AEM Accepts, published online ahead of print on 6 August 2010
Appl. Environ. Microbiol. doi:10.1128/AEM.01484-10

Development of *pyrF*-based genetic system for targeted gene deletion in *Clostridium thermocellum* and creation of a *pta* mutant

Running Title: *C. thermocellum pyrF pta* mutant

Shital A. Tripathi[a,c], Daniel G. Olson[a,b,c], D. Aaron Argyros[a,c], Bethany B. Miller[a,c], Trisha F. Barrett[a,c], Daniel M. Murphy[a,c], Jesse D. McCool[a,c,d], Anne K. Warner[a,c], Vineet B. Rajgarhia[a,c,e], Lee R. Lynd[a,b,c], David A. Hogsett[a,c], and Nicky C. Caiazza[a,c,*]

Author affiliations:

[a]Mascoma Corporation, 67 Etna Rd, Suite 300, Lebanon, NH 03755

[b]Thayer School of Engineering, Dartmouth College, 8000 Cummings Hall, Hanover, NH 03755

[c]BioEnergy Science Center (BESC), Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, TN 37831

Current address:

[d]Lonza Biologics, Inc., 97 South Street, Hopkinton, MA 01748

[e]Bio Architecture Lab, Inc., 604 Bancroft Way, Suite C, Berkeley, CA 94710

[*]To whom correspondence should be addressed. Nicky C. Caiazza. Mascoma Corporation, 67 Etna Rd, Suite 300, Lebanon, NH 03755. Phone: 603-676-3320. Fax: 603-676-3321. E-mail: ncaiazza@mascoma.com.

Abstract

We report development of a genetic system for making targeted gene knockouts in *Clostridium thermocellum,* a thermophilic anaerobic bacterium that rapidly solubilizes cellulose. A toxic uracil analog, 5-fluoroorotic acid (5-FOA), was used to select for deletion of the *pyrF* gene. The Δ*pyrF* strain is a uracil auxotroph that could be restored to prototroph via ectopic expression of *pyrF* from a plasmid, providing a positive genetic selection. Furthermore, 5-FOA was used to select against plasmid encoded *pyrF* creating a negative selection for plasmid loss. This technology was used to delete a gene involved in organic acid production, namely *pta* that encodes the enzyme phosphotransacetylase. The *C. thermocellum* Δ*pta* strain did not produce acetate. These results are the first examples of targeted homologous recombination and metabolic engineering in *C. thermocellum*, a microbe that holds an exciting and promising future in the biofuel industry and development of sustainable energy resources.

Introduction

Conversion of cellulosic biomass using saccharolytic fermentative microorganisms without the addition of purified cellulase and hemicellulase enzymes is a promising approach for low-cost production of renewable fuels and chemicals (22-23). Thermophilic, cellulolytic bacteria are one departure point for development of microorganisms with the requisite capabilities for such consolidated bioprocessing (cbp), with *Clostridium thermocellum* being exemplary in this regard. As reviewed elsewhere (6, 22), *C. thermocellum* is a Gram-positive organism able to ferment cellulose and products of cellulose solubilization to ethanol, acetic acid, lactic acid, formic acid, hydrogen, and CO_2. *C. thermocellum* appears to be a cellulose-utilizing specialist (6, 8), and produces a multi-enzyme cellulose-solubilizing complex termed a cellulosome (2-3, 9).

Metabolic engineering is required in order to increase the yield of ethanol or other desired products from mixed-product fermentation such as that carried out by *C. thermocellum*. Comprehensive work directed to this end has been carried out with genetically tractable organisms such as *Escherichia coli*, resulting in high or near theoretical yields achieved for ethanol (35-36), other native products (21, 25), and non-native products (7, 12). In these organisms, genetic systems involving both positive and negative selection markers have been employed in order to facilitate reuse of the same marker and to develop marker-free strains. One prominent system in the category involves use of the gene encoding the enzyme orotidine 5-phosphate decarboxylase (PyrF) (4, 11, 20, 27-29, 39). PyrF participates in *de novo* pyrimidine biosynthesis but is also a target for the antimetabolite 5-fluoroorotic acid (5-FOA) (4). Thus, cells lacking

pyrF are uracil auxotrophs and resistant to 5-FOA, creating an opportunity whereby ectopic expression of *pyrF* can be selected or counter selected (4).

Reliable genetic tractability has been elusive for *Clostridium* species. Prior to this report the only *Clostridia* in which gene deletion via homologous recombination has been demonstrated are *Clostridium acetobutylicum, Clostridium perfringens*, and *Clostridium septicum*. In the former organism, the use of a replicating plasmid for transformation followed by selection and screening for plasmid segregation resulted in a single clone, that when analyzed contained a disruption in the gene of interest but not by the expected recombination events (13). The later two organisms have either an unusually high transformation frequency or feasibility for acquiring DNA from *E. coli* via conjugation allowing the use of suicide plasmids (1, 16, 19). By comparison, the recently reported method of *C. thermocellum* transformation consists of a complex and cumbersome electroporation protocol using a custom pulse delivery system (37-38). In our hands, efficiency of the *C. thermocellum* electrotransformation system does not compare with that of typical model organisms and does not enable the use of non-replicating plasmids as a means of gene manipulation. Alternatively, group II intron technology has been used to inactivate gene targets in *Clostridia* that were previously characterized as genetically intractable, but systems described to date have a temperature restriction that make such approaches incompatible with thermophilic *Clostridia* (14-15, 34).

