Appl Microbiol Biotechnol
DOI 10.1007/s00253-016-7294-1

BIOENERGY AND BIOFUELS

Clostridium thermocellum releases coumaric acid during degradation of untreated grasses by the action of an unknown enzyme

Christopher D. Herring[1,2,3] · **Philip G. Thorne**[2] · **Lee R. Lynd**[1,2,3]

Received: 14 October 2015 / Revised: 20 December 2015 / Accepted: 26 December 2015

Abstract *Clostridium thermocellum* is an anaerobic thermophile with the ability to digest lignocellulosic biomass that has not been pretreated with high temperatures. Thermophilic anaerobes have previously been shown to more readily degrade grasses than wood. Part of the explanation for this may be the presence of relatively large amounts of coumaric acid in grasses, with linkages to both hemicellulose and lignin. We found that *C. thermocellum* and cell-free cellulase preparations both release coumaric acid from bagasse and switchgrass. Cellulase preparations from a mutant strain lacking the scaffoldin *cipA* still showed activity, though diminished. Deletion of all three proteins in *C. thermocellum* with ferulic acid esterase domains, either singly or in combination, did not eliminate the activity. Further work will be needed to identify the novel enzyme(s) responsible for the release of coumaric acid from grasses and to determine whether these enzymes are important factors of microbial biomass degradation.

Keywords Coumaric acid esterase · Cellulosic ethanol · Cellulolytic microorganisms · Analytical chemistry

✉ Christopher D. Herring
chrisherringfish@gmail.com

[1] Thayer School of Engineering, Dartmouth College, 14 Engineering Drive, Hanover, NH 03755, USA

[2] Mascoma Corporation, 67 Etna Road, Lebanon, NH 03766, USA

[3] Enchi Corporation, 610 Lincoln Street, Waltham, MA 02451, USA

Introduction

In most processes developed for conversion of lignocellulose into fuels or chemicals, an intensive chemical or physical pretreatment step is included to increase the effectiveness of added cellulase enzymes (Lynd et al. 1996, 2002; Yang and Wyman 2008). Thermophilic anaerobic bacteria have an often-overlooked ability to ferment recalcitrant biomass that has not been pretreated (Yang et al. 2009; Kataeva et al. 2013; Talluri et al. 2013; Basen et al. 2014; Chung et al. 2014; Izquierdo et al. 2014; Paye et al. 2015). Strictly speaking, the material in the above cited studies was sterilized by autoclaving, which could be considered a minimal form of pretreatment. The impetus of this study was to begin to determine the biochemical basis for degradation of minimally pretreated material in *Clostridium thermocellum*. The observation of coumaric acid in cultures of *C. thermocellum* growing on autoclaved sugarcane bagasse led to an investigation of this phenomenon and the role of genes encoding coumaric acid esterases.

Roughly 40 % of plant biomass is cellulose, which is compounded with lignin and hemicellulose, making up roughly 25 and 20 % of plant biomass, respectively. The recalcitrance of lignocellulose is due largely to the presence of lignin, which has a highly complex, variable structure, with few enzymes that depolymerize it (Bugg et al. 2011; Paliwal et al. 2012; Pu et al. 2013). To the best of our knowledge, no enzymes have been described that depolymerize lignin anaerobically, yet anaerobic organisms like *C. thermocellum* have evolved into specialists in the degradation of lignocellulose in the environment. The phenolic compound coumaric acid (Fig. 1) is ester-linked to both hemicellulose and lignin as well as ether-linked to lignin (Scalbert 1985; Lam et al. 2001; Sun et al. 2001; Wong 2006; Masarin et al. 2011). Coumaric and ferulic acids are abundant plant constituents known as

Published online: 14 January 2016

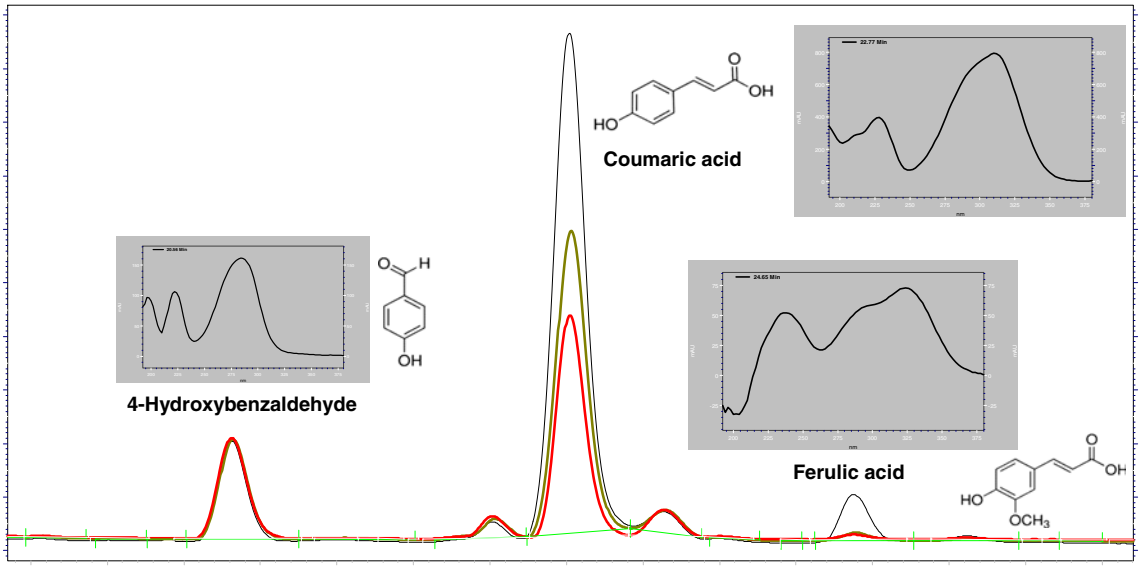

Fig. 1 HPLC chromatogram from analysis of phenolic components in an enzymatic hydrolysis of 10 g/L bagasse. The *black line* was hydrolyzed with *C. thermocellum* cellulase, the *green line* with boiled cellulase, and the *red line* had no enzyme addition. The identities and structures of the three peaks are given, along with the absorption spectra (color figure online)

