Appl Microbiol Biotechnol
DOI 10.1007/s00253-015-6610-5

Development of a regulatable plasmid-based gene expression system for *Clostridium thermocellum*

Elizabeth B. Mearls[1,2,4] · **Daniel G. Olson**[1,2] · **Christopher D. Herring**[1,3] · **Lee R. Lynd**[1,2]

Received: 30 October 2014 / Revised: 7 April 2015 / Accepted: 15 April 2015

Abstract *Clostridium thermocellum* can rapidly solubilize cellulose and produces ethanol as an end product of its metabolism. As such, it is a candidate for bioethanol production from plant matter. In this study, we developed an inducible expression system for *C. thermocellum* based on its native *celC* operon. We enhanced expression over the native operon structure by placing the repressor gene, *glyR3*, immediately after the *celC* promoter, and expressing the target gene after *glyR3*. Upon the addition of the inducer substrate, laminaribiose, an approximately 40-fold increase in gene expression was obtained using the test gene *spo0A*. Furthermore, induction of the sporulation histidine kinase, *clo1313_1942*, increased sporulation frequency by approximately 10,000-fold relative to an uninduced control. We have also shown that the laminaribiose (β1-3-linked carbon source) utilization pathway is not catabolite repressed by cellobiose, a β1-4-linked carbon source frequently used for *C. thermocellum* cultivation in laboratory conditions. Selective expression of target genes has the potential to inform metabolic engineering strategies as well as increase fundamental understanding of *C. thermocellum* biology.

✉ Lee R. Lynd
lee.r.lynd@dartmouth.edu

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

[2] BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

[3] Mascoma Corporation, 67 Etna Rd., Lebanon, NH 03766, USA

[4] Present address: Department of Biological Sciences, Mount Holyoke College, South Hadley, MA 01075, USA

Keywords Laminaribiose · Inducible promoter · Biofuels · Spore formation

Introduction

Clostridium thermocellum is an anaerobic bacterium that has the ability to solubilize plant biomass components rapidly using both a tethered enzyme complex known as the cellulosome (Bayer et al. 2004), and free cellulases (Johnson et al. 1982). *C. thermocellum* also produces ethanol as a by-product of its metabolism (McBee 1950), making it an organism of interest for the sustainable production of ethanol fuel. The cellulosome and cellulase systems of *C. thermocellum* have been thoroughly characterized in strain ATCC 27405 (American Type Culture Collection) (Ellis et al. 2012; Lynd et al. 2002). However, the potential to increase ethanol yields and titers in this strain through genetic manipulation is limited by its low and variable transformation rate (Olson and Lynd 2012). For this reason, metabolic engineering efforts have been focused on strain DSMZ 1313 (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Germany), for which it has been possible to make several genetic modifications. Methods for making markerless gene deletions have been developed (Argyros et al. 2011; Olson and Lynd 2012; Tripathi et al. 2010). Although much has been learned from these studies, we have encountered limitations when trying to investigate the function of genes involved in central metabolism. In addition to central metabolism, there are other aspects of *C. thermocellum* biology that need to be characterized at the molecular level in order to optimize growth dynamics and end product formation. This includes the process of spore formation, which causes growth arrest in vegetative cells and has been shown to negatively impact product yields in other clostridial systems (Qureshi et al. 1988, 2004). In line with this,

Published online: 21 May 2015

the abolishment of spore formation in combination with other metabolic engineering strategies has been effectively used to improve product yields and substrate utilization for *Clostridium cellulolyticum* (Li et al. 2014), and downregulation of the sporulation master regulator gene *spo0A* was also partially responsible for improved robustness in the *Populus* hydrolysate mutant of *C. thermocellum* strain ATCC 27405 (Linville et al. 2013). These studies have demonstrated the value of reducing spore formation during growth in order to maximize product formation and growth. However, the ability to produce spores is a valuable trait to retain as spores are easily stored and can be used as hearty sources of inocula. Hence, it would be useful to have the ability to manipulate spore formation during growth while still retaining the strain's sporulation capability. A regulatable genetic system would be ideal for this purpose.

Several inducible expression systems have been developed for *Clostridium* species. These include systems regulated by xylose for *Clostridium perfringens* (Nariya et al. 2011), lactose for *Clostridium ljungdahlii* and *Clostridium acetobutylicum* (Banerjee et al. 2014; Heap et al. 2007), tetracycline for *Clostridium difficile* (Fagan and Fairweather 2011), anhydrotetracycline for *C. acetobutylicum* (Dong et al. 2012), and most recently arabinose for *C. cellulolyticum* (Zhang et al. 2015). However, none of these inducible systems have been adapted for use in *C. thermocellum*. Furthermore, mechanisms for the transport of lactose, arabinose, or xylose have not been identified or characterized in *C. thermocellum*, which preferentially transports longer chain cellodextrins as opposed to simple carbon sources (Nataf et al. 2009; Zhang and Lynd 2005). Therefore, at least a few of these potential inducer molecules are less than ideal candidates for the creation of an efficient inducible system for *C. thermocellum*, as systems for their uptake into the cell would need to be better characterized and potentially optimized.

An alternative approach for the construction of an inducible expression system would be to adapt an existing inducible promoter from the host organism in question. The *celC* operon, which encodes two free cellulase genes, *celC* and *licA*, and one repressor protein gene, *glyR3*, was identified and characterized in *C. thermocellum* ATCC 27405 (and the corresponding strain DSMZ 1237) (Fuchs et al. 2003; Newcomb et al. 2007). Newcomb et al. reported the presence of a specific palindromic operator sequence in the *celC* promoter region responsible for the regulation of the operon via GlyR3 binding. They also determined that GlyR3 ceases to bind the operator in the presence of laminaribiose, but not in the presence of β-1-4-linked sugars (Newcomb et al. 2007). Expression of the *celC* operon was also induced by growth on laminarin, with peak expression occurring in late exponential phase (Newcomb et al. 2011). Expression levels of *celC* during growth on laminaribiose have not been previously determined; however, in vitro studies using cell lysate suggested that it is possible for expression to occur within 60 min after addition of laminaribiose (Newcomb et al. 2007). Laminaribiose is proposed to enter the cell via the laminaribiose-binding protein (Lbp) ABC transport system (Nataf et al. 2009). The promoter region of Lbp does not contain the palindromic sequence responsible for GlyR3 binding and is therefore unlikely to be regulated by GlyR3. To date, this is the only regulated operon that has been characterized in *C. thermocellum*.

The effect of spore formation on product yields has not been evaluated for *C. thermocellum*, but because spore formation arrests vegetative growth, there is potential for sporulation to negatively affect metabolic activity. Based on existing studies, clostridial sporulation is initiated when histidine kinases directly phosphorylate the sporulation master response regulator, Spo0A, resulting in Spo0A~P (Steiner et al. 2011; Underwood et al. 2009). Accumulation of Spo0A~P, the direct result of kinase activity (12), commits the cell to spore formation. In previous work, we found that expression of the histidine kinase Clo1313_1942 induced sporulation to levels 100-fold greater than the wild type when expressed under the native *C. thermocellum gapDH* promoter (Mearls and Lynd 2014). This result leads us to believe that varying the expression levels of this kinase would directly impact sporulation levels in a titratable manner. As kinase activity is not only an indicator of transcriptional output but also a functional output, we determined that assaying the sporulation response after induction of Clo1313_1942 would be an effective way to validate our inducible construct, as well as provide additional insight into the mechanics of spore formation in *C. thermocellum*.