The only *C. thermocellum* mutant characterized genetically was isolated following a random mutagenesis and enrichment for cells that did not adhere to cellulose (43). The random mutagenesis approach is limited in the sense that it does not lend itself well to reverse genetics as many desired mutations lack selectable or screenable

phenotypes. For instance, attempts have been made, with little success, to isolate saccharolytic thermophiles containing lesions in the *pta-ack* operon responsible for acetate production by selective enrichment using anti-metabolites (26). In contrast, the creation of a *Thermoanaerobacterium saccharolyticum* Δ*pta-ack* strain has been achieved using selectable markers that serve as a proxy for the events leading to targeted gene deletion (32). Motivated by the potential of microbial cellulose processing and the attributes of *C. thermocellum,* we undertook to develop a gene deletion system based on the *pyrF* gene.

Materials and methods

Media, culture conditions and reagents. *C. thermocellum* DSM 1313 was cultured in DSM 122 broth (37) with addition of 50 mM MOPS and 10 mM sodium citrate, herein referred to as rich medium. Cellobiose or Avicel was added at 5 g/liter unless otherwise noted and 0.8% Difco agar was used for solid media preparation. A uracil-free version of the rich medium, herein referred to as defined medium, omitted yeast extract and contained 3 µg/ml pyridoxamine hydrochloride, 0.3 µg/ml biotin, 0.6 µg/ml p-aminobenzoic acid, and 0.3 µg/ml Vitamin B12. Alternatively, MJ medium (17) which contains urea instead of ammonium sulfate as a nitrogen source was used as a uracil-free chemically defined medium. No difference was seen between MJ and the uracil-free version of DSM 122 when performing genetic selections. *C. thermocellum* was grown anaerobically at 55ºC unless otherwise noted. *E. coli* was grown aerobically at 37 ºC in LB Broth (Fischer Scientific). *Saccharomyces cerevisiae* was grown in YPD medium or synthetic uracil dropout medium (MP Biomedicals). For *C. thermocellum* thiamphenicol

(Tm) was used in a range from 3-48 µg/ml as indicated at 55ºC (at 60ºC and above, non-specific growth was observed indicating a potential decrease in the stability of Tm at elevated temperatures). For *E. coli* 20 µg/ml chloroamphenicol (Cm) and 15 µg/ml ampicillin (Ap) were used. When appropriate, 40 µg/ml uracil was used to supplement *C. thermocellum* growth. Where indicated, 500 µg/ml 5-Fluoroorotic acid (5-FOA) was added (Zymo Research). All reagents, unless noted, were purchased from Sigma-Aldrich.

Molecular cloning and plasmid construction. Primer design for amplification of DNA from *C. thermocellum* 1313 was based on the available *C. thermocellum* ATCC 27405 genome (http://genome.jgi-psf.org/cloth/cloth.home.html). The oligonucleotides and the plasmids/strains used in this study are listed in Table 1 and Table 2, respectively. The 5'and 3' flanking regions (~1 kb) of *pyrF* and *pta* were amplified and assembled using yeast gap repair cloning to create gene deletion plasmids (30). The *pyrF* and *pta* deletion vectors (pMU769 and pMU1162, respectively) contained *cat* (chloramphenicol acetyl-transferase) expressed from the *C. thermocellum gapDH* promoter (*gapDH*p) positioned between the 5' and 3' flanking regions (depicted in Fig. 3A). The *pyrF* complementing construct (pMU612) contained *pyrF* expressed from the *C. thermocellum* cellobiose phosphorylase (*cbp*) promoter (*cbp*p). All DNA manipulations and cloning procedures were performed as per Maniatis *et al*. (24). For maps of vectors and details on vector constructions, see appendix Fig. A1.

Transformation of *C. thermocellum* and selection of mutants. The transformation protocol used was modified substantially from the previously reported protocol (37). The

pulse generator was custom built and utilized a solid-state insulated-gate bipolar transistor (IGBT) instead of a power tetrode, as the high voltage switch (Infineon, part no. FZ200R65KF2). The device was charged with a high-voltage power supply from Emco (part no. F101). The charge was stored in an 8 kV 32 µF capacitor made by General Atomics (part no. 39742). Pulse duration and interval was controlled by an arbitrary function generator (Tektronix, part no. AFG3101). All manipulations were done under anaerobic conditions. Cultures were grown to mid-log phase (OD_{600} = 0.4-0.8) in rich medium and harvested by centrifugation (2200 x g for 12-14 min). Cells were washed twice in autoclaved, deionized water and the final pellet was resuspended in 200 µl deionized water. For each transformation, 20 µl of cell suspension was added, along with 1-8 µg of plasmid DNA, to a 0.1 cm gap electorporation cuvette (Fisher Scientific). A series of 60 square pulses were applied to the sample (see appendix figure A2 depicting pulse diagram). The period of the pulses was 300 µs and the amplitude was 1.9 kV, resulting in an applied field strength of 19 kV/cm. After pulsing, cells were recovered overnight (15-18 h) at 51°C in 3-5 ml rich medium. For liquid selection, recovered cultures were inoculated (10% v/v) into either rich medium supplemented with 3-6 µg/ml thiamphenicol (Tm) or uracil-free MJ medium when selecting for uracil prototrophy. For selecting transformants on solid medium, the recovery cultures were plated by mixing them in rich medium containing agar and 3-6 µg/ml Tm or MJ medium containing agar followed by a short incubation at room temperature in the anaerobic chamber to allow medium solidification before incubation of the plates at 55 ºC. To select Δ*pyrF* mutants, transformants were grown in 3 µg/ml Tm. The cultures were then diluted to approximately 10^8 cells/ml and 100 µl of the diluted culture was plated as agar

suspensions in rich medium containing 5-FOA. 5-FOA resistant colonies were screened by PCR using primers X02004 and X02005 (Table 1), which anneal outside of the regions of homology used to delete *pyrF* (Fig. 1A). More than 99% of the colonies screened were *pyrF* mutants that resulted from homologous recombination events as opposed to spontaneous mutations (see results for more details). The *pyrF*::*gapDH*p-*cat* mutants were isolated as described in the result section and the same primer set was used to screen Δ*pyrF*::*gapDH*p-*cat* mutants.