hydroxycinnamic acids. They are highly prevalent in grasses but less prevalent or undetectable in wood (Hartley and Harris 1981; Sun et al. 2001). The ester- and ether-linked hydroxycinnamic acids can be distinguished analytically based on their liberation by mild or harsh alkaline hydrolysis conditions, respectively (Scalbert 1985; Wong 2006). Different grasses contain more or less of the hydroxycinnamic acids; for example, bagasse contains 5–9 % coumaric acid while maize bran contains 3 % ferulic acid (Saulnier et al. 1995; Masarin et al. 2011). Cross-links between hydroxycinnamic acids, lignin, and carbohydrates are implicated in recalcitrance (Grabber et al. 1998; Lam et al. 2003; Siqueira et al. 2011).

It is largely unknown how cellulolytic anaerobes process the hydroxycinnamic acids and other biomass components other than cellulose and hemicellulose. Lignin is considered an insoluble and unreactive component in most industrial lignocellulose bioprocesses—it blocks access to the cellulose, so it is physically and chemically disrupted by pretreament to increase accessibility. It would be a mistake though to assume that phenolic components are treated the same way during biological lignocellulose degradation in nature. There are a range of possible scenarios for phenolic components during anaerobic degradation: (1) they remain as insoluble residue after the microbes remove the carbohydrate, (2) all constituents are uniformly solubilized, and (3) some phenolic constituents are solubilized and some are not. Relevant to this issue, Kataeva et al. (2013) noted that proportions of constituents were unchanged after fermentation of untreated switchgrass and poplar, while Paye et al. (2015) found that the fraction of carbohydrate decreased with increasing overall solubilization, suggesting an enrichment of non-carbohydrates in the residue.

In studies of biomass degradation in the bovine rumen, it was noted that up to half of the lignin in grass was solubilized as lignin-carbohydrate complexes (Gaillard and Richards 1975). Further studies suggested that lignin solubilized in the rumen was precipitated later in the bovine gut (Neilson and Richards 1978). Coumaric acid is known to be released by a group of enzymes called the ferulic acid esterases (FAEs; EC 3.1.1.73) (Crepin et al. 2004; Benoit et al. 2008; Koseki et al. 2009). These enzymes are active on the ester bonds of hydroxycinnamic acids and have been identified in a variety of fungi and bacteria. An FAE from *Thermobacilus xylanilyticus* was shown to release phenolic acids from non-pretreated biomass (Rakotoarivonina et al. 2011) and a commercial enzyme mix containing FAE (Depol 740 L from Biocatalysts Ltd., Wales, UK) has been shown to act as an effective form of enzymatic pretreatment while generating a potentially valuable coproduct (Akin 2007). *C. thermocellum* digests cellulose using multienzyme complexes known as the cellulosome, for which the central "scaffoldin" structure is encoded by the gene *cipA* (Olson et al. 2013). Two FAE domains have been described in *C. thermocellum*, both as components of cellulosomal xylanases (Blum et al. 2000), and a third domain is evident in the genome sequence (see Results). To our knowledge, no enzymes have been described that break the ether bond of coumaric acid to lignin.

The goal of this study was to search for non-carbohydrate biomass components solubilized during anaerobic digestion of untreated biomass. After identifying coumaric acid and related hydroxycinnamic acids in culture broth, we then investigated the release of these components in fermentations and enzymatic reactions. We hypothesized that genes encoding FAEs would be necessary for release of coumaric acid.

Materials and methods

Fermentations

Sugarcane bagasse, obtained from the University of Florida (Gainesville, FL, USA), was washed and milled in a Wiley knife mill to pass through a 2-mm screen. Switchgrass, obtained from the University of Wisconsin (Madison, WI, USA), was also milled to 2 mm. Sigmacell and Avicel purified cellulose were obtained from Sigma-Aldrich (St. Louis, MO, USA). Substrates were placed in serum bottles at 10-g/L final concentration with deionized water or water plus MOPS, flushed with ultrapure nitrogen, then autoclaved for 20 min. MTC medium components (Baskaran et al. 1995; Hogsett 1995) were then added. Inocula, grown in CTFUD (Olson and Lynd 2012) were used at 10 % of the final volume. Bottles were incubated shaking at 55 °C for up to 5 days, but as few as 2 days in later experiments comparing strains where complete solubilization was not necessary. Residual solids were recovered by pouring bottle contents into preweighed 50-ml conical tubes, then rinsing the bottles with water and adding the rinsate to the same tubes. The tubes were centrifuged, decanted, and then dried at 50–60 °C before reweighing.

Enzymatic solubilizations

C. thermocellum cellulases were prepared by growing strain ATCC 27405 on 5 g/L Avicel until early stationary phase. The culture broth was left to settle overnight at 4 °C, then decanted. The supernatant was concentrated fivefold using a 500-kDa filter in tangential flow filtration, then frozen until needed. Before use, cellulosome preparations were centrifuged briefly then filter sterilized. Hydrolyses were performed identically to fermentations (i.e., in sterile anaerobic serum bottles) except that no inoculum was used.