Inducible expression systems are extremely useful molecular tools that provide a means to express recombinant proteins, evaluate the function of unknown genes or proteins in a cell, regulate expression of essential genes, and map various pathways. In this study, we sought to develop a plasmid-based inducible expression system for *C. thermocellum* based on the *celC* operon and to use this system to understand how variable expression of the known sporulation kinase Clo1313_1942 impacts the sporulation response.

Materials and methods

Strains and culturing conditions Yeast and bacterial strains are listed in Table 1. *Saccharomyces cerevisiae* InvSc1 was grown in YPD liquid media and plated on SD-Ura medium (Sunrise Science Products, San Diego, CA, USA) for the selection of Ura3+ plasmids. *Escherichia coli* strains were grown on LB medium supplied with chloramphenicol (15 μg/ml) for plasmid maintenance. The *C. thermocellum spo0A* deletion strain, M1726, was a gift from the Mascoma

Table 1 Strains and plasmids used in this study

Strain or plasmid	Description and or relevant characteristics	Reference, source, or construction
Microbial strains		
S. cerevisiae InvSc1	Uracil auxotroph	Invitrogen
E. coli Top10	Used for plasmid maintenance	Invitrogen
E. coli BL21 (DE3)	Used for plasmid maintenance and appropriate DNA methylation	New England Biolabs
C. thermocellum strains		
DSMZ 1313	WT	DSMZ (Germany)
M1354	Δ*hpt*	Argyros et al. AEM, 2011
M1726	Δ*hpt*Δ*spo0A*	Mascoma Corporation
LL1050	DSMZ 1313 containing pEBM130	This study
LL1051	M1726 containing pEBM113	This study
LL1052	M1726 containing pEBM107	This study
LL1088	DSMZ 1313 containing pEBM109	This study
Plasmids		
pAMG205	Cloning vector, contains oriColE1, CEN6, ARSH4, URA3, pGapD-cat, pNW33N replication origin	Guss et al. Biotech for Biofuels, 2012
pEBM107	Contains the *celC* promoter region, *spo0A*, and *glyR3* cloned into pAMG205 EcoRI site	This study
pEBM109	Contains the *celC* promoter region and *glyR3* cloned into pAMG205 EcoRI site	This study
pEBM113	Contains the *celC* promoter region, *glyR3*, and *spo0A* cloned into pAMG205 EcoRI site	This study
pEBM130	Resembles pEBM113 with *clo1313_1942* replacing *spo0A*	This study

Corporation and was constructed in *C. thermocellum* strain DSMZ 1313 (Feinberg et al. 2011) following the procedures as outlined by Argyros and coworkers (Argyros et al. 2011). Briefly, the plasmid pMU1810 (GenBank Accession # KC779546) was used to delete the gene encoding the ortholog of the *Bacillus subtilis* response regulator of sporulation, *spo0A* (*clo1313_1409*), in a *C. thermocellum* Δ*hpt* background (M1354) (Argyros et al. 2011). The Δ*hpt* background provides a method for generating markerless deletions via negative selection when the toxic purine analog, 8-azahypoxanthine, is supplied. Deletion of *clo1313_1409* in this background yielded the Δ*hpt* Δ*spo0A* genotype henceforth referred to as M1726. *C. thermocellum* was cultured in liquid media in an anaerobic chamber (Coy Laboratory Products, Grass Lake, MI) in modified DSM122 liquid media (Dror et al. 2003) supplemented with 50 mM MOPS and 10 mM sodium citrate. Growth media contained cellobiose (Sigma), laminaribiose (Megazyme), or both as specified. *C. thermocellum* strains were transformed via electroporation according to established protocols (Olson and Lynd 2012), and 12 μg/ml thiamphenicol was used for selection and cultivation.

Testing for catabolite repression DSMZ 1313 was grown in the presence of 10 mM cellobiose, 5 mM cellobiose+5 mM laminaribiose, or 10 mM laminaribiose to check for catabolite repression. Growth was monitored every 3 min via optical density (OD) reads at 600 nm in an anaerobic spectrophotometer. Data shown for each sample represents the average of three replicate experiments.

Recombinant plasmid construction All plasmids were constructed via yeast gap repair cloning (Shanks et al. 2006, 2009), and *E. coli* was used as a host strain for plasmid propagation. To construct pEBM113 (GenBank Accession # KC920664), pAMG205 (Guss et al. 2012) was cut once with *Eco*RI restriction endonuclease (New England Biolabs (NEB), Ipswich, MA) and gel purified using Zymoclean™ Gel DNA Recovery Kit (Zymo Research, Orange, CA, USA). The DNA regions and target genes used for cloning pEBM113 included a previously described 200 bp region upstream of *celC* containing the *celC* promoter region (Newcomb et al. 2007), *glyR3* (*clo1313_0396*), and *spo0A* (*clo1313_1409*). These DNA regions were amplified from *C. thermocellum* genomic DNA using high-fidelity Phusion polymerase (NEB) with an initial denaturation step of 98° for 3' followed by 30 cycles of 98° for 30 s, 65° for 30 s, 72° for 60 s, with a final extension step of 72° for 5'. DNA oligos used to amplify DNA regions were obtained from Integrated DNA Technologies (IDT) and are listed in Table 2. Oligos used in the cloning reaction included 30–40 bp of overlapping homology to either the cut ends of pAMG205 or the adjacent DNA region in the cloning reaction, and oligos flanking *spo0A* also introduced unique restriction sites *Sph*I and *Not*I, which allow

