragen AGJ, Beldman G (1996) Acetyl esterases of Aspergillus niger: purification and mode of action on pectins. In: Visser J, Voragen AGJ (eds) Pectins and Pectinases. Elsevier, pp 793–798

74. Bonnin E, Clavurier K, Daniel S, Kauppinen S, Mikkelsen JDM, Thibault J-F (2008) Pectin acetylesterases from Aspergillus are able to deacetylate homogalacturonan as well as rhamnogalacturonan. Carbohydr Polym 74:411–418. doi:10.1016/j.carbpol.2008.03.014

75. Abbott DW, Boraston AB (2008) Structural biology of pectin degradation by Enterobacteriaceae. Microbiol Mol Biol Rev 72:301–316. doi:10.1128/MMBR.00038-07

76. Luis AS, Briggs J, Zhang X, Farnell B, Ndeh D, Labourel A, Baslé A, Cartmell A, Terrapon N, Stott K, Lowe EC, McLean R, Shearer K, Schückel J, Venditto I, Ralet M-C, Henrissat B, Martens EC, Mosimann SC, Abbott DW, Gilbert HJ (2017) Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic Bacteroides. Nat Microbiol 3:210. doi:10.1038/s41564-017-0079-1

77. Chatterjee A, McEvoy JL, Chambost JP, Blasco F, Chatterjee AK (1991) Nucleotide sequence and molecular characterization of pnlA, the structural gene for damage-inducible pectin lyase of Erwinia carotovora subsp. carotovora 71. J Bacteriol 173:1765–1769. doi:10.1128/jb.173.5.1765-1769.1991

78. Martens-Uzunova ES, Zandleven JS, Benen JAE, Awad H, Kools HJ, Beldman G, Voragen AGJ, Van Den Berg JA, Schaap PJ (2006) A new group of exo-acting family 28 glycoside hydrolases of Aspergillus niger that are involved in pectin degradation. Biochem. J. 400:43–52

79. Shevchik VE, Condemine G, Robert-Baudouy J, Hugouvieux-Cotte-Pattat N (1999) The exopolygalacturonate lyase PelW and the oligogalacturonate lyase Ogl, two cytoplasmic enzymes of pectin catabolism in Erwinia chrysanthemi 3937. J Bacteriol 181:3912–3919

80. Nasser W, Chalet F, Robert-Baudouy J (1990) Purification and characterization of extracellular pectate lyase from Bacillus subtilis. Biochimie 72:689–695

81. Soriano M, Diaz P, Pastor FIJ (2006) Pectate lyase C from Bacillus subtilis: a novel endo-cleaving enzyme with activity on highly methylated pectin. Microbiology 152:617–25. doi:10.1099/mic.0.28562-0

82. Thompson J, Pikis A, Rich J, Hall BG, Withers SG (2013) α-Galacturonidase(s): a new class of Family 4 glycoside hydrolases with strict specificity and a unique CHEV active site motif. FEBS Lett 587:799–803. doi:10.1016/j.febslet.2013.02.004

83. Searle-van Leeuwen MJF, van den Broek LAM, Schols HA, Beldman G, Voragen AGJ (1992) Rhamnogalacturonan acetylesterase: a novel enzyme from Aspergillus aculeatus, specific for the deacetylation of hairy (ramified) regions of pectins. Appl Microbiol Biotechnol 38:347–349. doi:10.1007/BF00170084

84. Kauppinen S, Christgau S, Kofod L V, Halkier T, Dorreich K, Dalboge H (1995) Molecular cloning and characterization of a rhamnogalacturonan acetylesterase from Aspergillus aculeatus. Synergism between rhamnogalacturonan degrading enzymes. J Biol Chem 270:27172–27178

85. Mølgaard A, Kauppinen S, Larsen S (2000) Rhamnogalacturonan acetylesterase elucidates the structure and function of a new family of hydrolases. Structure 8:373–383. doi:10.1016/S0969-2126(00)00118-0

86. Martinez-Martinez I, Navarro-Fernandez J, Daniel Lozada-Ramirez J, Garcia-Carmona F, Sanchez-Ferrer A (2008) YesT: a new rhamnogalacturonan acetyl esterase from Bacillus subtilis. Proteins 71:379–388. doi:10.1002/prot.21705

87. Navarro-Fernandez J, Martinez-Martinez I, Montoro-Garcia S, Garcia-Carmona F, Takami H, Sanchez-Ferrer A (2008) Characterization of a new rhamnogalacturonan acetyl esterase from Bacillus halodurans C-125 with a new putative carbohydrate binding domain. J Bacteriol 190:1375–1382. doi:10.1128/JB.01104-07

88. Itoh T, Ochiai A, Mikami B, Hashimoto W, Murata K (2006) A novel glycoside hydrolase family 105: the structure of family 105 unsaturated rhamnogalacturonyl hydrolase complexed with a disaccharide in comparison with family 88 enzyme complexed with the disaccharide. J Mol Biol 360:573–585. doi:10.1016/j.jmb.2006.04.047

89. Ochiai A, Itoh T, Kawamata A, Hashimoto W, Murata K (2007) Plant cell wall degradation by saprophytic Bacillus subtilis strains: gene clusters responsible for rhamnogalacturonan depolymerization. Appl Environ Microbiol 73:3803–13. doi:10.1128/AEM.00147-07

90. Iwai M, Kawakami T, Ikemoto T, Fujiwara D, Takenaka S, Nakazawa M, Ueda M, Sakamoto T (2015) Molecular characterization of a Penicillium chrysogenum exo-rhamnogalacturonan lyase that is structurally distinct from other polysaccharide lyase family proteins. Appl Microbiol Biotechnol 99:8515–8525. doi:10.1007/s00253-015-6600-7

91. Laatu M, Condemine G (2003) Rhamnogalacturonate lyase RhiE is secreted by the out system in Erwinia chrysanthemi. J Bacteriol 185:1642–1649. doi:10.1128/jb.185.5.1642-1649.2003

92. Latarullo MBG, Tavares EQP, Maldonado GP, Leite DCC, Buckeridge MS (2016) Pectins, endopolygalacturonases, and bioenergy. Front Plant Sci 7:1–7. doi:10.3389/fpls.2016.01401

93. Schneider DM, Henriksen ED, Boland WE, Edwards MC, Doran-Peterson J (2014) Tipula abdominalis: a natural biorefinery with novel microbial enzymes useful for pectin-rich biomass deconstruction. In: Sun J, Ding S-Y, Doran-Peterson J (eds) Biological Conversion of Biomass for Fuels and Chemicals: Exploration from Natural Utilization Systems. The Royal Society of Chemistry, Cambridge, pp 282–297

94. Sinsabaugh RL, Linkins AE, Benfield EF (1985) Cellulose digestion and assimilation by three leaf-shredding aquatic insects. Ecology 66:1464–1471

95. Cook DM, Henriksen ED, Upchurch R, Doran-Peterson JB (2007) Isolation of polymer-degrading bacteria and characterization of the hindgut bacterial community from the detritus-feeding larvae of Tipula abdominalis (Diptera: Tipulidae). Appl Environ Microbiol 73:5683–6. doi:10.1128/AEM.00213-07

96. Cook DM, Doran-Peterson J (2010) Mining diversity of the natural biorefinery housed within Tipula abdominalis larvae for use in an industrial biorefinery for production of lignocellulosic ethanol. Insect Sci 17:303–312. doi:10.1111/j.1744-7917.2010.01343.x

97. Boland WE, DeCrescenzo Henriksen E, Doran-Peterson J (2010) Characterization of two Paenibacillus amylolyticus strain 27C64 pectate lyases with activity on highly methylated pectin. Appl Environ Microbiol 76:6006–6009. doi:10.1128/AEM.00043-10

98. Takeda M, Iohara K, Shinmaru S, Suzuki I, Koizumi J-I (2000) Purification and properties of an enzyme capable of degrading the sheath of Sphaerotilus natans. Appl Environ Microbiol 66:4998–5004. doi:10.1128/AEM.66.11.4998-5004.2000.Updated

99. Soriano M, Blanco A, Díaz P, Pastor FIJ (2000) An unusual pectate lyase from a Bacillus sp. with high activity on pectin: Cloning and characterization. Microbiology 146:89–95

100. Ying XX, Chen LN, Yu ML, Xue Q, Wang Z (2012) Purification and characterization of a bioscouring pectate lyase from Paenibacillus sp. WZ008 with high activity on pectin. Adv Mater Res 441:457–461. doi:10.4028/www.scientific.net/AMR.441.457

101. Ko C-H, Tsai C-H, Tu J, Tang S-H, Liu C-C (2011) Expression and thermostability of Paenibacillus campinasensis BL11 pectate lyase and its applications in bast fibre processing. Ann Appl Biol 158:218–225. doi:10.1111/j.1744-7348.2010.00456.x

102. Sakiyama CCH, Paula EM, Pereira PC, Borges AC, Silva DO (2001) Characterization of pectin lyase produced by an endophytic strain isolated from coffee cherries. Lett Appl Microbiol 33:117–121. doi:10.1046/j.1472-765x.2001.00961.x

103. Li X, Wang H, Zhou C, Ma Y, Li J, Song J (2014) Cloning, expression and characterization of a pectate lyase from Paenibacillus sp. 0602 in recombinant Escherichia coli. BMC Biotechnol 14:18. doi:10.1186/1472-6750-14-18

CHAPTER 2

WHOLE GENOME SEQUENCE OF *PAENIBACILLUS AMYLOLYTICUS* 27C64 REVEALS A DIVERSE SET OF CARBOHYDRATE-ACTIVE ENZYMES AND COMPLETE PECTIN DECONSTRUCTION SYSTEM [1]

[1] Keggi, C and J Doran-Peterson. 2019. *Journal of Industrial Microbiology and Biotechnology*. 46:1-11.

2.1 ABSTRACT

A draft genome of *Paenibacillus amylolyticus* 27C64 was assembled and a total of 314 putative CAZymes in 108 different families were identified. Comparison to well-studied polysaccharide-degrading organisms revealed that *P. amylolyticus* 27C64 has as many or more putative CAZymes than most of these organisms. Four different pectic substrates and xylan supported growth but cellulose was not utilized. Measurement of enzyme activities in culture supernatants revealed low levels of cellulase activity, high levels of xylanase activity, and pectinase activities that adapted to the specific polysaccharides provided. Relative expression levels of each putative pectinase in cells grown with and without three different pectic substrates were evaluated with RT-qPCR and distinct sets of genes upregulated in response to homogalacturonan, methylated homogalacturonan, and rhamnogalacturonan I were identified. It is also noted that this organism's pectinolytic system differs from other well-studied systems and contains enzymes which are of value for further study.

2.2 INTRODUCTION

Plant cell wall polysaccharides have the potential to be a cheap and ubiquitous carbon source for the production of biofuels and chemicals. Pectin is the most structurally complex of these cell wall polysaccharides and is present in all parts of the cell wall but is most prevalent in the outermost layers [1]. The two most abundant types of pectin are homogalacturonan (HG) and rhamnogalacturonan I (RG-I). HG is comprised of an α-1,4 linked galacturonic acid (GalA) backbone that may be methylated (sometimes extensively) or acetylated but does not have side chains. In contrast, RG-I has a backbone of GalA–

rhamnose (Rha) disaccharide repeats, has variable arabinose and galactose-rich side chains, and may also be acetylated. Other types of pectic polysaccharides are less abundant than HG and RG-I but are based on the HG-type backbone [1]. The enzymes which deconstruct these polysaccharides, collectively referred to as pectinases, are just as complex as the substrates they work on and are found in all domains of life. Pectinases are broadly divided into two groups: esterases and depolymerases. Esterases remove methyl and acetyl groups decorating the backbone while depolymerases break glycosidic bonds. Depolymerases are further divided based on whether they rely on a hydrolytic mechanism (hydrolases) or β-elimination (lyases), their substrate specificity, and primary products [2]. These enzymes have numerous applications. The use of acidic fungal pectinases to improve fruit juice extraction yields and to clarify the juice is well established [2, 3]. Alkaline pectinases can be used to process plant fibers for the textiles industry, to improve paper production by acting as a biobleaching agent, or to reduce the cationic demand of paper pulp. These alkaline enzymes are also useful in various food processes such as fermenting coffee beans and tea leaves or extracting certain vegetable oils [4, 5].

Of particular interest is the use of pectinases for the production of biofuels and chemicals. Concerns over using edible crops for fuel production, land-use changes from energy crop production, and challenges associated with using woody feedstocks such as the production of compounds inhibitory to fermenting organisms [6] make the use of minimally lignified agricultural waste products attractive. Many such waste products including sugar beet pulp, citrus pulp and peels, and apple pomace are viable feedstocks for ethanol production but are pectin rich and require pectinases for efficient saccharification [7].

Despite the associated challenges, the use of low-pectin lignocellulosic energy crops such as poplar and switchgrass will likely still be necessary for large scale biofuel production. Unfortunately, efficient and cost-effective enzymatic saccharification of lignocellulosic substrates to fermentable monosaccharides remains a key challenge to their widespread use. In recent years, substantial progress has been made in improving biochemical conversion of lignocellulosic substrates but most of this work has focused on understanding the roles that cellulose, hemicelluloses, and lignin play in biomass recalcitrance [8]. The role of pectin has been largely ignored because of its lower abundance in these substrates. However, recent work has demonstrated that pectin also plays a key role in the recalcitrance of lignocellulosic substrates. For example, expression of enzymes which reduce the amount of demethylated homogalacturonan (HG) in *Arabidopsis thaliana*, tobacco, or wheat improved saccharification of these plants and reduced the need for pretreatment [9, 10]. Improved saccharification has also been demonstrated in woody aspen biomass when a native pectate lyase was overexpressed [11]. Likewise, knockouts or knockdowns of genes involved in pectin biosynthesis in *A. thaliana* [12] or *Populus deltoides* [13] have improved saccharification efficiency. Modification of a biocatalyst has also demonstrated the importance of pectin to biomass recalcitrance: deletion of a pectinase gene cluster in *Caldicellulosiruptor bescii,* a thermophilic anaerobe capable of growth on unpretreated biomass, results in a growth defect on whole plant substrates [14]. This recent evidence that pectin is a barrier to efficient saccharification agrees with older work showing that pectinases are important virulence factors for plant pathogens [15].

Despite these developments, there is still much to learn about pectinases. Although broadly distributed across all domains of the tree of life, subsets of pectinases like those associated with plant pathogens [16], involved in deconstruction of plant material in the human gut [17, 18], and fungal polygalacturonases useful for fruit juice clarification [3] have been studied intensively, and the full diversity of pectin deconstructing enzymes has yet to be explored. In particular, RG-I deconstruction is less well studied than HG deconstruction. A 2015 survey of RG-I deconstructing enzymes was only able to identify 23 characterized examples in the literature [19]. New catalytic activities responsible for pectin deconstruction are still being discovered including one family established in 2006 with the discovery of two novel enzymes from *Bacillus subtilis* [20] and more recently the discovery of several novel enzymes responsible for rhamnogalacturonan II (RG-II) deconstruction in *Bacteroides thetaiotaomicron* [18]. An examination of sequences in the Carbohydrate-Active Enzyme Database (CAZy, http://www.cazy.org/) [21] reveals that while some pectinase families (GH28, PL1) have thousands of sequences identified and more than a hundred listed as characterized, the majority of the remaining families (PL2, PL9, PL10, PL11, PL22, GH105) have ten or fewer characterized sequences listed (PL3 has 24). Even in the well-studied GH28 family, a recent article has highlighted the fact that most of these sequences belong to the genus *Aspergillus* or family *Enterobacteriaceae* and highlights the need to study a more diverse set [22].

One organism with the potential to contribute to understanding of pectinolytic enzymes is *Paenibacillus amylolyticus* 27C64. Originally isolated from the hindgut of an aquatic cranefly (*Tipula abdominalis*) larvae, it is a member of a microbial consortium responsible for the saccharification of conditioned leaf litter [23]. Of the organisms isolated

from this gut system, this Gram-positive bacterium displayed the widest array of plant cell wall deconstruction capabilities including cellulose, hemicellulose, and pectinase activities [24]. Two pectate lyases identified through a genomic library screen, PelA and PelB, have demonstrated broad substrate specificities [25] and PelB was able to replace a commercial pectinase mixture in fermentations of pectin-rich cull peaches [26]. Additionally, *P. amylolyticus* is taxonomically divergent from phylogenetic groups whose pectinolytic systems have been studied in detail.

2.3 METHODS

2.3.1 Media and Culture Conditions

P. amylolyticus was grown aerobically in tryptic soy broth without dextrose (TSB; 17 g pancreatic digest of casein, 3 g enzymatic digest of soybean meal, 5 g NaCl, 2.5 g dipotassium phosphate per liter) at 37°C with shaking unless stated otherwise. Tryptic soy agar (15 g pancreatic digest of casein, 5 g papaic digest of soybean meal, 5 g NaCl, 15 g agar per liter) was used for solid media. All media were purchased from Becton-Dickinson (Franklin Lakes, NJ). Polygalacturonic acid (PGA), citrus pectin, apple pectin, galacturonic acid (GalA), glucose, crystalline cellulose (Sigmacell), oat spelt xylan, and $CaCl_2$ were purchased from Sigma-Aldrich (St. Louis, MO). RG-I from soy was purchased from Megazyme (Bray, County Wicklow, Ireland).

2.3.2 Genome Sequencing, Assembly, and Annotation

Total genomic DNA was isolated from *P. amylolyticus* 27C64 using a Promega Wizard Genomic DNA Purification Kit (Madison, WI). Library preparation and sequencing was performed by the Georgia Genomics Facility (Athens, GA) using the

Illumina NextSeq 500 platform (San Diego, CA) with paired-end 150 base pair reads. The reads were quality trimmed using Trimmomatic [27] and assembled with the SPAdes assembler [28] by the Quantitative Biology Consulting Group (QBCG) at the University of Georgia (Athens, GA). The draft genome was filtered to exclude contigs with an average read coverage less than two or which returned no hits in a BLAST+ search against NCBI's non-redundant nucleotide database [29]. The assembled draft genome was annotated with Rapid Annotation using Subsystem Technology (RAST) [30, 31].

Pairwise average nucleotide identity (ANI) between this draft genome and every other *Paenibacillus* genome available through the Joint Genome Institute's Integrated Microbial Genomics (IMG) website (https://img.jgi.doe.gov/) [32] was calculated using their pairwise ANI tool [33].

2.3.3 Carbohydrate-Active Enzyme Identification and Annotation

Carbohydrate-active enzymes (CAZymes) were identified using two tools: the CAZy Analysis Toolkit (CAT) [34] and the Database for Carbohydrate-Active Enzyme Annotation (dbCAN) [35]. Both tools identify CAZymes and assign them to a family in the Carbohydrate-Active Enzyme (CAZy; http://www.cazy.org/) database [21].

To identify putative pectinases, an E-value cutoff of 0.01 was used for CAT with default cutoffs for dbCAN and the merged results were filtered to exclude all CAZy families not known to contain pectinases. Genes on this narrowed list were manually reviewed based on evidence from BLAST+ searches [29] against the CAZy database and NCBI non-redundant protein database, by evaluating conserved protein domains with CD-Search [36], and by performing multiple sequence alignments against characterized members of each CAZy family using Multiple Sequence Comparison by Log-Expectation

(MUSCLE) [37]. Initial putative functions were assigned based on the list of known activities in the CAZy family to which they belong. Subcellular localization of each putative pectinase was predicted with SignalP 4.0 [38], PSORTb [39], and Phobius [40].

Only dbCAN results were used to generate a summary of CAZymes in the genome because of the relatively low false-positive rate. Genes that did not match a catalytic domain and had only a carbohydrate binding domain (CBM) or S-layer homology domain (SLH) were excluded. Any hits to CE10 family were also excluded since this family is now defunct.

2.3.4 Growth Curves

TSB was made per the manufacturer's instructions and MOPS minimal media was made as previously described [41]. Both types of media were supplemented with a final concentration of 0.1 mM $CaCl_2$, adjusted to a final pH of 7.0, and filter sterilized through a 0.22 μM PES membrane (Stericup, EMD Millipore, Burlington, MA). All carbohydrates were added to a final concentration of 0.4% w/v, the substrate concentration that was identified as optimal for PGA. For growth curves with soluble pectic substrates, five milliliters of each medium evaluated was inoculated to an OD_{600} of 0.01 with a washed *P. amylolyticus* cell suspension. Cell suspensions were prepared by pelleting 1 mL of each of four overnight cultures, washing once in sterile saline, and resuspending in 500 μL of saline. The OD_{600} of each resuspension was measured in duplicate. Three hundred microliters of inoculated media were dispensed into each of nine wells with one uninoculated blank well and growth at 37°C with constant maximal shaking was monitored at 600 nm using a Bioscreen C (Growth Curves USA, Piscataway, NJ). Growth in each well was normalized with the appropriate blank, measurements of technical replicates were

averaged, and mean OD and standard deviation were calculated for the four biological replicates. After growth, technical replicates were pooled, mixed, pelleted, and cell-free supernatant was harvested for enzyme assays.

Media with insoluble substrates (cellulose and xylan) were prepared by autoclaving the appropriate amount of each substrate in 12.5 mL of distilled water in a 125-mL shake flask then adding 12.5 mL of 2X TSB or 2X MOPS minimal media prepared as for the soluble substrates but in half the final volume. Inoculation to an initial OD of 0.01 was performed exactly as above. Growth was monitored by spread plating cells diluted into sterile saline onto TSA. Colonies were counted after incubating overnight at 37°C.

2.3.5 Measurement of Enzyme Activities

Pectate lyase activity was measured as previously described [42]. Briefly, 5 µL of culture supernatant was added to 1 mL of substrate solution (100 mM Tris-HCl pH 8.5, 0.2% w/v PGA, 0.5 mM $CaCl_2$) which had been pre-equilibrated to 25°C and mixed. The absorbance at 232 nm was monitored for two minutes and the slope was used with the published molar extinction coefficient to calculate product formation. Activity is reported in international units of enzyme activity (IU) per mL (1 unit = 1 µmol product formed in 1 minute). Pectin lyase assays were conducted similarly but with a different substrate solution (100 mM Tris-HCl, pH 8.0, 0.2% w/v 85% methoxy citrus pectin, 1 mM EDTA) as previously described [42]. Rhamnogalacturonan lyase activity was measured with this assay as well using a RG-I substrate solution (50 mM Tris-HCl, pH 7.5, 0.05% w/v RG-I from potato, 2 mM $CaCl_2$) and one unit of activity was defined as the amount of enzyme that increased absorbance by 1.0 in one minute [43]. RG-I from potato was purchased from

Megazyme (County Wicklow, Ireland) and 85% methoxy citrus pectin was from Sigma (St. Louis, MO).

Xylanase activity was measured using a miniaturized version of the Nelson-Somogyi reducing sugar assay with the standard copper and arsenomolybdate reagents [44, 45]. For each assay 50 µL of xylan substrate solution (2% w/v oat spelt xylan solubilized overnight at 37°C with shaking in 20 mL of 0.2 N NaOH, 100 mM MOPS pH 7, 0.02% w/v sodium azide) was mixed with 50 µL of culture supernatant and incubated for 15 minutes at 40°C. Reactions were stopped with 200 µL of copper reagent, boiled for 10 minutes, cooled, and 200 µL of color reagent was added. Samples were diluted fivefold and the absorbance was measured at 500 nm. Sugar release was determined relative to a xylose standard curve. Enzyme activity was calculated in IU/mL after subtracting the values of substrate only and supernatant only blanks for each sample. Oat spelt xylan was purchased from Sigma (St. Louis, MO).

Cellulase activity was measured similarly except a carboxymethyl cellulose (CMC) substrate solution (1% w/v CMC, 100 mM MOPS, pH 7, 0.02% w/v sodium azide) was used, reactions were allowed to proceed for 4 hours, and glucose was used as the standard. CMC was purchased from Sigma (St. Louis, MO).

2.3.6 RT-qPCR

Total RNA was isolated from cells grown in TSB without pectin and in TSB supplemented with 0.4% w/v of either PGA, apple pectin, or RG-I from soy at 12.75 hours after inoculation with the ZR Fungal/Bacterial RNA Miniprep kit (Zymo Research, Irvine, CA). Ten µg of RNA was treated with DNase using the TURBO DNA-free Kit (Thermo Fisher, Waltham, MA). RNA was used as a PCR template to confirm that no genomic DNA

was remaining. One microgram of each sample was reverse transcribed with the iScript cDNA Synthesis Kit (Bio-Rad, Hercules, CA) and diluted 1:10 with nuclease-free water. Two microliters of diluted cDNA was used as template in 20-µL qPCR reactions with Luna Universal qPCR Master Mix (New England Biolabs, Ipswich, MA) and run on a StepOne Plus instrument (Applied Biosystems, Foster City, CA). All conditions were repeated with four biological replicates. Oligonucleotide primers were purchased from Integrated DNA Technologies (Coralville, IA).

Six potential reference genes (*ftsZ*, *rpoD*, *era*, *adk*, *gyrA*, and *gap* homologues) were evaluated and ranked using the geNorm algorithm [46] and two most stable genes, *ftsZ* and *rpoD,* were selected as endogenous control genes. Quantification thresholds and Cq values were calculated by the StepOne Plus software. Fold change in expression was calculated using the $2^{-\Delta\Delta Cq}$ method [47].

2.4 RESULTS AND DISCUSSION

2.4.1 Genome Sequencing, Assembly, and Annotation

The draft genome consisted of 103 contigs totaling 6.77 Mb in length with 5922 protein coding sequences and 93-fold average read coverage. The 20 largest contigs accounted for >90% of the total genome length while the largest 33 contigs account for 99% of the genome length (N50 = 325,177 bp, L50 = 7). The remaining small contigs accounting for ~ 1% of genome length were composed primarily of repetitive rRNA and tRNA elements.

Pairwise average nucleotide identity comparison of this draft genome against all 217 public *Paenibacillus* genomes available through IMG revealed no other genome above

the recommended species cutoff (ANI ≥ 96.5 and AF ≥ 0.6) [33]. The most closely related genome, *Paenibacillus phyllostachyicola* BL9 (*Paenibacillus polysaccharolyticus* BL9T) [48], had an ANI of 93.2% with an AF of 0.84. The ANI of the next most related genome, *Paenibacillus sp.* PAMC 26794, was much lower at 79.99% ANI and an AF value of 0.72. Interestingly, the only *P. amylolyticus* genome already published, *P. amylolyticus* Heshi-A3 [49], was the eighth most related with an ANI value of 79.82 and an AF of 0.70, well below the recommended speciation cutoff. These data suggest that this isolate is not the same species as any of the other IMG publicly available genomes.

2.4.2 CAZyme Identification and Annotation

Analysis of the whole genome for CAZymes revealed that 314 (5.3%) of the 5922 predicted coding sequences matched one or more CAZy family. These hits included matches to all five CAZy catalytic groups as well as to CBM domains with glycoside hydrolases being the most represented group (Figure 2.1a). These genes represented 110 different CAZy families including CBMs, 87 of which are catalytic families (Figure 2.1b). A total of 193 glycoside hydrolases (GH) or polysaccharide lyases (PL), catalytic activities responsible for deconstructing polysaccharides, were identified in 53 different GH families and 8 PL families (Figure 2.1b). Seven SLH domains, sometimes important for CAZyme attachment to the cell surface, are not a CAZy family but were also identified by dbCAN. Although the exact number of false positives or unidentified CAZymes within these results cannot be determined precisely, it has been previously demonstrated that dbCAN correctly identifies 99.3% of the true CAZymes in the well-studied *C. thermocellum* genome and that 89.4% of hits generated were true CAZymes [35]. A majority of the hits generated by dbCAN in this genome were very good matches: 227 or 72.3% of these 314 putative

CAZymes had an E-value less than 1 x 10^{-20} and 138 or 44.0% had an E-value less than 1 x 10^{-50}. This analysis validated previous culture-based work which showed that *P. amylolyticus* 27C64 has a diverse array of polysaccharide deconstruction capabilities [24]. In fact, *P. amylolyticus* 27C64 has two to three times as many predicted GH and PL enzymes as cellulolytic thermophiles such as *Caldicellulosiruptor bescii* and *Ruminiclostridium thermocellum* (*Clostridium thermocellum*) which have been of interest for the conversion of lignocellulosic material to biofuels (Figure 2.1c). It also has more predicted GH and PL enzymes than well-studied pectin degrading bacteria (*Dickeya chrysanthemi, Dickeya dadantii, Pectobacterium carotovorum*, and *Bacillus subtilis*). Within its own genus, it tops both *Paenibacillus polymyxa* and the most closely related genome *P. polysaccharolyticus*. Of all the bacterial genomes selected for comparison, only the human gut associated bacterium *Bacteroides thetaiotaomicron* had more predicted GH and PL enzymes.