Knockout mutations

Strains and plasmids are listed in Table 1. Gene knockouts were made essentially as described in (Olson and Lynd 2012), using strain 1313 with genome sequence accession number CP002416. Knockout plasmids were made using the method of Gibson (2011). Plasmid pDGO68 was digested with *Eco*RV, then mixed in a Gibson reaction with PCR products of upstream, downstream, and internal homology regions of each targeted gene, primers for which are given in Table 2. Plasmid pDGO68 is derived from and is essentially the same as pAMG258 (Olson and Lynd 2012), with minor modifications to enhance Gibson assembly. The Gibson products were transformed into BL21 star chemically competent cells (Invitrogen, Carlsbad, CA), and resulting plasmids were digested with *Aat*II and sequenced with primers X22190-X22193 to verify proper construction. The knockout plasmids were transformed into *C. thermocellum* strain M1354 (or into the subsequent single and double FAE knockouts); then the series of selections were performed as described (Olson and Lynd 2012), verifying clones at each step by PCR using external, internal, and *repB* (X22996-X22997) primers.

HPLC

High-performance liquid chromatography (HPLC) separations and compound identifications were performed using a Thermo-Fisher (Waltham, MA, USA) P2000 gradient pump, AS3000 autosampler, UV3000 scanning UV detector (192–389 nm), with a YMC-Carotenoid (C30) column, 3 μm, 4.6 mm × 150 mm at 0.3 mL/min with the following gradient of 0.1 % formic acid (A) and acetonitrile (B): t0min-97 %A, t5min-80 %A, t27min-70 %A, t33min-40 %A, T33.2 min-97 %A. Compound identities were confirmed by retention time and spectral matching.

Analytical hydrolyses

Mild base hydrolysis was conducted with 0.5 M potassium hydroxide at 60 °C for 1 h to release coumaric acid from ester bonds. Strong base hydrolysis was performed by increasing the molarity of base to 4 M potassium hydroxide and repeating the incubation at 60 °C for an additional hour to release coumaric acid from ether bonds.

Results

Looking for solubilized lignin components, we analyzed fermentation broth of *C. thermocellum* growing on sugarcane bagasse by HPLC using a C30 column designed to analyze carotenoids. Figure 1 shows a typical chromatographic separation and the identities of the main peaks. A large peak representing coumaric acid is observed, along with lesser peaks of ferulic acid and 4-hydroxybenzaldehyde. Table 3 shows the quantity of coumaric acid measured in various fermentations. In an early experiment (rows 1–3), no coumaric acid was observed in a fermentation of purified cellulose (Sigmacell), while 77 mg/L was measured in a bagasse fermentation. Only 7 mg/L was observed in a control bagasse bottle with no inoculum. The amount of bagasse and cells recoverable following fermentation was less in the inoculated bottle; there was 33 % weight loss compared to 7 % in the control. The weight loss in the

Table 1 Strains and plasmids

Strains	KO1[a]	KO2[a]	KO3[a]	KO4[a]	Reference
E. coli BL21star					Invitrogen (Carlsbad, CA, USA)
M1354	*hpt*				Argyros et al. (2011)
M1442	*T. saccharolyticum* engineered ethanologen				Lee et al. (2011)
M1668	*hpt*	*cipA*			Olson et al. (2013)
M5296	*hpt*	1313_1305			This study
M5297	*hpt*	1313_2635			This study
M5298	*hpt*	1313_2858			This study
M5676	*hpt*	1313_1305	1313_2635		This study
M5678	*hpt*	1313_1305	1313_2858		This study
M5680	*hpt*	1313_2635	1313_2858		This study
LL1099	*hpt*	1313_1305	1313_2635	1313_2858	This study
Plasmids					
pMU3602	KO plasmid for gene 1313_1305				This study
pMU3603	KO plasmid for gene 1313_2635				This study
pMU3604	KO plasmid for gene 1313_2858				This study
pDGO68	Backbone plasmid for KO plasmid construction				This study
pAMG258	Backbone plasmid for KO plasmid construction				Olson and Lynd (2012)

[a] Strains of *C. thermocellum* contained between one and four knockout mutations, given here as KO1, KO2, KO3, and KO4

control should be attributed to small amounts of solubilization as well as incomplete recovery of the solids. We measured the bagasse to contain 8.2 % coumaric acid by weight. If 33 % of the total material was uniformly solubilized, then we would expect to observe 271 mg/L coumaric acid. Since we only measured 77 mg/L, the results indicate that either coumaric acid is unstable or partly consumed under the tested conditions or that coumaric acid is not solubilized in proportion to overall biomass solubilization.

In another experiment (Table 3, rows 4–8), bagasse or switchgrass were fermented with either *C. thermocellum*, an ethanologen strain of *Thermoanaerobacterium saccharolyticum* (M1442), or a coculture of both. With just *C. thermocellum*, more coumaric acid was released from bagasse than from switchgrass. No coumaric acid or ferulic acid was observed with either *T. saccharolyticum* or the coculture, though other peaks were observed. To determine the fate of coumaric acid in cultures containing *T. saccharolyticum*, the organism was inoculated into media containing 50 mg/L added coumaric or ferulic acid. Within 2 h, both compounds were eliminated and new peaks with shorter retention times appeared and remained relatively constant for the remainder of growth (data not shown). The spectra of these new peaks were consistent with hydrogenation of coumaric and ferulic acid to phloretic and hydroferulic acid, respectively (Barthelmebs et al. 2000).