To select *pta*::*gapDH*p-*cat* mutants, the Δ*pyrF* strain transformed with pMU1162 was grown in 5 ml of rich medium supplemented with 6-12 µg/ml Tm or in MJ medium for about 14-16 hours. Various volumes of the cultures (ranging from 10 µl to 1 ml) were plated as agar suspensions of rich medium containing 5-FOA and 48 µg/ml Tm. Resistant colonies were screened by PCR using primers X02051 and X05980 (Table 1), which anneal outside of the regions of homology used to delete *pta* (Fig. 3A).

Fermentation conditions and analytical Methods. Batch fermentations in anaerobic tubes with wild-type, Δ*pyrF*, and Δ*pta*::*gapDH*p-*cat* were performed at 55ºC in rich medium under a nitrogen atmosphere utilizing cellobiose or Avicel as the primary carbon source. The fermentation products were analyzed using high-performance liquid chromatography (HPLC) as previously described (32). Growth rate measurements were performed in a 200 µl volume in a 96-well plate at 55ºC. The optical density at 600 nm was read by a Powerwave XS platereader customized by the manufacturer to incubate up to 68°C (BioTek). The plates were shaken continuously and read at three minute intervals. Each sample was measured in quadruplicate. The specific growth rate (µ) was

determined by measuring the slope of the natural log-transformed OD readings. A two hour sliding window of OD readings between 0.08 and 1.00 were used for determination of maximum rate μ_{max}. In all cases, the R-squared value was greater than 0.99.

Results

Isolation of a *C. thermocellum* Δ*pyrF* strain

A schematic depicting events leading to deletion of *pyrF* is shown in Figure 1A. The current transformation efficiency of *C. thermocellum* does not allow use of non-replicating plasmids for genetic manipulation. Thus, to delete *pyrF*, wild type (wt) *C. thermocellum* was transformed with a replicating allelic exchange vector, pMU482, using Thiamphenicol (Tm) selection (step 1). Single colonies representing transformants were propagated in liquid medium with Tm selection prior to plating on 5-FOA (step 2). As a control, the wt strain harboring an empty vector encoding Tm resistance, pMU102, was subjected to the same selective pressures. After 2 days of incubation a substantial number of colony-forming units (cfu's) resulted with pMU482, whereas only occasional spontaneous 5-FOA resistant mutants arose with pMU102.

5-FOA resistant colonies were screened by PCR using primers that anneal outside of the regions of homology used to delete *pyrF* (Fig. 1A). As expected, in the presence of pMU482, a number of 5-FOA resistant colonies exhibited double recombination events leading to the deletion of *pyrF*. These events were easily detected using diagnostic PCR in which the expected amplicon for wt is 3.1 kb and that for the Δ*pyrF* strain is 2.3 kb (Fig 1A and B). The *pyrF* locus was sequenced to confirm gene deletion.

Characterization and complementation of the *C. thermocellum* Δ*pyrF* strain

A *pyrF* complementing plasmid, pMU612, was constructed that contained the native *C. thermocellum pyrF* gene under control of the native *C. thermocellum* cellobiose phosphorylase (*cbp*) promoter (*cbp*p). Either pMU612 or pMU102 was transformed into the Δ*pyrF* strain to evaluate known *pyrF* phenotypes, for example 5-FOA sensitivity/resistance and uracil auxotrophy/prototrophy. Plasmid free versions of the wt strain and the Δ*pyrF* strain were used as controls. All strains were tested on four variations of solid media: A rich medium, with and without supplementation of 5-FOA, and a uracil-free defined medium, with and without supplementation of uracil. The cfu's are listed in Table 3. As expected, the wild type strain was sensitive to 5-FOA whereas the Δ*pyrF* strain was resistant. The occurrence of spontaneous 5-FOA resistance in the wt background was 0.3%. Additionally, growth of the Δ*pyrF* strain was severely impaired in the defined medium lacking uracil and addition of uracil supported normal growth. These results indicated that *pyrF* can be used as both a positive and a negative selection marker in *C. thermocellum.*

Comparisons made with the Δ*pyrF* strain containing either pMU102 or pMU612 further validated the use of *pyrF* based selections. The Δ*pyrF* strain containing pMU102 was resistant to 5-FOA and a uracil auxotroph mimicking the parental Δ*pyrF* strain (Table 3). The Δ*pyrF* strain containing pMU612 was complemented with respect to uracil prototrophy, as indicated by its ability to grow on a defined medium lacking uracil (Table 3). This confirms that *pyrF* is functionally expressed from the *cbp* promoter on pMU612. Surprisingly, the complemented strain was as resistant to 5-FOA as the Δ*pyrF* strain. Since the occurrence of spontaneous 5-FOA resistant mutants was low, the likely

explanation for this finding is that selection of plasmid loss in the presence of 5-FOA is unusually high. However, an alternative possibility is that the absence of antibiotic selection might be responsible for plasmid loss. To investigate this, colonies representing the Δ*pyrF* strain harboring pMU612 that appeared in the presence of 5-FOA were screened for the presence of plasmid. Twenty colonies were tested and none contained pMU612. In contrast, nineteen out of twenty colonies representing the Δ*pyrF* strain harboring pMU102 plated on 5-FOA contained the control plasmid, indicating that lack of antibiotic selection alone is insufficient to achieve high rates of plasmid loss. Collectively, the results in Table 3 indicated that *pyrF* can be used as a negative selection to identify plasmid loss in *C. thermocellum*.