To evaluate the extent to which the richness of putative CAZymes in the genome might be related to the deconstruction of plant cell wall polysaccharides specifically, the presence of putative cellulases, xylanases, and pectinases was examined. After removing any RAST hits to cellulases that were not assigned to a CAZy family by dbCAN, a total of three putative endoglucanases and six putative β-glucosidases were identified. Similarly, 17 endo-xylanases, six β-xylosidases, and two acetyl xylan esterases were identified. Pectinases were difficult to identify in the same manner because of their diversity, inaccurate assignment by RAST, and inconsistent naming in databases. Instead, hits to CAZy families relevant to pectin deconstruction were manually reviewed to identify 28 putative pectinases including a variety of lyases, hydrolases, and esterases related to both

HG and RG-I deconstruction. The presence of numerous putative plant cell wall polysaccharide deconstruction enzymes is consistent with this organism's role as a member of an insect hindgut community responsible for degrading conditioned leaf litter.

2.4.3 Growth on Plant Cell Wall Polysaccharides

To evaluate whether this CAZyme diversity allowed this organism to use plant cell wall polysaccharides as a carbon source, *P. amylolyticus* was grown with various polysaccharides as a sole and supplemental carbon source (Figure 2.2). Improved growth was observed when TSB was supplemented with each of four different pectins (Figure 2.1a) and xylan (Figure 2.1b) but not when crystalline cellulose was added. Only three of the four pectins could be utilized as a sole carbon source in MOPS minimal media, with PGA failing to support growth. Xylan also supported growth as a sole carbon source but cellulose did not.

At the end of each growth curve, culture supernatant was harvested and assayed for relevant enzymatic activity. Supernatant from TSB or MOPS cultures with glucose was also included in these assays. Since *P. amylolyticus* was unable to utilize cellulose for growth, it was unsurprising that only very low cellulase activity was observed overall. In contrast, strong xylanase activity was observed when *P. amylolyticus* was grown in TSB, in TSB with xylan, and in MOPS with xylan. Glucose completely repressed xylanase production in both media backgrounds.

Similarly, all three pectinase activities measured (pectate lyase, pectin lyase, and RG lyase) were completely repressed by glucose. Pectate lyases, which cleave demethylated HG, were expressed at high levels when RG-I, citrus pectin, or apple pectin were supplied for growth in either TSB or MOPS. The much lower amount of activity

observed after growth in TSB with PGA was unexpected since pectate lyases would be primarily responsible for deconstruction of PGA in an organism without extracellular polygalacturonases. This suggests that a pectin degradation product responsible for inducing these pectate lyases is absent from the homogeneous PGA substrate. This low level of induction likely explains the lack of growth on PGA as a sole carbon source. RG lyase activity could be observed when RG-I was supplied for growth or when whole citrus and apple pectins (which include some RG-I) were present. PGA did not induce these enzymes. In contrast to the pectate and RG lyase activities, pectin lyase activity appeared at similar levels in the TSB control condition as it did in all other conditions that allowed for growth except TSB with glucose. Since pectin lyases cleave only methylated pectin, it is possible that this activity is responsible for initial release of methylated fragments from the web of cell wall polysaccharides and is constitutively expressed. It is likely that this activity, along with very low levels of pectate or RG lyase activities, releases the degradation products that induce the rest of the system. The fact that this organism's deconstruction system adapts to various pectic polysaccharides indicates that different parts of the pectinolytic system are responding to different inducers.

2.4.4 Differential Expression of Putative Pectinases

The number and diversity of putative pectinases made this system in particular seem of value for further study, but more information was necessary to identify the role of each gene in the overall system. The adaptability of the pectinolytic system to different substrates made it possible to better define the role of each individual element by comparing the relative expression of each gene when cells were grown with and without different pectic polysaccharides. RNA was isolated from cells grown in PGA and RG-I

from soy because they should reliably induce the HG and RG-I deconstruction systems specifically. Apple pectin was also included because the HG it contains is partially methylated, unlike PGA.

In total, 18 of the 28 putative pectinases were upregulated on one or more pectic substrates (Figure 2.3). The pectin lyase, which appeared to be constitutively expressed in the growth experiments, was among the ten genes not upregulated on any substrate but is included in the putative pectinolytic system. So too is the putative pectin acetylesterase which cannot be excluded by RT-qPCR alone since none of the substrates used in this experiment had highly acetylated HG. The other eight genes not upregulated in any condition are likely to have some other activity unrelated to pectin deconstruction (Appendix A, Figure A1). Growth with PGA largely resulted in several expected changes: three extracellular pectate lyases (*pamy_4343, pamy_2972, pamy_1763*) and one cytoplasmic polygalacturonase (*pamy_82*) were upregulated. The isomerases responsible for initiating catabolism of saturated and unsaturated GalA were also upregulated (*pamy_4442, pamy_520*). The increased expression of one GH105 family hydrolase (*pamy_1066*) on PGA suggests that this enzyme specifically removes an unsaturated GalA from the end of HG oligosaccharides instead of from RG-I fragments, an activity which has only recently been demonstrated within that family [17].

Growth on apple pectin resulted in the upregulation of the all the PGA-induced enzymes as well as the putative pectin methylesterase (*pamy_4273*), one RG lyase (*pamy_4178*), and a second GH105 hydrolase (*pamy_4272*). Upregulation of the pectin methylesterase on this substrate was anticipated given the partial methylation of the substrate, but the increased GH105 hydrolase production (which is likely cotranscribed

with the pectin methylesterase) suggests that this enzyme is a HG-specific unsaturated galacturonidase like *pamy_1066* but is specific for methylated substrates. Likewise, upregulation of the putative RG lyase *pamy_4178* was unexpected, but the potential utility of this enzyme for deconstructing methylated HG substrates is unclear.

Growth with RG-I resulted in lower upregulation of the PGA-inducible genes, likely because the RG-I substrate is not completely separated from the other pectic polysaccharides. However, this substrate produced high fold change values for expected RG-I deconstructing enzymes including two extracellular RG lyases (*pamy_4308, pamy_3149*) and one additional GH105 hydrolase (*pamy_5356*). Cytoplasmic rhamnosidases (*pamy_408, pamy_1805, pamy_2439, pamy_2460*) were less highly upregulated and it is possible that one or more of these enzymes is not specific to RG-I deconstruction. The one putative cytoplasmic galacturonidase (*pamy_922*) was upregulated by a similar amount. Interestingly, a fourth pectate lyase (*pamy_4669*) which was not upregulated on PGA or apple pectin did appear somewhat upregulated in response to RG-I.

2.4.5 Summary of the putative pectinolytic system

A summary of the putative pectinolytic enzymes is provided (Table 1) as well as a diagram of the system based on the bioinformatic and RT-qPCR analyses (Figure 2.4). Overall this system includes a diverse set of pectinases from 13 different CAZy catalytic families (plus one CBM family) including some of the less well studied pectate lyase families, the recently described GH105 hydrolases, and both known RG lyase families. In addition, the existence of a GH105 hydrolase specifically upregulated on methylated HG (*pamy_4272*) has not been described in the literature and it is possible that this enzyme has

a novel activity. Also, the presence of an intracellular pectin methylesterase (*pamy*_4273) is unusual since demethylation is generally thought of as an early step in deconstruction. This perhaps explains earlier observations that PelA and PelB retain activity on methylated substrates [25]; this flexibility would be necessary in a system which lacks an extracellular methylesterase. It is likely that the other two pectate lyases in this system also share this trait. Identification of additional enzymes with such broad substrate specificities is potentially valuable because they may allow for the creation of simplified pectinase mixtures which are as effective as more complex cocktails.

The RG-I deconstruction system also contains unique features. Most notably, the extracellular enzyme Pamy_4308 contains both RG lyase and RG acetylesterase domains and is likely able to cleave acetylated RG-I directly. A single enzyme with both of these catalytic domains has not yet been described, and RG lyases characterized to date have been inhibited by acetylation [19]. The presence of a putative cytoplasmic RG acetylesterase is unusual for the same reason that the intracellular pectin methylesterase was unexpected and likewise makes sense in the context of an extracellular RG lyase which may act on acetylated substrates. As in the case of the HG deconstruction system, identification of RG depolymerases which are uninhibited by acetylation may help improve and simplify pectinase mixtures.

2.5 CONCLUSION

P. amylolyticus 27C64 is distinct from other published *Paenibacillus* genomes and has a large and diverse set of putative CAZymes putting it on par with well-studied polysaccharide deconstructing organisms. Its pectinolytic system allows it to utilize four

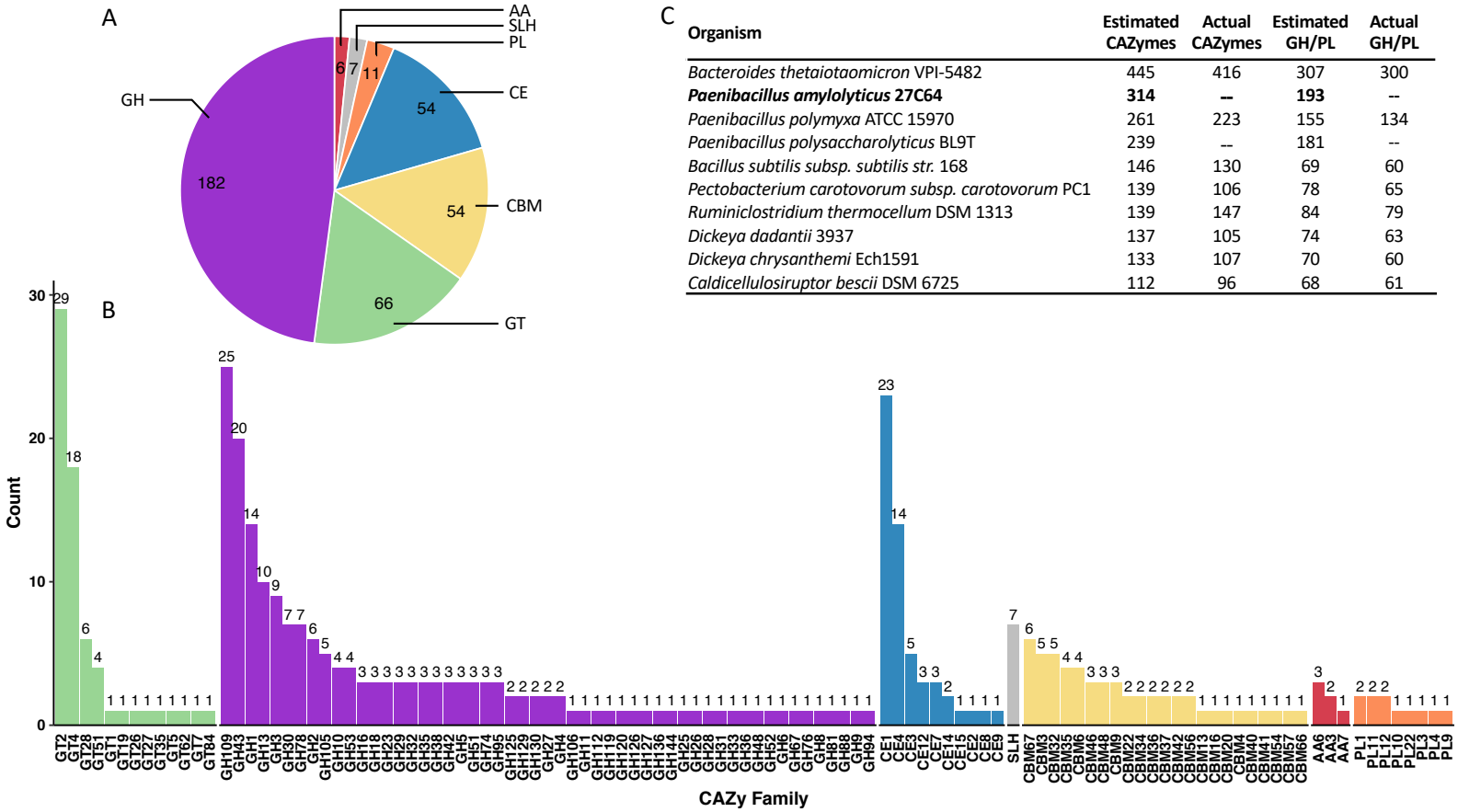
different pectic substrates as carbon sources and adapts to the specific substrate available. This system differs from other well-studied systems because it appears to rely on cytoplasmic deesterification of substrates and has enzymes with potentially novel functions and domain architectures. Direct biochemical study of this pectinolytic system may lead to improved understanding of the full diversity of natural pectinolytic systems.

2.6 ACKNOWLEDGEMENTS

The authors would like to thank Rebecca Gardner for her work refining pectinase assays, Resham Gehani for her work on growth curves, and the undergraduate Biochemistry and graduate Biomanufacturing and Bioprocessing students who have helped with this work.

Table 2.1 Putative pectinases of *P. amylolyticus* 27C64

Gene	CAZy Family	Putative Assignment	dbCAN E-value	Size (kDa)	Signal Peptide
Homogalacturonan Deconstruction					
pamy_4060	CE12	Pectin acetylesterase	2.0E-52	41.9	Yes
pamy_2278	PL1_8	Pectin lyase	5.5E-91	38.7	Yes
pamy_4343	PL1	Pectate lyase (PelB)	3.2E-41	51.0	Yes
pamy_2972	PL3_1	Pectate lyase (PelA)	1.1E-72	23.4	Yes
pamy_1763	PL9	Pectate lyase	8.2E-135	47.8	Yes
pamy_4178	PL4_1	RG lyase	8.2E-199	57.7	Yes
pamy_1066	GH105	Unsat. galacturonidase	3.5E-121	42.2	No
pamy_4272	GH105	Unsat. galacturonidase	1.8E-118	42.1	No
pamy_4273	CE8	Pectin methylesterase	1.6E-96	38.4	No
pamy_82	GH28	Polygalacturonase	1.7E-78	58.7	No
Rhamnogalacturonan I Deconstruction					
pamy_4669	PL10	Pectate lyase	4.9E-99	34.9	Yes
pamy_4308	PL11|CE12|CBM37	RG lyase & RG acetylesterase	2.0E-287	188.2	Yes
pamy_3149	PL11_1	RG lyase	6.7E-153	86.1	Yes
pamy_1595	CE12	RG acetylesterase	4.8E-77	49.2	No
pamy_5356	GH105	Unsat. RG hydrolase	1.9E-78	39.2	No
pamy_408	GH78	RG rhamnohydrolase	1.3E-185	102.3	No
pamy_1805	GH78	RG rhamnohydrolase	6.0E-79	61.7	No
pamy_2439	GH78	RG rhamnohydrolase	5.9E-172	106.8	No
pamy_2460	GH106	RG rhamnohydrolase	2.1E-122	105.2	No
pamy_922	GH4	α-galacturonidase	3.2E-68	48.7	No



C

Organism	Estimated CAZymes	Actual CAZymes	Estimated GH/PL	Actual GH/PL
Bacteroides thetaiotaomicron VPI-5482	445	416	307	300
***Paenibacillus amylolyticus* 27C64**	**314**	**--**	**193**	**--**
Paenibacillus polymyxa ATCC 15970	261	223	155	134
Paenibacillus polysaccharolyticus BL9T	239	--	181	--
Bacillus subtilis subsp. subtilis str. 168	146	130	69	60
Pectobacterium carotovorum subsp. carotovorum PC1	139	106	78	65
Ruminiclostridium thermocellum DSM 1313	139	147	84	79
Dickeya dadantii 3937	137	105	74	63
Dickeya chrysanthemi Ech1591	133	107	70	60
Caldicellulosiruptor bescii DSM 6725	112	96	68	61

Figure 2.1 Summary of dbCAN results. All predicted protein sequences from *P. amylolyticus* 27C64 were submitted to the dbCAN webserver and compared to hidden Markov models of each CAZy family. Hits below the published cutoffs were excluded. Hits are summarized by CAZy catalytic activity (A) and CAZy family (B). Identical analyses were carried out on other bacterial genomes for comparison (C). The dbCAN results are presented in the "Estimated" columns while the number of CAZymes recognized by the CAZy database is presented as "Actual" columns for those organisms present in the database.

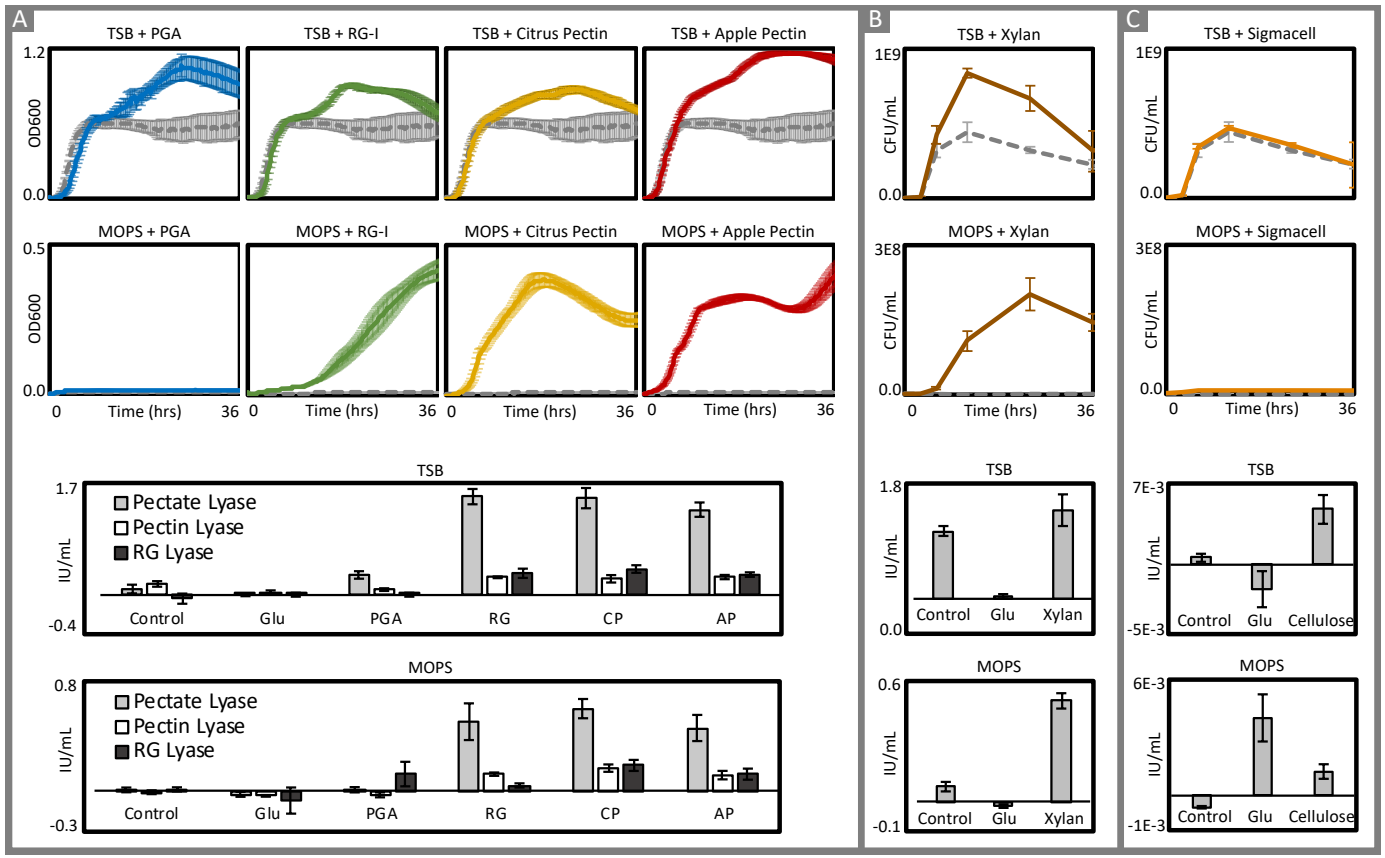


Figure 2.2 Growth of *P. amylolyticus* 27C64 on four different pectins (A), oat spelt xylan (B), and crystalline cellulose (C). Upper panels depict growth on each substrate as either a supplemental carbon source in TSB without dextrose or as a sole carbon source in MOPS minimal media. Dashed gray lines indicate the control condition (no polysaccharide added). Growth was measured as the OD at 600 nm for soluble substrates and based on plate counts of serial dilutions for insoluble substrates. Lower panels depict the amount of pectinase (A), xylanase (B), or cellulase (C) activity in cell-free supernatants harvested from each growth condition at the end of the growth period. All experiments were conducted with four biological replicates and error bars indicate the standard deviation. *Glu* glucose, *PGA* polygalacturonic acid, *RG* rhamnogalacturonan I, *CP* citrus pectin, *AP* apple pectin.

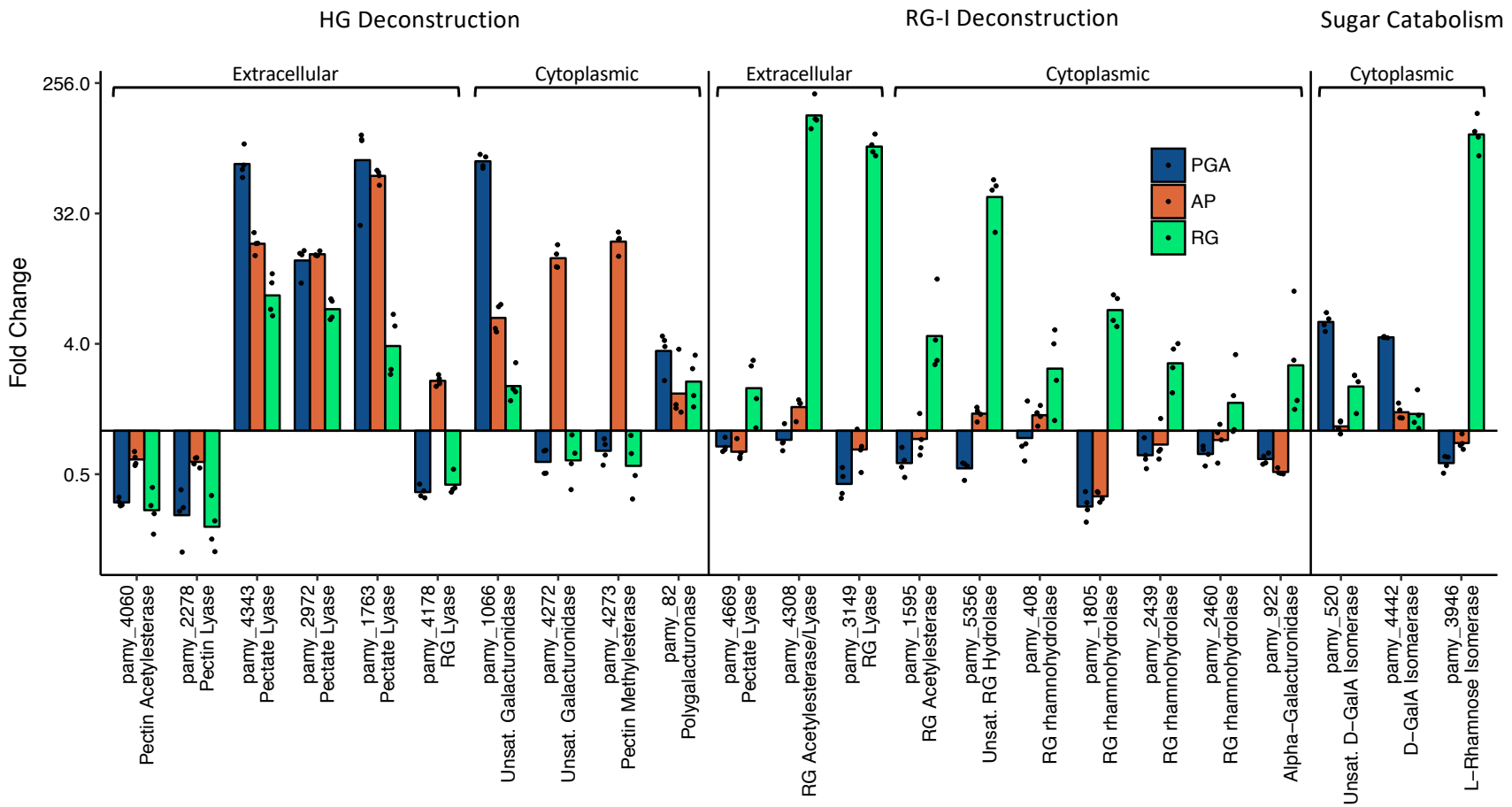


Figure 2.3 Differential expression of putative pectinases when grown on three different pectic substrates. Points indicate the fold change in expression as compared to the control condition (TSB without pectin) for individual biological replicates and columns indicate the mean fold change values of all four biological replicates. Fold change values were calculated with the ΔΔCt method using ftsZ and rpoD as reference genes. *HG* homogalacturonan, *RG* rhamnogalacturonan I, *PGA* polygalacturonic acid, *AP* apple pectin.

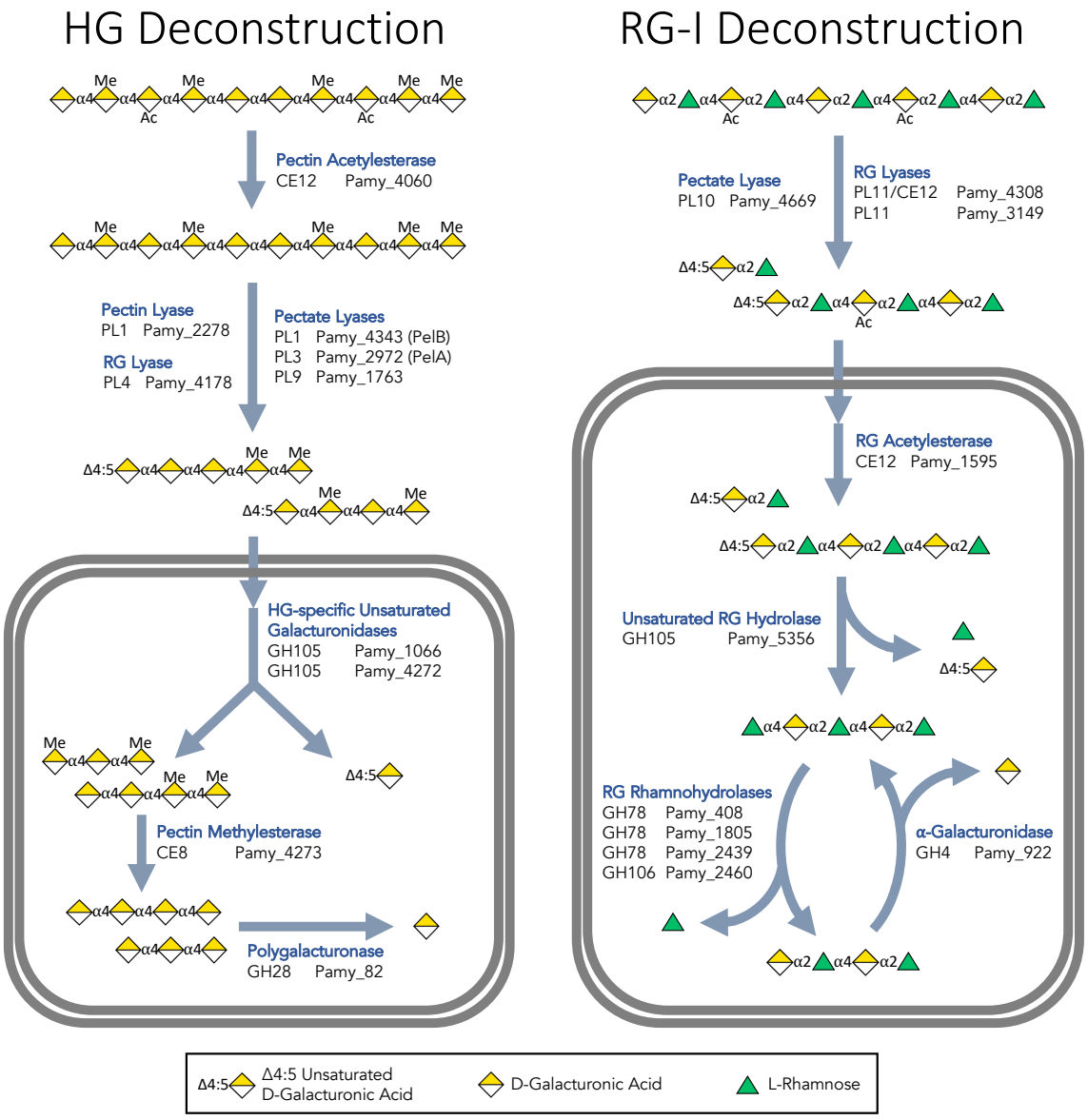


Figure 2.4 Summary of putative pectinolytic system. Symbols are based on the standardized symbol nomenclature for glycan notation. CAZy family assignments are provided to the left of each gene. *HG* homogalacturonan, *RG-I* rhamnogalacturonan-I, *GalA* D-galacturonic acid, *Rha* L-rhamnose.