To determine whether the release of coumaric acid was dependent on cellulosome structure, a mutant strain of *C. thermocellum* lacking the scaffoldin CipA was grown on bagasse (Table 3, row 9). Coumaric acid release and weight loss were less in the mutant, but greater than an uninoculated control, indicating that coumaric acid release is not strictly dependent on cellulosomal complexing. Since overall growth of this mutant is less on cellulose (Olson et al. 2013), the lower coumaric acid release may simply reflect less cell growth rather than diminished coumaric acid release per se.

The most likely enzymes responsible for release of coumaric acid in *C. thermocellum* are the FAEs. Two have been documented in the literature as domains of xylanases (genes 1313_1305 and 1313_2635). A third FAE domain is present in gene 1313_2858, which is roughly half the size of the other two genes and contains just a cellulose binding module, FAE and dockerin domain. All three genes carry dockerins, indicating their possible inclusion in cellulosomes.

The three FAE domains were deleted, individually and in combination. The deletions were targeted to only the FAE domains so the other functions of those proteins could remain intact. Figure 2 shows the results of a PCR confirming all three deletions in triple FAE knockout strain LL1099. The fragment internal to the three deletions does not amplify in strain LL1099, while the primers located external to the flanking homology regions amplify a product smaller than that from the parent strain. Table 3, rows 10–20, shows that the amount of coumaric acid released and weight loss of

Table 2 Oligonucleotide primers

Primer number	Description[a]	Sequence[b]
X22166	Gene 1305 5′ flank F, 582 bp, up to aa 825	GACATTAACCTATAAAAATAGGCGTATCACGAGGCcagcgtagtggcatcgttta
X22167	Gene 1305 5′ flank R, 582 bp, up to aa 825	CTCAATCATGTTTTTCTTTCtgcggacaagggtttaaata
X22168	Gene 1305 3′ flank F, 624 bp, immediately after CDS	TTAAACCCTTGTCCGCAgaaagaaaaacatgattgagtttgtaa
X22169	Gene 1305 3′ flank R, 624 bp, immediately after CDS	CGGGCAAAAAAATCTTTTCCATGcCTATTCCCACccctttcccctgtcttcttc
X22170	Gene 1305 in. F, 520 bp, in last of gene	CCCCATGCTTTAATACATACCTGGCCCAGTAGTTacggtgaacttgagcctttg
X22171	Gene 1305 in. R, 520 bp, in last of gene	GCATTATCCCTGATTTTTTCACTACTATTAGacgtatccccaccagtgagt
X22172	Gene 1305 external F, 2051 bp in WT	AATCCCAGAACGTTTGTTCC
X22173	Gene 1305 external R, 2051 bp in WT	TCGTAACAAGACCGTCATAACC
X22174	Gene 2635 5′ flank F, 699 bp, 1 bp before start codon	GACATTAACCTATAAAAATAGGCGTATCACGAGGCgcagcaggtgactgtaccaa
X22175	Gene 2635 5′ flank R, 699 bp, 1 bp before start codon	CTGCCATTTGAAGGAAAtttctcctccttttgatttatga
X22176	Gene 2635 3′ flank F, 569 bp, starting at aa 271	ATCAAAAAGGAGGAGAAAtttccttcaaatggcagatg
X22177	Gene 2635 3′ flank R, 569 bp, starting at aa 271	CGGGCAAAAAAATCTTTTCCATGcCTATTCCCACaaagagcctctcccgtaagc
X22178	Gene 2635 in. F, 551 bp, in first 800 bp of gene	CCCCATGCTTTAATACATACCTGGCCCAGTAGTTaggccggcaagagtttattt
X22179	Gene 2635 in. R, 551 bp, in first 800 bp of gene	GCATTATCCCTGATTTTTTCACTACTATTAGatgttgttggcaacgcaata
X22180	Gene 2635 ext F, 2217 bp in WT	GATTGGCCTTGTCCCAAATA
X22181	Gene 2635 ext R, 2217 bp in WT	TGCCGCTCCTATTTACATCC
X22182	Gene 2858 5′ flank F, 603 bp, ends 51 bp into gene	GACATTAACCTATAAAAATAGGCGTATCACGAGGCgagtagccagcgctacaagc
X22183	Gene 2858 5′ flank R, 603 bp, ends 51 bp into gene	CTCCTCCCTGGATAAGCtgctataagcaccaggcaaa
X22184	Gene 2858 3′ flank F, 584 bp starting at aa 254	GCCTGGTGCTTATAGCAgcttatccagggaggaggac
X22185	Gene 2858 3′ flank R, 584 bp starting at aa 254	CGGGCAAAAAAATCTTTTCCATGcCTATTCCCACttatatttccgtcgccgttt
X22186	Gene 2858 in. F, 540 bp, in first 607 bp of gene	CCCCATGCTTTAATACATACCTGGCCCAGTAGTTtaaacatttcatcggcagca
X22187	Gene 2858 in. R, 540 bp, in first 607 bp of gene	GCATTATCCCTGATTTTTTCACTACTATTAGaggtgttcggagctgaagaa
X22188	Gene 2858 external F, 2002 bp in WT	AAAAGGGGCAGCAAGTTTTA
X22189	Gene 2858 external R, 2002 bp in WT	TCCGCATTTATCAGGCTTGT
X22190	pDGO68 insert confirmation	CACCTGACGTCTAAGAAA
X22191	pDGO68 insert confirmation	TCTTTTCCTCTCTTTCGG
X22192	pDGO68 insert confirmation	TTTAAACCCGCTGATCCT
X22193	pDGO68 insert confirmation	GTTGTCTAACTCCTTCCT
X22996	*repB* from pMU102 Ctherm	ATCGCTAATGTTGATTACTTTGAACTTCTGC
X22997	*repB* from pMU102 Ctherm	TGAATCAAATCGCCTTCTTCTGTGTC

[a] Sizes given in base pairs (bp) are those of the generated PCR product

[b] Lowercase and uppercase letters are used to distinguish regions of the oligonucleotide that match either of the two sequences being joined by Gibson assembly

bagasse was essentially unchanged in the deletion strains.