Table 2 Primers used in this study

Primer name	Gene/DNA region	Sequence 5′-3′	Reference
Yeast cloning primers:			
CelCF.YH	*celC* promoter	gtcggaagactgggccttttgttttggtaccgaagcatgcaattaaattagaatgaac	This study
CelCR.YH	*celC* promoter	cgacccgaatcttatttgatgtcaactcgagctcgagtttttcctcctgaaatattgtg	This study
spo0AF.YH	*spo0A* (*clo1313_1409*)	acacaatatttcaggaggaaaaactcgagctcgagttgacatcaaataagattcgg	This study
spo0AR.YH	*spo0A* (*clo1313_1409*)	ttcactggtcatcgtccacctcccctagatctagatctttatgtagccattttcatttcc	This study
GlyR3F.YH	*glyR3* (*clo1313_0396*)	atggctacataaagatctagatctaggggaggtggacgatgaccagtgaagaaatagcaa	This study
Glyr3R.YH	*glyR3* (*clo1313_0396*)	gatcagcgggtttaaacgctgaggcgcgccgaacccgggtcagaattccaaagccctcttg	This study
pEM130F	*clo1313_1942*	ccaagagggctttggaattctgaaggagggcatgcatgcaagaaccggtatacacaattc	This study
pEM130R	*clo1313_1942*	gaggcgcgccgaattgcggccgctcaatagtatatatcagaaaattctatatgcacattg	This study
CelCF.YH.EV	*celC* promoter	gccttttgttttggtaccggaattcgagctcgcatgcaattaaattagaatgaacgcgcg	This study
CelCR.YH.EV	*celC* promoter	ctggaaacaccacataattttgctatttcttcactggtcatcgtccacctcccctgtcg	This study
GlyR3F.YH.EV	*glyR3* (*clo1313_0396*)	tttcaggaggagcggccgctgtcgacaggggaggtggacgatgaccagtgaagaaatagc	This study
GlyR3R.YH.EV	*glyR3* (*clo1313_0396*)	aaacgctgaggcgcgccgaattgaattcgagctccccgggtcaaaattccaaagccctc	This study
qRT-PCR primers:			
F: celC-F	*celC* (*clo1313_0395*)	cgggaacatattgcctttgaac	Newcomb et al., PNAS 2007
R: celC-R	*celC* (*clo1313_0395*)	ggtggaatcaatttccctgattg	Newcomb et al., PNAS 2007
F: X0352	*recA* (*clo1313_1163*)	gttgcggtaaatctcgatattgtaaa	This study
R: X0353	*recA* (*clo1313_1163*)	ggccaatcttctgaccgttg	This study
F: spo0AF	*spo0A* (*clo1313_1409*)	gagggatgccataataatggtagtc	This study
R: spo0AR	*spo0A* (*clo1313_1409*)	ctataaattccgaatttgtgggtttacc	This study

for the removal of *spo0A*. Cloning was performed via yeast gap repair cloning using a modified Lazybones protocol (Elble 1992; Shanks et al. 2006), in which the amplified fragments and plasmid backbone were assembled into pEBM113 via yeast homologous recombination. To construct pEBM130 (GenBank Accession # KC788287), pEBM113 was cut with *Sph*I and *Not*I, and gel purified to extract only the plasmid backbone and remove *spo0A*. The gene encoding a putative histidine kinase, *clo1313_1942*, was amplified with oligos with homology to the cut ends of pEBM113, and the vector was again assembled via yeast gap repair cloning. To construct an empty vector control (pEBM109, GenBank Accession # KC779543) harboring only the *celC* promoter and *glyR3*, pAMG205 was cut with *Eco*RI. The *celC* promoter region (as described above) and *glyR3* were amplified with primers that included 30–40 bp of overlapping homology to either the cut ends of pAMG205 or the adjacent DNA region in the cloning reaction, then assembled via yeast gap repair cloning. Plasmids were isolated from yeast using Zymoprep Yeast Plasmid Miniprep II Kit (Zymo Research) and were introduced via electroporation into *E. coli* Top10 (Invitrogen, Carlsbad, CA). Subsequently, plasmids were transformed via chemical competence into *E. coli* BL21 (Cat# C2523, NEB) before *C. thermocellum* electro-transformation. Plasmids were isolated from *E. coli* using QIAGEN QIAprep Spin Miniprep Kit (QIAGEN, Valencia, CA, USA).

Laminaribiose induction Strain M1726/pEBM113 (LL1051) was cultured overnight on liquid media containing 10 mM cellobiose, subcultured to an optical density at 600 nm (OD_{600}) of 0.1, and allowed to grow to late exponential phase (OD 0.7–0.9) before being induced. During the induction procedure, cultures were split into two aliquots, and either laminaribiose or cellobiose (uninduced control) was added to each aliquot. Unless otherwise stated, 10 mM (final concentration) of laminaribiose was used for induction. When less than 10 mM laminaribiose was used, the culture was supplemented with the appropriate amount of cellobiose so that total additional carbon source was always 10 mM. Samples for gene expression analysis were harvested into RNAprotect Reagent (QIAGEN) at 1, 2, 3, 6, or 9 h after induction as noted. Cell pellets were stored at −20° for up to 1 week before RNA extraction. Three independent biological replicates were obtained for all analyses.

RNA isolation and gene expression measurements RNA was isolated from all samples using RNeasy Mini Kit (QIAGEN). An on-column DNase (QIAGEN) treatment was used to eliminate DNA contamination; however, an additional DNase treatment followed by a secondary cleanup procedure was necessary to remove all traces of DNA for subsequent analysis via quantitative RT-PCR.

All DNA oligos used for RT-PCR are listed in Table 2. DNA oligos designed to target a 70-bp region of *recA*

(*clo1313_1163*) and 180-bp region of *spo0A* (*clo1313_1409*) were created and searched against the *C. thermocellum* chromosome for uniqueness using the BLAST algorithm. For XD352 (*clo1313_1163*), the next highest BLAST hit had an expectation value of 0.58, and for XD353 (*clo1313_1163*), the next highest BLAST hit had an expectation value of 0.35. For *spo0AF* (*clo1313_1409*), the next highest BLAST hit had an expectation value of 0.15, and for *spo0AR* (*clo1313_1409*), the next highest BLAST hit expectation value was 0.18. Additionally, no secondary melt peaks were detected during melt-curve analysis using Bio-Rad CFX manager software, version 2.1. Primer amplification efficiency was determined by using nine tenfold dilutions of genomic DNA to generate a standard curve for all primer pairs. The amplification efficiency for *spo0A*, *celC*, and *recA* primer pairs were 91, 104, and 112 %, respectively. The cDNA library and subsequent qPCR reaction was performed in a one-step reaction using iScript cDNA Synthesis Kit (Bio-Rad) on CFX96 qPCR machine (Bio-Rad Corporation) with an annealing temperature of 50 °C and other cycling parameters as suggested by the iScript datasheet. Cq values were determined using Bio-Rad CFX manager software, and normalized gene expression levels were calculated using the ΔΔCq method with *recA* serving as the reference gene. All samples were run in duplicate, and at least three replicates were obtained for each sample.

Induction of spore formation and spore quantification DSM1313 expressing pEBM130 (LL1050) was grown on laminaribiose and cellobiose at different concentrations to evaluate spore formation as a result of *clo1313_1942* expression. Ratios of laminaribiose to cellobiose included 0:10, 2:8, 5:5, 8:2, and 10:0 mM. All cultures were grown for 24 h. Spores were quantified after 24 h by heat-treating cultures at 100 °C for 30 min as described previously (Mearls et al. 2012). Spores were quantified via serial dilution followed by plating in modified DSM122 media containing 8 % agar. Three separate biological replicates were obtained, and spore numbers are reported as colony-forming units/ml of culture (CFU/ml).