2.7 REFERENCES

1. Mohnen D (2008) Pectin structure and biosynthesis. Curr Opin Plant Biol 11:266–277. doi:10.1016/j.pbi.2008.03.006

2. Jayani RS, Saxena S, Gupta R (2005) Microbial pectinolytic enzymes: a review. Process Biochem 40:2931–2944. doi:10.1016/j.procbio.2005.03.026

3. Kashyap DR, Vohra PK, Chopra S, Tewari R (2001) Applications of pectinases in the commercial sector: a review. Bioresour Technol 77:215–227. doi:10.1016/S0960-8524(00)00118-8

4. Hoondal G, Tiwari R, Tewari R, Dahiya N, Beg Q (2002) Microbial alkaline pectinases and their industrial applications: a review. Appl Microbiol Biotechnol 59:409–418. doi:10.1007/s00253-002-1061-1

5. Kohli P, Gupta R (2015) Alkaline pectinases: a review. Biocatal Agric Biotechnol 4:279–285. doi:10.1016/j.bcab.2015.07.001

6. Wang S, Sun X, Yuan Q (2018) Strategies for enhancing microbial tolerance to inhibitors for biofuel production: a review. Bioresour Technol 258:302–309. doi:https://doi.org/10.1016/j.biortech.2018.03.064

7. Edwards MC, Doran-Peterson J (2012) Pectin-rich biomass as feedstock for fuel ethanol production. Appl Microbiol Biotechnol 95:565–75. doi:10.1007/s00253-012-4173-2

8. Chundawat SPS, Beckham GT, Himmel ME, Dale BE (2011) Deconstruction of lignocellulosic biomass to fuels and chemicals. Annu Rev Chem Biomol Eng 2:121–145. doi:10.1146/annurev-chembioeng-061010-114205

9. Lionetti V, Francocci F, Ferrari S, Volpi C, Bellincampi D, Galletti R, D'Ovidio R, De Lorenzo G, Cervone F (2010) Engineering the cell wall by reducing de-methyl-esterified homogalacturonan improves saccharification of plant tissues for bioconversion. Proc Natl Acad Sci 107:616–621. doi:10.1073/pnas.0907549107

10. Tomassetti S, Pontiggia D, Verrascina I, Reca IB, Francocci F, Salvi G, Cervone F, Ferrari S (2015) Controlled expression of pectic enzymes in Arabidopsis thaliana enhances biomass conversion without adverse effects on growth. Phytochemistry 112:221–230. doi:10.1016/j.phytochem.2014.08.026

11. Biswal AK, Soeno K, Gandla M, Immerzeel P, Pattathil S, Lucenius J, Serimaa R, Hahn MG, Moritz T, Jönsson LJ, Israelsson-Nordström M, Mellerowicz EJ (2014) Aspen pectate lyase PtxtPL1-27 mobilizes matrix polysaccharides from woody tissues and improves saccharification yield. Biotechnol Biofuels 7:11. doi:10.1186/1754-6834-7-11

12. Francocci F, Bastianelli E, Lionetti V, Ferrari S, De Lorenzo G, Bellincampi D, Cervone F (2013) Analysis of pectin mutants and natural accessions of Arabidopsis highlights the impact of de-methyl-esterified homogalacturonan on tissue saccharification. Biotechnol Biofuels 6:163. doi:10.1186/1754-6834-6-163

13. Biswal AK, Hao Z, Pattathil S, Yang X, Winkeler K, Collins C, Mohanty SS, Richardson EA, Gelineo-Albersheim I, Hunt K, Ryno D, Sykes RW, Turner GB, Ziebell A, Gjersing E, Lukowitz W, Davis MF, Decker SR, Hahn MG, Mohnen D (2015) Downregulation of GAUT12 in Populus deltoides by RNA silencing results in reduced recalcitrance, increased growth and reduced xylan and pectin in a woody biofuel feedstock. Biotechnol Biofuels 8:41. doi:10.1186/s13068-015-0218-y

14. Chung D, Pattathil S, Biswal AK, Hahn MG, Mohnen D, Westpheling J (2014) Deletion of a gene cluster encoding pectin degrading enzymes in Caldicellulosiruptor bescii reveals an important role for pectin in plant biomass recalcitrance. Biotechnol Biofuels 7:147. doi:10.1186/s13068-014-0147-1

15. Collmer A, Keen NT (1986) The role of pectic enzymes in plant pathogenesis. Annu Rev Phytopathol 24:383–409. doi:10.1146/annurev.py.24.090186.002123

16. Abbott DW, Boraston AB (2008) Structural biology of pectin degradation by Enterobacteriaceae. Microbiol Mol Biol Rev 72:301–316. doi:10.1128/MMBR.00038-07

17. Luis AS, Briggs J, Zhang X, Farnell B, Ndeh D, Labourel A, Baslé A, Cartmell A, Terrapon N, Stott K, Lowe EC, McLean R, Shearer K, Schückel J, Venditto I, Ralet M-C, Henrissat B, Martens EC, Mosimann SC, Abbott DW, Gilbert HJ (2017) Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic Bacteroides. Nat Microbiol 3:210. doi:10.1038/s41564-017-0079-1

18. Ndeh D, Rogowski A, Cartmell A, Luis AS, Baslé A, Gray J, Venditto I, Briggs J, Zhang X, Labourel A, Terrapon N, Buffetto F, Nepogodiev S, Xiao Y, Field RA, Zhu Y, O'Neill MA, Urbanowicz BR, York WS, Davies GJ, Abbott DW, Ralet M-C, Martens EC, Henrissat B, Gilbert HJ (2017) Complex pectin metabolism by gut bacteria reveals novel catalytic functions. Nature 544:65–70. doi:10.1038/nature21725

19. Silva IR, Jers C, Meyer AS, Mikkelsen JD (2015) Rhamnogalacturonan I modifying enzymes: an update. N Biotechnol 33:41–54. doi:10.1016/j.nbt.2015.07.008

20. Itoh T, Ochiai A, Mikami B, Hashimoto W, Murata K (2006) A novel glycoside hydrolase family 105: the structure of family 105 unsaturated rhamnogalacturonyl hydrolase complexed with a disaccharide in comparison with family 88 enzyme complexed with the disaccharide. J Mol Biol 360:573–585. doi:10.1016/j.jmb.2006.04.047

21. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:490–495. doi:10.1093/nar/gkt1178

22. Latarullo MBG, Tavares EQP, Maldonado GP, Leite DCC, Buckeridge MS (2016) Pectins, endopolygalacturonases, and bioenergy. Front Plant Sci 7:1–7. doi:10.3389/fpls.2016.01401

23. Schneider DM, Henriksen ED, Boland WE, Edwards MC, Doran-Peterson J (2014) Tipula abdominalis: a natural biorefinery with novel microbial enzymes useful for pectin-rich biomass deconstruction. In: Sun J, Ding S-Y, Doran-Peterson J (eds) Biological Conversion of Biomass for Fuels and Chemicals: Exploration from Natural Utilization Systems. The Royal Society of Chemistry, Cambridge, pp 282–297

24. Cook DM, Henriksen ED, Upchurch R, Doran-Peterson JB (2007) Isolation of polymer-degrading bacteria and characterization of the hindgut bacterial community from the detritus-feeding larvae of Tipula abdominalis (Diptera: Tipulidae). Appl Environ Microbiol 73:5683–6. doi:10.1128/AEM.00213-07

25. Boland WE, DeCrescenzo Henriksen E, Doran-Peterson J (2010) Characterization of two Paenibacillus amylolyticus strain 27C64 pectate lyases with activity on highly methylated pectin. Appl Environ Microbiol 76:6006–6009. doi:10.1128/AEM.00043-10

26. Edwards MC, Williams T, Pattathil S, Hahn MG, Doran-Peterson J (2014) Replacing a suite of commercial pectinases with a single enzyme, pectate lyase B, in Saccharomyces cerevisiae fermentations of cull peaches. J Ind Microbiol Biotechnol 41:679–686. doi:10.1007/s10295-013-1394-z

27. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–20. doi:10.1093/bioinformatics/btu170

28. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V, Sirotkin A V, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–77. doi:10.1089/cmb.2012.0021

29. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:421. doi:10.1186/1471-2105-10-421

30. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O (2008) The RAST server: rapid annotations using subsystems technology. BMC Genomics 9:75. doi:10.1186/1471-2164-9-75

31. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R (2014) The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res 42:D206-14. doi:10.1093/nar/gkt1226

32. Chen IMA, Markowitz VM, Chu K, Palaniappan K, Szeto E, Pillay M, Ratner A, Huang J, Andersen E, Huntemann M, Varghese N, Hadjithomas M, Tennessen K, Nielsen T, Ivanova NN, Kyrpides NC (2017) IMG/M: Integrated genome and metagenome comparative data analysis system. Nucleic Acids Res 45:D507–D516. doi:10.1093/nar/gkw929

33. Varghese NJ, Mukherjee S, Ivanova N, Konstantinidis KT, Mavrommatis K, Kyrpides NC, Pati A (2015) Microbial species delineation using whole genome sequences. Nucleic Acids Res 43:6761–6771. doi:10.1093/nar/gkv657

34. Park BH, Karpinets T V, Syed MH, Leuze MR, Uberbacher EC (2010) CAZymes Analysis Toolkit (CAT): web service for searching and analyzing carbohydrate-active enzymes in a newly sequenced organism using CAZy database. Glycobiology 20:1574–84. doi:10.1093/glycob/cwq106

35. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y (2012) dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res 40:W445-51. doi:10.1093/nar/gks479

36. Marchler-Bauer A, Bryant SH (2004) CD-Search: protein domain annotations on the fly. Nucleic Acids Res 32:W327-31. doi:10.1093/nar/gkh454

37. Edgar RC (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. doi:10.1093/nar/gkh340

38. Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8:785–6. doi:10.1038/nmeth.1701

39. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ, Brinkman FSL (2010) PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics 26:1608–15. doi:10.1093/bioinformatics/btq249

40. Käll L, Krogh A, Sonnhammer ELL (2004) A combined transmembrane topology and signal peptide prediction method. J Mol Biol 338:1027–1036. doi:10.1016/j.jmb.2004.03.016

41. Neidhardt FC, Bloch PL, Smith DF (1974) Culture medium for enterobacteria. J Bacteriol 119:736–747

42. Collmer A, Ried JL, Mount MS (1988) Assay methods for pectic enzymes. Methods Enzymol 161:329–335. doi:https://doi.org/10.1016/0076-6879(88)61037-8

43. Ochiai A, Itoh T, Kawamata A, Hashimoto W, Murata K (2007) Plant cell wall degradation by saprophytic Bacillus subtilis strains: gene clusters responsible for rhamnogalacturonan depolymerization. Appl Environ Microbiol 73:3803–13. doi:10.1128/AEM.00147-07

44. Somogyi M (1952) Notes on sugar determination. J Biol Chem 195:19–23

45. Nelson N (1944) A photometric adaptation of the Somogyi method for the determination of glucose. J Biol Chem 153:375–380

46. Vandesompele J, De Preter K, Pattyn ilip, Poppe B, Van Roy N, De Paepe A, Speleman rank (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3:34–1. doi:10.1186/gb-2002-3-7-research0034

47. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. Methods 25:402–408. doi:10.1006/meth.2001.1262

48. Madhaiyan M, Poonguzhali S, Saravanan VS, Duraipandiyan V, Al-Dhabi NA, Kwon S-W, Whitman WB (2017) Paenibacillus polysaccharolyticus sp. nov., a xylanolytic and cellulolytic bacteria isolated from leaves of Bamboo Phyllostachys aureosulcata. Int J Syst Evol Microbiol 67:2127–2133. doi:10.1099/ijsem.0.001901

49. Akuzawa S, Nagaoka J, Kanekatsu M, Kubota E, Ohtake R, Suzuki T, Kanesaki Y (2016) Draft genome sequence of Paenibacillus amylolyticus Heshi-A3, isolated from fermented rice bran in a Japanese fermented seafood dish. Genome Announc 4:e00218-16. doi:10.1128/genomeA.00218-16

CHAPTER 3

THE HOMOGALACTURONAN DECONSTRUCTION SYSTEM OF *PAENIBACILLUS AMYLOLYTICUS* 27C64 REQUIRES NO EXTRACELLULAR PECTIN METHYLESTERASE AND HAS SIGNIFICANT INDUSTRIAL POTENTIAL [1]

[1] Keggi, C and J Doran-Peterson. To be submitted to *Applied and Environmental Microbiology*.

3.1 ABSTRACT

Paenibacillus amylolyticus 27C64, a Gram-positive bacterium with diverse plant cell wall polysaccharide deconstruction capabilities, was previously isolated from an insect hindgut. Previous work suggested that this organism's pectin deconstruction system differs from known systems because its sole pectin methylesterase is cytoplasmic, not extracellular. In this work, we have characterized the specific roles of key extracellular pectinases involved in homogalacturonan deconstruction, including four pectate lyases and one pectin lyase. We show that one newly characterized pectate lyase, PelC, has a novel substrate specificity with a lower K_m for highly methylated pectins than for polygalacturonic acid. PelC works synergistically with PelB, a high-turnover exo-pectate lyase that releases Δ4,5-unsaturated trigalacturonate as its major product, likely by releasing internal stretches of demethylated homogalacturonan which PelB can degrade. We also show that the sole pectin lyase has a high K_{cat} value and rapidly depolymerizes methylated substrates. Three cytoplasmic GH105 hydrolases were screened for the ability to remove terminal unsaturated galacturonic acid residues from oligogalacturonides, and we found that two are involved in homogalacturonan deconstruction. This work confirms that efficient homogalacturonan deconstruction in *P. amylolyticus* 27C65 does not require extracellular pectin methylesterase activity. Three of the extracellular lyases studied in this work are also thermostable, function well over a broad pH range, and have significant industrial potential.

3.2 IMPORTANCE

Pectin is an important structural polysaccharide found in most plant cell walls. In the environment, pectin degradation is part of the decomposition process that turns over dead plant material and is important to organisms that feed on plants. Industrially, pectinases are used to improve the quality of fruit juices and can also be used to process coffee cherries or tea leaves. These enzymes may also prove useful in reducing the environmental impact of paper and cotton manufacturing. This work is significant because it focuses on a Gram-positive bacterium that is evolutionarily distinct from other well-studied pectin-degrading organisms and differs from known systems in key ways. Most importantly, a simplified extracellular deconstruction process in this organism is able to break down pectins without first removing the methyl groups that inhibit other systems. Moreover, some of the enzymes described here have the potential to improve industrial processes that rely on pectin deconstruction.

3.3 INTRODUCTION

Pectins are a complex family of galacturonic acid (GalA) rich polysaccharides present in virtually all plant cell walls, structures that store the majority of the 55 billion tons of carbon fixed by plants in terrestrial ecosystems each year [1]. Pectic polysaccharides are abundant in the primary cell walls and middle lamellae of most plants where they play a number of vital roles [2, 3]. The pectin network is interwoven with and covalently linked to other matrix polysaccharides [4–6], cellulose [7], and lignin [8], and is structurally important. In *Arabidopsis*, disruption of normal pectin biosynthesis results in dwarfism and brittle stems [9]. Pectin is the main component of the outermost cell wall

layer, the middle lamella, where it is responsible for cell-cell adhesion [2, 3, 10]. Pectin has also been implicated in a number of other important plant functions, including cell defense and development [10]. This family of polysaccharides includes several structurally distinct types of pectin, with homogalacturonan (HG) and rhamnogalacturonan I (RG-I) being the two most abundant [10]. HG accounts for approximately 65% of pectin and is comprised of a linear 1,4-α-D-GalA backbone that can be methylated at the C-6 carboxyl or acetylated at the O-2 or O-3 positions. While methylation is often extensive, in some cases occurring on more than 90% of the GalA residues, acetylation is typically minimal, except in the pectins of a few specific plant species [10]. RG-I has an alternating GalA-rhamnose backbone that may also be acetylated and has variable arabinose and galactose rich side chains attached to 20-80% of the rhamnosyl residues [10].

The mechanisms of microbial pectin deconstruction have been of interest over the last few decades since pectinases are important virulence factors for some Gram-negative plant pathogens and are also useful for fruit juice processing. Specifically, three genera of plant pathogens within the *Enterobacteriaceae,* including *Pectobacterium, Dickeya,* and *Erwinia,* have well-studied pectin deconstruction systems [11], while enzymes from various *Aspergillus* species that are active in acidic conditions are commonly used to improve fruit juice yields, reduce the viscosity of the extracted juice, and clarify the final product [12, 13]. More recently, pectin has been identified as an important contributor to biomass recalcitrance in feedstocks used for lignocellulosic ethanol production, such as pectin-rich agricultural wastes [14], herbaceous plant tissues [15–17], and woody plant material [18–20]. Additionally, deletion of a pectinase gene cluster from an organism that can normally grow on unmodified plant material, *Caldicellulosiruptor bescii*, results in a

growth defect when *Arabidopsis*, switchgrass, or poplar wood are supplied as growth substrates [21]. This suggests that efficient pectin deconstruction is necessary to eliminate or reduce the severity of thermochemical pretreatment steps. Better understanding of microbial pectin deconstruction, especially in organisms phylogenetically distinct from *Enterobacteriaceae* and *Aspergillus*, may improve the efficiency and reduce the cost of plant cell wall deconstruction for biofuel or chemicals production, and may also identify novel enzymes with other industrial uses. Bacterial enzymes often have alkaline optimum pH ranges and are of particular interest for retting and degumming plant fibers in textiles manufacturing, removing of pulp from coffee cherries, and processing tea leaves [13, 22]. These bacterial enzymes also facilitate the removal of non-cellulosic polysaccharides from paper pulp or cotton, thereby reducing the need for strongly basic processing steps and eliminating problems associated with alkaline wastewater [13, 22].

Paenibacillus amylolyticus 27C64 is a Gram-positive bacterium that falls into a taxonomic group distinct from the other well studied pectinolytic bacteria. It was originally isolated from the microbial hindgut community of an aquatic crane fly (*Tipula abdominalis*) larvae where it is partly responsible for the breakdown of plant material ingested by the insect [23]. This isolate displayed a wide range of plant cell wall polysaccharide deconstruction capabilities when screened on differential media [23], and two of its pectate lyases have been characterized [24]. These two enzymes, PelA and PelB, retained an unusual amount of activity on highly methylated pectins that are normally poor substrates for pectate lyases [24]. Subsequent analysis of the carbohydrate-active enzymes (CAZymes) present in the genome revealed a remarkable diversity of CAZymes, including ten enzymes believed to be capable of deconstructing HG [25]. The regulation of these

putative pectinases when *P. amylolyticus* was grown with different pectic substrates provided some insight into the functional role of each enzyme. A pectin acetylesterase that removes O-2 or O-3 acetylation (Pae), four pectate lyases that cut at unmethylated GalA residues (PelA, PelB, PelC, and PelD), and one pectin lyase that cuts at methylated sites (Pnl) seemed to be responsible for extracellular HG deconstruction (Table 3.1). In the cytoplasm, two GH105 hydrolases (YteR and YteZ), which likely remove terminal unsaturated GalA residues from HG-derived oligosaccharides, were identified along with a pectin methylesterase (Pem; removes C-6 methylation) and a polygalacturonase (Peh; cleaves unmethylated substrates by hydrolysis) (Table 3.1). Protein names, in lieu of locus tags generated by annotation software that were referenced previously, are assigned here for clarity and are based on sequence identity to biochemically similar enzymes where possible. Despite some differences in their regulation [25], significant ambiguity in the precise biochemical role of each enzyme exists. For example, the utility of four catalytically redundant pectate lyases was unclear.

It was previously noted that this system apparently lacks an extracellular pectin methylesterase, a finding that stands in stark contrast to Gram-negative [11] and fungal [26] systems, and perhaps explains the high activity of PelA and PelB on methylated substrates [24]. In this study, we confirm that the extracellular pectinases of *P. amylolyticus* 27C64 are able to break down methylated HG without the aid of an extracellular pectin methylesterase, and we provide a model of HG deconstruction that may be applicable to other Gram-positive bacteria. We also identify elements of the system that have significant industrial potential.

3.4 RESULTS

3.4.1 Protein Purification

All five extracellular lyases (PelA, PelB, PelC, PelD, Pnl) had their polyhistidine-rubredoxin affinity tags removed and were purified until no other bands were clearly visible on a Coomassie-stained SDS-PAGE gel (Appendix B, Figure B1). Polyhistidine-rubredoxin tags were not removed from the cytoplasmic hydrolases (YteR, YteZ, YesR), and very few contaminants were present on SDS-PAGE gels (Appendix B, Figure B1). The relative mobility of the proteins on SDS-PAGE gels were consistent with their theoretical molecular weights.

3.4.2 Characterization of Pectic Substrates

The commercial substrates used had very different degrees of methylation (DM), ranging from 0 - 81.7% (Appendix B, Table B1). Most DM values were consistent with the product labeling if a degree of esterification value was specified, except in one case (CP_90). Acetylation was minimal across all the substrates, with the highest degree of acetylation (DA) reaching 3.4% for AP_Sigma. Most substrates were high-molecular weight pectins, although CP_ICN and RG_P-A had low-molecular weight peaks as well. CP_90 was notably different from the other citrus pectins as it included a broad mix of high and low molecular weight fragments.

3.4.3 Determination of Optimum Reaction Conditions

The three lyases newly characterized in this study were PelC, PelD, and Pnl. All three worked optimally in a neutral to alkaline pH range (Appendix B, Figure B2). PelC had its optimal activity at pH 10.0 and 55°C, with 0.1-1.0 mM $CaCl_2$ included in the reaction buffer. Optimal activity for PelD was observed at pH 9.0 and 45°C in the presence

of 1 mM $CaCl_2$. Optimal activity for Pnl was observed between pH 9 and 10 at a temperature of 55°C. Although 1 mM $CaCl_2$ slightly increased the activity of Pnl, it was not essential and addition of 1 mM EDTA did not affect activity. In contrast, activities of PelC and PelD could not be detected when 1 mM EDTA was included in the reaction buffer (data not shown). The optimum reaction conditions for PelA and PelB have been described [24], but the temperature optimum for PelB activity was reevaluated since new equipment enabled measurements at higher temperatures. The new optimum temperature for PelB activity is 70°C.

3.4.4 Substrate Specificity

The activity of each extracellular lyase was evaluated on 17 different commercially available pectins, and multiple linear models were used to determine the impact of methylation, acetylation, molecular weight, and GalA content on activity (Appendix B, Table B2). All of the enzymes were specific to homogalacturonan and had no measurable activity on rhamnogalacturonan. Overall, substrate methylation had the greatest impact on activity. Initial reaction rates for PelA, PelB and PelD were highest on moderately methylated pectins, while PelC and Pnl displayed more activity on highly methylated substrates than on minimally or moderately methylated substrates (Figure 3.1a). Except for PelC, the activity of each enzyme was significantly higher on lower molecular weight substrates ($p < 0.001$), but the magnitude of this effect was greatest for PelA (Figure 3.1c). In fact, substrate molecular weight had a greater impact on activity of PelA than substrate methylation. Acetylation did have an impact on the activity of PelA and PelD ($p < 0.01$), but the magnitude of the effect was small and was not a major predictor of activity. The GalA content of each substrate was included as a model term to control for its effect, and

with the exception of PelC, all of the enzymes displayed significantly higher activity with higher GalA ($p < 0.001$).

3.4.5 Kinetics

Three high molecular weight pectins with varying degrees of methylation were selected to determine the kinetic values of each enzyme: PGA-C (0% methylated), CP_Sigma (55% methylated), and CP_85-C (80% methylated). Generally, the four pectate lyases had similar K_{cat} values on each substrate, but K_m values differed (Table 3.2). PelA had low K_m values on PGA-C and CP_Sigma, but its K_m on CP_85-C was more than an order of magnitude higher. PelB, PelC, and PelD had the lowest K_m on CP_Sigma. However, while PelB and PelD had K_m values for CP_85-C that were an order of magnitude higher than those for PGA-C, PelC had a lower K_m for CP_85-C than it did for PGA-C. The pectin lyase, Pnl, had much higher affinity for methylated substrates than PGA-C, and its turnover rate was hindered by low methylation. PelA and PelD were inhibited by high substrate concentrations of PGA and CP_Sigma. Some deviation from the expected Michaelis-Menten curve was observed for PGA with PelB, PelC, and PelD (Appendix B, Figure B3). It is possible that this deviation is an artefact of substrate binding to assay materials since PGA is anionic and supplied in its acidic (not salt) form, unlike the other two substrates. Alternatively, processivity of these enzymes on completely demethylated substrates may cause a deviation from the Michaelis-Menten model. In either case, the K_m values still effectively described the relative affinity of each enzyme for the three different substrates.

3.4.5 Analysis of Reaction Products and Mode of Action

The products formed by the action of each extracellular lyase on the same three pectins used for kinetic assays were analyzed after allowing the reactions to run to completion. All four pectate lyases released Δ4,5-unsaturated trigalacturonate as their major product on PGA-C, but also generated Δ4,5-unsaturated digalacturonate. When these enzymes acted on methylated substrates, a mix of higher molecular weight oligogalacturonides and lower overall product release was observed. However, PelB was hindered by methylation to a much greater extent than the other pectate lyases were. The pectin lyase released Δ4,5-unsaturated digalacturonate as the major product on CP_85-C, but an unsaturated trisaccharide as the major product on CP_Sigma and PGA.

Intermediate reaction products of PGA digestion reactions with each pectate lyase or CP_85-C digestion with the pectin lyase were evaluated at five time points. Pnl initially released exclusively very large products, which were later degraded to primarily unsaturated digalacturonate (Figure 3.3). In contrast, PelA and PelD initially produced a mix of short oligogalacturonides and some larger products that were later degraded to unsaturated tri- and di-saccharides. PelC released mostly unsaturated trigalacturonate, but larger products were observed in the first time point. PelB consistently released a mixture of unsaturated tri- and di-galacturonate.

3.4.6 Synergy Between Extracellular Lyases

The ability of enzyme combinations to completely deconstruct PGA-C, CP_Sigma, and CP_85-C were evaluated using very low enzyme concentrations (1 nM). Under these conditions, PelA had very little activity compared to the other enzymes. PelB accounted for most of the PGA-C deconstruction, while Pnl was the most effective on highly

methylated CP_85-C (Figure 3.4). In contrast, CP_Sigma (55% methylated) required both pectate and pectin lyases for complete deconstruction. On this moderately methylated substrate, PelB and PelC worked synergistically, achieving a higher percent deconstruction together than either did separately.