This lack of effect after deletion of the FAEs is possibly due to an unknown enzyme catalyzing the hydrolysis of ether linkages. To determine how much ether-linked coumaric acid is present in our substrate, washed bagasse substrate was subjected to mild base hydrolysis to release coumaric acid from ester bonds and strong base hydrolysis release courmaric acid from ether bonds. Eighty percent of the coumaric acid came from ester bonds and 20 % from ether bonds. Since only a minority of the coumaric acid is ether-linked, a putative etherase is unlikely to be responsible for coumaric acid release in strain LL1099, and a search for other proteins with esterase activity may be more fruitful.

Anaerobic enzymatic solubilization reactions were performed to analyze the coumaric acid releasing activity. Cellulase preparations from *C. thermocellum* (including cellulosomes and free cellulases) catalyzed the release of coumaric acid (Table 4, rows 1, 5, 7). Filter-sterilized enzyme was still active, indicating that cells are not necessary. Boiling greatly diminished the activity (row 2 and Fig. 1).

The cellulosome dissociation method of Morag et al. (1996) was tested in order to begin to identify an enzyme responsible for the activity. Table 4, rows 6–9, shows that the dissociation method greatly reduced the activity. Since a *cipA*(−) strain retains some activity (Table 3, row 9), cellulosomal association is not required for activity. A more likely explanation for the

Table 3 Fermentations

Row	Substrate	Strain (*C. thermocellum* unless noted)	Notes	Hours	Coumaric acid (mg/L)	Std. dev.	Weight loss	Std. dev.
1	Sigmacell	WT 1313		120	0		95 %	
2	Bagasse	WT 1313		120	68.4	14.6	32 %	2 %
3	Bagasse	None		120	9.2	3.0	8 %	1 %
4	Switchgrass	WT 1313	pH 7.0	120	17.8		40 %	
5	Switchgrass	*T. saccharolyticum* M1442	pH 6.6	120	0		20 %	
6	Switchgrass	Coculture	pH 7.0	120	0		34 %	
7	Bagasse	*T. saccharolyticum* M1442	pH 6.6	120	0		6 %	
8	Bagasse	Coculture	pH 7.0	120	0		30 %	
9	Bagasse	*cipA*(−) strain M1668		120	39.1		14 %	
10	Bagasse	M5296	FAE 1305 KO	96	81.4		37 %	
11	Bagasse	M5297	FAE 2635 KO	96	76.8		30 %	
12	Bagasse	M5298	FAE 2858 KO	96	84.1		33 %	
13	Bagasse	M1354	Parent	96	89.7		33 %	
14	Bagasse	M5676	FAE 1305 and 2635 KO	72	44.3	0.8	27 %	2 %
15	Bagasse	M5678	FAE 1305 and 2858 KO	72	56.8	1.8	33 %	3 %
16	Bagasse	M5680	FAE 2635 and 2858 KO	72	46.6	5.3	27 %	1 %
17	Bagasse	M1354	Parent	72	52	0.5	30 %	4 %
18	Bagasse	LL1099	FAE 1305, 2635, and 2858 KO	48	31.5	2.8	28 %	1 %
19	Bagasse	M1354	Parent	48	41.5	4.2	31 %	2 %
20	Bagasse	None		48	12.4	0.4	12 %	6 %

lack of activity after dissociation is that coumaric acid-releasing enzymes were inactivated during the dissociation procedure.

Crude size fractionation with molecular weight cutoff spin filters was performed with broth from the *cipA*(−) strain M1668 (Table 4, rows 10–16). The fraction of protein >10 kDa showed marginally more activity than those >50 or 100 kDa. A fraction >10 kDa but <50 kDa was indistinguishable from a no-enzyme control. While identification of coumaric acid-releasing enzymes in M1668 culture supernatant may be possible, the low activity of M1668 broth makes it challenging, particularly if the activity requires multiple proteins.

Discussion

In this study, we sought to identify lignin components solubilized during anaerobic fermentation of various grass feedstocks. The cross-linking phenolic compound coumaric acid was identified in culture broth, and efforts were undertaken to determine the

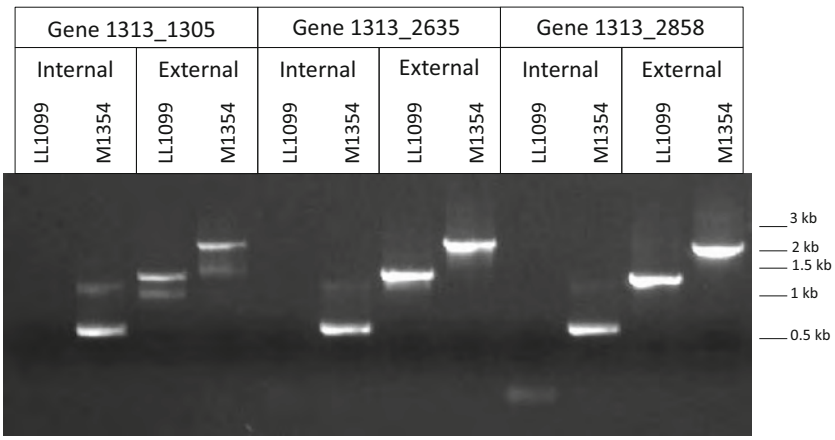

Fig. 2 PCR verification of deletions in strain LL1099. PCR primers located internal to the deletion or external to the cloned flanking regions, as listed in Table 2, were used to amplify genomic DNA from LL1099 (the triple FAE knockout strain) and M1354 (its parent)