Optimization of the *cat* marker for use in single copy on the chromosome

In order to create a marked mutation, a positive selection was needed to select for a chromosomal integration event and a negative selection was needed to select for loss of the replicating knock out plasmid. The latter component can be achieved using the Δ*pyrF* strain and ectopic expression of *pyrF* from a plasmid. To achieve the former, the *cat* marker, which provides Tm resistance at thermophilic temperatures from a multi-copy plasmid (33), was investigated for its ability to provide Tm resistance when harbored in single copy on the chromosome at the *pyrF* locus. An allelic replacement vector was constructed, pMU769, to delete the *pyrF* gene and replace it with *cat* controlled by the native glyceraldehyde 3-phosphate dehydrogenase (*gapDH*) promoter of *C. thermocellum*. The vector contained *gapDH*p-*cat* elements positioned between 5' and 3' *pyrF* flanking DNA. To replace *pyrF* with *gapDH*p-*cat*, *C. thermocellum* transformants containing pMU769 were subjected to two simultaneous selections in liquid, rich

medium. Thiamphenicol was used to select for the plasmid encoded *gapDH*p-*cat*, while 5-FOA was used to select against chromosomal *pryF*. Recovered cultures were evaluated by PCR using primers that anneal upstream and downstream of *pyrF*. Using these conditions, replacing the *pyrF* gene with *gapDH*p-*cat* would increase the PCR amplicon size by ~300 bp as compared to the wt. The result of this is shown in figure 1B, comparing lanes 2 and 4. The strain represented in lane 4 is a plasmid free, Δ*pyrF*::*gapDH*p-*cat* strain that was resistant to Tm concentrations up to 48 µg/ml. This informed us that *cat* expressed from the *gapDH* promoter was functional in a single copy on the *C. thermocellum* chromosome and could be used as marker for allele replacement.

Deletion of the *C. thermocellum pta* gene using *pyrF* and *cat* selection

Mixed acid fermentation of *C. thermocellum* involves co-production of lactic acid, acetic acid, formic acid, and ethanol (Fig. 2). For *C. thermocellum* strain DSM 1313 acetic acid is the major co-product that needs to be eliminated to create a strain with increased ethanol yield. The production of acetic acid from acetyl-CoA involves two enzymatic activities that are catalyzed by Pta and Ack.

The scheme used to replace *pta* with *cat* expressed from the *gapDH* promoter in the *C. thermocellum* Δ*pyrF* background is shown in figure 3A. MJ medium lacking uracil was used to select Δ*pyrF* clones restored to uracil prototrophy as a result of being transformed with pMU1162 (Fig. 3A step 1). Single colonies representing transformants were propagated in liquid medium with Tm selection prior to plating on Tm plus 5-FOA (Fig. 3A Step 2). Colonies resistant to both Tm and 5-FOA were screened by PCR using primers that anneal outside of the regions of homology used to delete *pta* (Fig. 3A). More than 75% of the colonies screened indicated the occurrence of homologous

recombination events at the *pta* locus. Clones in which *pta* was replaced by the *gapDH*p-*cat* cassette were discernable by a 0.5 kb increase in the size of the amplicon. For simplicity the *pta* mutants generated in the Δ*pyrF* background strain are designated as Δ*pta*::*gapDH*p-cat strain from here after (excluding the background strain – Δ*pyrF* genotype). The expected amplicon was 3.3 kb for wt and 3.8 kb for the Δ*pta*::*gapDH*p-cat strain (Fig 3A and B). The *pta* locus was sequenced to confirm allele replacement.

Growth analysis and fermentation profile of the Δ*pta*::*gapDH*p-*cat* strain

The growth of the Δ*pta*::*gapDH*p-*cat* strain was compared to the wt and Δ*pyrF* strains in rich medium, with and without uracil supplementation. Although initial rates of growth of the Δ*pyrF* and wt strains were similar (Fig. 3C), the Δ*pyrF* strain slowed abruptly at an OD of ~0.7, while the wt continued to grow until it reached an OD of ~1.6 (Fig 3C), suggesting that the rich medium was uracil-limited. Supplementing the medium with an additional 40 μg/ml uracil eliminated the growth defect of the Δ*pyrF* strain and resulted in a growth curve that was indistinguishable from the wt strain (Fig 3C). Even with additional uracil supplementation to compensate for the Δ*pyrF* mutation, the maximum specific growth rate (μ_{max}) of the Δ*pta*::*gapDH*p-*cat* strain was about one-third lower than that of either the wt or the Δ*pyrF* strains and the final OD was also reduced (Fig. 3C). This indicates that the growth defect of the Δ*pta*::*gapDH*p-*cat* strain is distinct from the growth defect of the Δ*pyrF* strain and is a result of the *pta* mutation.

End product analysis was performed on batch fermentations started at pH 7.0 with 5 g/l cellobiose as the primary carbon source under anaerobic conditions with a nitrogen atmosphere and 80 ml working volume. As show in figure 4A, after 48 hours of fermentation the wt and Δ*pyrF* strain produced about 1 g/l of acetic acid whereas the

acetic acid production of the Δ*pta*::*gapDH*p-*cat* was indistinguishable from background levels (average 0.03 g/l). All three strains produced comparable amounts of ethanol and lactic acid. Due to the growth defect of the Δ*pta*::*gapDH*p-*cat* strain a 96 hour sample point was taken but acetate levels did not change, measuring 0.031 g/l. The average dry cell mass for wt, Δ*pyrF*, and Δ*pta*::*gapDH*p-*cat* strains were 0.54 g, 0. 54 g and 0.35 g, respectively indicating that the Δ*pta*::*gapDH*p-*cat* strain made about one-third less biomass compared to the wt and Δ*pyrF* strain. This was expected based on the growth defect of the Δ*pta*::*gapDH*p-*cat* strain observed previously (Fig. 3C).