Results

Lack of catabolite repression by cellobiose and conditions for maximum induction of the *celC* promoter

Cellobiose and laminaribiose are both glucose dimers that differ in their β-linkages (Fig 1a). While *C. thermocellum* can grow on both β1-4 and β1-3-linked carbon sources, it was unknown whether it would preferentially utilize one of these carbon sources over the other when grown in the presence of both. Catabolite repression via another substrate could reduce expression from our system if the organism preferred to use the other substrate first. We verified that the *celC* operon was not catabolite repressed by cellobiose. Cultures were grown in the presence of cellobiose, laminaribiose, or a combination of both substrates to ascertain their growth patterns and check for diauxic growth when both substrates were present. Growth curves for all cultures were highly similar, and

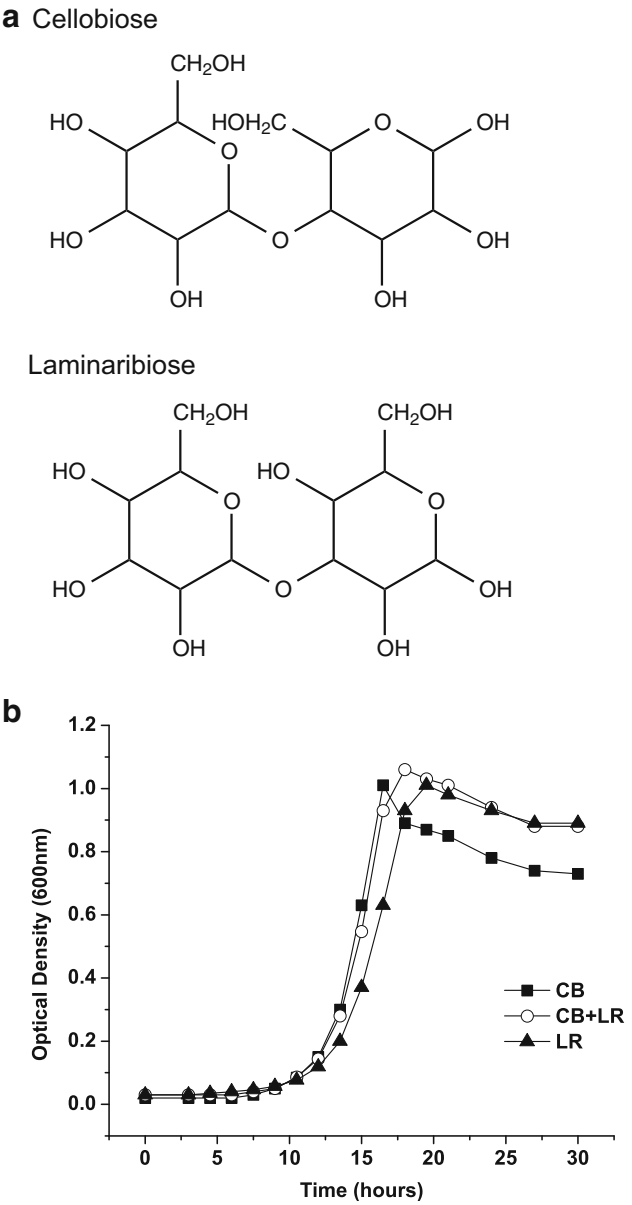


Fig. 1 Growth curves for DSM1313 in cellobiose vs. laminaribiose. **a** The chemical structure of cellobiose (β-1-4 linkage) and laminaribiose (β-1-3 linkage). **b** DSMZ 1313 was grown in the presence of 10 mM cellobiose (*CB*; *squares*), 5 mM cellobiose+5 mM laminaribiose (*CB+LR*; *open circles*), and 10 mM laminaribiose (*LR*; *triangles*). Data shown for each sample represents the average of three replicate experiments. No significant difference in growth was observed between conditions at any time point, as was determined by an unpaired Student's *t* test assuming unequal variance

diauxic growth was not observed in the cultures that contained both substrates (Fig. 1b).

Next, we wanted to measure how quickly laminaribiose induction of the *celC* operon occurs and how long it is sustained in vivo; we induced *C. thermocellum* cells in exponential phase (OD 0.5–0.7) with 10 mM laminaribiose and measured the expression of *celC* via quantitative RT-PCR. Our results indicate that expression of *celC* occurs within an hour of induction (Fig. 2). The highest levels of induction measured were between 1 and 2 h after induction, and expression dropped significantly after 3 h. In early experiments, we noted that the addition of excess laminaribiose (up to 20 mM) did not significantly increase induction, and likewise did not significantly extend the time of induction past the 2-h mark (data not shown). Therefore, we chose 10 mM laminaribiose as our maximum amount of inducer substrate for all remaining experiments. We also chose the 2-h time point as a reference point in subsequent experiments, as expression levels were the most consistent.

Reorganization of the *celC* operon enhances expression from a plasmid

As the *celC* operon is self-regulated through GlyR3, we hypothesized that all the machinery needed for effective expression and repression lies within the operon and the promoter. We therefore sought to clone the important segments of this operon in such a way as to resemble the native chromosomal structure (Fig. 3a) with the following modifications: (1) the *celC* gene was replaced with a reference gene, *spo0A*, and (2) *licA* was not cloned. We chose to use *spo0A* as a reference gene as a *spo0A* deletion mutant (M1726) was available, which obviated the need to distinguish chromosomal versus episomal expression. This initial construct was named pEBM107, and this construct was transformed into the *spo0A* deletion strain.

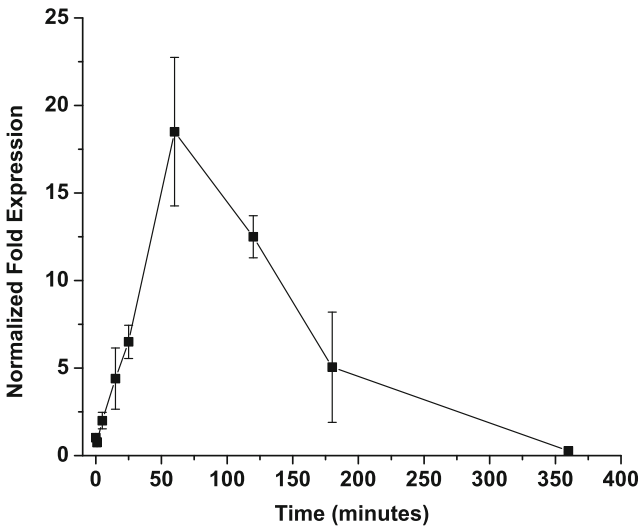


Fig. 2 Time course of *celC* expression after laminaribiose induction measured by qRT-PCR. DSM1313 cells in late exponential phase were induced with 10 mM laminaribiose and harvested at 0, 1, 5, 15, 25, 60, 120, 180, and 360 min after induction to determine *celC* expression over time. All samples were compared against an uninduced control supplemented with 10 mM cellobiose. *Error bars* represent one standard deviation, $n=3$