Table 4 Anaerobic, enzymatic solubilization of biomass

Row	Substrate	Enzyme	Hours	Coumaric acid (mg/L)	Std dev.	Weight loss	Std dev.
1	Bagasse	WT *C. thermocellum* enzyme	96	26.9	1.5	15 %	2 %
2	Bagasse	Boiled WT *C. thermocellum* enzyme	96	17.1	0.6	8 %	2 %
3	Bagasse	No enzyme	96	11.3	0.5	9 %	2 %
4	Switchgrass	WT *C. thermocellum* enzyme	18	8.6		33.6 %	
5	Bagasse	WT *C. thermocellum* enzyme	18	13.8		27.1 %	
6	Bagasse	None	120	12.8		17 %	
7	Bagasse	WT *C. thermocellum* enzyme	120	42.3		37 %	
8	Bagasse	Filter-sterilized WT *C. thermocellum* enzyme	120	28.7		28 %	
9	Bagasse	Dissociated, filter-sterilized WT *C. thermocellum* enzyme	120	11.7		11 %	
10	Bagasse	None	72	9.1		n/d	
11	Bagasse	M1668 broth >10 kDa	72	13.0		n/d	
12	Bagasse	M1668 broth >50 kDa	72	10.3		n/d	
13	Bagasse	M1668 broth >100 kDa	72	10.2		n/d	
14	Bagasse	None	120	15.2		n/d	
15	Bagasse	M1668 broth > 10 kDa and < 50 kDa	120	13.7		n/d	
16	Bagasse	Filter-sterilized WT *C. thermocellum* enzyme	120	23.9		n/d	

mechanism responsible for its release. The activity was determined to be enzymatic, heat-labile and at least partly independent of the cellulosome. The most likely candidates for the responsible enzymes, the FAEs, were deleted in *C. thermocellum* but the activity was not eliminated. A recent study (Tan et al. 2013) identified the affinity of various extracellular binding proteins for coumaric acid. While those proteins do not have homologues in *C. thermocellum*, the methods may be useful in identifying novel candidates for the undiscovered coumaric releasing enzyme.

The observation of coumaric acid being released during fermentation of grass suggested that at least some phenolic components were being solubilized. In previous studies, switchgrass was solubilized more than wood (Yang et al. 2009; Paye et al. 2015). Incidentally, hydroxycinnamic acids are present in much higher amounts in grass than in wood (Hartley and Harris 1981; Sun et al. 2001) and form cross-links between lignin and hemicellulose (Grabber et al. 1998; Lam et al. 2001, 2003; Siqueira et al. 2011). It seems plausible that coumaric acid-releasing activity may be a key component of the ability of *C. thermocellum* to digest minimally pretreated lignocellulose. FAEs such as Depol 740L (Biocatalysts Ltd., Wales, UK) are commercially available, but there has only been limited investigation of their efficacy with untreated substrates or their ability to replace conventional pretreatment (Akin 2007).

The commercial production of cellulosic ethanol is challenging given the large amount of capital investment required (Gnansounou and Dauriat 2010). Pretreatment is a major component of capital costs and affects the process economics (Kumar and Murthy 2011), and we expect that elimination of pretreatment entirely would greatly improve the cost of ethanol production. The generation of other coproducts is also likely to improve the economics (Gnansounou and Dauriat 2010). Coumaric acid has been recognized as a potentially valuable product as an antioxidant, antimicrobial, and as a precursor to flavoring agents (Mathew and Abraham 2004; Wong 2006; Benoit et al. 2008; Koseki et al. 2009; Shin et al. 2011). Copolymers of the "phytomonomers" coumaric and caffeic acid have been shown to have high mechanical strength and could be developed as rapidly biodegradable bioplastics (Kaneko et al. 2006). Further study to determine the genetic and enzymatic basis of coumaric acid release will be useful to understand microbial depolymerization of untreated lignocellulose and may help facilitate the development of economically and environmentally sustainable fuels.

Acknowledgments The authors wish to thank Julie Paye for helpful discussion, William Kenealy, and Abigail Foster for providing *C. thermocellum* cellulase and Dan Olson for providing strain M1668.

Compliance with ethical standards

Funding This research was funded by Mascoma Corporation and the BioEnergy Science Center, a US Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. This manuscript has been authored by a contractor of the US Government under contract DE-AC05-00OR22725.

Conflict of interest The authors of this study have been employed by Mascoma Corporation and/or Enchi Corporation, which have held a financial interest in technology related to *C. thermocellum* and *T. saccharolyticum*.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Akin DE (2007) Grass lignocellulose: strategies to overcome recalcitrance. Appl Biochem Biotech 137–140:3–15

Argyros DA, Tripathi SA, Barrett TF, Rogers SR, Feinberg LF, Olson DG, Foden JM, Miller BB, Lynd LR, Hogsett DA, Caiazza NC (2011) High ethanol titers from cellulose by using metabolically engineered thermophilic, anaerobic microbes. Appl Environ Microbiol 77:8288–8294. doi:10.1128/AEM.00646-11

Barthelmebs L, Divies C, Cavin JF (2000) Knockout of the *p*-coumarate decarboxylase gene from *Lactobacillus plantarum* reveals the existence of two other inducible enzymatic activities involved in phenolic acid metabolism. Appl Environ Microbiol 66:3368–3375. doi:10.1128/AEM.66.8.3368-3375.2000