3.4.7 Role of Cytoplasmic GH105 Hydrolases

The three GH105 hydrolases that were previously shown to be upregulated on different pectic substrates [25] were screened for the ability to remove terminal Δ4,5-unsaturated GalA residues from the product mixtures generated by digestion of PGA-C with PelB and PelC or by Pnl digestion of CP_85-C. Both YteR and YteZ, enzymes that were upregulated on substrates containing demethylated HG and methylated HG, respectively, were active on the products released from PGA, but had no activity on the methylated fragments (Figure 3.5a). YesR, the GH105 enzyme that was upregulated on RG-I only, did not have activity on any of the HG digestion products. HPSEC analysis of these GH105 digestion reactions confirms that breakdown of the unsaturated trisaccharide released by PelC and PelB is the source of the decrease in absorbance used to assay the enzymes (Figure 3.5b).

3.5 DISCUSSION

3.5.1 The Role of Each Extracellular Depolymerase

Prior to this work, the differential regulation of each enzyme in this system had been investigated and it was known that PelA, PelB, and PelC were upregulated whenever pectins were used as a carbon source, that PelD was only upregulated when RG-I was supplied for growth, and that Pnl was constitutively expressed [25]. An extracellular pectin

acetylesterase that had a high degree of sequence similarity to characterized enzymes had also been identified. Despite some differences in regulation, the specific roles of each of the five extracellular depolymerases remained unclear, especially for the three pectate lyases that had similar expression patterns (PelA, PelB, and PelC). Correlation between pectin characteristics and the activity of each enzyme revealed some key differences. As is expected for a pectin lyase, Pnl had the most activity on highly methylated substrates, a reflection of the fact that these enzymes only cut efficiently at methylated sites. It was previously noted that PelA and PelB have unusually high activities on methylated substrates [24], an unexpected characteristic of this class of enzymes, which typically only cut at unmethylated sites. With the larger set of substrates used in this study, we confirmed the previous finding and added that PelB has significantly higher activity on moderately methylated substrates than on completely demethylated ones (Figure 3.1a and 3.1c). This is true for PelD as well, an enzyme newly described in this study. PelC, the other newly described enzyme, was even more unusual in that it has a higher initial reaction rate on the highly methylated pectins than it does on demethylated substrates. Taken together, these results reveal that this system is well adapted to working on methylated substrates even though four out of the five enzymes cannot cut at methylated sites. All of the enzymes, save PelC, also showed a significant preference for lower molecular weight substrates, but this effect was most pronounced in PelA where substrate molecular weight predicts activity better than substrate methylation. The strength of this preference led to PelA having much higher activity on CP_90, the lowest M_n substrate, than on any other pectin (Figure 3.1a and 3.1b).

Kinetic experiments told a similar, but more complete story, with PelA, PelB, and PelD all having much lower K_m values on PGA-C and CP_Sigma than on highly methylated CP_85-C (Table 3.2). In other words, these three pectate lyases were not hindered by moderate methylation, functioning effectively on all but the most highly methylated substrates. Between the unmethylated PGA and moderately methylated CP_Sigma, PelB and PelD had a slightly higher affinity for 55% methylation than for 0%, but PelA behaved in the opposite manner. The K_m values for PelC, however, were very unusual. This enzyme did not just retain activity on highly methylated substrates, it was actually hindered by a lack of methylation. The K_m on 55% methylated pectin was 10-fold lower than on the 85% methylated substrate, and was the highest on demethylated PGA. This implies that both methylated and unmethylated sites are important for PelC binding, and suggests that this enzyme specifically targets the junction of methylated and demethylated regions.

Analysis of reaction products showed that Δ4,5-unsaturated trigalacturonate is the major product released by all four pectate lyases with a lesser amount of unsaturated digalacturonate also present (Figure 3.2). The Pnl produced a similar mixture with relatively more digalacturonate. Tracking these samples over time revealed that Pnl is a true endo-type pectin lyase as it initially produced very large fragments (Figure 3.3). PelA, PelC, and PelD also cut at internal sites, but produced a mixture of short oligosaccharides at earlier time points instead of primarily very large fragments. However, PelB produced a homogenous product consistent with an exo-type enzyme. This conclusion is supported by the very low amount of product released from 80% methylated citrus pectin; if PelB cannot

access internal unmethylated stretches the degree of reaction completion would be severely reduced.

Collectively these data identify specific roles for each enzyme. PelB, as an exo-type enzyme with a very high K_{cat} is effective at rapidly saccharifying unmethylated stretches of HG. PelC, with a two order of magnitude lower K_{cat} than PelB, is not as effective at breaking down demethylated stretches of GalA residues but can bind at the interface of methylated and unmethylated regions to free internal stretches of unmethylated residues for PelB. PelA had the lowest K_m overall on PGA-C and may be useful when substrate concentrations are very low. It is also very effective on smaller substrates. Pnl, as the only enzyme able to cut at demethylated sites, is responsible for the deconstruction of methylated regions and works rapidly with a high K_{cat} on 80% methylated pectin. PelD, which was only upregulated on RG-I, does not have a clear role in HG deconstruction based on these experiments.

3.5.2 Synergy Between Extracellular Enzymes

It is clear that PelB is largely responsible for the deconstruction of demethylated homogalacturonan since it is able to break down unmethylated polygalacturonic acid almost as well as all five enzymes combined (Figure 3.4). Likewise, Pnl seems chiefly responsible for depolymerization of highly methylated substrates. However, on moderately methylated CP_Sigma, the combination of PelB and Pnl is not as effective as all five enzymes together. PelC, which can make internal demethylated regions available to the exo-type PelB, is required for complete deconstruction. Under the conditions used (low enzyme concentration, 0.25 mg/mL substrate) both PelA and PelD could be dropped from the all-enzyme mixture without compromising the rate of deconstruction or the final degree

of completion. For PelD, this is likely because it is involved in removing HG regions from rhamnogalacturonan I or II. In contrast, PelA is most likely involved in HG deconstruction based on its regulation, but may be important for solubilization of fragments from cell walls or when substrate concentrations are very low, both possibilities that were not explored in this work.

3.5.3 Cytoplasmic Breakdown of Pectic Oligosaccharides

Previous findings that YteR was upregulated when any pectin was supplied for growth, but that YteZ was only upregulated when methylated substrates were supplied led to the hypothesis that YteZ may be specific to methylated substrates [25]. YesR was only upregulated on RG-I, and it was assumed to be an unsaturated rhamnogalacturonyl hydrolase. Here we see that both YteR and YteZ were active on unsaturated HG-derived oligosaccharides, but as expected, YesR was not (Figure 3.5a). Interestingly both active enzymes only break down the demethylated fragments despite the fact that YteZ was upregulated on methylated substrates. Such redundancy is necessary because this catabolic function is required even if only very highly methylated HG, which would not induce YteR, is present. Because both of these hydrolases work on the same substrate, methylated fragments would have to first be demethylated by the putative cytoplasmic pectin methylesterase (Pem) [25]. A putative cytoplasmic polygalacturonase (Peh) that is also encoded in the genome likely breaks down the saturated demethylated digalacturonate remaining after the action of YteR and YteZ.

3.5.4 Significance of Findings

The clearest difference between this homogalacturonan deconstruction system and those described in other organisms is that efficient and complete HG breakdown does not

require an extracellular pectin methylesterase. Instead, the only pectin methylesterase here appears to be cytoplasmic. This has created selective pressure for a pectin lyase with a high K_{cat} and for the unique substrate specificity of PelC. Paired with the very high turnover rate of PelB, these three enzymes represent a simplified set of pectinolytic enzymes that can completely degrade pectins and have significant industrial potential. In fact, since most natural pectins are moderately or highly methylated, Pnl alone is likely sufficient for many applications where only initial size reduction is required, such as removing pectic haze in beverages, reducing the viscosity of pectin-rich solutions, eliminating the foaming potential of powdered tea beverages, or removing pectinaceous pulp from coffee cherries and ramie fibers [12, 13]. When evaluating potential synergy between these enzymes, we noted that reactions with PelB, Pnl, or PelBC+Pnl were nearly half complete within three hours despite the low concentration of substrate (0.25 mg/mL initial) and the very low enzyme concentration (1 nM). This one nanomolar concentration equates to just one gram of each purified enzyme per 21,000-28,000 liters. Additionally, these experiments were carried out in simple buffer conditions at room temperature, but both PelC and Pnl are fairly thermostable with an optimum temperature of 55°C, while PelB is very thermostable with maximal activity at 70°C. Pnl also retains activity at cooler temperatures, working at 13% of its maximal activity at 10°C. All three of these enzymes can function across a broad pH range from just below neutral to very alkaline. Put together, these traits make each of these pectinases great candidates for use with different process requirements.

These findings also have broader implications for bacterial physiology. Most of the HG deconstruction enzymes known in Gram-positive bacteria have been studied in isolation and not in the context of a system [27]. This model of HG deconstruction in *P.*

amylolyticus is therefore an important contribution to the overall understanding of diverse pectin deconstruction systems. Also, the unusually high activity on methylated pectin that was previously noted in PelA and PelB [24] has been specifically noted in another pectate lyase from *Bacillus subtilis* [28], and in a homologue of PelB from another species within the genus *Paenibacillus* [29]. This suggests that the model presented here may be conserved in other Gram-positive bacteria and that this work may be able to serve as a resource to understand related systems.

Future work should directly evaluate the utility of enzymes from this system, especially PelB, PelC, and Pnl, in neutral or alkaline industrial processes like fiber processing, coffee and tea fermentations, paper biobleaching, and cotton bioscouring. The conservation of this HG deconstruction system in other bacteria should also be explored, as similar systems may be important to environmental biomass degradation.

3.6 MATERIALS AND METHODS

3.6.1 Strains and Culture Conditions

P. amylolyticus 27C64 was previously isolated from an insect hindgut [23] and was grown at 37°C with shaking in tryptic soy broth without dextrose (17 g tryptone, 3 g soytone, 5 g NaCl, 2.5 g NaK_2PO_4 per liter) or on tryptic soy agar (15 g tryptone, 5 g soytone, 5 g NaCl, 15 g agar per liter). *Escherichia coli* strains were growth at 37°C with shaking in LB (10 g tryptone, 5 g yeast extract, 10 g NaCl per liter) or on LB agar (includes 15 g agar per liter) unless indicated otherwise. When antibiotics were necessary to maintain plasmid selection, ampicillin was used at a concentration of 100 µg/mL and chloramphenicol was used at 40 µg/mL (Sigma-Aldrich, St. Louis, MO). Terrific broth was

used to express some of the genes (24 g yeast extract, 12 g tryptone, 4 mL glycerol, 2.31 g KH_2PO_4, 12.54 g K_2HPO_4 per liter).

3.6.2 Cloning

Genes included in this study were analyzed for signal peptides using SignalP 4.0 [30], Phobius [31], and PSORTb 3.0 [32] as previously described [25]. Each gene was amplified without its signal peptide (if present) in a 50 μL PCR reaction with high-fidelity Phusion DNA polymerase according to the manufacturer's instructions (New England Biolabs, Ipswich, MA). Template genomic DNA was isolated from *P. amylolyticus* 27C64 with a Wizard Genomic DNA Purification Kit (Promega, Madison, WI) and oligonucleotide primers were provided by Integrated DNA Technologies (Coralville, IA). Primers included 20 bp of homology to a pET-21a-derived plasmid that included an N-terminal tag comprised of an 8x His sequence, the *Pyrococcus furiosis* rubredoxin gene, a FLAG tag, and a tobacco etch virus (TEV) protease cleavage site. Genes were inserted into an AatII site immediately downstream of the TEV recognition sequence such that no non-native amino acids would remain after tag removal. Sequence and ligation independent cloning was used to assemble the plasmids as previously described except that PCR products were gel-purified in lieu of DpnI treatment and spin-column purification [33]. Plasmids were transformed into chemically competent *E. coli* DH5α prepared as described [34], isolated with the ZR Classic plasmid purification kit (Zymo Research, Irvine, CA), and sequenced to confirm that the insert was correct (Eton Bioscience, Research Triangle, NC).

3.6.3 Protein Expression and Purification

Plasmids were transformed into different strains of *E. coli* for expression: BL21(DE3) was used for PelA, PelB, PelC, and Pnl, while Rosetta 2(DE3) (MilliporeSigma, Burlington, MA) was used for PelD. *E. coli* KRX (Promega, Madison, WI) was used for YteR, YteZ, and YesR. Each of the five lyases were grown in 100 mL of terrific broth in 500 mL baffled Erlenmeyer flasks at 37°C with shaking. When cultures reached an optical density of 0.8-1.0, they were transferred to shaker set at 19°C, allowed to equilibrate for 30 minutes, then induced with 1 mM IPTG (Gold Biotechnology, St. Louis, MO) and left overnight. The strains carrying the three cytoplasmic hydrolases were grown in 10 mL of ZYP-5052 autoinduction media [35] supplemented with 0.1% rhamnose (Sigma-Aldrich, St. Louis, MO) plus 1 mM IPTG and allowed to grow at 25°C with shaking for 24 hours. Cells from each culture were harvested by centrifugation and stored at -20°C until needed.

Cell paste was resuspended in five milliliters of equilibration buffer per gram of cell paste (50 mM phosphate pH 7.6, 300 mM NaCl) and sonicated on ice. Lysate was clarified by centrifugation at 17,000 rpm using a JA-20 rotor in a Beckman floor centrifuge (Brea, CA) then filtered through a 0.45 μm PES membrane filter (MilliporeSigma, Burlington, MA) and bound to Ni-NTA resin. The larger samples were purified using an AKTA Start (GE Life Sciences, Chicago, IL) system with a five-milliliter bed of Ni-NTA resin (Sigma-Aldrich, St. Louis, MO) in a GE XK16/40 column. Smaller samples were purified using Ni-NTA spin-columns (Qiagen, Hilden, Germany). All samples were washed with ten or more column volumes of equilibration buffer with 10 mM imidazole and eluted with elution buffer (50 mM phosphate pH 7.6, 300 mM NaCl, 250 mM

imidazole, 10% glycerol). Hydrolases were dialyzed into 25 mM Tris pH 7.2 with 25 mM NaCl and stored at 4°C (the tags were not removed). Lyases were dialyzed into 10 mM Tris pH 7.5 with 10 mM NaCl. Tag cleavage reactions were set up by bringing the buffer strength up to 50 mM, adding 1 mM TCEP, and adding the amount of TEV protease recommended for the P1' amino acid of the construct [36]. After tag cleavage overnight at 4°C, uncut protein and free tag were removed by passing the samples over Ni-NTA resin equilibrated with 50 mM phosphate pH 8.0 with 300 mM NaCl, then samples were dialyzed into 25 mM Tris pH 7.2 with 25 mM NaCl and 10% glycerol and frozen at -20°C. The purity of all samples was evaluated using SDS-PAGE as described [37].

3.6.4 Enzyme Assays

Lyase activity was measured by monitoring the increase in absorbance at 232 nm as unsaturated products were formed [38]. Enzyme dilutions were stored on ice and five microliters was added to 995 µL of substrate solution just before collecting data. Reactions above or below room temperature were allowed to equilibrate at the target temperature for five minutes in a Peltier-thermostatted cuvette holder before the enzyme was added. A Cary 60 spectrophotometer (Agilent, Santa Clara, CA) was used to monitor the absorbance for one minute and determine the slope. Activity in International Units (IU; µmol/min) was calculated using published molar extinction coefficients [38]. Substrate solutions for determining the optimum reaction conditions of pectate lyases contained 0.2% polygalacturonic acid (PGA) and 0.5 mM $CaCl_2$ in a 100 mM buffer. Acetate buffers were used for pH values less than six, HEPES from pH six to seven, Tris from pH seven to nine, and glycine-NaOH above pH nine. Optimum calcium concentrations and temperatures were determined at each enzyme's optimal pH value. Pectin lyase optimization used the

same buffers but contained 0.05% esterified citrus pectin (Sigma-Aldrich, St. Louis, MO) and no calcium. For substrate specificity experiments, 0.5 mg/mL of each substrate was dissolved in 50 mM Tris pH 8.0 with 1 mM $CaCl_2$. Reaction progress curves contained only 0.25 mg/mL of each substrate.

GH105 hydrolase activity was monitored as a decrease in absorbance at 232 nm. As the terminal unsaturated carbohydrates are released, the ring-opening reaction reduces the absorbance [39]. To generate unsaturated oligogalacturonides, 0.5 mg/mL PGA or esterified citrus pectin in 50 mM phosphate buffer pH 6.9 was digested with PelB and PelC or the Pnl respectively. Digests included enzymes at a 10 nM concentration and were allowed to proceed for 24 hours at room temperature before being heat inactivated by boiling for five minutes. Hydrolase assay reactions contained two microliters of each enzyme in 398 microliters of substrate solution and were monitored in 1 mm path length quartz cuvettes due to the high starting absorbance.

For all assays, specific activity was calculated by dividing the activity in IU/mL by the concentration of the enzyme solution added to the assay. Enzyme concentrations were determined by measuring the absorbance at 280 nm using a NanoDrop Lite (ThermoFisher, Waltham, MA) and the molar extinction coefficients predicted by ProtParam [40].

3.6.5 Kinetics

Activity of each enzyme was determined in triplicate at a range of substrate concentrations from 0.01 to 2.0 mg/mL. Reactions were performed at room temperature in 50 mM Tris pH 8.0 with 1 mM $CaCl_2$ and monitored for one minute. Enzyme concentrations in the reactions were 5 nM for PelB, 10 nM for PelD and Pnl, and 50 nM

for PelA and PelC. Non-linear least squares minimization fitting to the Micahelis-Menten model in R was used to calculate the kinetic parameters.

3.6.6 Analysis of Pectins

The degree of methylation and degree of acetylation of each pectin was determined by releasing the esterified groups with a NaOH-isopropanol mixture and quantifying the amount of acetate and methanol released using an Aminex HPX-87H HPLC column (Bio-Rad, Hercules, CA) [41]. Galacturonic acid content was determined with the *m*-hydroxybiphenyl colorimetric uronic acid assay with galacturonic acid as the external standard [42]. This assay was scaled down linearly so the volumes were appropriate for a PCR plate that could be heated in an aluminum heat block, and samples were transferred a standard microtiter plate to read the absorbance. The molecular weight values (M_p, M_n, M_w) and polydispersity index (PI) of each pectin were determined using a high-performance size exclusion chromatography method as described [43]. Shodex P-82 narrow-polydispersity pullulan molecular weight standards (Showa Denko, Tokyo, Japan) were used to generate a calibration curve.

3.6.7 Statistical Analysis

JMP Pro 14 was used to calculate multiple linear regressions that described the activity of each lyase as a function of substrate characteristics. Enzyme activities were normalized as a percent of the maximum activity for each enzyme. To enable comparison of parameter estimates, values of independent variables were standardized by subtracting the mean of each variable and dividing by the standard deviation. None of the enzymes examined were active on rhamnogalacturonan substrates, which were enzymatically treated to remove HG, so these compounds were excluded from the analysis.

3.6.8 Analysis of Reaction Products

Products from pectin digestion reactions were separated using a Shimadzu Prominence HPLC system (Kyoto, Japan) equipped with a TSKgel G3000PW_{XL} high-performance size exclusion chromatography column and a TSKgel PW_{XL} guard column (Tosoh, Tokyo, Japan). The mobile phase was 50 mM phosphate buffer, pH 6.9, with 100 mM NaCl, and the column oven was maintained at 40°C. Products were detected as the absorbance at 232 nm, and the UV detector cell was maintained at 40°C. A galacturonic acid standard from Sigma (St. Louis, MO), and di- and tetra-galacturonic acid standards from Toronto Research Chemicals (Toronto, Canada) were used as markers.

3.7 ACKNOWLEDGMENTS

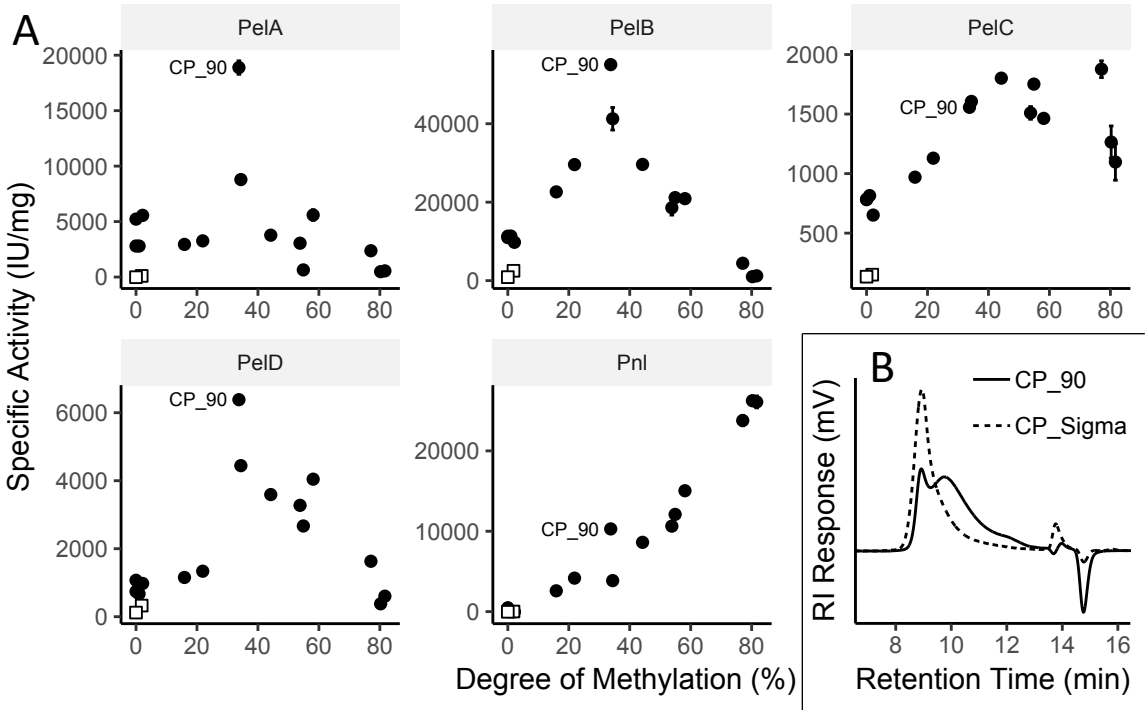
The authors would like to thank Rebecca Gardner for work expressing PelC and PelD and identifying their optimum reaction conditions. We would also like to thank the staff at the University of Georgia's Bioexpression and Fermentation Facility, especially Dr. David Blum, for all of their help with protein purification.

Table 3.1 Genes characterized and discussed in this study.

Gene	Protein	CAZy Assignment	Product	Size (kDa)
Extracellular proteins				
pamy_2972	PelA	PL3	Pectate lyase	20.7
pamy_4343	PelB	PL1	Pectate lyase	48.0
pamy_1763	PelC	PL9	Pectate lyase	43.3
pamy_4669	PelD	PL10	Pectate lyase	31.6
pamy_2278	Pnl	PL1	Pectin lyase	35.2
pamy_4060	Pae	CE12	Pectin acetylesterase	38.1
Cytoplasmic proteins				
pamy_1066	YteR	GH105	Unsaturated galacturonidase	42.2
pamy_4272	YteZ	GH105	Unsaturated galacturonidase	42.1
pamy_5356	YesR	GH105	Unsaturated RG hydrolase	39.2
pamy_4273	Pem	CE8	Pectin methylesterase	51.0
pamy_82	Peh	GH28	Polygalacturonase	58.7

Table 3.2 Kinetic parameters of each extracellular lyase on three pectic substrates.

Enzyme	Parameter	PGA-C	CP_Sigma	CP_85-C
PelA	K_m (mg/mL)	0.24 ± 0.06	0.43 ± 0.10	8.44 ± 2.58
	V_{max} (IU/mg)	39.7 ± 6.38	14.7 ± 2.42	32.1 ± 8.28
	K_{cat} (s^{-1})	4.94 ± 0.79 x 10^4	1.83 ± 0.30 x 10^4	3.99 ± 1.03 x 10^4
	K_{cat} / K_m	2.04 x 10^5	4.22 x 10^4	4.73 x 10^3
PelB	K_m (mg/mL)	2.48 ± 0.95	0.55 ± 0.09	50.1 ± 140
	V_{max} (IU/mg)	458 ± 110	302 ± 19.6	745 ± 2024
	K_{cat} (s^{-1})	1.32 ± 0.32 x 10^6	8.69 ± 0.56 x 10^5	2.14 ± 5.83 x 10^6
	K_{cat} / K_m	5.30 x 10^5	1.57 x 10^6	4.28 x 10^4
PelC	K_m (mg/mL)	1.02 ± 0.23	0.04 ± 0.00	0.41 ± 0.06
	V_{max} (IU/mg)	15.1 ± 1.55	9.68 ± 0.471	12.8 ± 0.610
	K_{cat} (s^{-1})	3.94 ± 0.41 x 10^4	2.52 ± 0.12 x 10^4	3.32 ± 0.16 x 10^4
	K_{cat} / K_m	3.84 x 10^4	5.49 x 10^5	8.02 x 10^4
PelD	K_m (mg/mL)	3.05 ± 1.95	0.89 ± 0.55	33.6 ± 159
	V_{max} (IU/mg)	304 ± 131	185 ± 81.1	531 ± 2400
	K_{cat} (s^{-1})	5.77 ± 2.48 x 10^5	3.51 ± 1.54 x 10^5	1.01 ± 4.55 x 10^6
	K_{cat} / K_m	1.89 x 10^5	3.94 x 10^5	2.99 x 10^4
Pnl	K_m (mg/mL)	7.61 ± 9.26	0.07 ± 0.00	0.13 ± 0.00
	V_{max} (IU/mg)	11.91 ± 11.91	59.27 ± 0.676	144.6 ± 1.894
	K_{cat} (s^{-1})	2.52 ± 2.52 x 10^4	1.25 ± 0.01 x 10^5	3.05 ± 0.04 x 10^5
	K_{cat} / K_m	3.30 x 10^3	1.57 x 10^6	2.21 x 10^6



C	PelA	PelB	PelC	PelD	Pnl
DM	-1.1	1.1	16.6***	5.8**	33.9***
DM²	-20.1***	-34.4***	-16.9***	-28.1***	7.3***
DA	7.6**	1.2	-1.1	13.1***	0.6
Mn	-27.3***	-17.3***	1.1	-17.1***	-5.3***
GalA	29.6***	14.1***	1.9	26.9***	8.4***

Figure 3.1 Correlation between key characteristics of commercial pectins and the activity of each extracellular lyase. (A) Scatterplots show the specific activity of each enzyme as a function of substrate methylation, the most important substrate characteristic (n = 3). Open squares represent rhamnogalacturonan I substrates. (B) Size exclusion elution profiles are provided for a typical high molecular weight pectin, CP_Sigma, as well as the lowest molecular weight pectin, CP_90 (inset). (C) Standardized parameter estimates from linear models comparing the relative activity of each enzyme to the degree of methylation (DM), DM^2, degree of acetylation (DA), number-average molecular weight (Mn), and galacturonic acid content (GalA), of each pectin. Asterisks denote significant values (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

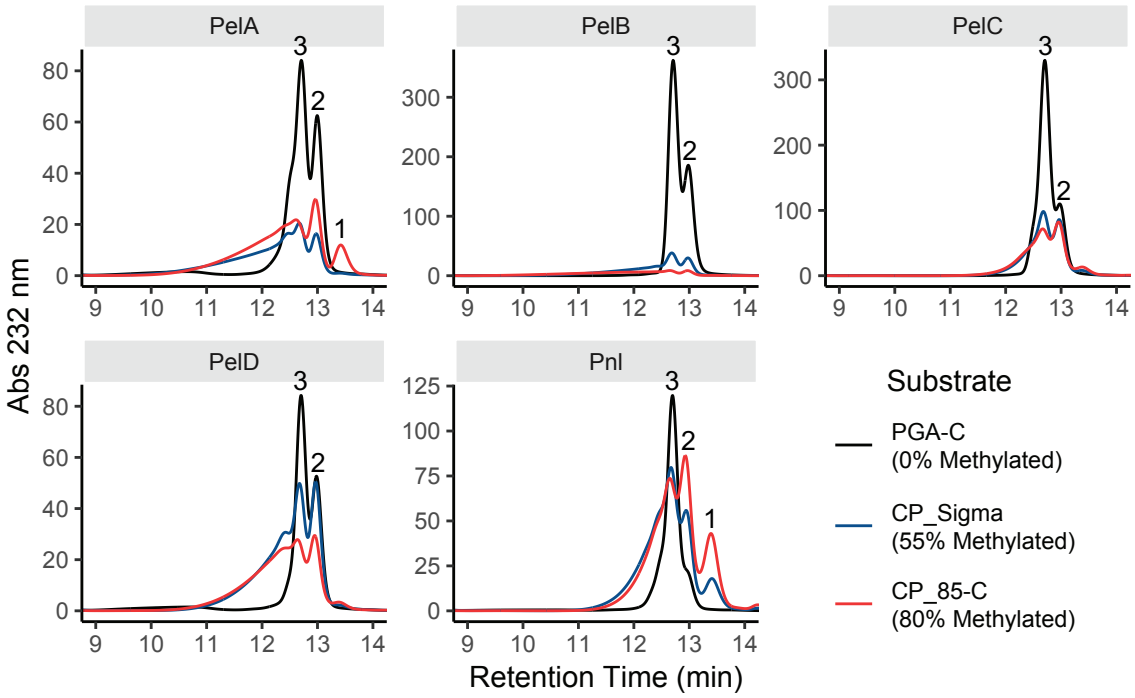


Figure 3.2 Analysis of products released by the action of each extracellular lyase on three different pectins. Reactions were allowed to run to completion then digestion products were separated with high-performance size-exclusion chromatography and detected as absorbance at 232 nm. Numbers above major peaks indicate the associated degree of polymerization and were determined using mono-, di-, and tetra-galacturonic acid standards.