The elimination of the acetic acid pathway is expected to cause a decrease in ATP gained per mole of glucose consumed (22, 32). Cellulosome synthesis is considered to be an ATP expensive process. Therefore, we examined the cellulolytic capacity of the Δ*pta*::*gapDH*p-*cat* strain and determined the fermentation product profile on Avicel, a refined, crystalline cellulose. Batch fermentations were performed starting at pH 7.0 using 5 g/l Avicel as the primary carbon source under anaerobic conditions with a nitrogen atmosphere and 80 ml working volume. As seen in figure 4B, after 48 hours of fermentation the wt and the parental Δ*pyrF* strain both produced about 1 g/l of acetic acid whereas the Δ*pta*::*gapDH*p-*cat* strain produced ~0.03 g/l acetic acid. All three strains produced comparable levels of ethanol and lactic acid. Even after 96 h of fermentation, the Δ*pta*::*gapDH*p-*cat* strain did not produce more than 0.03 g/l acetate. Avicel is a solid substrate and non-hydrolyzed material can be estimated by measuring the dry weight of the fermentation. Table 4 shows that the average dry weights for each strain at the end of 96 hours of fermentation are comparable and this translates into ~89% Avicel hydrolysis. However, the fermentation dry weight does not distinguish residual Avicel from dry cell

mass. For Avicel fermentations, the latter can be estimated by assuming that the dry cell mass produced from 5 g/l Avicel is similar to that produced from 5 g/l cellobiose (see above). Using this logic, all strains showed > 95% solubilization of Avicel. Therefore, despite a growth defect the cellulose hydrolysis capacity of the Δ*pta*::*gapDH*p-*cat* strain was comparable to the wt strain.

Discussion

To create a gene deletion system for *C. thermocellum*, which has low transformation efficiency and requires replicating plasmids, we needed to develop multiple positive and negative selectable markers to allow selection of both allele replacement and loss of the replicating plasmid used to deliver the allelic replacement cassette. Key factors responsible for achieving this result involved exploiting the native capacity of the host in question. To this end we utilized the native *gapDH* promoter to express *cat* and confer resistance to thiamphenicol. In addition, *C. thermocellum* was sensitive to the antimetabolite 5-FOA allowing us to create, as in other model systems, a native dual selection marker around *pyrF* (4). The use of native elements when developing a genetic system serves to constrain the experimental space by eliminating functional uncertainty, allowing focus on other areas such as transformation.

In this report we have simplified electroporation-based transformation of *C. thermocellum* and obtained reliable and reproducible results. In contrast to the original protocol (37-38) the generation of competent cells has been simplified such that the isoniacin incubation, cellobiose solution wash and custom built electroporation cuvettes are no longer required. Best results are obtained with a custom pulse generator but

success can be achieved with a commercial device (the latter being the scope of a separate study). Furthermore, the gene deletion system we developed utilizes replicating plasmids and does not rely on high efficiency transformation. Development of the *pyrF* based genetic selections, in addition to the reliable electrotransformation protocol, removes a significant barrier to routine transformation and genetic manipulation of *C. thermocellum*.

The genetic tools described in this report were applied in the context of metabolic engineering to delete *pta* and create a *C. thermocellum* strain that does not produce acetate. This is a major metabolic perturbation and like various bacteria in which the acetate production pathway has been eliminated, the *C. thermocellum pta* mutant displayed growth defects (10, 18, 32, 42). Adaptation and evolutionary approaches have been shown to ameliorate this effect but have not been explored in this study (10, 18, 32, 42). Thus, it is unclear if the growth defects associated with the *C. thermocellum* Δ*pta* strain are stable or part of a cellular adjustment process associated with re-directing carbon and energy flux. Possible explanations for growth limitation include redox imbalance and/or perturbation in acetyl-CoA flux (5, 40). The fate of acetyl-CoA is related to the pool of reduced ferredoxin, which if used to make hydrogen results in acetate production (31). Elimination of the ability to make acetate requires the cell to utilize alternate pathways to oxidize ferredoxin. If such pathways are present, strain adaptation and evolution should improve growth. However, if they are lacking from the system, further metabolic engineering would be required to provide such an activity. The other significant phenotype of the *C. thermocellum* Δ*pta* strain was, in spite of a growth defect, that the cellulolytic capability was not compromised. This implies that ATP yield

associated with acetate production is not essential for cellulosome production and is consistent with an ATP conserving model of oligosaccharide uptake (41). According to this model, upon cellulose hydrolysis, *C. themocellum* preferentially assimilates cellodextrins up to 4 subunits as opposed to cellobiose, which is the primary cellulose hydrolysis end-product of other cellulolytic organisms such as *Trichoderma reesei*. This allows more hexose units to be transported into the cell per unit ATP. Additionally, *C. thermocellum* expressed cellobiose and cellodextrin phosphorylases that can use inorganic phosphate to phosphorolytically cleave *β*-glucan bonds generating glucose-1-phosphate in the process. These aspects of cellulose utilization allows *C. thermocellum* to conserve ATP per hexose monomer consumed and more than compensate for the ATP expensive process of cellulase synthesis.

Unlike the *T. saccharolyticum pta-ack* mutant, the *C. thermocellum* Δ*pta* strain did not exhibit significantly increased ethanol production (32). Moreover, the *C. thermocellum* Δ*pta* strain showed notable difference in lactic acid production when fermenting Avicel vs. cellobiose as compared to the wild type strain (Fig. 4). These results might be an indication of fundamental differences between the carbon flux through pyruvate in these two organisms. Future effort in developing a *C. thermocellum* homoethanologenic strain would involve deletion of lactate dehydrogenase, carbon flux analysis, and adaptation/evolution efforts. Nevertheless, the genetic system developed here demonstrates a step towards engineering a *C. thermocellum* homoethanologenic strain.