Basen M, Rhaesa AM, Kataeva I, Prybol CJ, Scott IM, Poole FL, Adams MWW (2014) Degradation of high loads of crystalline cellulose and of unpretreated plant biomass by the thermophilic bacterium *Caldicellulosiruptor bescii*. Bioresour Technol 152:384–392. doi:10.1016/j.biortech.2013.11.024

Baskaran S, Ahn HJ, Lynd LR (1995) Investigation of the ethanol tolerance of *Clostridium thermosaccharolyticum* in continuous culture. Biotechnol Prog 11:276–281. doi:10.1021/bp00033a006

Benoit I, Danchin EGJ, Bleichrodt RJ, de Vries RP (2008) Biotechnological applications and potential of fungal feruloyl esterases based on prevalence, classification and biochemical diversity. Biotechnol Lett 30:387–396

Blum DL, Kataeva IA, Li XL, Ljungdahl LG (2000) Feruloyl esterase activity of the *Clostridium thermocellum* cellulosome can be attributed to previously unknown domains of XynY and XynZ. J Bacteriol 182:1346–1351. doi:10.1128/JB.182.5.1346-1351.2000

Bugg T, Ahmad M, Hardiman EM, Rahmanpour R (2011) Pathways for degradation of lignin in bacteria and fungi. Nat Prod Rep 28:1883. doi:10.1039/c1np00042j

Chung D, Cha M, Guss AM, Westpheling J (2014) Direct conversion of plant biomass to ethanol by engineered *Caldicellulosiruptor bescii*. Proc Natl Acad Sci U S A 111:8931–6. doi:10.1073/pnas.1402210111

Crepin VF, Faulds CB, Connerton IF (2004) Functional classification of the microbial feruloyl esterases. Appl Microbiol Biotechnol 63:647–652

Gaillard BD, Richards GN (1975) Presence of soluble lignin-carbohydrate complexes in the bovine rumen. Carbohydr Res 42:135–145. doi:10.1016/S0008-6215(00)84106-3

Gibson DG (2011) Enzymatic assembly of overlapping DNA fragments. Methods Enzymol 498:349–361. doi:10.1016/B978-0-12-385120-8.00015-2

Gnansounou E, Dauriat A (2010) Techno-economic analysis of lignocellulosic ethanol: a review. Bioresour Technol 101:4980–4991. doi:10.1016/j.biortech.2010.02.009

Grabber JH, Hatfield RD, Ralph J (1998) Diferulate cross-links impede the enzymatic degradation of non-lignified maize walls. J Sci Food Agric 77:193–200. doi:10.1002/(SICI)1097-0010(199806)77:2<193::AID-JSFA25>3.0.CO;2-A

Hartley RD, Harris PJ (1981) Phenolic constituents of the cell walls of dicotyledons. Biochem Syst Ecol 9:189–203

Hogsett DA (1995) Cellulose hydrolysis and fermentation by *Clostridium thermocellum* for the production of ethanol. PhD Thesis, Dartmouth College, Hanover, NH, USA

Izquierdo JA, Pattathil S, Guseva A, Hahn MG, Lynd LR (2014) Comparative analysis of the ability of *Clostridium clariflavum* strains and *Clostridium thermocellum* to utilize hemicellulose and unpretreated plant material. Biotechnol Biofuels 7:1–8. doi:10.1186/s13068-014-0136-4

Kaneko T, Thi TH, Shi DJ, Akashi M (2006) Environmentally degradable, high-performance thermoplastics from phenolic phytomonomers. Nat Mater 5:966–970. doi:10.1038/nmat1778

Kataeva I, Foston MB, Yang S-J, Pattathil S, Biswal AK, Ii FLP, Basen M, Rhaesa AM, Thomas TP, Azadi P, Olman V, Saffold TD, Mohler KE, Lewis DL, Doeppke C, Zeng Y, Tschaplinski TJ, York WS, Davis M, Mohnen D, Xu Y, Ragauskas AJ, Ding S-Y, Kelly RM, Hahn MG, Adams MWW (2013) Carbohydrate and lignin are simultaneously solubilized from unpretreated switchgrass by microbial action at high temperature. Energy Environ Sci 6:2186–2195. doi:10.1039/c3ee40932e

Koseki T, Fushinobu S, Ardiansyah, Shirakawa H, Komai M (2009) Occurrence, properties, and applications of feruloyl esterases. Appl Microbiol Biotechnol 84:803–810

Kumar D, Murthy GS (2011) Impact of pretreatment and downstream processing technologies on economics and energy in cellulosic ethanol production. Biotechnol Biofuels 4:27. doi:10.1186/1754-6834-4-27

Lam TBT, Kadoya K, Iiyama K (2001) Bonding of hydroxycinnamic acids to lignin: ferulic and *p*-coumaric acids are predominantly linked at the benzyl position of lignin, not the β-position, in grass cell walls. Phytochemistry 57:987–992. doi:10.1016/S0031-9422(01)00052-8

Lam TBT, Iiyama K, Stone BA (2003) Hot alkali-labile linkages in the walls of the forage grass *Phalaris aquatica* and *Lolium perenne* and their relation to in vitro wall digestibility. Phytochemistry 64:603–607. doi:10.1016/S0031-9422(03)00301-7

Lee JM, Venditti RA, Jameel H, Kenealy WR (2011) Detoxification of woody hydrolyzates with activated carbon for bioconversion to ethanol by the thermophilic anaerobic bacterium *Thermoanaerobacterium saccharolyticum*. Biomass Bioenergy 35:626–636. doi:10.1016/j.biombioe.2010.10.021

Lynd LR, Elamder RT, Wyman CE (1996) Likely features and costs of mature biomass ethanol technology. Appl Biochem Biotechnol 57–58:741–761

Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev 66:506–577. doi:10.1128/MMBR.66.3.506-577.2002

Masarin F, Gurpilhares DB, Baffa DC, Barbosa MH, Carvalho W, Ferraz A, Milagres AM (2011) Chemical composition and enzymatic digestibility of sugarcane clones selected for varied lignin content. Biotechnol Biofuels 4:55. doi:10.1186/1754-6834-4-55

Mathew S, Abraham TE (2004) Ferulic acid: an antioxidant found naturally in plant cell walls and feruloyl esterases involved in its release and their applications. Crit Rev Biotechnol 24:59–83. doi:10.1080/07388550490491467

Morag E, Yaron S, Lamed R, Kenig R, Shoham Y, Bayer EA (1996) Dissociation of the cellulosome of *Clostridium thermocellum* under nondenaturing conditions. J Biotechnol 51:235–242. doi:10.1016/S0168-1656(96)01601-X

Neilson MJ, Richards GN (1978) The fate of the soluble lignin-carbohydrate complex produced in the bovine rumen. J Sci Food Agric 29:513–519

Olson DG, Lynd LR (2012) Transformation of *Clostridium thermocellum* by electroporation. Methods Enzymol 510:317–330. doi:10.1016/B978-0-12-415931-0.00017-3

Olson DG, Giannone RJ, Hettich RL, Lynd LR (2013) Role of the CipA scaffoldin protein in cellulose solubilization, as determined by targeted gene deletion and complementation in *Clostridium thermocellum*. J Bacteriol 195:733–739. doi:10.1128/JB.02014-12

Paliwal R, Rawat AP, Rawat M, Rai JPN (2012) Bioligninolysis: recent updates for biotechnological solution. Appl Biochem Biotechnol 167:1865–1889. doi:10.1007/s12010-012-9735-3

Paye J, Guseva A, Hammer S, Gjersing E, Davis M, Davison B, Olstad J, Donohoe B, Nguyen T, Wyman C, Pattathil S, Hahn M, Lynd L (2015) Biological lignocellulose solubilization: comparative evaluation of biocatalysts and enhancement via cotreatment. Biotechnol Biofuels (in press)

Pu Y, Hu F, Huang F, Davison BH, Ragauskas AJ (2013) Assessing the molecular structure basis for biomass recalcitrance during dilute acid and hydrothermal pretreatments. Biotechnol Biofuels 6:15. doi:10.1186/1754-6834-6-15

Rakotoarivonina H, Hermant B, Chabbert B, Touzel JP, Remond C (2011) A thermostable feruloyl-esterase from the hemicellulolytic bacterium *Thermobacillus xylanilyticus* releases phenolic acids from non-pretreated plant cell walls. Appl Microbiol Biotechnol 90:541–552. doi:10.1007/s00253-011-3103-z

Saulnier L, Vigouroux J, Thibault JF (1995) Isolation and partial characterization of feruloylated oligosaccharides from maize bran. Carbohydr Res 272:241–253. doi:10.1016/0008-6215(95)00053-V

Scalbert A (1985) Ether linkage between phenolic acids and lignin fractions from wheat straw. Phytochemistry 24:1359–1362

Shin SY, Han NS, Park YC, Kim MD, Seo JH (2011) Production of resveratrol from *p*-coumaric acid in recombinant *Saccharomyces cerevisiae* expressing 4-coumarate:coenzyme A ligase and stilbene synthase genes. Enzyme Microb Technol 48:48–53. doi:10.1016/j.enzmictec.2010.09.004

Siqueira G, Milagres AM, Carvalho W, Koch G, Ferraz A (2011) Topochemical distribution of lignin and hydroxycinnamic acids in sugar-cane cell walls and its correlation with the enzymatic hydrolysis of polysaccharides. Biotechnol Biofuels 4:7. doi:10.1186/1754-6834-4-7

Sun RC, Sun XF, Zhang SH (2001) Quantitative determination of hydroxycinnamic acids in wheat, rice, rye, and barley straws, maize stems, oil palm frond fiber, and fast-growing poplar wood. J Agric Food Chem 49:5122–5129. doi:10.1021/jf010500r

Talluri S, Raj SM, Christopher LP (2013) Consolidated bioprocessing of untreated switchgrass to hydrogen by the extreme thermophile *Caldicellulosiruptor saccharolyticus* DSM 8903. Bioresour Technol 139:272–279. doi:10.1016/j.biortech.2013.04.005

Tan K, Chang C, Cuff M, Osipiuk J, Landorf E, Mack JC, Zerbs S, Joachimiak A, Collart FR (2013) Structural and functional characterization of solute binding proteins for aromatic compounds derived from lignin: *p*-coumaric acid and related aromatic acids. Proteins Struct Funct Bioinform 81:1709–1726. doi:10.1002/prot.24305

Wong DWS (2006) Feruloyl esterase: a key enzyme in biomass degradation. Appl Biochem Biotechnol 133:87–112. doi:10.1385/ABAB:133:2:87

Yang B, Wyman CE (2008) Pretreatment: the key to unlocking low-cost cellulosic ethanol. Biofuels Bioprod Biorefining 2:26–40

Yang SJ, Kataeva I, Hamilton-Brehm SD, Engle NL, Tschaplinski TJ, Doeppke C, Davis M, Westpheling J, Adams MWW (2009) Efficient degradation of lignocellulosic plant biomass, without pretreatment, by the thermophilic anaerobe "*Anaerocellum thermophilum*" DSM 6725. Appl Environ Microbiol 75:4762–4769. doi:10.1128/AEM.00236-09