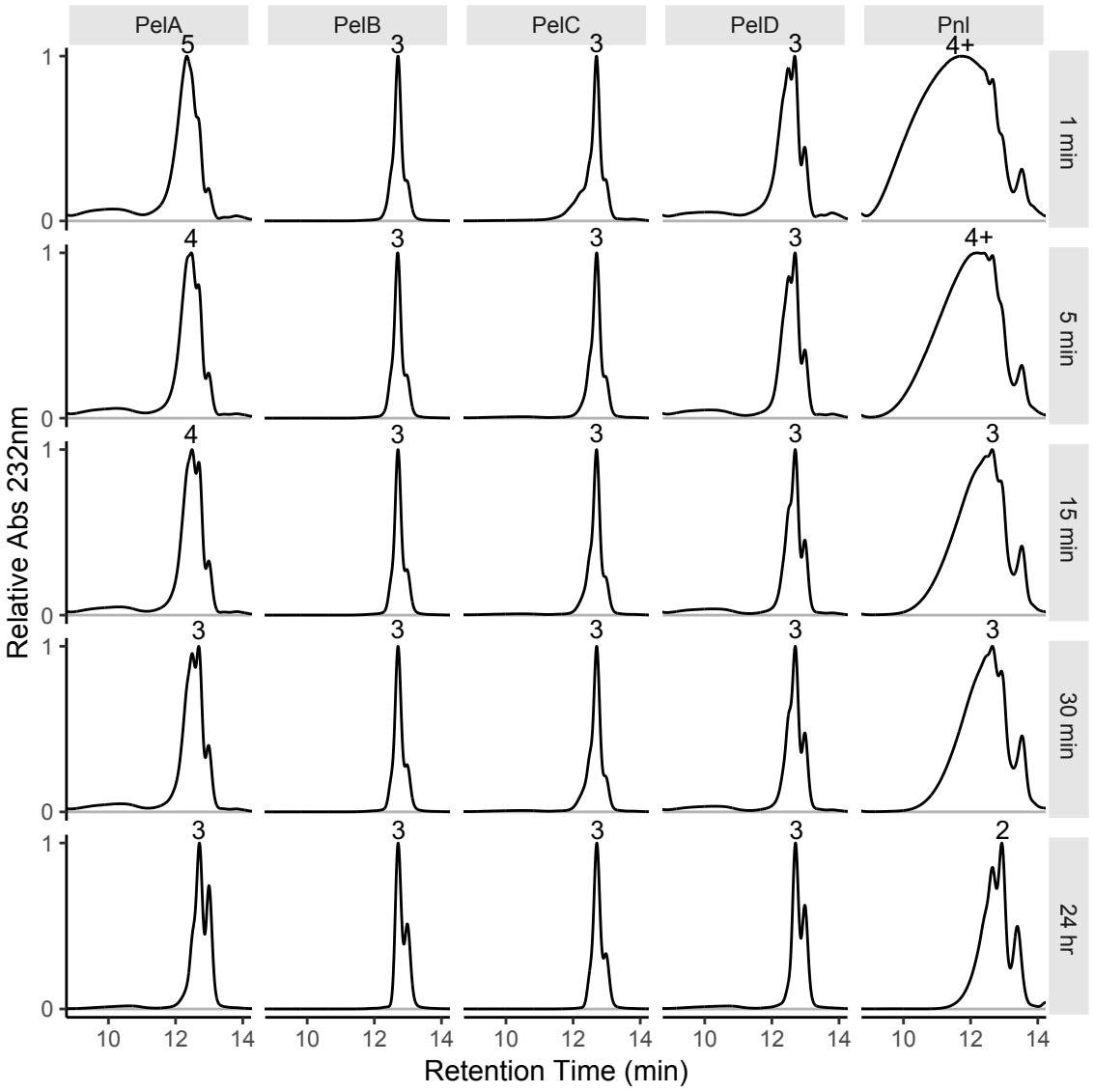


Figure 3.3 Mode of action of extracellular pectate lyases. Intermediate reaction products produced by the action of PelA, PelB, PelC, and PelD on polygalacturonic acid (0% methylated) or Pnl on 80% methylated citrus pectin were analyzed by high-performance size exclusion chromatography. Absorbance at 232 nm was used to detect products. All chromatograms were scaled to fill the plot area so that product distributions can be clearly visualized. Grey bars on the right indicate the timepoint in minutes. Peak labels indicate the degree of polymerization associated with the major product at a given timepoint and were determined by comparison to mono-, di-, and tetra-galacturonic acid standards. Grey arrows indicate mixtures of larger oligosaccharides while the black arrow is highlighting a shoulder of larger products.

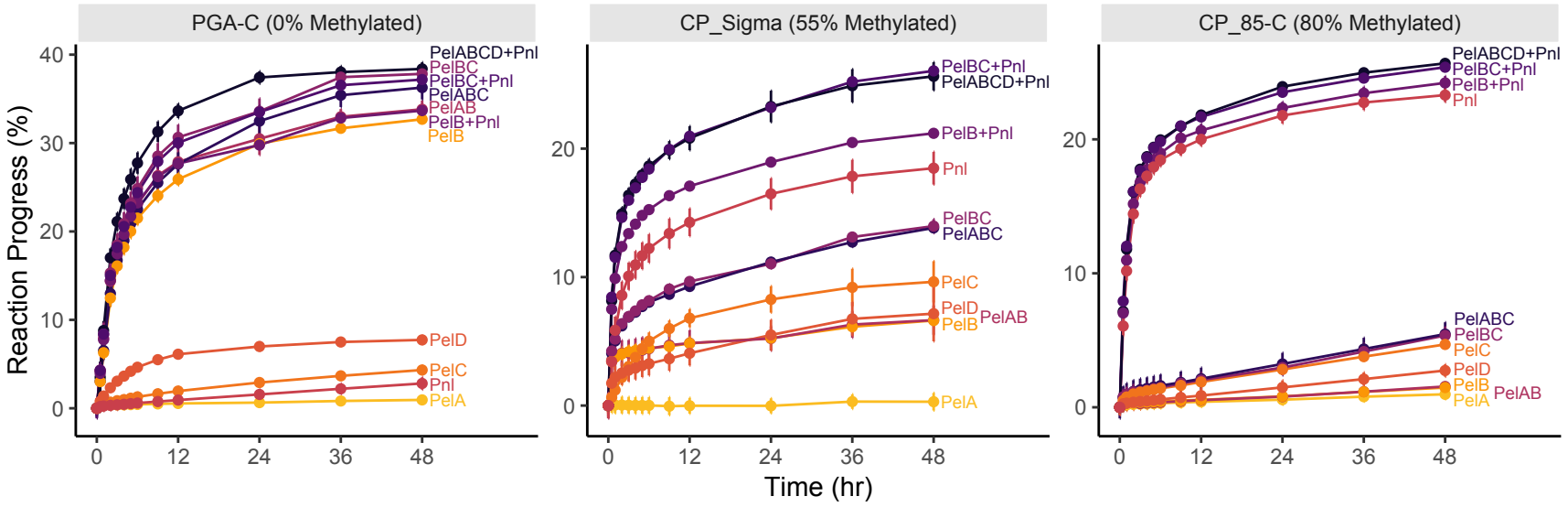


Figure 3.4 Effect of extracellular lyase combinations on the digestion of three different pectins. Each enzyme was included at a 1 nM concentration and the reactions were monitored by tracking the absorbance at 232 nm. The y-axis represents reaction progress as a percentage of the theoretical maximum absorbance that was calculated using the galacturonic acid content of each substrate and the molar extinction coefficient of the products. Thirty-three percent progress is the maximum achievable assuming all available galacturonic acid is broken down into a trimer.

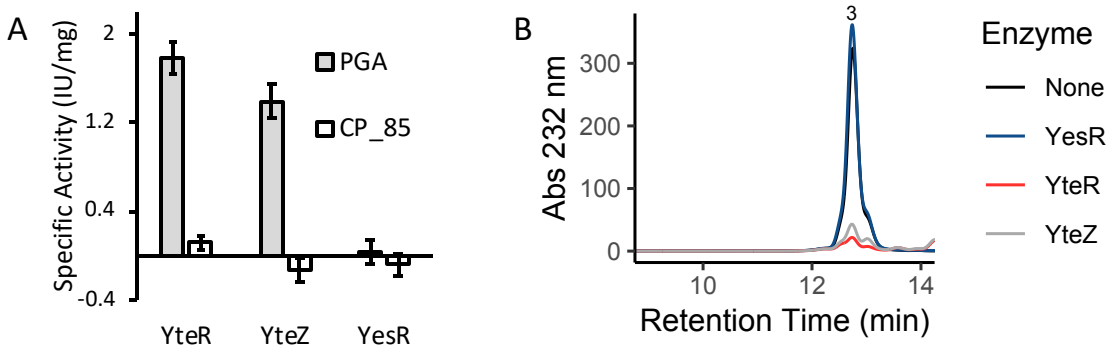


Figure 3.5 Activity of the three GH105 hydrolases on unsaturated oligogalacturonides. PGA (0% methylation) was digested with PelB and PelC while CP_85-C (80% methylation) was digested with Pnl to generate mixed oligogalacturonides. Digested substrates were heat inactivated and cooled before adding each GH105 enzyme. (A) Columns represent the specific activity of each enzyme, measured as the decrease in absorbance at 232 nm (n = 3). (B) High-performance size exclusion elution profiles of reactions allowed to run to completion were monitored at 232 nm as well. The number above the major peak indicates its degree of polymerization as determined by comparison to oligogalacturonide standards.

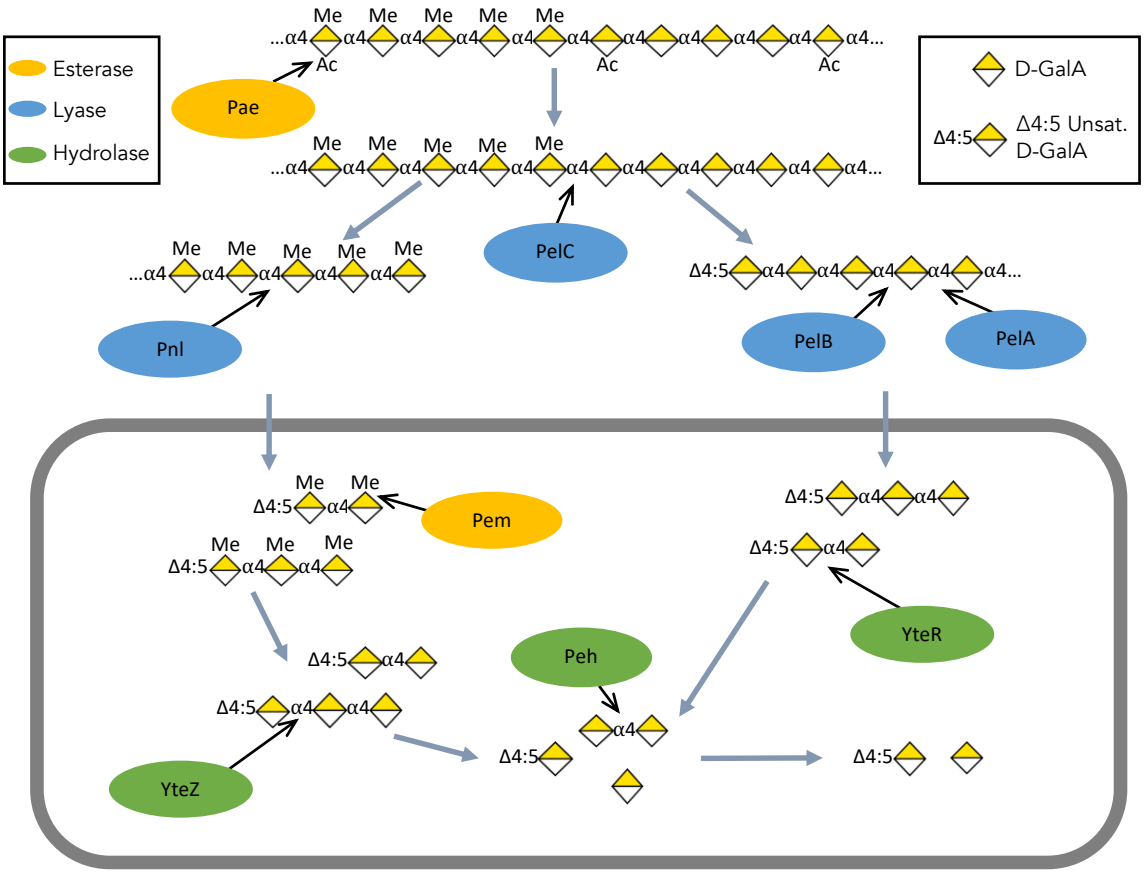


Figure 3.6 Emended model of homogalacturonan deconstruction in *Paenibacillus amylolyticus* 27C64. Black arrows indicate the sites targeted by each enzyme. Carbohydrates are represented using the standard nomenclature for glycans.

3.8 REFERENCES

1. Running SW, Nemani RR, Heinsch FA, Zhao M, Reeves M, Hashimoto H (2004) A continuous satellite-derived measure of global terrestrial primary production. Bioscience 54:547–560

2. Ding S, Himmel ME (2008) Anatomy and ultrastructure of maize cell walls: an example of energy plants. In: Himmel ME (ed) Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy. Blackwell, pp 38–60

3. Mohnen D, Bar-peled M, Somerville C (2008) Cell wall polysaccharide synthesis. In: Himmel ME (ed) Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy. Blackwell, pp 94–187

4. Tan L, Eberhard S, Pattathil S, Warder C, Glushka J, Yuan C, Hao Z, Zhu X, Avci U, Miller JS, Baldwin D, Pham C, Orlando R, Darvill A, Hahn MG, Kieliszewski MJ, Mohnen D (2013) An Arabidopsis cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. Plant Cell 25:270–287. doi:10.1105/tpc.112.107334

5. Popper ZA, Fry SC (2005) Widespread occurrence of a covalent linkage between xyloglucan and acidic polysaccharides in suspension-cultured angiosperm cells. Ann Bot 96:91–99. doi:10.1093/aob/mci153

6. Bunzel M, Ralph J, Steinhart H (2005) Association of non-starch polysaccharides and ferulic acid in grain amaranth (Amaranthus caudatus L.) dietary fiber. Mol Nutr Food Res 49:551–559. doi:10.1002/mnfr.200500030

7. Broxterman SE, Schols HA (2018) Interactions between pectin and cellulose in primary plant cell walls. Carbohydr Polym 192:263–272. doi:https://doi.org/10.1016/j.carbpol.2018.03.070

8. Meshitsuka G, Lee ZZ, Nakano J, Eda S (1982) Studies on the nature of lignin - carbohydrate bonding. J Wood Chem Technol 2:251–267. doi:10.1080/02773818208085134

9. O'Neill MA, Ishii T, Albersheim P, Darvill AG (2004) Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. Annu Rev Plant Biol 55:109–139. doi:10.1146/annurev.arplant.55.031903.141750

10. Mohnen D (2008) Pectin structure and biosynthesis. Curr Opin Plant Biol 11:266–277. doi:10.1016/j.pbi.2008.03.006

11. Abbott DW, Boraston AB (2008) Structural biology of pectin degradation by Enterobacteriaceae. Microbiol Mol Biol Rev 72:301–316. doi:10.1128/MMBR.00038-07

12. Kashyap DR, Vohra PK, Chopra S, Tewari R (2001) Applications of pectinases in the commercial sector: a review. Bioresour Technol 77:215–227. doi:10.1016/S0960-8524(00)00118-8

13. Hoondal G, Tiwari R, Tewari R, Dahiya N, Beg Q (2002) Microbial alkaline pectinases and their industrial applications: a review. Appl Microbiol Biotechnol 59:409–418. doi:10.1007/s00253-002-1061-1

14. Edwards MC, Doran-Peterson J (2012) Pectin-rich biomass as feedstock for fuel ethanol production. Appl Microbiol Biotechnol 95:565–75. doi:10.1007/s00253-012-4173-2

15. Lionetti V, Francocci F, Ferrari S, Volpi C, Bellincampi D, Galletti R, D'Ovidio R, De Lorenzo G, Cervone F (2010) Engineering the cell wall by reducing de-methyl-esterified homogalacturonan improves saccharification of plant tissues for bioconversion. Proc Natl Acad Sci 107:616–621. doi:10.1073/pnas.0907549107

16. Francocci F, Bastianelli E, Lionetti V, Ferrari S, De Lorenzo G, Bellincampi D, Cervone F (2013) Analysis of pectin mutants and natural accessions of Arabidopsis highlights the impact of de-methyl-esterified homogalacturonan on tissue saccharification. Biotechnol Biofuels 6:163. doi:10.1186/1754-6834-6-163

17. Tomassetti S, Pontiggia D, Verrascina I, Reca IB, Francocci F, Salvi G, Cervone F, Ferrari S (2015) Controlled expression of pectic enzymes in Arabidopsis thaliana enhances biomass conversion without adverse effects on growth. Phytochemistry 112:221–230. doi:10.1016/j.phytochem.2014.08.026

18. Biswal AK, Hao Z, Pattathil S, Yang X, Winkeler K, Collins C, Mohanty SS, Richardson EA, Gelineo-Albersheim I, Hunt K, Ryno D, Sykes RW, Turner GB, Ziebell A, Gjersing E, Lukowitz W, Davis MF, Decker SR, Hahn MG, Mohnen D (2015) Downregulation of GAUT12 in Populus deltoides by RNA silencing results in reduced recalcitrance, increased growth and reduced xylan and pectin in a woody biofuel feedstock. Biotechnol Biofuels 8:41. doi:10.1186/s13068-015-0218-y

19. Biswal AK, Soeno K, Gandla M, Immerzeel P, Pattathil S, Lucenius J, Serimaa R, Hahn MG, Moritz T, Jönsson LJ, Israelsson-Nordström M, Mellerowicz EJ (2014) Aspen pectate lyase PtxtPL1-27 mobilizes matrix polysaccharides from woody tissues and improves saccharification yield. Biotechnol Biofuels 7:11. doi:10.1186/1754-6834-7-11

20. Biswal AK, Atmodjo MA, Li M, Baxter HL, Yoo CG, Pu Y, Lee Y-C, Mazarei M, Black IM, Zhang J-Y, Ramanna H, Bray AL, King ZR, LaFayette PR, Pattathil S, Donohoe BS, Mohanty SS, Ryno D, Yee K, Thompson OA, Rodriguez M, Dumitrache A, Natzke J, Winkeler K, Collins C, Yang X, Tan L, Sykes RW, Gjersing EL, Ziebell A, Turner GB, Decker SR, Hahn MG, Davison BH, Udvardi MK, Mielenz JR, Davis MF, Nelson RS, Parrott WA, Ragauskas AJ, Neal Stewart C, Mohnen D (2018) Sugar release and growth of biofuel crops are improved by downregulation of pectin biosynthesis. Nat Biotechnol. doi: 10.1038/nbt.4067

21. Chung D, Pattathil S, Biswal AK, Hahn MG, Mohnen D, Westpheling J (2014) Deletion of a gene cluster encoding pectin degrading enzymes in Caldicellulosiruptor bescii reveals an important role for pectin in plant biomass recalcitrance. Biotechnol Biofuels 7:147. doi:10.1186/s13068-014-0147-1

22. Kohli P, Gupta R (2015) Alkaline pectinases: a review. Biocatal Agric Biotechnol 4:279–285. doi:10.1016/j.bcab.2015.07.001

23. Cook DM, Henriksen ED, Upchurch R, Doran-Peterson JB (2007) Isolation of polymer-degrading bacteria and characterization of the hindgut bacterial community from the detritus-feeding larvae of Tipula abdominalis (Diptera: Tipulidae). Appl Environ Microbiol 73:5683–6. doi:10.1128/AEM.00213-07

24. Boland WE, DeCrescenzo Henriksen E, Doran-Peterson J (2010) Characterization of two Paenibacillus amylolyticus strain 27C64 pectate lyases with activity on highly methylated pectin. Appl Environ Microbiol 76:6006–6009. doi:10.1128/AEM.00043-10

25. Keggi C, Doran-Peterson J (2019) Paenibacillus amylolyticus 27C64 has a diverse set of carbohydrate-active enzymes and complete pectin deconstruction system. J Ind Microbiol Biotechnol 46:1–11. doi:10.1007/s10295-018-2098-1

26. Jayani RS, Saxena S, Gupta R (2005) Microbial pectinolytic enzymes: a review. Process Biochem 40:2931–2944. doi:10.1016/j.procbio.2005.03.026

27. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:490–495. doi:10.1093/nar/gkt1178

28. Soriano M, Diaz P, Pastor FIJ (2006) Pectate lyase C from Bacillus subtilis: a novel endo-cleaving enzyme with activity on highly methylated pectin. Microbiology 152:617–25. doi:10.1099/mic.0.28562-0

29. Li X, Wang H, Zhou C, Ma Y, Li J, Song J (2014) Cloning, expression and characterization of a pectate lyase from Paenibacillus sp. 0602 in recombinant Escherichia coli. BMC Biotechnol 14:18. doi:10.1186/1472-6750-14-18

30. Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8:785–6. doi:10.1038/nmeth.1701

31. Käll L, Krogh A, Sonnhammer ELL (2004) A combined transmembrane topology and signal peptide prediction method. J Mol Biol 338:1027–1036. doi:10.1016/j.jmb.2004.03.016

32. Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ, Brinkman FSL (2010) PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics 26:1608–15. doi:10.1093/bioinformatics/btq249

33. Li MZ, Elledge SJ (2007) Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. Nat Methods 4:251–256. doi:10.1038/nmeth1010

34. Inoue H, Nojima H, Okayama H (1990) High efficiency transformation of Escherichia coli with plasmids. Gene 96:23–28. doi:0378-1119(90)90336-P [pii]

35. Studier FW (2005) Protein production by auto-induction in high-density shaking cultures. Protein Expr Purif 41:207–234. doi:10.1016/j.pep.2005.01.016

36. Kapust RB, Tozser J, Copeland TD, Waugh DS (2002) The P1' specificity of tobacco etch virus protease. Biochem Biophys Res Commun 294:949–955. doi:10.1016/S0006-291X(02)00574-0

37. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685. doi:10.1038/227680a0

38. Collmer A, Ried JL, Mount MS (1988) Assay methods for pectic enzymes. Methods Enzymol 161:329–335. doi:https://doi.org/10.1016/0076-6879(88)61037-8

39. Itoh T, Ochiai A, Mikami B, Hashimoto W, Murata K (2006) A novel glycoside hydrolase family 105: the structure of family 105 unsaturated rhamnogalacturonyl hydrolase complexed with a disaccharide in comparison with family 88 enzyme complexed with the disaccharide. J Mol Biol 360:573–585. doi:10.1016/j.jmb.2006.04.047

40. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM (ed) The Proteomics Protocols Handbook. Humana Press, Totowa, NJ, pp 571–607

41. Voragen A, Schols HA, Pilnik W (1986) Determination of the eegree of methylation and acetylation of pectins by HPLC. Food Hydrocoll 1:65–70. doi:10.1016/S0268-005X(86)80008-X

42. Blumenkrantz N, Asboe-Hansen G (1973) New method for quantitative determination of uronic acids. Anal Biochem 54:484–489. doi:https://doi.org/10.1016/0003-2697(73)90377-1

43. White GW, Katona T, Zodda JP (1999) The use of high-performance size exclusion chromatography (HPSEC) as a molecular weight screening technique for polygalacturonic acid for use in pharmaceutical applications. J Pharm Biomed Anal 20:905–912. doi:10.1016/S0731-7085(99)00083-7

CHAPTER 4

A MULTIDOMAIN PECTINASE FROM *PAENIBACILLUS AMYLOLYTICUS* 27C64 HAS BOTH RHAMNOGALACTURONAN LYASE AND ACETYLESTERASE ACTIVITIES [1]

[1] Keggi, C and J Doran-Peterson. To be submitted to *Applied and Environmental Microbiology*.

4.1 ABSTRACT

Paenibacillus amylolyticus 27C64, a Gram-positive bacterium originally isolated from the hindgut of a *Tipula abdominalis* larvae, has previously shown promise as a source of novel pectinases. This work first characterizes the function of a multi-domain pectinase, RglA, identified in the organism's genome. This sequence includes a rhamnogalacturonan acetylesterase domain, a rhamnogalacturonan lyase domain, and eight fibronectin type III like domains, five of which are probable carbohydrate binding domains. RglA was expressed in *E. coli* and purified for characterization. Both rhamnogalacturonan lyase and acetylesterase activities were highest under mildly alkaline conditions, and calcium was required for lyase activity. Response surface methodology was used to identify pH 7.75 and 38°C as the best condition when considering the activity of both domains. Evaluation of the lyase activity on 17 commercial pectins revealed that the enzyme had significantly higher activity on acetylated substrates, but lower activity on highly methylated ones. These features make this enzyme notably different from other published rhamnogalacturonan lyases.

4.2 IMPORTANCE

Pectin is an important component of plant cell walls, where it provides structural stability and is responsible for cell-cell adhesion. For this reason, pectinases can benefit many industrial processes that break down plant material. This includes fruit juice processing, textile fiber manufacturing, coffee and tea fermentation, and advanced biofuel production from lignocellulosic biomass. Discovery of new and different pectinases has the potential to improve these applications, especially when new enzymes target types of

pectin that are otherwise difficult to break down. In this work, we describe one such enzyme, RglA, which is able to degrade the acetylated rhamnogalacturonan I regions of pectin that are usually poor substrates for other rhamnogalacturonan lyases. We also describe the complex multidomain structure RglA and note that no enzyme with a similar structure has been previously characterized.

4.3 INTRODUCTION

Pectin is the most structurally complex plant cell wall polysaccharide and it has several critical functional roles in plants [1]. In the middle lamella, the outermost layer of the cell wall, pectin is the major component and is responsible for cell-cell adhesion [1–3]. In the primary cell wall, which lies directly beneath the middle lamella, pectin is an important structural polysaccharide [1–3] and forms covalent linkages to proteins [4], matrix polysaccharides [5, 6], lignin [7], and cellulose [8]. The two most abundant members of this family of galacturonic acid (GalA) rich polysaccharides are homogalacturonan and rhamnogalacturonan I (RG-I) [1]. While homogalacturonan has a linear GalA backbone and is free of sidechains, RG-I has an alternating GalA – rhamnose (Rha) backbone based on a repeating (-α-D-Gal*p*A-1,2-α-L-Rha*p*-1,4-) disaccharide [1]. In general, 20-80% of the rhamnosyl residues in this type of pectin have arabinose and galactose rich sidechains [1]. The GalA residues of either homogalacturonan or RG-I may be acetylated at the O-2 or O-3 positions, and homogalacturonans are often extensively methylated at the C-6 carboxyl [1].