The recalcitrance of cellulose is the major challenge to utilizing lignocellulosic biomass as a feedstock. In this regard, *C. thermocellum* is now a unique platform in that it

is highly cellulolytic and amenable to genetic manipulation. The work presented here is a foundation for future metabolic engineering that will enable *C. thermocellum* to serve as a biocatalyst for the production of cellulosic fuels and chemicals with high yield and titer. The same technology can be used to explore the regulation of cellulosome synthesis and the roles of various structural and functional components of this multi-protein complex, providing a better understanding of how this fascinating microbe hydrolyzes cellulose. Advancements in these areas, facilitated by the described genetic tools, opens the way for further development of *C. thermocellum* and related cbp organisms as biocatalysts in the conversion of lignocellulosic biomass to sustainable fuels and chemicals.

Acknowledgements

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

References

1. **Awad, M. M., A. E. Bryant, D. L. Stevens, and J. I. Rood.** 1995. Virulence studies on chromosomal alpha toxin and theeta toxin mutants constructed by allelic exchange provide genetic evidence for the essential role of alpha toxin in *Clostridium perfringens*-mediated gas gangrene. Mol Microbiol **15:**191-202.

2. **Bayer, E. A., J.-P. Belaich, Y. Shoham, and R. Lamed.** 2004. THE CELLULOSOMES: Multienzyme machines for degradation of plant cell wall polysaccharides. Annu Rev Microbiol **58:**521-554.

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

4. **Boeke, J. D., F. Croute, and G. R. Fink.** 1984. A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. Mol Gen Genet **197:**345-346.

5. **Chang, D.-E., S. Shin, J.-S. Rhee, and J.-G. Pan.** 1999. Acetate metabolism in a *pta* mutant of *Escherichia coli* W3110: importance of maintaining acetyl coenzyme A flux for growth and survival. J Bacteriol **181:**6656-6663.

6. **Demain, A. L., M. Newcomb, and J. H. D. Wu.** 2005. Cellulase, clostridia, and ethanol. Microbiol Mol Biol Rev **69:**124-154.

7. **Farmer, W. R., and J. C. Liao.** 2001. Precursor balancing for metabolic engineering of Lycopene production in *Escherichia coli*. Biotechnol Prog **17:**57-61.

8. **Freier, D., C. P. Mothershed, and J. Wiegel.** 1988. Characterization of *Clostridium thermocellum* JW20. Appl Environ Microbiol **54:**204-211.

9. **Gold, N. D., and V. J. J. Martin.** 2007. Global view of the *Clostridium thermocellum* cellulosome revealed by quantitative proteomic analysis. J Bacteriol **189:**6787-6795.

10. **Green, E. M., Z. L. Boynton, L. M. Harris, F. B. Rudolph, E. T. Papoutsakis, and G. N. Bennett.** 1996. Genetic manipulation of acid formation pathways by

gene inactivation in *Clostridium acetobutylicum* ATCC 824. Microbiology **142:**2079-2086.

11. **Haas, L. O., J. M. Cregg, and M. A. Gleeson.** 1990. Development of an integrative DNA transformation system for the yeast *Candida tropicalis*. J Bacteriol **172:**4571-4577.

12. **Hanai, T., S. Atsumi, and J. C. Liao.** 2007. Engineered synthetic pathway for isopropanol production in *Escherichia coli*. Appl Environ Microbiol **73:**7814-7818.

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

14. **Heap, J. T., S. A. Kuehnea, M. Ehsaana, S. T. Cartmana, C. M. Cooksleya, J. C. Scotta, and N. P. Minton.** 2009. The ClosTron: Mutagenesis in *Clostridium* refined and streamlined J Microbiol Methods **in press. Available online 3 November 2009**.

15. **Heap, J. T., O. J. Pennington, S. T. Cartman, G. P. Cartera, and N. P. Minton.** 2007. The ClosTron: A universal gene knock-out system for the genus *Clostridium*. J Microbiol Methods **70:**452-464.

16. **Heap, J. T., T. C. Stephen, J. P. Oliver, M. C. Clare, C. S. Jamie, B. Ben, A. B. David, and P. M. Nigel.** 2009. Development of genetic knock-out systems for clostridia, p. 179-198, Clostridia: molecular biology in post-genomic era. Caister Academic Press.

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

18. **Kakuda, H., K. Shiroishi, K. Hosono, and S. Ichihara.** 1994. Construction of Pta-Ack pathway deletion mutants of *Escherichia coli* and characteristic growth profiles of the mutants in a rich medium. Biosci Biotechnol Biochem **58:**2232-2235.

19. **Kennedy, C. L., E. O. Krejany, L. F. Young, J. R. O'Connor, M. M. Awad, R. L. Boyd, J. J. Emmins, D. Lyras, and J. I. Rood.** 2005. The alpha toxin of *Clostridium septicum* is essential for virulence. Molecular Microbiology **57:**1357-1366.

20. **Kondo, S., A. Yamagishi, and T. Oshima.** 1991. Positive selection for uracil auxotrophs of the sulfur-dependent thermophilic archaebacterium *Sulfolobus acidocaldarius* by use of 5-fluoroorotic acid. J Bacteriol **173:**7698-7700.

21. **Lin, H., G. N. Bennett, and K.-Y. San.** 2005. Metabolic engineering of aerobic succinate production systems in *Escherichia coli* to improve process productivity and achieve the maximum theoretical succinate yield. Metab Eng **7:**116-127.

22. **Lynd, L. R., P. J. Weimer, W. H. van Zyl, and I. S. Pretorius.** 2002. Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev **66:**506-577.

23. **Lynd, L. R., W. H. v. Zyl, J. E. McBride, and M. Laser.** 2005. Consolidated bioprocessing of cellulosic biomass: an update. Curr Opin Biotechnol **16:**577-583.

24. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press.

25. **Patnaik, R., and J. C. Liao.** 1994. Engineering of *Escherichia coli* central metabolism for aromatic metabolite production with near theoretical yield. Appl Environ Microbiol **60:**3903-3908.

26. **Rothstein, D. M.** 1986. *Clostridium thermosaccharolyticum* strain deficient in acetate production. J Bacteriol **165:**319-320.