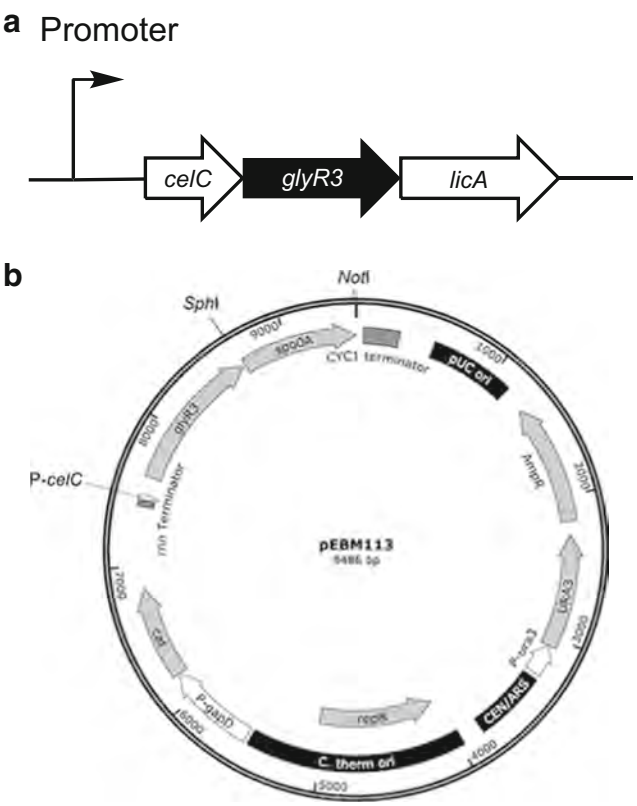
We tested the ability of laminaribiose to induce the expression of *spo0A* in this strain, but *spo0A* uninduced expression was high and increased no more than twofold upon induction, suggesting leaky expression (data not shown). To increase repression, a new vector was cloned where *glyR3* was situated before the target gene and right after the *celC* promoter. This construct was called pEBM113 (Fig. 3b) and was transformed into M1726 resulting in strain M1726/pEBM113 (LL1051). When laminaribiose was used to induce expression of *spo0A* in M1726/pEBM113, *spo0A* expression increased nearly 40-fold (Fig. 3c). *CelC* expression from the native locus served as an internal control, and also increased after induction roughly 15-fold compared to uninduced levels.

Modulation of gene expression

In order to explore the effect of toxic genes or to regulate specific levels of gene expression, it would be useful to be able to modulate the level of expression from the *celC* promoter. We tested whether or not varying the levels of laminaribiose would serve to vary gene expression levels with our construct. We induced late exponential phase M1726/pEBM113 (LL1051) cells with 0, 2, 5, 8, and 10 mM laminaribiose and measured gene the abolishment of spore formation of *spo0A* after 2 h. Our results indicate that gene expression of our target gene (*spo0A*) increased in a nearly linear fashion with the levels of laminaribiose (Fig. 4). However, unlike *spo0A*, *celC* expression leveled off when laminaribiose was supplied in excess of 8 mM. This is likely indicative of the maximum level of induction possible for the native operon.

Titration of the sporulation response via laminaribiose induction

In order to demonstrate a functional output for our system, we first attempted to induce spore formation in the *spo0A* deletion background by complementing the strain with pEBM113 grown in the presence of laminaribiose. However, even though we occasionally observed spore-like structures via microscopy, no viable spores were detected via heat treatment and subsequent viable count plating. This was not entirely unexpected, as previous studies have shown that spore formation is dependent on the cell reaching threshold levels of phosphorylated Spo0A protein, rather than being purely dependent on levels of *spo0A* expression (Burkholder and Grossman 2000; Chung et al. 1994).



◀ **Fig. 3** The *celC* operon, plasmid construct, and gene expression measurements. **a** The *celC* operon of *C. thermocellum* is comprised of the *celC* promoter region, two free cellulase genes, *celC* and *licA*, and a LacI-like repressor gene that negatively regulates the operon, *glyR3*. The *celC* promoter region and *glyR3* were included in subsequent plasmid constructions. **b** The second construct, pEBM113 (GenBank Accession # KC779545), reversed the positions of *spo0A* and *glyR3* and introduced unique restriction sites flanking *spo0A*. The *plasmid map* shows the positions of replication origins (*black*), promoters (*white*), terminators (*dark gray*), and protein-coding regions (*light gray*), which include *spo0A* and *glyR3*, as well as genes encoding chloramphenicol and thiamphenicol resistance (*cat*), orotidine 5-phosphate decarboxylase (*ura3*), and the gene for the *C. thermocellum* replication protein, *repB*. Full-sequence files have been uploaded to GenBank. Plasmid map was generated with SnapGene software (from GSL Biotech; available at snapgene.com). **c** The expression levels of the target gene, *spo0A* (*black bars*), and the chromosomal control gene, *celC* (*white bars*), were measured using qRT-PCR in M1726/pEBM113 (LL1051). Cultures were grown to exponential phase and then cultured with either 10 mM laminaribiose (*induced*) or 10 mM cellobiose (*uninduced*) for 2 h. Expression levels were normalized to the housekeeping gene, *recA*. *Error bars* are indicative of one standard deviation, $n=3$

To circumvent this problem, we constructed a new vector, pEBM130, in which *spo0A* is replaced by a sporulation histidine kinase, *clo1313_1942* (Mearls and Lynd 2014). This plasmid was introduced into the wild-type strain and cultured in either laminaribiose or cellobiose, then plated for spore counts. Spore formation increased in cultures grown on laminaribiose by approximately four orders of magnitude compared to that in cultures grown on cellobiose only (Fig. 5a). To see if spore formation could be modulated by varying laminaribiose concentrations, DSM1313 expressing pEBM130 (LL1050) was grown in the presence of 0, 2, 5, 8, and 10 mM laminaribiose for 24 h and assayed for viable spores. Our results indicate there is a correlation between spore count (CFU/ml) and the level of laminaribiose used for induction (Fig. 5b). Sporulation frequencies were comparable when an empty vector control DSM1313/pEBM109 (LL1088) was grown on either cellobiose or laminaribiose (Fig. 5c).

Discussion

Metabolic engineering of *C. thermocellum* has progressed markedly over the last few years, but still is limited by

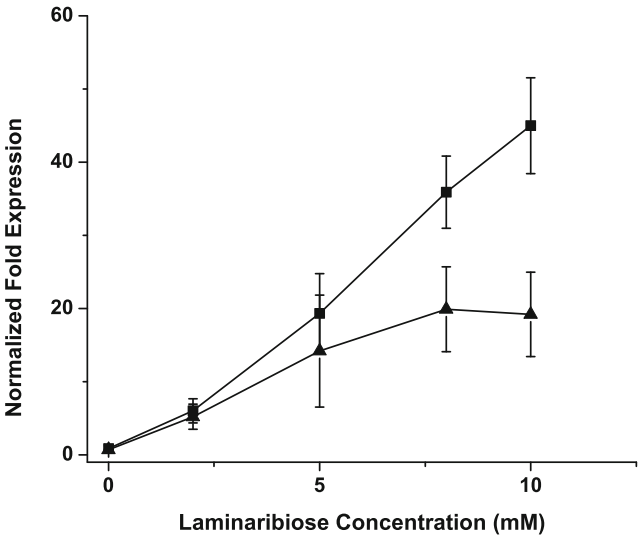


Fig. 4 Modulation of *spo0A* gene expression. M1726/pEBM113 (LL1051) was induced with varying concentrations of laminaribiose, and the expression levels of *spo0A* (*squares*) and *celC* (*triangles*) were measured via qRT-PCR. The *recA* gene was used as a reference. *Error bars* are indicative of one standard deviation, $n=3$