Historically, industrial interest in deconstruction of pectin by microbial pectinases has stemmed from the ability of homogalacturonan-deconstruction enzymes to improve

yields in fruit juice extraction processes, clarify the final product, and reduce its viscosity [9]. Apple juice, however, contains a significant amount of RG-I that is retained after treatment with HG-deconstruction enzymes and ultrafiltration [10]. This led to the discovery of the first RG-I deconstruction enzyme [11]. However, this enzyme along with other known rhamnogalacturonan hydrolases and rhamnogalacturonan lyases (RGL), are not able to cleave acetylated sites in RG-I [12] and therefore require a separate enzyme to deacetylate the substrate. Only one such rhamnogalacturonan acetylesterase (RGAE) has been thoroughly characterized using real pectic substrates [13–15], although other enzymes have been characterized on model substrates. Identification of RG lyases able to act directly on acetylated RG-I could help improve the efficiency of pectin deconstruction in applications where acetylation is common.

Development of more efficient pectin deconstruction systems has the potential to impact many industries in addition to fruit juice production. Pectinases are useful for removing pulp from coffee cherries, degumming and retting plants to produce textile fibers, tea fermentation, biobleaching of paper pulp, and bioscouring of cotton fibers [9, 16, 17]. In addition, efficient pectin deconstruction is important for the production of second-generation biofuels from sustainable non-food crops. Pectinases are necessary to achieve high ethanol yields when using pectin-rich agricultural wastes as a feedstock [18, 19] and also for the growth of an organism that can use whole plant material, including grass and poplar wood, as a carbon source [20]. It has also been shown that modifying pectin to reduce its recalcitrance to deconstruction can improve sugar release from lignocellulosic plant material even when pectin is a very minor component of the biomass overall [21–26].

Interest in furthering these industrial processes has largely driven attempts to better understand microbial pectin degradation including RG-I deconstruction.

One organism with the potential to improve the current understanding of RG-I deconstruction is *Paenibacillus amylolyticus* 27C64. Two pectate lyases have already been characterized in this Gram-positive bacterium [27] and analysis of its genome revealed that it has more to contribute [28]. One of the most promising features of the genome was the presence of a putative bifunctional RGL and RGAE. This enzyme, RglA, may be able to act directly on acetylated RG-I. This study seeks to understand the multidomain architecture of this sequence, confirm its bifunctional nature, and characterize each of its catalytic activities.

4.4 RESULTS

4.4.1 Sequence Analysis

Previous work exploring the CAZome of *P. amylolyticus* 27C64 identified *rglA*, predicted a signal peptide, and assigned the sequence to two CAZy families: CE12 and PL11 [28]. In this work, we conducted a more detailed assessment of the multi-domain structure of the enzyme (Figure 4.1). A search for conserved domains identified hits to catalytic protein families associated with RG lyase and RG acetylesterase activities as well as eight non-overlapping regions with homology to different fibronectin type III like (FNIII) domains (Figure 4.1). These domains have been assigned single-letter names based on similarity to each other. The Y1 and Y2 domains are so named because they share 37.2% and 30.4% amino acid sequence identity, respectively, with the putative CBM from a *B. subtilis* pectin acetylesterase, YxiM [29]. Names for the C1 and C2 domains are derived

from the fact that their top CD-Search hit was to an E_set superfamily domain (cd02848) typically associated with chitinases. The three M domains and the E domain stand simply for ‘middle’ and ‘end’ because of their relative positions in the sequence. These domains only returned hits to general FNIII superfamily domains and not to specific families associated with a particular catalytic activity.

The sequence regions associated with each catalytic domain were aligned to all characterized sequences within their assigned CAZy family (Appendix C, Table C1). The closest homologue to the RGL domain was YesW, an endo-RG lyase from *Bacillus licheniformis*, which shared 62.6% identity and 73.7% similarity. The two best characterized bacterial RG lyases, YesW and YesX from *B. subtilis*, were the next most related sequences with 61.2% and 57.5% identity respectively. The RGAE domain’s closest match was Rgae12A from *Hungateiclostridium thermocellum* at 50.2% identity and 62.5% similarity. The five characterized CE12 esterases from *Bacillus* were the next closest with 36.3-42.1% identity.

Aligning each domain sequence to the CAZy database provided some insight into the possible roles of each domain. Virtually all of the hits returned for the RGAE and RGL domains belonged to CE12 and PL11 as expected (Figure 4.1). Y1 and Y2 returned primarily hits to sequences from CE12. E, C1, and C2 domains returned results including many sequences from GH18. In contrast, hits to M1, M2, and M3 spanned a much wider range of domains; no one catalytic family was clearly associated with all three.

4.4.2 Purification

After all purification steps including tag removal, a total of 4.22 mg of protein was recovered. RG lyase specific activity was purified 534-fold from the lysate (Appendix C,

Table C2). The final sample contained some smaller fragments that could not be separated. These are likely either early termination products or the N-terminal portions of degraded full-length protein, as the bands shift downward on an SDS-PAGE gel after removal of the fusion tag. The magnitude of this shift corresponds to the known tag size (Appendix C, Figure C1).

4.4.3 Optimum Reaction Conditions

A one factor at a time approach was used to characterize the permissive pH, temperature, and calcium ranges for each catalytic activity, and each activity had a different optimum. RGL activity was highest at pH 8.5, 57.5°C, with 10-20 mM $CaCl_2$ (Figure 1a). EDTA completely inhibited this activity, and magnesium could not replace calcium. RGAE activity was highest at pH 7-9 and 35°C (Figure 1a). Esterase activity was not dependent on calcium, magnesium, or EDTA.

Response surface models built to describe both activities as a function pH (from 7-9) and temperature (from 30-50°C) accurately fit the data: there were no confounded model terms and lack of fit tests were insignificant. An absolute optimum for the RGAE domain was identified at pH 7.75 and 38°C. RGL activity, as expected, did not reach an absolute maximum within the characterized space, but had a local optimum at an axial point (pH 8, 45°C).

4.4.4 Substrate Specificity

Substrate specificity was evaluated by testing the lyase activity of RglA on 17 commercial pectins. The two substrates on which RglA had the highest activity were the two rhamnogalacturonan substrates, RG-I from potato, and rhamnogalacturonan from soy, both supplied by Megazyme (County Wickow, Ireland). The linear model describing

specific activity as a function of each pectin's DM, DA, M_n, and GalA content indicated that DM had a negative effect on activity (Figure 4.3a; $t(46) = -11.8$, $p = 1.6 \times 10^{-15}$), DA had a positive effect on activity (Figure 4.3b; $t(46) = 4.63$, $p = 3.0 \times 10^{-5}$), and M_n had a negative effect ($t(46) = -3.09$, $p = 0.0034$). GalA content did not affect activity ($p > 0.05$). The overall model had an adjusted R^2 value of 0.822.

4.5 DISCUSSION

4.5.1 Domain Identification and Organization

The two catalytic domains of RglA are arranged with the CE12 RGAE domain closer to the N-terminus of the protein and the PL11 RGL domain towards the C-terminus (Figure 4.1). The sequences of each domain are largely similar to the characterized sequences within their respective CAZy families. The RGL domain had the highest percent identity (62.6%) to YesW, a thermostable endo-RGL from *B. licheniformis* that has been characterized [30, 31]. YesW from *B. subtilis* is the second most similar characterized sequence (61.2% identity), and it too is an endo-RGL [32]. The RGAE domain was likewise very similar to several characterized CE12 esterases. The most similar sequence, Rgae12A from *H. thermocellum* (50.2% identity), has not be catalytically characterized but rather the specificity of an associated CBM was determined (which was not the source of the alignment to RglA) [33]. Many of the other similar sequences have not been thoroughly characterized and have not been tested with RG-I as a substrate. However, significant (30%) similarity to the very well characterized RGAE from *Aspergillus aculeatus,* Rha1, [13–15] is noted despite its taxonomic divergence. The four blocks of residues typically conserved in SNGH hydrolases like Rha1, including the key catalytic

triad residues, are conserved in RglA as well (Appendix C, Figure C2). Collectively, this sequence-based evidence suggests that the novelty of RglA lies not in the activity of its catalytic domains, but rather in the assembly of these domains into a bifunctional enzyme with eight non-catalytic FNIII-like regions.

Only two of the eight non-catalytic domains, Y1 and Y2, are homologous to a domain from a known pectinase. These domains are similar to the putative N-terminal CBM associated with a characterized pectin acetylesterase from *B. subtilis* [29]. Paired with the fact the vast majority of similar sequences in the CAZy database are associated with CE12 catalytic domains (Figure 4.1), it is likely that these regions are functionally important and act as CBMs. Given the association with acetylesterase domains specifically, it is possible that Y1 and Y2 are specific to an acetylated substrate. The E, C1, and C2 domains are also probable CBMs because most CAZy homologues are from GH18, a family that contains hydrolases specific to acetylated polysaccharides other than pectins (Figure 4.1) [34]. At least one of these related FNIII-like sequences from a GH18 chitinase has a demonstrated role in substrate binding [35]. In contrast to these CBM-like domains, the M1, M2, and M3 domains had homology to FNIII-like sequences scattered across numerous CAZy families with unrelated catalytic activities (Figure 4.1). The only consistent theme in homologous sequences was the presence of the FNIII-like primary fold, suggesting that these structures are not involved in substrate binding. They may instead be linkers responsible for maintaining the necessary spacing between the catalytic domains.

4.5.2 Optimum Reaction Conditions

Both activities present in RglA displayed reaction requirements typical of homologous enzymes; an alkaline optimum pH was preferred, and the lyase domain had a

strict calcium requirement but the acetylesterase domain did not. However, when the best reaction temperature was evaluated at pH 8.5, the optimum for the two activities differed by 22.5°C (Figure 4.1a). Modeling the activity of each domain from pH 7-9 at 30-50°C resulted in the identification of a new optimum at a pH of 7.75 and a temperature of 38°C with 5.7% higher RGAE activity (Figure 4.1c). RGL activity at this new optimum was 48% higher than at the optimum identified with the one factor at a time approach (Figure 4.1b).

4.5.3 Substrate Specificity

The average molecular weight, DM, and DA of pectins all had a significant impact on the reaction rate of RglA, but the effects of DM and DA were the most pronounced. Substrate M_n had only a small effect. In the case of acetylation, increasing DA had a positive effect on activity (Figure 4.3b). This is consistent with the presence of five putative CBM domains that are homologous to related domains in enzymes that act on acetylated polysaccharides. DM also had a significant, but negative, impact on activity (Figure 4.3a). The mechanism underlying this negative effect is less clear since a high DM value indicates highly methylated homogalacturonan, the major component of pectins, not highly methylated RG-I. While it is possible that the methylation is affecting activity directly, it may also be the case that extraction methods used to produce highly methylated pectins are extracting more homogalacturonan than RG-I, or that some other substrate characteristic that affects RglA activity is also correlated with DM.

4.5.4 Conclusion

This work has demonstrated that RglA indeed has both the rhamnogalacturonan lyase and rhamnogalacturonan acetylesterase activities previously predicted based on analysis of its sequence, and that both catalytic domains are homologous to other

characterized enzymes within their respective CAZy families. These domains were also biochemically similar to known enzymes. The acetylesterase domain has an alkaline pH optima similar to at least two other known RGAE enzymes from *Bacillus* [36, 37] and did not require calcium. The RGL domain likewise worked best in a mildly alkaline pH range and had a strict calcium requirement, just as three homologous *Bacillus* RG lyases do [30, 32]. In contrast, the bifunctional nature of the enzyme is unique. None of the PL11 or PL4 sequences listed as characterized in the CAZy database are also members of CE12 [34]. The presence of eight non-catalytic domains is also different from known enzymes. Although at least one CE12 pectin acetylesterase [29], two CE12 RG acetylesterases [33, 37], and one PL4 RG lyase [38] have been described with CBM domains, the presence of five putative CBMs along with three spacer domains is unusual. The assembly of the two catalytic domains into this complex structure results in an enzyme with more activity on acetylated pectins, substrates that are inhibitory to other RG lyases [12]. This ability may be helpful for improving industrial processes that require the deconstruction or removal of highly acetylated pectins, and future work should explore the utility of this enzyme in such applications.

4.6 METHODS

4.6.1 Sequence Analysis

Signal peptide prediction and CAZy family assignment of *pamy_4308*, here renamed *rglA* for clarity, was performed previously [28]. Pairwise comparison to other CAZy sequences was performed using EMBOSS Matcher, while multiple sequence alignments were performed with MUSCLE [39]. Domain locations were determined by

manually reviewing conserved domain search results [40] and the structures of each predicted domain by I-TASSER [41]. Domain sequences were compared to sequences from the CAZy database using the blast+ suite of applications [42].

4.6.2 Strains and Culture Conditions

P. amylolyticus 27C64, previously isolated from an insect hindgut [43], was cultured at 37°C on solid tryptic soy agar (15 g tryptone, 5 g soytone, 5 g NaCl, 15 g agar per liter) or at 37°C in tryptic soy broth without dextrose (17 g tryptone, 3 g soytone, 5 g NaCl, 2.5 g NaK_2PO_4 per liter). *E. coli* was grown in LB broth (10 g tryptone, 5 g yeast extract, 10 g NaCl per liter) or LB agar (also had 15 g agar per liter) at 37°C unless noted otherwise.

4.6.3 Cloning

Genomic DNA was isolated from *P. amylolyticus* 27C64 using a Wizard Genomic DNA Purification Kit (Promega, Madison, WI) and was used as template to amplify *rglA* with high-fidelity Phusion polymerase (New England Biolabs, Ipswich, MA). Oligonucleotide primers (Integrated DNA Technologies, Coralville, IA) targeted the entire gene without its signal peptide. PCR products were inserted into a pET-21a based plasmid with a tobacco etch virus (TEV) protease cleavable fusion tag comprised of 8x His tag and *Pyrococcus furiosis* rubredoxin sequence. Plasmid construction was performed using sequence and ligation independent cloning as previously described [44] except that gel purification of the PCR products was used instead of DpnI digestion and spin-column purification. Plasmid constructs were transformed into *E. coli* DH5α. Complete plasmids were isolated using the ZR Classic plasmid purification kit (Zymo Research, Irvine, CA) and sequenced to confirm the insert (Eton Bioscience, Research Triangle, NC).

4.6.4 Protein Expression

A two-liter glass BioStat A vessel (Sartorius, Göttingen, Germany) was autoclaved with 1.6 L of Davis minimal media (7 g K_2HPO_4, 2 g KH_2PO_4, 1 g $(NH_4)_2SO_4$, 1 g glucose, 0.5 g sodium citrate, 0.1 g $MgSO_4 \bullet 7H_2O$ per liter). After cooling, the pH was adjusted to 6.8, ampicillin was added to a concentration of 100 μg/mL, and 1X trace metals were added. Trace metals were made as previously described [45]. The media was inoculated with 16 mL of *E. coli* BL21(DE3) carrying the plasmid with the tagged full-length *rglA* gene and allowed to grow overnight at 25°C, pH 6.8, with 20% dissolved oxygen and constant stirring at 1000 rpm. The pH was adjusted automatically with 15% w/v ammonium hydroxide. After overnight growth, the temperature was adjusted to 37°C and an 8-hour exponential feed was started (feed solution: 500 g glucose, 12 g $MgSO_4 \bullet 7H_2O$ per liter with 1X trace metals, 100 μg/mL ampicillin). At the end of the feed window the temperature was lowered to 20°C and IPTG was added to a final concentration of 1 mM for induction along with 12.5 mL of 32% w/v yeast extract. Expression was allowed to continue overnight. Cells were harvested by centrifugation and frozen at -20°C until needed.

4.6.5 Protein Purification

Five grams of cell paste was thawed and resuspended in five milliliters of ice-cold equilibration buffer (50 mM phosphate pH 7.6, 300 mM NaCl) per gram of cell paste. Complete Ultra protease inhibitor with EDTA (Sigma-Aldrich, St. Louis, MO) was added to the recommended concentration. Cells were disrupted by sonication and ammonium sulfate was added to 15% saturation. The lysate was spun down for 30 minutes at 23,000 x g in a Beckman (Brea, CA) floor centrifuge, the clear supernatant was recovered, and

additional ammonium sulfate was added to achieve 25% saturation. The precipitated protein was pelleted for five minutes at 23,000 x g, resuspended in 10 mL of ice-cold equilibration buffer containing Complete Ultra protease inhibitor without EDTA (Sigma-Aldrich, St. Louis, MO), and filtered through a 0.45 μM PES membrane (MilliporeSigma, Burlington, MA). Elution buffer (50 mM phosphate pH 7.6, 300 mM NaCl, 250 mM imidazole, 10% glycerol) was added to a concentration of 4% by volume.

Lysate was bound to 50 mL of Ni-NTA resin (Sigma-Aldrich, St. Louis, MO) in a GE XK50/30 column using an AKTA Prime purification system (GE Life Sciences, Chicago, IL). The sample was washed with 250 mL of 4% v/v elution buffer, eluted with 100% elution buffer, and dialyzed into 10 mM Tris pH 7.5 with 10 mM NaCl. After recovery from dialysis tubing, a tag cleavage reaction was set up with supplemental Tris (buffer strength increased to 50 mM pH 7.5), 1 mM tris(2-carboxyethyl)phosphine, and the recommended amount of tobacco etch virus protease [46]. The tag cleavage reaction was allowed to proceed overnight at 4°C then the uncut protein and free tag were removed from the sample by passing it over Ni-NTA resin equilibrated with 50 mM phosphate buffer pH 8 with 300 mM NaCl. The final sample was dialyzed exhaustively into 25 mM Tris pH 7.2 with 25 mM NaCl, recovered, and supplemented with 10% glycerol by volume. Protein concentration was determined by measuring the absorbance at 280 nm with a NanoDrop Lite (ThermoFisher, Waltham, MA). All purification samples were evaluated for purity with SDS-PAGE as previously described [47].

4.6.6 Enzyme Assays

Rhamnogalacturonan lyase activity was determined as previously described [32]. Briefly, 1 mL of substrate solution containing 0.5 mg/mL RG-I from potato (Megazyme,

County Wicklow, Ireland) or other substrate as indicated was placed in a cuvette and 5 µL of enzyme was added directly from ice to start the reaction. The increase in absorbance at 232 nm was monitored continuously with a Cary 60 spectrophotometer (Agilent, Santa Clara, CA). Reactions at elevated temperatures were allowed to pre-equilibrate for 5 minutes in a temperature-controlled cuvette holder. Activity in international units (IU; 1 µmol / min) were determined using the published molar extinction coefficient of the product [32]. Substrate solutions to determine the optimum pH included 2 mM $CaCl_2$ in a 100 mM buffer. Acetate buffers were used for pH 4-6, MOPS for pH 6-7, Tris for pH 7-9, and Glycine for 9-11. These assays were conducted at room temperature. Optimum calcium and temperature testing used a 50 mM Tris buffer at pH 8.5, with calcium assays also occurring at room temperature.

RG acetylesterase activity was measured using the model colorimetric substrate *p*-nitrophenol acetate (pNPA). Reactions were set up using the same buffers as for RG lyase assays except no calcium was included. Ten microliters of 100 mM pNPA in methanol was kept on ice and added to the reaction immediately before beginning the assay along with the five microliters of enzyme. Activity was monitored as the increase in absorbance at the isosbestic point (348 nm) where the molar extinction coefficient did not depend on the ratio of *p*-nitrophenol to *p*-nitrophenolate. The molar extinction coefficient of *p*-nitrophenol was determined to be 5441 M^{-1} cm^{-1}, while pNPA has a molar extinction coefficient of 382.2 M^{-1} cm^{-1} at this wavelength. Product release was calculated using these molar extinction coefficients in the Beer-Lambert law. The total absorbance was equal to the sum of substrate and product absorbances assuming the concentration of remaining substrate was equal to the initial substrate concentration less the amount of product released. Since pNPA

is unstable in aqueous solutions, no-enzyme controls were collected in triplicate for each condition used to correct raw slope values.

4.6.7 Modeling Optimum Reaction Conditions

Response surface methodology was used to optimize the activity of both enzymatic domains simultaneously from pH 7-9 (in Tris buffer) at a temperature of 30-50°C. JMP 14 (SAS, Cary, NC) was used to generate a face-centered central composite experimental design with two replicates of the center point. Samples were run in the random order generated by the software, and a second set of replicates was added as a block. The model included all first order terms and interactions, second order temperature and pH terms, interaction of second order terms (pH^2 * temperature, $temperature^2$ * pH), and a block term.

4.6.8 Analysis of Pectic Substrates

Degree of methylation (DM) and degree of acetylation values (DA) of commercial pectins were determined as previously described [48]. Briefly, pectin samples were saponified in an isopropanol – NaOH solution for one hour, the insoluble pectin was pelleted, and filtered supernatant was separated by high performance liquid chromatography with an Aminex HPX-87H column (BioRad, Hercules, CA). External methanol and acetic acid standards were used for quantification. Percent methylation or acetylation was calculated as the molar ratio of methanol or acetate to GalA as determined by a colorimetric uronic acid assay [49]. This colorimetric assay was scaled down linearly to fit a 300 µL PCR plate, and samples were transferred to a microtiter plate to read the absorbance with a plate reader. GalA was used as the external standard. Molecular weight values of each substrate were determined with high-performance size exclusion chromatography with refractive index detection. A simple isocratic method relying on a

Tosoh TSKgel G3000PW$_{XL}$ column and guard column was used as previously described [50]. Shodex P-82 narrow-polydispersity pullulan samples (Showa Denko, Tokyo, Japan) were used as molecular weight standards.

4.6.9 Substrate Specificity

The ability of RglA to deconstruct 17 different commercial pectins was evaluated using the continuous assay described in section 4.5.6 'Enzyme Assays.' A linear model describing RGL activity as a function of the DM, DA, number-average molecular weight (M_n), and GalA content of each substrate was built using R.

4.7 ACKNOWLEDGMENTS

The authors would like to thank the staff at the University of Georgia Bioexpression and Fermentation Facility, especially Dr. David Blum and Dr. Michelle Lewis, for their assistance with protein expression and purification.

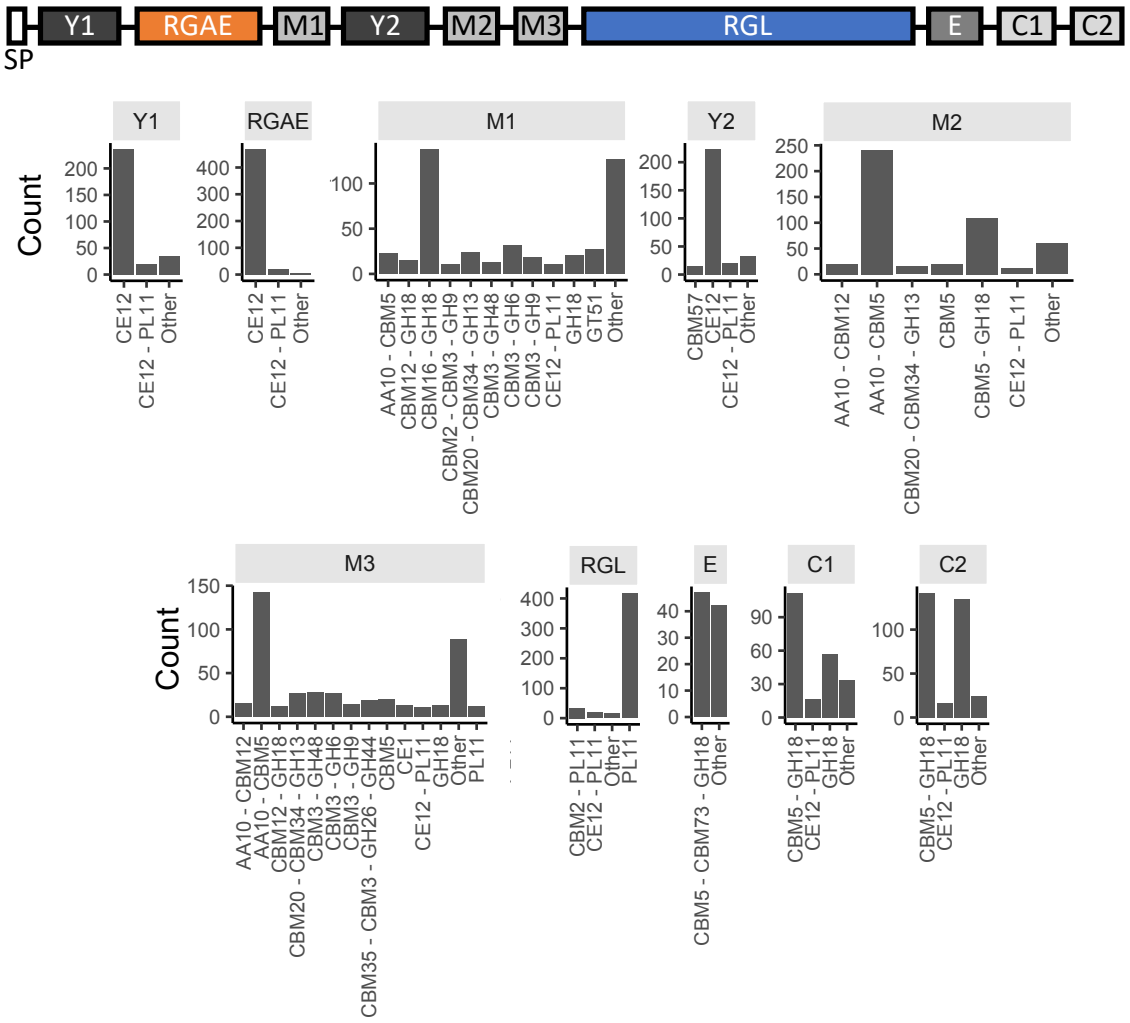


Figure 4.1 Multidomain structure of RglA. Schematic of domain order is drawn to scale. Plots summarize the results of protein blast+ searches of each domain sequence against the CAZy database. Up to 500 hits were returned per sequence. Bars indicate the number of hits returned to each CAZy family. *RGAE* - RG acetylesterase. *RGL* - RG lyase.

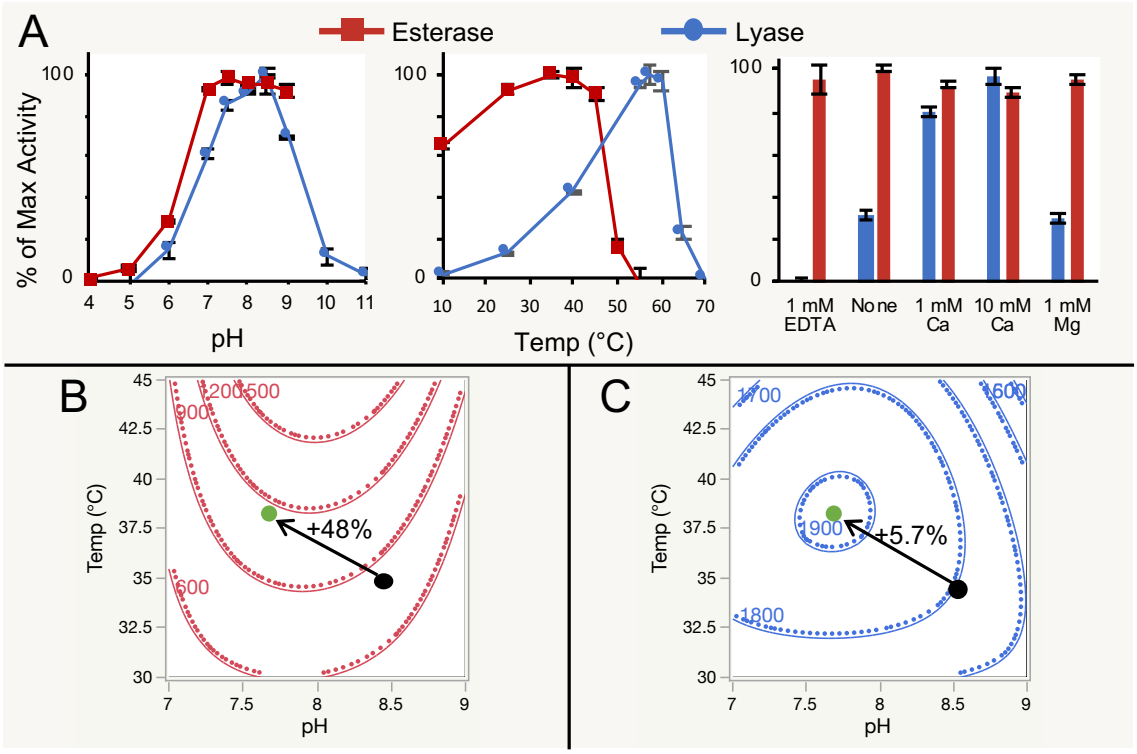


Figure 4.2 Optimum reaction conditions of RglA. Lyase activity was determined by monitoring product release directly from rhamnogalacturonan I while esterase activity was monitored using the model substrate *p*-nitrophenol acetate. One factor at a time optimization (A) was performed for pH first, temperature second, and calcium last. Contour plots represent the response surface models built for RG lyase (B) and RG acetylesterase activity (C). Dots on contour lines indicate the upward direction. Arrows indicate the change in optimum compared to the one factor at a time approach and percentage values indicate the associated increase in activity.