27. **Sato, T., T. Fukui, H. Atomi, and T. Imanaka.** 2005. Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon *Thermococcus kodakaraensis*. Appl Environ Microbiol **71:**3889-3899.

28. **Sato, T., T. Fukui, H. Atomi, and T. Imanaka.** 2003. Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. J Bacteriol **185:**210-220.

29. **Schneider, J. C., A. F. Jenings, D. M. Mun, P. M. McGovern, and L. C. Chew.** 2005. Auxotrophic markers *pyrF* and *proC* can replace antibiotic markers on protein production plasmids in high-cell-density *Pseudomonas fluorescens* fermentation. Biotechnol Prog **21:**343-348.

30. **Shanks, R. M. Q., N. C. Caiazza, S. M. Hinsa, C. M. Toutain, and G. A. O'Toole.** 2006. *Saccharomyces cerevisiae*-based molecular tool kit for manipulation of genes from Gram-negative bacteria. Appl Environ Microbiol **72:**5027-5036.

31. **Shaw, A. J., D. A. Hogsett, and L. R. Lynd.** 2009. Identification of the [FeFe]-hydrogenase responsible for hydrogen generation in *Thermoanaerobacterium saccharolyticum* and demonstration of increased ethanol yield via hydrogenase knockout. J Bacteriol**:**JB.00497-09.

32. **Shaw, A. J., K. K. Podkaminer, S. G. Desai, J. S. Bardsley, S. R. Rogers, P. G. Thorne, D. A. Hogsett, and L. R. Lynd.** 2008. Metabolic engineering of a thermophilic bacterium to produce ethanol at high yield. Proc Natl Acad Sci **105:**13769-13774.

33. **Soutschek-Bauer, E., W. Scholz, E. Grill, and W. L. Staudenbauer.** 1987. Thermostability and superhelicity of plasmid DNA in *Bacillus stearothermophilus*. Mol Gen Genet **209:**575-579.

34. **Tolonen, A. C., A. C. Chilaka, and G. M. Church.** 2009. Targeted gene inactivation in *Clostridium phytofermentans* shows that cellulose degradation requires the family 9 hydrolase Cphy3367. Mol Microbiol **74:**1300-1313.

35. **Trinh, C. T., and F. Srienc.** 2009. Metabolic engineering of *Escherichia coli* for efficient conversion of glycerol to ethanol. Appl Environ Microbiol **75:**6696-6705.

36. **Trinh, C. T., P. Unrean, and F. Srienc.** 2008. Minimal *Escherichia coli* cell for the most efficient production of ethanol from hexoses and pentoses. Appl Environ Microbiol **74:**3634-3643.

37. **Tyurin, M. V., S. G. Desai, and L. R. Lynd.** 2004. Electrotransformation of *Clostridium thermocellum*. Appl Environ Microbiol **70:**883-890.

38. **Tyurin, M. V., C. R. Sullivan, and L. R. Lynd.** 2005. Role of spontaneous current oscillations during high-efficiency electrotransformation of thermophilic anaerobes. Appl Environ Microbiol **71:**8069-8076.

39. **Yamagishi, A., T. Tanimoto, T. Suzuki, and T. Oshima.** 1996. Pyrimidine biosynthesis genes (*pyrE* and *pyrF*) of an extreme thermophile, *Thermus thermophilus*. Appl Environ Microbiol **62:**2191-2194.

40. **Yang, Y.-T., G. N. Bennett, and K.-Y. San.** 1999. Effect of inactivation of *nuo* and *ackA-pta* on redistribution of metabolic fluxes in *Escherichia coli*. Biotechnol Bioeng **65:**291-297.

41. **Zhang, Y.-H. P., and L. R. Lynd.** 2005. Cellulose utilization by *Clostridium thermocellum*: bioenergetics and hydrolysis product assimilation. Proc Natl Acad Sci **102:**7321-7325.

42. **Zhu, Y., X. Chen, T. Chen, and X. Zhao.** 2007. Enhancement of riboflavin production by overexpression of acetolactate synthase in a *pta* mutant of *Bacillus subtilis*. FEMS Microbiol Lett **266:**224-230.

43. **Zverlov, V. V., M. Klupp, J. Krauss, and W. H. Schwarz.** 2008. Mutations in the scaffoldin gene, *cipA*, of *Clostridium thermocellum* with impaired cellulosome formation and cellulose hydrolysis: insertions of a new transposable element, IS1447, and Implications for cellulase synergism on crystalline cellulose. J Bacteriol **190:**4321-4327.

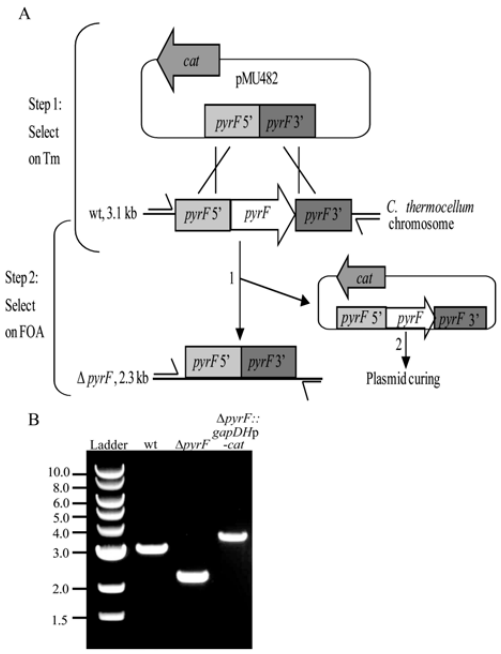

Fig. 1. Creation of a *C. thermocellum* Δ*pyrF* strain. (A) Relevant features of the *pyrF* deletion vector, pMU482, are the positive marker for transformation (*cat*), and the flanking regions of homology that target deletion of *pyrF* (*pyrF* 5' and *pyrF* 3'). The chromosomal *pyrF* locus contains the *pyrF* gene in the context of the flanking DNA. Step 1: Plating on Tm allows selection of transformants harboring pMU482. Step 2: plating on 5-FOA selects for: 1) deletion of *pyrF* form chromosome and insertion into pMU482. 2) Loss of pMU482 containing *pyrF*. Primers used for PCR are indicated by the one-sided arrows and the size of the expected amplicon are shown to the left. (B) DNA gel showing colony PCR results using primers X02204 and X02205 on wild type (lane 2), Δ*pyrF* (lane 3), and Δ*pyrF*::*gapDH*p-*cat* (lane 4). The numbers on the left indicate the band size in Kb for the NEB 1 KB ladder used as a marker (lane 1).