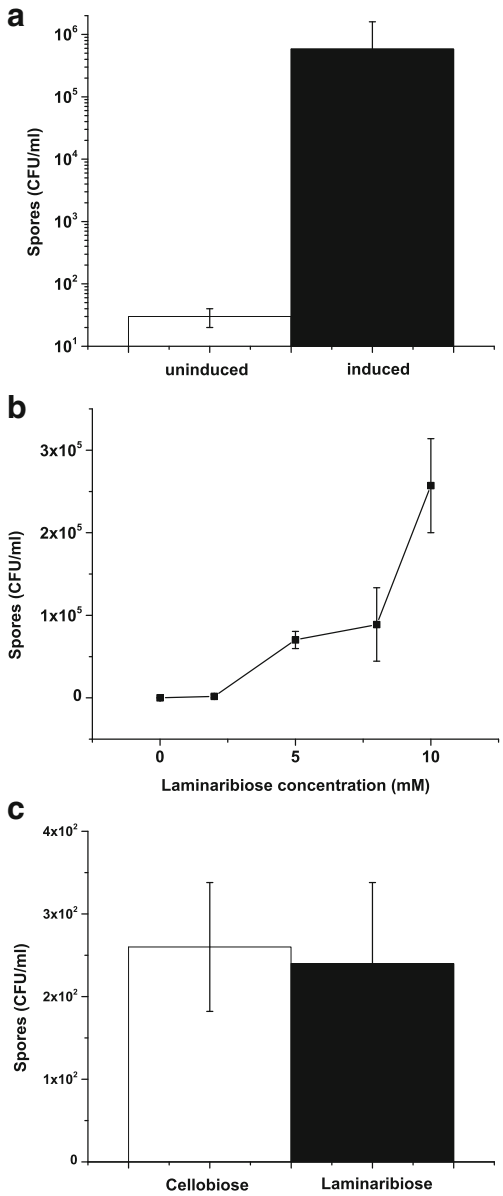


Fig. 5 Spore formation increases with laminaribiose concentration. **a** Strain DSM1313/pEBM130 (LL1050) was cultivated in 10 mM cellobiose (*uninduced*, *white bars*) or in 10 mM laminaribiose (induced, black bars) for 24 h, then heat treated and plated to quantify viable spores. **b** Spores were quantified as above following growth in the presence of varying levels of laminaribiose. Cellobiose was added in all cases to bring the total initial sugar concentration to 10 mM. **c** An empty vector control strain, DSM1313/pEBM109 (LL1088) was grown in the presence of either 10 mM cellobiose (*white bars*) or 10 mM laminaribiose (*black bars*). Spores were quantified after heat treatment. *Error bars* are indicative of one standard deviation, $n=3$ in all experiments

available genetic tools. Inducible gene expression systems, such as the *lac* operon-based system of *E. coli*, have been adapted for selective gene expression in a variety of microorganisms. However, as *C. thermocellum* is not known to import lactose (or its derivatives such as IPTG) into the cell, construction of a *lac*-based inducible system was not considered tractable. In this study, we describe the construction and function of an inducible expression system for *C. thermocellum* based on the native *celC* operon. This system allows for the expression of a target gene in the presence of laminaribiose, a β1-3-linked carbohydrate.

Initial determination of *celC* expression after the addition of laminaribiose to a late exponential phase culture indicated that the maximum expression of *celC* occurs between 1 and 2 h after induction, and declines significantly after 3 h. This decline in gene expression was expected, as the two cellulases found in the *celC* operon, *CelC* and LicA, potentially act to degrade laminaribiose. While CelC enzymatic functions have not been characterized with respect to laminaribiose, *celC* contains active sites similar to those found in fungal cellulases and to *celA*, an endoglucanase from *C. thermocellum* with demonstrated activity on β-1-4 and β-1-3 sugars (Petre et al. 1986; Schwarz et al. 1988). LicA exhibits specific activity towards β-1-3 sugars including barley β-glucan and laminarin (Fuchs et al. 2003). Frequently, sugar inducible expression systems exhibit autocatalysis, in which the genes encoding the transport of the sugar into the cell are under the transcriptional control of the sugar itself. Unlike the autocatalytic *lac* operon of *E. coli* (Novick and Weiner 1957; Siegele and Hu 1997), the genes encoded within the *celC* operon do not appear to play a direct role in laminaribiose transport into the cell (Nataf et al. 2009). Therefore, deletion of the native operon or its individual components has the potential to extend the duration of induction from the plasmid construct by slowing the degradation of laminaribiose in the cell without affecting laminaribiose transport. Such a modification is unlikely to prevent laminaribiose catabolism entirely due to the presence of other β-1-3 endoglucanases (Schwarz et al. 1986), but such work has the potential to improve the system and/or provide insight into laminaribiose utilization in *C. thermocellum*. For example, the mechanism of lactose utilization in *E. coli* was elucidated by studies in which the enzyme responsible for lactose catabolism, *lacZ* (β-galactosidase), was deleted. In the absence of LacZ, lactose fails to induce the *lac* operon because LacZ is also responsible for converting lactose to the inducer molecule, allolactose (Mulller-Hill 1996). The exact mechanism of the laminaribiose/GlyR3 interaction is currently unknown, and while we did not find it necessary to delete or modify the native *celC* operon of *C. thermocellum* for the purposes of this work, deletion of CelC and LicA could provide insight into this mechanism or help to understand their respective roles in laminaribiose catabolism.

Several plasmid-based inducible expression systems described for *E. coli* including those based on arabinose (Newman and Fuqua 1999), tetracycline (Skerra 1994), and lactose (Yansura and Henner 1984) function through the incorporation of the operator/promoter region and the repressor gene of the respective system. Our initial construct, pEBM107, was based on the native chromosomal arrangement of the *celC* operon and incorporated the *celC* promoter

region (as described by Newcomb et al. 2007), as well as the repressor gene *glyR3*. This plasmid failed to produce significant levels of expression after induction, and initial data suggested poor repression (data not shown). There are many approaches for improving repression in an inducible system, one of which is to enhance or stabilize expression of the repressor (Gruber et al. 2008; Lutz and Bujard 1997; Mulller-Hill et al. 1968). We took this approach by cloning the repressor, *glyR3*, before the target gene. This modification was made to ensure that the repressor would be expressed sufficiently for the repression of the operon in the absence of laminaribiose. The new construct, pEBM113, exhibited an approximately 40-fold increase in target gene expression upon induction and displayed low levels of uninduced expression.

We also found that gene expression levels increased in a linear fashion with the concentration of laminaribiose added. The Lbp ABC transporter system (*clo1313_2342-clo1313_2347*) is predicted to be responsible for laminaribiose import (Nataf et al. 2009). The *lbp* operon is not located near the *celC* operon in the chromosome, and is not predicted to be under the regulation of GlyR3 (Nataf et al. 2009). However, the genetic regulation of the *lbp* operon has not been explored in detail. Available transcriptomic and proteomic data show that the *lbp* operon is moderately expressed during continuous growth on cellulose and cellobiose (Riederer et al. 2011) and throughout batch growth on cellobiose (Raman et al. 2009), which suggests it is constitutively expressed. However, to date, the expression and regulation of this operon have not been evaluated in the presence of laminaribiose or other β1-3-linked sugars.