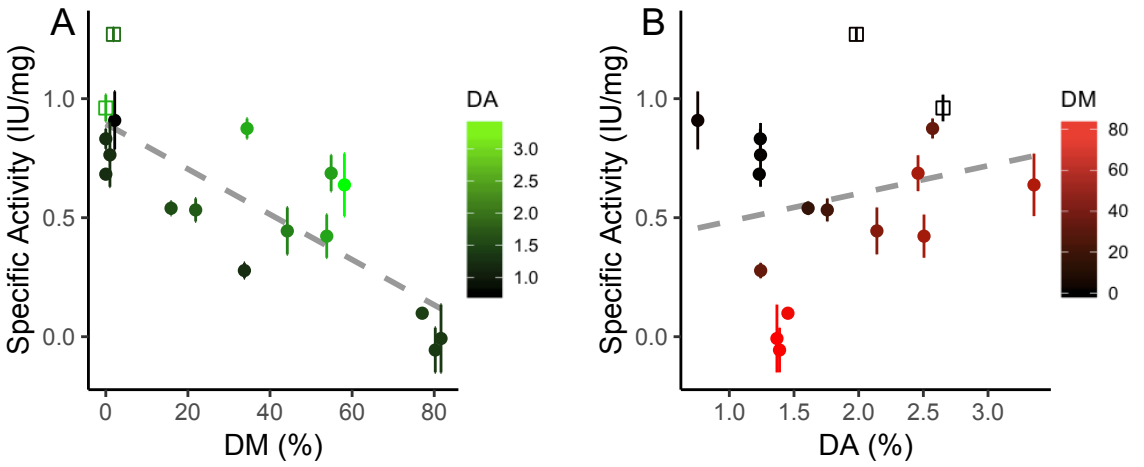


Figure 4.3 Depolymerase activity of RglA as a function of pectin methylation (A) and acetylation (B). Dashed grey lines indicate the correlation between activity and methylation or acetylation. Closed circles indicate citrus or apple pectins while open boxes are rhamnogalacturonan substrates. Each point and error bar represent the mean and standard deviation of three replicates.

4.8 REFERENCES

1. Mohnen D (2008) Pectin structure and biosynthesis. Curr Opin Plant Biol 11:266–277. doi:10.1016/j.pbi.2008.03.006

2. Mohnen D, Bar-peled M, Somerville C (2008) Cell wall polysaccharide synthesis. In: Himmel ME (ed) Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy. Blackwell, pp 94–187

3. Ding S, Himmel ME (2008) Anatomy and ultrastructure of maize cell walls: an example of energy plants. In: Himmel ME (ed) Biomass Recalcitrance: Deconstructing the Plant Cell Wall for Bioenergy. Blackwell, pp 38–60

4. Tan L, Eberhard S, Pattathil S, Warder C, Glushka J, Yuan C, Hao Z, Zhu X, Avci U, Miller JS, Baldwin D, Pham C, Orlando R, Darvill A, Hahn MG, Kieliszewski MJ, Mohnen D (2013) An Arabidopsis cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. Plant Cell 25:270–287. doi:10.1105/tpc.112.107334

5. Popper ZA, Fry SC (2005) Widespread occurrence of a covalent linkage between xyloglucan and acidic polysaccharides in suspension-cultured angiosperm cells. Ann Bot 96:91–99. doi:10.1093/aob/mci153

6. Bunzel M, Ralph J, Steinhart H (2005) Association of non-starch polysaccharides and ferulic acid in grain amaranth (Amaranthus caudatus L.) dietary fiber. Mol Nutr Food Res 49:551–559. doi:10.1002/mnfr.200500030

7. Meshitsuka G, Lee ZZ, Nakano J, Eda S (1982) Studies on the nature of lignin - carbohydrate bonding. J Wood Chem Technol 2:251–267. doi:10.1080/02773818208085134

8. Broxterman SE, Schols HA (2018) Interactions between pectin and cellulose in primary plant cell walls. Carbohydr Polym 192:263–272. doi:https://doi.org/10.1016/j.carbpol.2018.03.070

9. Kashyap DR, Vohra PK, Chopra S, Tewari R (2001) Applications of pectinases in the commercial sector: a review. Bioresour Technol 77:215–227. doi:10.1016/S0960-8524(00)00118-8

10. Schols HA, Posthumus MA, Voragen AGJ (1990) Structural features of hairy regions of pectins isolated from apple juice produced by the liquefaction process. Carbohydr Res 206:117–129. doi:https://doi.org/10.1016/0008-6215(90)84011-I

11. Schols HA, Geraeds CCJM, Searle-van Leeuwen MF, Kormelink FJM, Voragen AGJ (1990) Rhamnogalacturonase: a novel enzyme that degrades the hairy regions of pectins. Carbohydr Res 206:105–115. doi:10.1016/0008-6215(90)84010-R

12. Silva IR, Jers C, Meyer AS, Mikkelsen JD (2015) Rhamnogalacturonan I modifying enzymes: an update. N Biotechnol 33:41–54. doi:10.1016/j.nbt.2015.07.008

13. Searle-van Leeuwen MJF, van den Broek LAM, Schols HA, Beldman G, Voragen AGJ (1992) Rhamnogalacturonan acetylesterase: a novel enzyme from Aspergillus aculeatus, specific for the deacetylation of hairy (ramified) regions of pectins. Appl Microbiol Biotechnol 38:347–349. doi:10.1007/BF00170084

14. Kauppinen S, Christgau S, Kofod L V, Halkier T, Dorreich K, Dalboge H (1995) Molecular cloning and characterization of a rhamnogalacturonan acetylesterase from Aspergillus aculeatus. Synergism between rhamnogalacturonan degrading enzymes. J Biol Chem 270:27172–27178

15. Mølgaard A, Kauppinen S, Larsen S (2000) Rhamnogalacturonan acetylesterase elucidates the structure and function of a new family of hydrolases. Structure 8:373–383. doi:10.1016/S0969-2126(00)00118-0

16. Hoondal G, Tiwari R, Tewari R, Dahiya N, Beg Q (2002) Microbial alkaline pectinases and their industrial applications: a review. Appl Microbiol Biotechnol 59:409–418. doi:10.1007/s00253-002-1061-1

17. Jayani RS, Saxena S, Gupta R (2005) Microbial pectinolytic enzymes: a review. Process Biochem 40:2931–2944. doi:10.1016/j.procbio.2005.03.026

18. Edwards MC, Doran-Peterson J (2012) Pectin-rich biomass as feedstock for fuel ethanol production. Appl Microbiol Biotechnol 95:565–75. doi:10.1007/s00253-012-4173-2

19. Edwards MC, Williams T, Pattathil S, Hahn MG, Doran-Peterson J (2014) Replacing a suite of commercial pectinases with a single enzyme, pectate lyase B, in Saccharomyces cerevisiae fermentations of cull peaches. J Ind Microbiol Biotechnol 41:679–686. doi:10.1007/s10295-013-1394-z

20. Chung D, Pattathil S, Biswal AK, Hahn MG, Mohnen D, Westpheling J (2014) Deletion of a gene cluster encoding pectin degrading enzymes in Caldicellulosiruptor bescii reveals an important role for pectin in plant biomass recalcitrance. Biotechnol Biofuels 7:147. doi:10.1186/s13068-014-0147-1

21. Lionetti V, Francocci F, Ferrari S, Volpi C, Bellincampi D, Galletti R, D'Ovidio R, De Lorenzo G, Cervone F (2010) Engineering the cell wall by reducing de-methyl-esterified homogalacturonan improves saccharification of plant tissues for bioconversion. Proc Natl Acad Sci 107:616–621. doi:10.1073/pnas.0907549107

22. Francocci F, Bastianelli E, Lionetti V, Ferrari S, De Lorenzo G, Bellincampi D, Cervone F (2013) Analysis of pectin mutants and natural accessions of Arabidopsis highlights the impact of de-methyl-esterified homogalacturonan on tissue saccharification. Biotechnol Biofuels 6:163. doi:10.1186/1754-6834-6-163

23. Biswal AK, Soeno K, Gandla M, Immerzeel P, Pattathil S, Lucenius J, Serimaa R, Hahn MG, Moritz T, Jönsson LJ, Israelsson-Nordström M, Mellerowicz EJ (2014) Aspen pectate lyase PtxtPL1-27 mobilizes matrix polysaccharides from woody tissues and improves saccharification yield. Biotechnol Biofuels 7:11. doi:10.1186/1754-6834-7-11

24. Biswal AK, Hao Z, Pattathil S, Yang X, Winkeler K, Collins C, Mohanty SS, Richardson EA, Gelineo-Albersheim I, Hunt K, Ryno D, Sykes RW, Turner GB, Ziebell A, Gjersing E, Lukowitz W, Davis MF, Decker SR, Hahn MG, Mohnen D (2015) Downregulation of GAUT12 in Populus deltoides by RNA silencing results in reduced recalcitrance, increased growth and reduced xylan and pectin in a woody biofuel feedstock. Biotechnol Biofuels 8:41. doi:10.1186/s13068-015-0218-y

25. Tomassetti S, Pontiggia D, Verrascina I, Reca IB, Francocci F, Salvi G, Cervone F, Ferrari S (2015) Controlled expression of pectic enzymes in Arabidopsis thaliana enhances biomass conversion without adverse effects on growth. Phytochemistry 112:221–230. doi:10.1016/j.phytochem.2014.08.026

26. Biswal AK, Atmodjo MA, Li M, Baxter HL, Yoo CG, Pu Y, Lee Y-C, Mazarei M, Black IM, Zhang J-Y, Ramanna H, Bray AL, King ZR, LaFayette PR, Pattathil S, Donohoe BS, Mohanty SS, Ryno D, Yee K, Thompson OA, Rodriguez M, Dumitrache A, Natzke J, Winkeler K, Collins C, Yang X, Tan L, Sykes RW, Gjersing EL, Ziebell A, Turner GB, Decker SR, Hahn MG, Davison BH, Udvardi MK, Mielenz JR, Davis MF, Nelson RS, Parrott WA, Ragauskas AJ, Neal Stewart C, Mohnen D (2018) Sugar release and growth of biofuel crops are improved by downregulation of pectin biosynthesis. Nat Biotechnol. doi: 10.1038/nbt.4067

27. Boland WE, DeCrescenzo Henriksen E, Doran-Peterson J (2010) Characterization of two Paenibacillus amylolyticus strain 27C64 pectate lyases with activity on highly methylated pectin. Appl Environ Microbiol 76:6006–6009. doi:10.1128/AEM.00043-10

28. Keggi C, Doran-Peterson J (2019) Paenibacillus amylolyticus 27C64 has a diverse set of carbohydrate-active enzymes and complete pectin deconstruction system. J Ind Microbiol Biotechnol 46:1–11. doi:10.1007/s10295-018-2098-1

29. Bolvig PU, Pauly M, Orfila C, Scheller HV, Schnorr K (2003) Sequence analysis and characterisation of a novel pectin acetyl esterase from Bacillus subtilis. In: Voragen F, Schols H, Visser R (eds) Advances in Pectin and Pectinase Research. Springer Netherlands, Dordrecht, pp 315–330

30. Silva IR, Larsen DM, Meyer AS, Mikkelsen JD (2011) Identification, expression, and characterization of a novel bacterial RGI Lyase enzyme for the production of bio-functional fibers. Enzyme Microb Technol 49:160–166. doi:10.1016/j.enzmictec.2011.04.015

31. Silva IR, Jers C, Otten H, Nyffenegger C, Larsen DM, Derkx PMF, Meyer AS, Mikkelsen JD, Larsen S (2014) Design of thermostable rhamnogalacturonan lyase mutants from Bacillus licheniformis by combination of targeted single point mutations. Appl Microbiol Biotechnol 98:4521–4531. doi:10.1007/s00253-013-5483-8

32. Ochiai A, Itoh T, Kawamata A, Hashimoto W, Murata K (2007) Plant cell wall degradation by saprophytic Bacillus subtilis strains: gene clusters responsible for rhamnogalacturonan depolymerization. Appl Environ Microbiol 73:3803–13. doi:10.1128/AEM.00147-07

33. Montanier C, van Bueren AL, Dumon C, Flint JE, Correia MA, Prates JA, Firbank SJ, Lewis RJ, Grondin GG, Ghinet MG, Gloster TM, Herve C, Knox JP, Talbot BG, Turkenburg JP, Kerovuo J, Brzezinski R, Fontes CMGA, Davies GJ, Boraston AB, Gilbert HJ (2009) Evidence that family 35 carbohydrate binding modules display conserved specificity but divergent function. Proc Natl Acad Sci 106:3065–3070. doi:10.1073/pnas.0808972106

34. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:490–495. doi:10.1093/nar/gkt1178

35. Songsiriritthigul C, Pantoom S, Aguda AH, Robinson RC, Suginta W (2008) Crystal structures of Vibrio harveyi chitinase A complexed with chitooligosaccharides: Implications for the catalytic mechanism. J Struct Biol 162:491–499. doi:10.1016/j.jsb.2008.03.008

36. Martinez-Martinez I, Navarro-Fernandez J, Daniel Lozada-Ramirez J, Garcia-Carmona F, Sanchez-Ferrer A (2008) YesT: a new rhamnogalacturonan acetyl esterase from Bacillus subtilis. Proteins 71:379–388. doi:10.1002/prot.21705

37. Navarro-Fernandez J, Martinez-Martinez I, Montoro-Garcia S, Garcia-Carmona F, Takami H, Sanchez-Ferrer A (2008) Characterization of a new rhamnogalacturonan acetyl esterase from Bacillus halodurans C-125 with a new putative carbohydrate binding domain. J Bacteriol 190:1375–1382. doi:10.1128/JB.01104-07

38. McDonough MA, Kadirvelraj R, Harris P, Poulsen J-CN, Larsen S (2004) Rhamnogalacturonan lyase reveals a unique three-domain modular structure for polysaccharide lyase family 4. FEBS Lett 565:188–194. doi:10.1016/j.febslet.2004.03.094

39. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. doi: 10.1093/nar/gkz268

40. Marchler-Bauer A, Bryant SH (2004) CD-Search: protein domain annotations on the fly. Nucleic Acids Res 32:W327-31. doi:10.1093/nar/gkh454

41. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015) The I-TASSER Suite: protein structure and function prediction. Nat. Methods 12:7–8

42. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics 10:421. doi:10.1186/1471-2105-10-421

43. Cook DM, Henriksen ED, Upchurch R, Doran-Peterson JB (2007) Isolation of polymer-degrading bacteria and characterization of the hindgut bacterial community from the detritus-feeding larvae of Tipula abdominalis (Diptera: Tipulidae). Appl Environ Microbiol 73:5683–6. doi:10.1128/AEM.00213-07

44. Li MZ, Elledge SJ (2007) Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. Nat Methods 4:251–256. doi:10.1038/nmeth1010

45. Studier FW (2005) Protein production by auto-induction in high-density shaking cultures. Protein Expr Purif 41:207–234. doi:10.1016/j.pep.2005.01.016

46. Kapust RB, Tozser J, Copeland TD, Waugh DS (2002) The P1' specificity of tobacco etch virus protease. Biochem Biophys Res Commun 294:949–955. doi:10.1016/S0006-291X(02)00574-0

47. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685. doi:10.1038/227680a0

48. Voragen A, Schols HA, Pilnik W (1986) Determination of the eegree of methylation and acetylation of pectins by HPLC. Food Hydrocoll 1:65–70. doi:10.1016/S0268-005X(86)80008-X

49. Blumenkrantz N, Asboe-Hansen G (1973) New method for quantitative determination of uronic acids. Anal Biochem 54:484–489. doi:https://doi.org/10.1016/0003-2697(73)90377-1

50. White GW, Katona T, Zodda JP (1999) The use of high-performance size exclusion chromatography (HPSEC) as a molecular weight screening technique for polygalacturonic acid for use in pharmaceutical applications. J Pharm Biomed Anal 20:905–912. doi:10.1016/S0731-7085(99)00083-7

CHAPTER 5

CONCLUSION

The overarching goal of this work at its inception was to explore the plant cell wall polysaccharide deconstruction capabilities of *Paenibacillus amylolyticus* 27C64. This organism was originally isolated from the gut system of an aquatic crane fly larvae, an organism which acts as a natural biorefinery, breaking down plant material in a stepwise manner that parallels the industrial process we sought to improve. Previous work with this organism had already noted the presence of extracellular enzymes capable of breaking down different plant cell polysaccharides but none of the underlying systems had been characterized. Only two enzymes of the many that were likely produced by the organism had been cloned and studied directly. However, both of these enzymes had an unusually broad substrate specificity, a feature which hinted that this organism may approach pectin deconstruction differently than others. Collectively, these features suggested that *P. amylolyticus* could be a valuable source of novel plant cell wall deconstruction enzymes.

Whole genome sequencing confirmed that *P. amylolyticus* had a rich array of hundreds of carbohydrate active enzymes including cellulases, hemicellulases, and pectinases (Chapter 2). This assessment of the CAZome, paired with growth of the organism on polysaccharides as sole carbon sources, revealed a specialization in deconstructing non-cellulosic polysaccharides. We found that the pectin deconstruction system was particularly extensive and responded to the specific type of pectin supplied for

growth. We also noted that there were several potentially novel enzymes that warranted direct biochemical characterization.

The homogalacturonan deconstruction system was unusual in that the extracellular depolymerases were unhindered by substrate methylation. These enzymes are able to reduce highly polymerized pectin to a mixture of oligosaccharides which are only demethylated after import into the cytoplasm. This key feature likely resulted in selective pressures that led to the development of industrially useful enzymes. PelC, an endo-pectate lyase, was more active on methylated substrates than on unmethylated ones, despite the fact that it cannot cut at methylated sites. This unusual substrate specificity is most likely the result of PelC binding at the junction of methylated and demethylated regions. This enzyme worked synergistically with PelB, a fast exo-pectate lyase, by releasing internal stretches of demethylated homogalacturonan for PelB to saccharify. We also noted that the only pectin lyase, Pnl, had a very high turnover rate. Put together, the efficient deconstruction system driven by these enzymes is simpler than the fungal pectinase mixtures that are commercially available today. Enzymes from this system have the potential to improve a wide variety of industrial sectors from food processing to textiles manufacturing. In retrospect, the fact that PelA and PelB had unusually high activity on methylated substrates was an early hint that the extracellular HG deconstruction system did not require a pectin methylesterase at all. Given that isolated pectate lyases with substrate specificities similar to PelA and PelB have been described in other Gram-positive bacteria, it is likely that the model of HG deconstruction presented in this work is conserved in other bacterial genera.

Deconstruction of rhamnogalacturonan I in *P. amylolyticus* also differed from characterized systems. One rhamnogalacturonan lyase identified in the genome, RglA, was a bifunctional enzyme which was also capable of deacetylating its substrate. This enzyme has a complex structure with eight non-catalytic domains, five of which are potential carbohydrate binding domains. Much like the homogalacturonan deconstruction system, this enzyme would allow *P. amylolyticus* to directly degrade the acetylated form of rhamnogalacturonan I that is typically resistant to cleavage by similar lyases. This enzyme could likewise be useful in industrial applications where highly acetylated pectin is abundant.

After mining the genome of this organism for novel CAZymes, it is clear that early indications that *P. amylolyticus* was a valuable source of CAZymes were accurate. We were able to successfully identify and characterize novel pectinases contained within the genome, and in doing so contributed to a more complete understanding of CAZyme diversity. Ultimately, exploring the capabilities of an organism which was taxonomically divergent from other model pectinolytic organisms led to improved understanding of pectin deconstruction and discovery of enzymes that may prove useful in numerous industrial applications.

APPENDIX A

SUPPLEMENTARY DATA FOR CHAPTER TWO

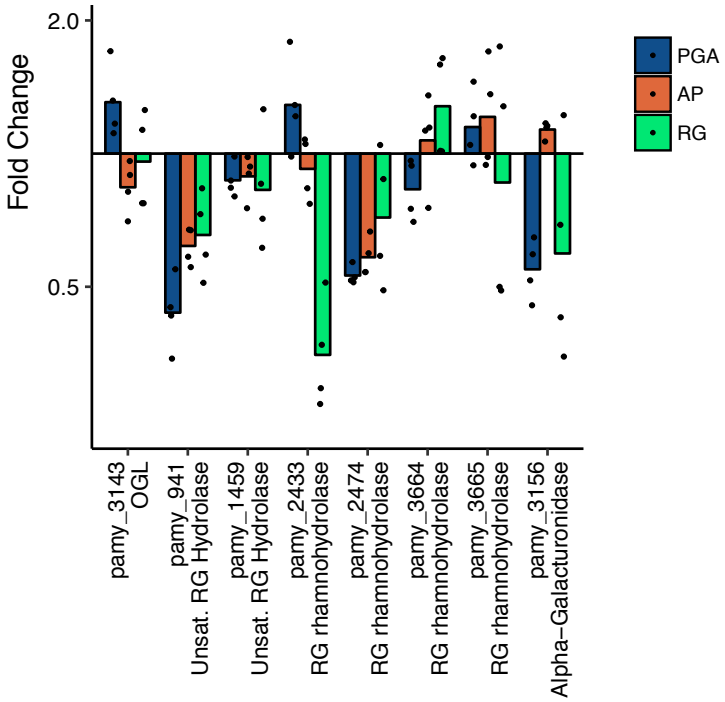


Figure A1 Differential expression of genes eliminated as potential pectinases when grown on three different pectic substrates. Points indicate the fold change in expression as compared to the control condition (TSB without pectin) for individual biological replicates and columns indicate the mean fold change values of all four biological replicates. Fold change values were calculated with the ΔΔCt method using ftsZ and rpoD as reference genes. *HG* homogalacturonan, *RG* rhamnogalacturonan I, *PGA* polygalacturonic acid, *AP* apple pectin.

APPENDIX B

SUPPLEMENTARY DATA FOR CHAPTER THREE

Table B1 Inventory and characteristics of pectins used in this study.

Code	Substrate	Manu-facturer	Product No.	Lot No.	DM (%)	DA (%)	GalA (wt %)	Mp (kDa)	Mn (kDa)	Mw (kDa)	PI
Polygalacturonic acids											
PGA-A	Polygalacturonic acid, orange	Sigma	P3889	BCBB4735	0.0	1.2	89	172	89.9	144	1.61
PGA-B	Polygalacturonic acid, orange	Sigma	P3889	BCBD8707V	2.2	0.8	89	172	90.5	144	1.60
PGA-C	Polygalacturonic acid, orange	Sigma	P3889	BCBV4304	0.0	1.2	90	169	106	148	1.40
PGA-D	Polygalacturonic acid, orange	Sigma	P3889	BCBV4304	1.0	1.2	90	169	107	148	1.38
Rhamnogalacturonans											
RG_P-A	Rhamnogalacturonan I, potato	Megazyme	P-RHAM1	141102a	1.9	2.0	56	3.04	38.2	97.7	2.56
RG_S	Rhamnogalacturonan, soy	Megazyme	P-RHAGN	20202b	0.0	2.7	45	203	173	291	1.68
Unmodified pectins											
CP_ICN	Pectin, citrus	ICN	102587	4213A	34.4	2.6	65	0.35	70.8	153	2.16
CP_Sigma	Pectin, citrus peel	Sigma	P9135	SLBS8828	54.9	2.5	81	166	124	175	1.42
AP_Sigma	Pectin, apple	Sigma	93854	BCBS3576	58.1	3.4	80	170	108	161	1.49
CP_20_34-A	Pectin, 20-34% esterified, citrus	Sigma	P9311	077K1583	21.9	1.8	74	160	89.2	153	1.71
CP_20_34-B	Pectin, 20-34% esterified, citrus	Sigma	P9311	040M1178V	15.9	1.6	75	95.1	81.3	136	1.67
CP_55_70-A	Pectin, 55-70% esterified, citrus	Sigma	P9436	018K1649	44.2	2.1	85	170	121	202	1.67
CP_55_70-B	Pectin, 55-70% esterified, citrus	Sigma	P9436	051M1376V	53.8	2.5	76	170	129	203	1.58
CP_85_A	Pectin, 85% esterified, citrus	Sigma	P9561	110M1409V	77.1	1.5	83	160	90.3	136	1.51
CP_85_B	Pectin, 85% esterified, citrus	Sigma	P9561	BCBS3782V	81.7	1.4	87	164	96.7	143	1.48
CP_85_C	Pectin, 85% esterified, citrus	Sigma	P9561	BCBS3782	80.3	1.4	91	164	97.5	145	1.49
CP_90	Pectin, 90% esterified, citrus	Sigma	P9561	037K1272	33.8	1.2	98	170	74.8	128	1.71

Table B2 Results of linear models describing HG lyase activity as a function of pectin characteristics.

	PelA			PelB			PelC			PelD			Pnl		
	Est.	***t***	***p***	**Est.**	***t***	***p***	**Est.**	***t***	***p***	**Est.**	***t***	***p***	**Est.**	***t***	***p***
DM	-1.1	-0.66	0.5113	1.1	0.93	0.3566	16.6	8.4	<.0001*	5.8	3.06	0.0040*	33.9	58.25	<.0001*
DM^2	-20.1	-8.23	<.0001*	-34.4	-20.18	<.0001*	-16.9	-5.83	<.0001*	-28.1	-10.06	<.0001*	7.3	8.53	<.0001*
DA	7.6	2.86	0.0068*	1.2	0.64	0.5251	-1.1	-0.34	0.7353	13.1	4.33	0.0001*	0.6	0.67	0.5087
Mn	-27.3	-10.32	<.0001*	-17.3	-9.41	<.0001*	1.1	0.35	0.7293	-17.1	-5.66	<.0001*	-5.3	-5.79	<.0001*
GalA	29.6	9.74	<.0001*	14.1	6.66	<.0001*	1.9	0.53	0.6008	26.9	7.75	<.0001*	8.4	7.88	<.0001*
R^2	0.860			0.948			0.761			0.867			0.992		

All tests had 39 degrees of freedom. Est. = Parameter estimate.