We used the newly developed inducible promoter system to explore elements of the sporulation mechanism in *C. thermocellum*. Previously, we had shown that expression of *spo0A* from a constitutive promoter complemented the Δ*spo0A* strain by restoring sporulation and producing viable spores in *C. thermocellum* (Mearls and Lynd 2014) However, expressing *spo0A* constitutively from the inducible promoter failed to complement the Δ*spo0A* strain by restoring viable spores. Interestingly, while viable spores were not detected, immature spore-like structures were observed by microscopy. We can only speculate on the cause for the failed complementation, but it would appear that *spo0A* expression levels are critical not only for the initial onset of spore formation but also for the completion of the process and the formation of a fully mature spore. While this was an unexpected result, it may illustrate an additional use for the inducible promoter system in uncovering critical timing events or evaluating the necessity of sustained gene expression in certain regulatory mechanisms.

We have shown previously that the histidine kinase Clo1313_1492 is required for sporulation in *C. thermocellum* and increases sporulation frequencies by roughly 100-fold when constitutively overexpressed in the wild-type background. Based on work done in *B. subtilis* (wherein overexpression of sporulation kinase KinA increased sporulation frequency (Eswaramoorthy et al. 2010)), we predicted that the sporulation response of *C. thermocellum* could be titrated by the expression of *clo1313_1492*. As predicted, overexpression of the sporulation histidine kinase *clo1313_1942* increased spore formation, and this response was modulated by laminaribiose concentrations. Interestingly, maximum expression of *clo1313_1492* from the inducible system achieved tenfold higher sporulation frequencies than when it was cloned under our *gapDH*-driven constitutive system (Mearls and Lynd 2014), despite the constitutive system providing a roughly twofold higher level of gene expression (unpublished data). We determined that laminaribiose had no impact on sporulation frequencies compared to cellobiose. Therefore, these differences must be due to the combination of induced expression and some other factor or set of factors occurring within the cells at the time of induction, which act together to increase the sporulation response. Here, we have demonstrated that our inducible gene expression system can produce functional output (spores) and can also serve as a platform for studying the sporulation pathway in *C. thermocellum*.

The ability to modulate spore formation by growth in the presence of laminaribiose suggests that this system can be used to induce and modulate expression of different target genes. This new tool may be useful in understanding the roles of many *C. thermocellum* genes, or in introducing and expressing genes from other organisms. The ability to selectively induce and regulate gene expression is desirable when expressing genes that are potentially toxic or when dynamic rather than steady-state expression is most useful.

Acknowledgments We would like to thank Dr. Adam Guss for his suggestions and input on the manuscript.

This research was supported by a grant from the BioEnergy Science Center (BESC), Oak Ridge National Laboratory, a US Department of Energy (DOE) BioEnergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Portions of this research were performed during an internship at the Mascoma Corporation. We would like to thank the Mascoma Corporation for their generous gift of strain M1726 and for their support during the duration of this work.

References

Argyros D, Tripathi S, Barrett T, Rogers S, Feinberg L, Olson D, Foden J, Miller B, Lynd L, Hogsett D, Caiazza N (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

Banerjee A, Leang C, Ueki T, Nevin KP, Lovley DR (2014) Lactose-inducible system for metabolic engineering of *Clostridium ljungdahlii*. Appl Environ Microbiol 80:2410–2416. doi:10.1128/AEM.03666-13

Bayer EA, Belaich JP, Shoham Y, Lamed R (2004) The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. Annu Rev Microbiol 58:521–554. doi:10.1146/annurev.micro.57.030502.091022

Burkholder W, Grossman, AD (2000) Prokaryotic development. American Society of Microbiology, Washington, D.C

Chung JD, Stephanopoulos G, Ireton K, Grossman AD (1994) Gene expression in single cells of *Bacillus subtilis*: evidence that a threshold mechanism controls the initiation of sporulation. J Bacteriol 176: 1977–1984

Dong H, Tao W, Zhang Y, Li Y (2012) Development of an anhydrotetracycline-inducible gene expression system for solvent-producing *Clostridium acetobutylicum*: a useful tool for strain engineering. Metab Eng 14:59–67. doi:10.1016/j.ymben.2011.10.004

Dror TW, Rolider A, Bayer EA, Lamed R, Shoham Y (2003) Regulation of expression of scaffoldin-related genes in *Clostridium thermocellum*. J Bacteriol 185:5109–5116

Elble R (1992) A simple and efficient procedure for transformation of yeasts. BioTechniques 13:18–20

Ellis LD, Holwerda EK, Hogsett D, Rogers S, Shao X, Tschaplinski T, Thorne P, Lynd LR (2012) Closing the carbon balance for fermentation by *Clostridium thermocellum* (ATCC 27405). Bioresour Technol 103:293–299. doi:10.1016/j.biortech.2011.09.128

Eswaramoorthy P, Duan D, Dinh J, Dravis A, Devi SN, Fujita M (2010) The threshold level of the sensor histidine kinase KinA governs entry into sporulation in *Bacillus subtilis*. J Bacteriol 192:3870–3882. doi:10.1128/JB.00466-10

Fagan RP, Fairweather NF (2011) *Clostridium difficile* has two parallel and essential Sec secretion systems. J Biol Chem 286:27483–27493. doi:10.1074/jbc.M111.263889

Feinberg L, Foden J, Barrett T, Davenport KW, Bruce D, Detter C, Tapia R, Han C, Lapidus A, Lucas S, Cheng JF, Pitluck S, Woyke T, Ivanova N, Mikhailova N, Land M, Hauser L, Argyros DA, Goodwin L, Hogsett D, Caiazza N (2011) Complete genome sequence of the cellulolytic thermophile *Clostridium thermocellum* DSM1313. J Bacteriol 193:2906–2907. doi:10.1128/JB.00322-11

Fuchs KP, Zverlov VV, Velikodvorskaya GA, Lottspeich F, Schwarz WH (2003) Lic16A of *Clostridium thermocellum*, a non-cellulosomal, highly complex endo-beta-1,3-glucanase bound to the outer cell surface. Microbiology 149:1021–1031

Gruber DF, Pieribone VA, Porton B, Kao HT (2008) Strict regulation of gene expression from a high-copy plasmid utilizing a dual vector system. Protein Expr Purif 60:53–57. doi:10.1016/j.pep.2008.03.014

Guss AM, Olson DG, Caiazza NC, Lynd LR (2012) Dcm methylation is detrimental to plasmid transformation in *Clostridium thermocellum*. Biotechnol Biofuels 5:30. doi:10.1186/1754-6834-5-30

Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP (2007) The ClosTron: a universal gene knock-out system for the genus *Clostridium*. J Microbiol Methods 70:452–464. doi:10.1016/j.mimet.2007.05.021

Johnson EA, Sakajoh M, Halliwell G, Madia A, Demain AL (1982) Saccharification of complex cellulosic substrates by the cellulase system from *Clostridium thermocellum*. Appl Environ Microbiol 43(5):1125–1132

Li Y, Xu T, Tschaplinski TJ, Engle NL, Yang Y, Graham DE, He Z, Zhou J (2014) Improvement of cellulose catabolism in *Clostridium cellulolyticum* by sporulation abolishment and carbon alleviation. Biotechnol Biofuels 7:25. doi:10.1186/1754-6834-7-25

Linville JL, Rodriguez L, Land M, Syed MH, Engle NL, Tschaplinski TJ, Mielenz JR, Cox CD (2013) Industrial robustness: understanding the mechanism of tolerance for the Populus hydrolysate-tolerant mutant strain of *Clostridium thermocellum*. PLoS One 8, e78829. doi:10.1371/journal.pone.0078829

Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. Nucleic Acids Res 25:1203–1210

Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev 66:506–577, **table of contents**

McBee R (1950) The anaerobic thermophilic cellulolytic bacteria. Bacteriol Rev 14:51–63

Mearls EB, Lynd LR (2014) The identification of four histidine kinases that influence sporulation in *Clostridium thermocellum*. Anaerobe 28:109–119. doi:10.1016/j.anaerobe.2014.06.004

Mearls EB, Izquierdo JA, Lynd LR (2012) Formation and characterization of non-growth states in *Clostridium thermocellum*: spores and L-forms. BMC Microbiol 12:180. doi:10.1186/1471-2180-12-180

Mulller-Hill B (1996) The *lac* operon: a short history of a genetic paradigm. Walter de Gruyter, Berlin

Mulller-Hill B, Crapo L, Gilbert W (1968) Mutants that make more *lac* repressor. Proc Natl Acad Sci U S A 59:1259–1264

Nariya H, Miyata S, Kuwahara T, Okabe A (2011) Development and characterization of a xylose-inducible gene expression system for *Clostridium perfringens*. Appl Environ Microbiol 77:8439–8441. doi:10.1128/AEM.05668-11

Nataf Y, Yaron S, Stahl F, Lamed R, Bayer EA, Scheper TH, Sonenshein AL, Shoham Y (2009) Cellodextrin and laminaribiose ABC transporters in *Clostridium thermocellum*. J Bacteriol 191:203–209. doi:10.1128/JB.01190-08

Newcomb M, Chen CY, Wu JH (2007) Induction of the *celC* operon of *Clostridium thermocellum* by laminaribiose. Proc Natl Acad Sci U S A 104:3747–3752. doi:10.1073/pnas.0700087104

Newcomb M, Millen J, Chen CY, Wu JH (2011) Co-transcription of the *celC* gene cluster in *Clostridium thermocellum*. Appl Microbiol Biotechnol 90:625–634. doi:10.1007/s00253-011-3121-x

Newman JR, Fuqua C (1999) Broad-host-range expression vectors that carry the L-arabinose-inducible *Escherichia coli araBAD* promoter and the araC regulator. Gene 227:197–203

Novick A, Weiner M (1957) Enzyme induction as an all-or-none phenomenon. Proc Natl Acad Sci U S A 43:553–556

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

Petre D, Millet J, Longin R, Beguin P, Girard H, Aubert JP (1986) Purification and properties of the endoglucanase C of *Clostridium thermocellum* produced in *Escherichia coli*. Biochimie 68:687–695

Qureshi N, Paterson AHJ, Maddox IS (1988) Model for continuous production of solvents from whey permeate in a packed bed reactor using cells of *Clostridium acetobutylicum* immobilized by adsorption onto bonechar. Appl Microbiol Biotechnol 29:323–328

Qureshi N, Lai LL, Blaschek HP (2004) Scale-up of a high productivity continuous biofilm reactor to produce butanol by adsorbed cells of *Clostridium beijerinckii*. Food Bioprod Process 82:164–173

Raman B, Pan C, Hurst GB, Rodriguez M Jr, McKeown CK, Lankford PK, Samatova NF, Mielenz JR (2009) Impact of pretreated Switchgrass and biomass carbohydrates on *Clostridium thermocellum* ATCC 27405 cellulosome composition: a quantitative proteomic analysis. PLoS One 4, e5271. doi:10.1371/journal.pone.0005271

Riederer A, Takasuka TE, Makino S, Stevenson DM, Bukhman YV, Elsen NL, Fox BG (2011) Global gene expression patterns in *Clostridium thermocellum* as determined by microarray analysis of chemostat cultures on cellulose or cellobiose. Appl Environ Microbiol 77:1243–1253. doi:10.1128/AEM.02008-10

Schwarz WH, Grabnitz F, Staudenbauer WL (1986) Properties of a *Clostridium thermocellum* endoglucanase produced in *Escherichia coli*. Appl Environ Microbiol 51:1293–1299

Schwarz WH, Schimming S, Rucknagel KP, Burgschwaiger S, Kreil G, Staudenbauer WL (1988) Nucleotide sequence of the celC gene

encoding endoglucanase C of *Clostridium thermocellum*. Gene 63: 23–30

Shanks RM, Caiazza NC, Hinsa SM, Toutain CM, O'Toole GA (2006) *Saccharomyces cerevisiae*-based molecular tool kit for manipulation of genes from gram-negative bacteria. Appl Environ Microbiol 72: 5027–5036. doi:10.1128/AEM.00682-06

Shanks RM, Kadouri DE, MacEachran DP, O'Toole GA (2009) New yeast recombineering tools for bacteria. Plasmid 62:88–97. doi:10.1016/j.plasmid.2009.05.002

Siegele DA, Hu JC (1997) Gene expression from plasmids containing the *araBAD* promoter at subsaturating inducer concentrations represents mixed populations. Proc Natl Acad Sci U S A 94:8168–8172

Skerra A (1994) Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in *Escherichia coli*. Gene 151:131–135

Steiner E, Dago AE, Young DI, Heap JT, Minton NP, Hoch JA, Young M (2011) Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in *Clostridium acetobutylicum*. Mol Microbiol 80:641–654. doi:10.1111/j.1365-2958.2011.07608.x

Tripathi SA, Olson DG, Argyros DA, Miller BB, Barrett TF, Murphy DM, McCool JD, Warner AK, Rajgarhia VB, Lynd LR, Hogsett DA, Caiazza NC (2010) Development of *pyrF*-based genetic system for targeted gene deletion in *Clostridium thermocellum* and creation of a *pta* mutant. Appl Environ Microbiol 76:6591–6599. doi:10.1128/AEM.01484-10

Underwood S, Guan S, Vijayasubhash V, Baines SD, Graham L, Lewis RJ, Wilcox MH, Stephenson K (2009) Characterization of the sporulation initiation pathway of *Clostridium difficile* and its role in toxin production. J Bacteriol 191:7296–7305. doi:10.1128/JB.00882-09

Yansura D, Henner D (1984) Use of the *Escherichia coli* lac repressor and operator to control gene expression in *Bacillus subtilis*. Proc Natl Acad Sci U S A 81:439–443

Zhang YH, Lynd LR (2005) Regulation of cellulase synthesis in batch and continuous cultures of *Clostridium thermocellum*. J Bacteriol 187:99–106. doi:10.1128/JB.187.1.99-106.2005

Zhang J, Liu Y, Cui GX, Cui Q (2015) A novel arabinose-inducible genetic operation system developed for *Clostridium cellulolyticum*. Biotechnol Biofuels 8:36