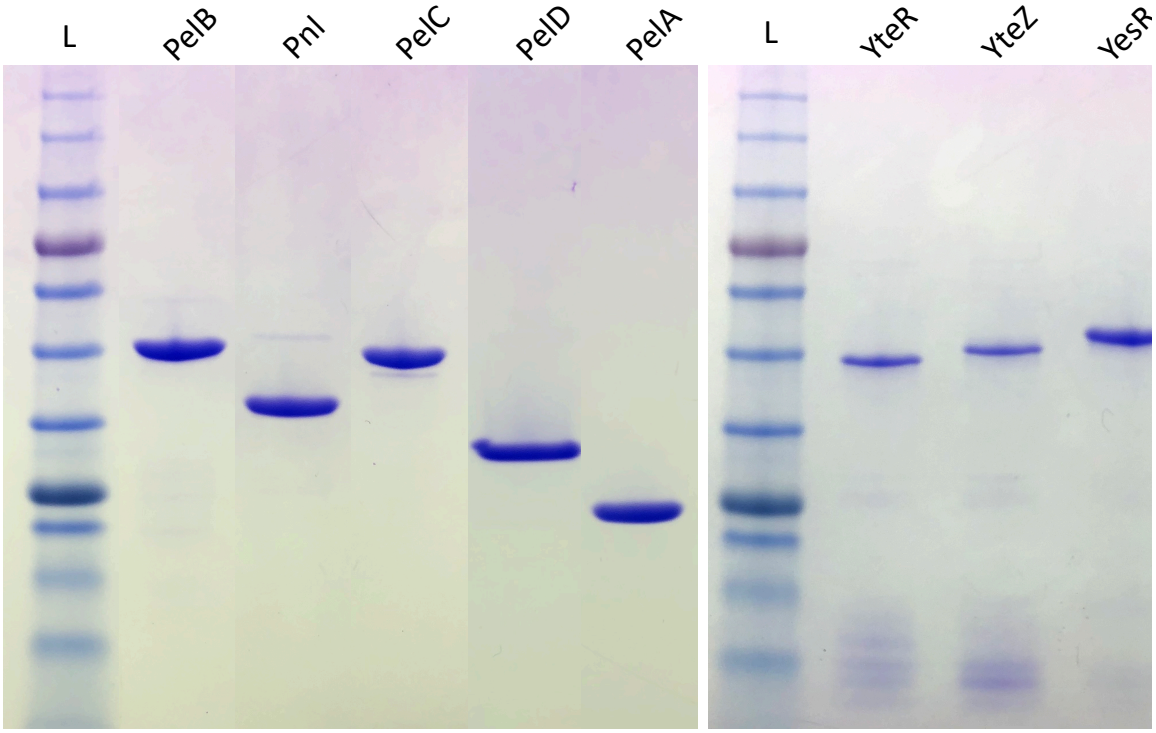
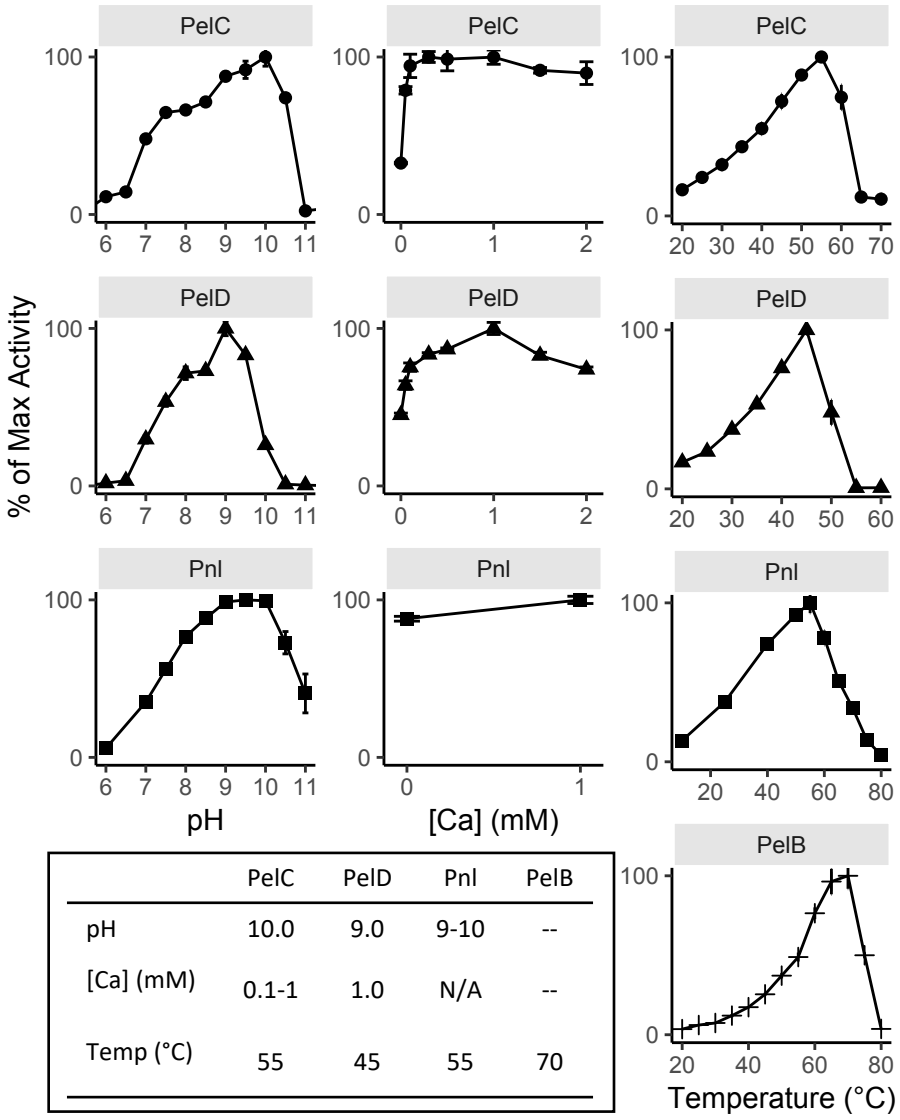


Figure B1. SDS-PAGE analysis of purified proteins included in this study. Intermediate purification steps have been cropped out of the leftmost six lanes but the position of each band relative to the ladder has been preserved. Gold Bio Bluestain2 was used as the molecular weight marker (indicated by lanes marked 'L').



	PelC	PelD	PnI	PelB
pH	10.0	9.0	9-10	--
[Ca] (mM)	0.1-1	1.0	N/A	--
Temp (°C)	55	45	55	70

Figure B2. Optimum reaction conditions for three new pectic lyases and updated temperature profile for PelB. Each point and error bar represent the mean and standard deviation of three replicates. Individual optima are summarized in the table (inset). The optimal values were determined sequentially: pH first, calcium second, and temperature last. The temperature profile of PelB was re-evaluated because new equipment made accurate measurement at elevated temperatures possible.

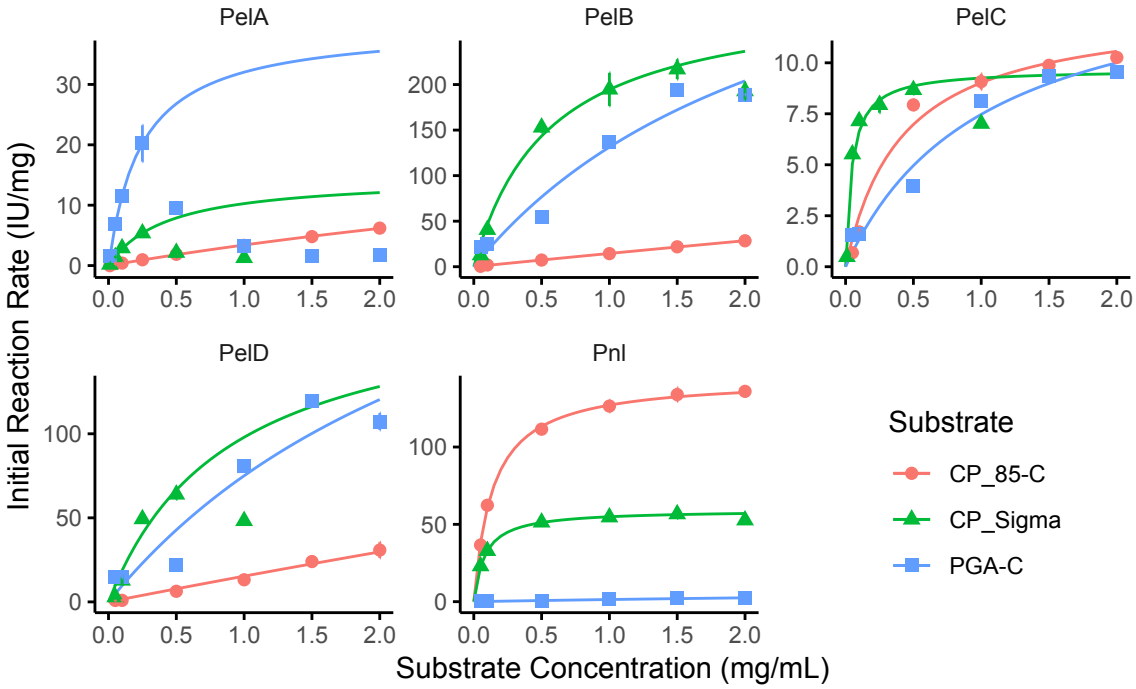


Figure B3. Activity of each extracellular lyase on different concentrations of three different pectins. Each point represents the initial reaction rate (n = 3). Kinetic parameters were determined by non-linear least-squares minimization fitting to the Michaelis-Menten equation. Lines represent theoretical activity based on calculated K_m and V_{max} values.

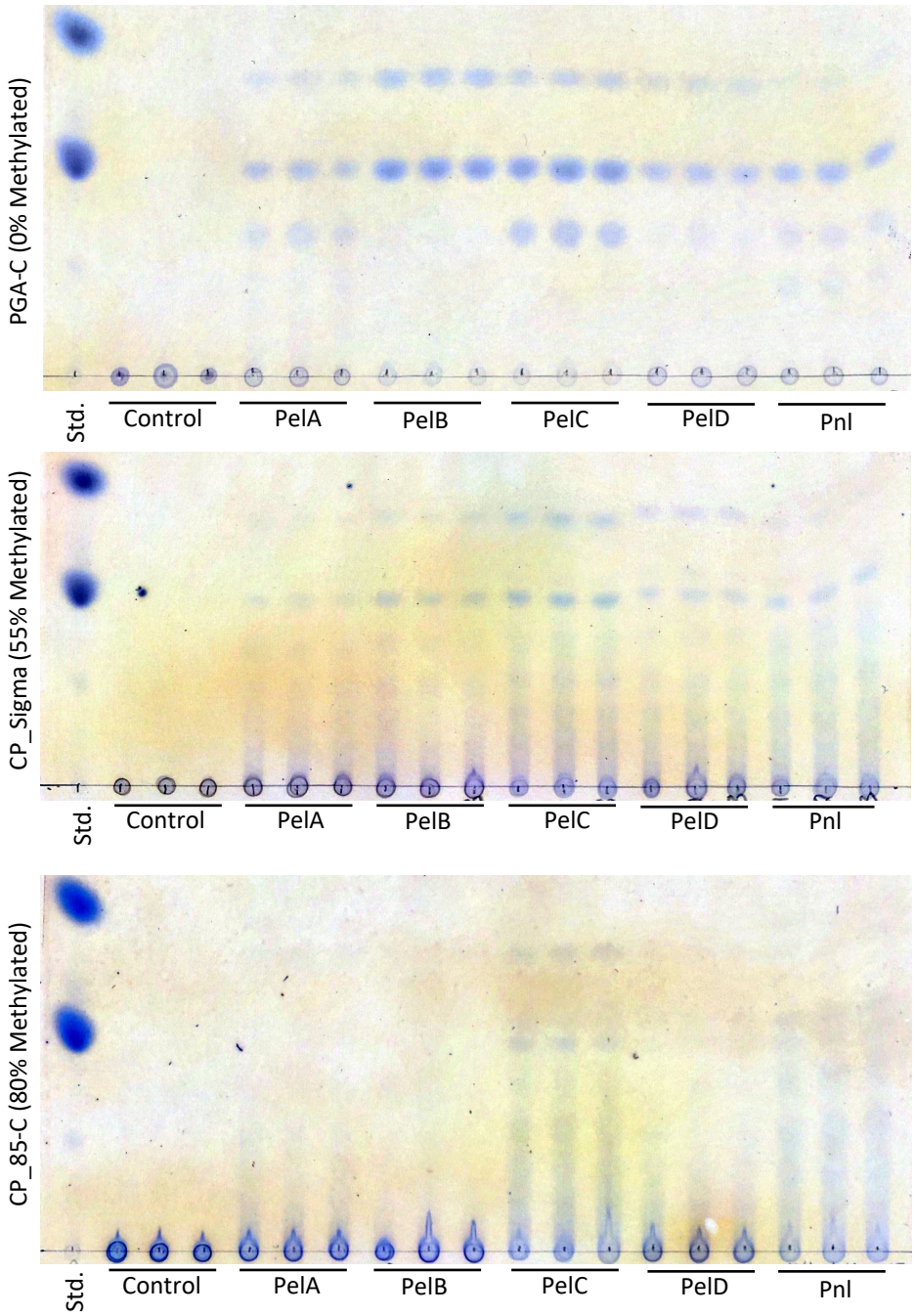


Figure B4. Thin-layer chromatography analysis of digestion products released from three different pectins by each extracellular lyase. All reactions were allowed to run to completion (24 hours). The leftmost sample on each plate is a mixed mono- and di-galacturonic acid standard. Labels below each sample identify replicate digestion reactions, and labels on the left of each plate indicate the substrate used for those reactions.

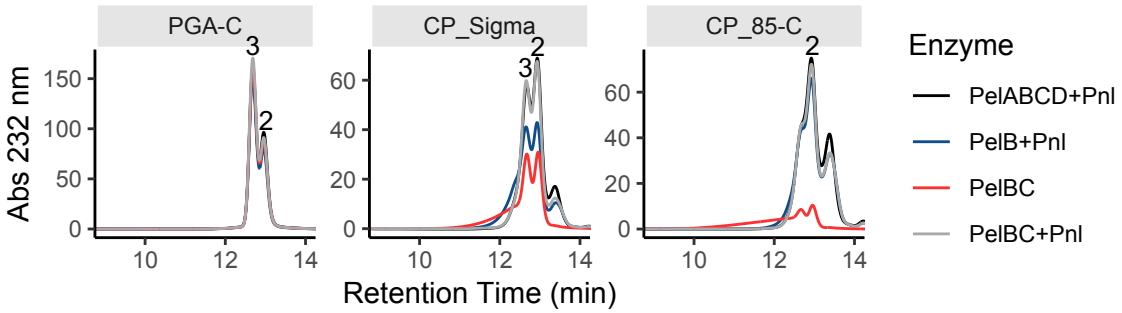


Figure B5. Products released by selected lyase combinations on three different pectins. Products were separated with high-performance size exclusion chromatography and elution profiles were monitored as the absorbance at 232 nm. Labels over peaks represent the degree of polymerization associated with that peak and were determined by comparison to oligogalacturonide standards.

APPENDIX C

SUPPLEMENTARY DATA FOR CHAPTER FOUR

Table C1 Similarity of RglA catalytic domains to characterized sequences within their respective CAZy families.

Organism	Protein	Function	GenBank Accession	Length (bp)	Identity (%)	Similarity (%)
Rhamnogalacturonan lyase domain (PL11)						
Bacillus licheniformis	YesW	RG endolyase (EC 4.2.2.23)	AAU40280.1	639	62.6	73.7
Bacillus subtilis	YesW	RG endolyase (EC 4.2.2.23)	CAB12524.1	642	61.2	72.7
Bacillus subtilis	YesX	RG exolyase (EC 4.2.2.24)	CAB12525.1	654	57.5	69.3
Ruminiclostridium cellulolyticum	Rgl11Y	RG endolyase (EC 4.2.2.23)	ACL75117.1	638	54.9	70.8
Cellvibrio japonicus	Rgl11A	RG endolyase (EC 4.2.2.23)	ACE83155.1	658	54.4	66.7
Bacteroides thetaiotaomicron	BT4175	RG endolyase (EC 4.2.2.23)	AAO79280.1	664	47.6	62
Rhamnogalacturonan acetylesterase domain (CE12)						
Hungateiclostridium thermocellum	Rgae12A	RG acetylesterase	ABN54336.1	253	50.2	62.5
Bacillus licheniformis	BliPAE	Pectin acetylesterase	AAU42913.1	259	42.1	57.1
Bacillus sp. KCCM10143	--	Cephalosporin C deacetylase	AAF25818.1	253	41.1	57.3
Bacillus subtilis	YesY	Pectin acetylesterase	CAB12526.1	254	40.9	57.5
Bacillus halodurans	Rgae	Acetyl xylan esterase (EC 3.1.1.72)	BAB04834.1	263	36.9	53.2
Bacillus subtilis	YesT	Acetyl xylan esterase (EC 3.1.1.72)	CAB12521.1	262	36.3	55
Dickeya dadantii	PaeY	Pectin acetylesterase	CAA70971.1	311	29.9	43.1
Aspergillus aculeatus	Rha1	RG acetylesterase	CAA61858.1	291	29.6	40.5
Aspergillus nidulans	AN2528.2	RG acetylesterase	EAA64633.1	287	29.6	42.9
Uncultured organism	--	Pectinase (partial)	ADD61757.1	374	28.1	40.9

Table C2 Purification of RglA.

Sample	Volume (mL)	RGL Activity (IU/mL)[a]	Total Activity (IU)	[Protein] (mg/mL)[b]	Total Protein (mg)	Specific Activity (IU/mg)	Yield (%)	Fold Purification	Purity[c]
Lysate	30.0	0.103	3.09	61.3	1838	0.00168	(100)	(1)	0.0842
<15% AS[d]	30.0	0.0107	0.321	--	--	--	--	--	--
15-25% AS[d]	10.0	0.259	2.59	24.3	243	0.0107	13.2	6.35	0.534
>25% AS[d]	28.0	0.0180	0.503	--	--	--	--	--	--
Ni-NTA Elution	8.4	0.439	3.69	2.10	17.7	0.209	0.960	124	10.5
-IMAC Elution	7.0	0.00564	0.0395	0.250	1.75	0.0226	0.0953	13.4	1.13
-IMAC Flowthrough	11.75	0.323	3.79	0.360	4.22	0.897	0.230	534	44.6

[a]As determined by a spectrophotometric assay assuming a molar extinction coefficient of 5200 M^{-1} cm^{-1}

[b]Based on the absorbance at 280 nm assuming a 1 mg/mL solution has an absorbance of 1 for all samples prior to the Ni-NTA elution. Subsequent samples used the ProtParam predicted absorbance values of 1.547 (with the tag) or 1.534 (without the tag) as appropriate

[c]Purity of the final sample was estimated based on ImageJ analysis of a SDS-PAGE gel. Value is % full-length target.

[d]AS - ammonium sulfate. Ranges indicate samples that precipited within the cutoff percent AS saturation at 0 °C

Table C3 Inventory and characteristics of pectins used in this study.

Code	Substrate	Manu-facturer	Product No.	Lot No.	DM (%)	DA (%)	GalA (wt %)	Mp (kDa)	Mn (kDa)	Mw (kDa)	PI
Polygalacturonic acids											
PGA-A	Polygalacturonic acid, orange	Sigma	P3889	BCBB4735	0.0	1.2	89	172	89.9	144	1.61
PGA-B	Polygalacturonic acid, orange	Sigma	P3889	BCBD8707V	2.2	0.8	89	172	90.5	144	1.60
PGA-C	Polygalacturonic acid, orange	Sigma	P3889	BCBV4304	0.0	1.2	90	169	106	148	1.40
PGA-D	Polygalacturonic acid, orange	Sigma	P3889	BCBV4304	1.0	1.2	90	169	107	148	1.38
Rhamnogalacturonans											
RG_P-A	Rhamnogalacturonan I, potato	Megazyme	P-RHAM1	141102a	1.9	2.0	56	3.04	38.2	97.7	2.56
RG_S	Rhamnogalacturonan, soy	Megazyme	P-RHAGN	20202b	0.0	2.7	45	203	173	291	1.68
Unmodified pectins											
CP_ICN	Pectin, citrus	ICN	102587	4213A	34.4	2.6	65	0.35	70.8	153	2.16
CP_Sigma	Pectin, citrus peel	Sigma	P9135	SLBS8828	54.9	2.5	81	166	124	175	1.42
AP_Sigma	Pectin, apple	Sigma	93854	BCBS3576	58.1	3.4	80	170	108	161	1.49
CP_20_34-A	Pectin, 20-34% esterified, citrus	Sigma	P9311	077K1583	21.9	1.8	74	160	89.2	153	1.71
CP_20_34-B	Pectin, 20-34% esterified, citrus	Sigma	P9311	040M1178V	15.9	1.6	75	95.1	81.3	136	1.67
CP_55_70-A	Pectin, 55-70% esterified, citrus	Sigma	P9436	018K1649	44.2	2.1	85	170	121	202	1.67
CP_55_70-B	Pectin, 55-70% esterified, citrus	Sigma	P9436	051M1376V	53.8	2.5	76	170	129	203	1.58
CP_85_A	Pectin, 85% esterified, citrus	Sigma	P9561	110M1409V	77.1	1.5	83	160	90.3	136	1.51
CP_85_B	Pectin, 85% esterified, citrus	Sigma	P9561	BCBS3782V	81.7	1.4	87	164	96.7	143	1.48
CP_85_C	Pectin, 85% esterified, citrus	Sigma	P9561	BCBS3782	80.3	1.4	91	164	97.5	145	1.49
CP_90	Pectin, 90% esterified, citrus	Sigma	P9561	037K1272	33.8	1.2	98	170	74.8	128	1.71

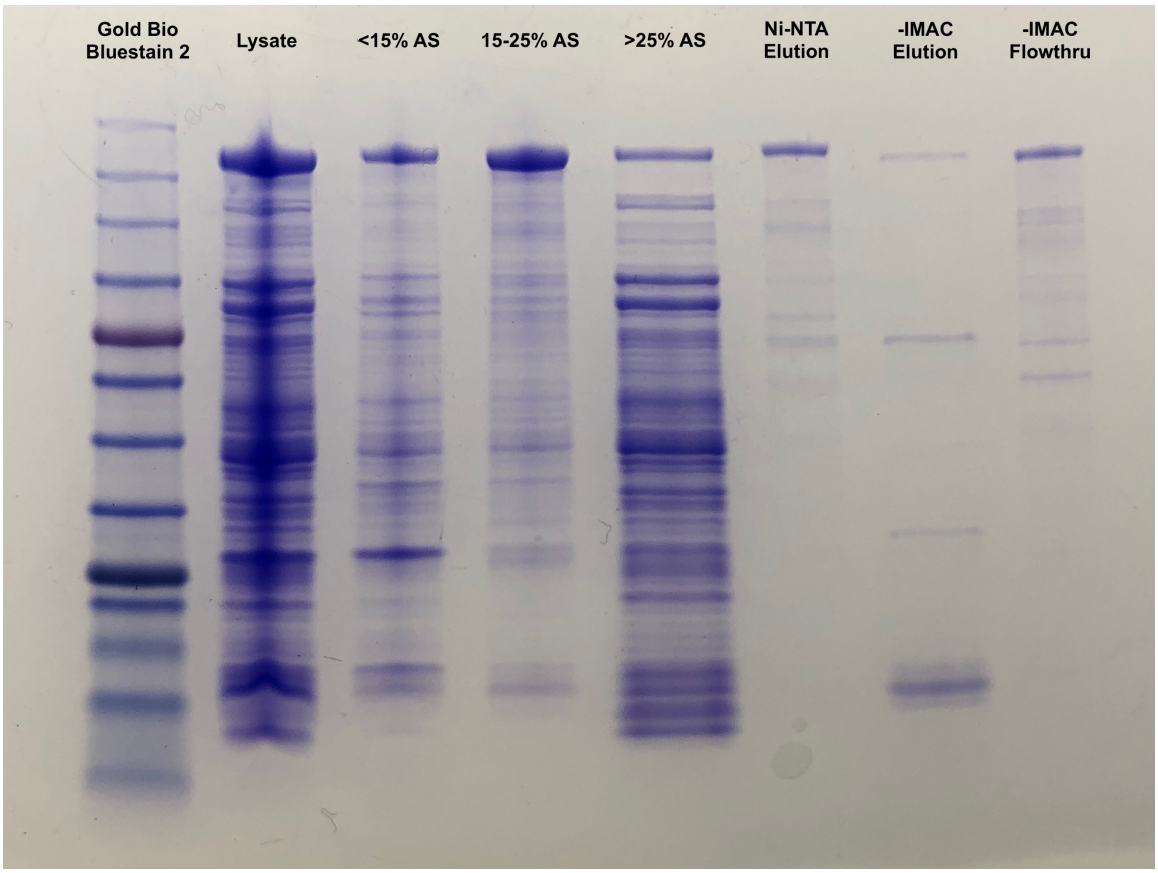


Figure C1 SDS-PAGE of RglA purification samples. Percentages for ammonium sulfate indicate the percent saturation at 0 °C. The -IMAC samples refer to the removal of uncut protein and free tag. The final sample on the right-most band (-IMAC flowthrough).

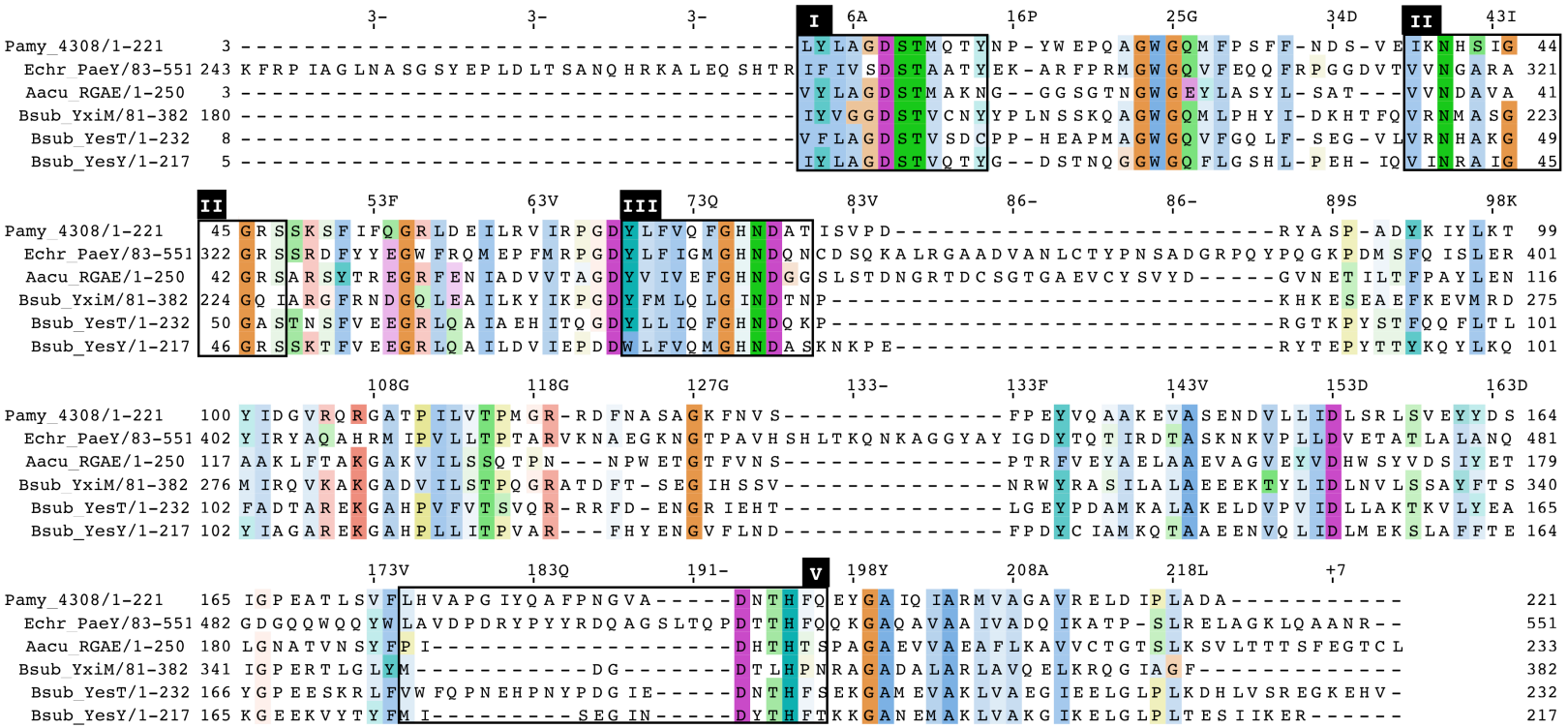


Figure C2 Alignment of RglA with characterized CE12 sequences. Boxes indicate key conserved regions previously identified.