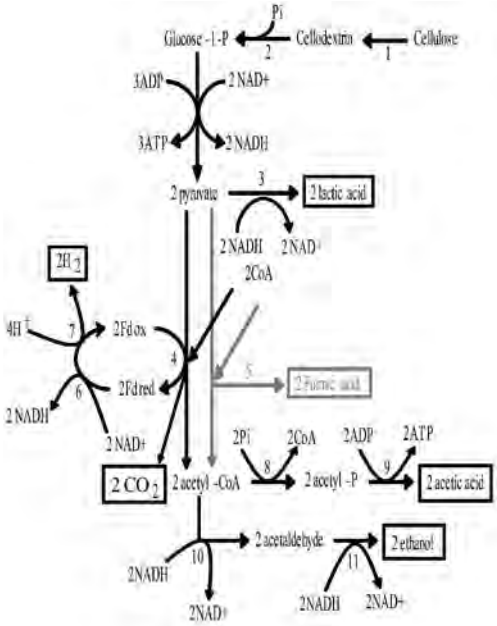

Fig. 2. Mixed acid fermentation of *C. thermocellum*. The primary fermentation products are highlighted in black boxes. Formic acid, highlighted in a grey box, is a fermentation end product that has been observed in C. thermocellum, but was not under the conditions used in this study. Key intermediates and enzymatic steps are noted: 1. extracellular cellulose hydrolysis. 2. intracellular phosphorolytic cleavage of cellodextrins by cellobiose phosphorylase. 3. lactate dehydrogenase, 4. pyruvate ferredoxin oxidoreductase, 5. pyruvate formate lyase, 6. ferredoxin NAD+(H) oxidoreductase, 7. ferredoxin dependent hydrogenase, 8. phosphotransacetylase, 9. acetate kinase, 10. acetaldehyde dehydrogenase, 11. alcohol dehydrogenase.

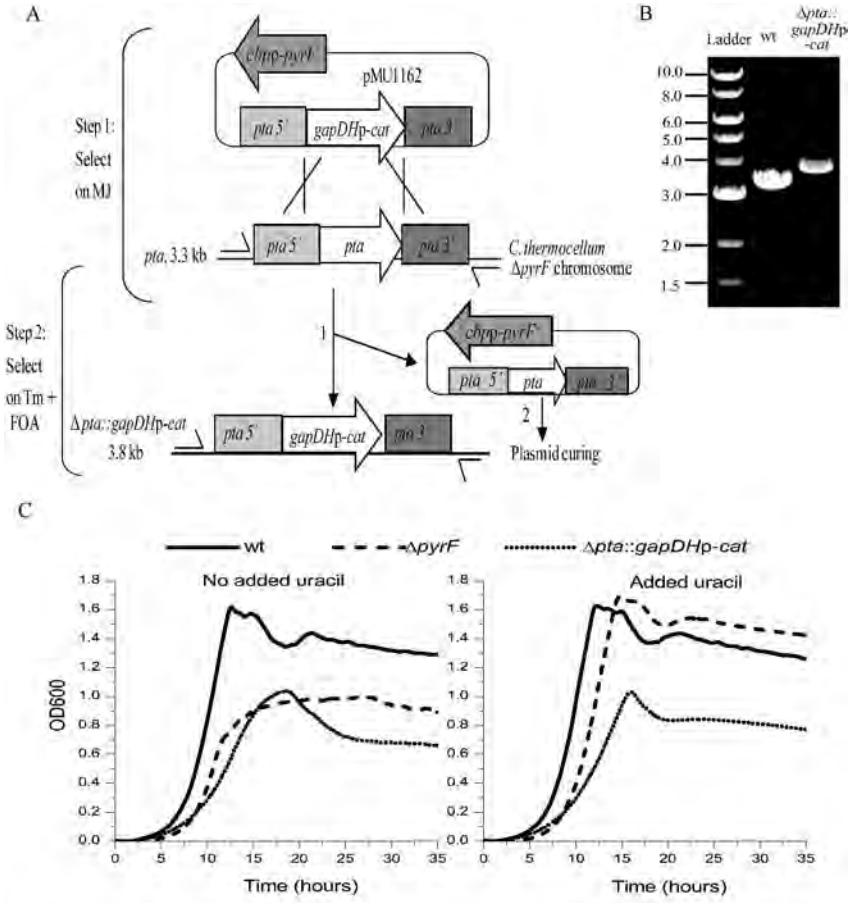

Fig. 3. Creation of an *C. thermocellum* Δ*pta*::*gapDH*p-*cat* strain. (A) Relevant features of the pta deletion vector, pMU1162, are the dual selection cassette for transformation and plasmid loss (*cbp*p-*pyrF*), and the allelic replacement cassette containing *gapDH*p-*cat* between the flanking regions (*pta* 5' and *pta* 3). The chromosomal pta locus contains the *pta* gene in the context of the flanking DNA. Step 1: Plating on MJ allows selection of transformants harboring pMU1162. Step 2: Plating on Tm plus 5-FOA selects 1. Replacement of pta with the *gapDH*p-*cat* cassette and insertion of *pta* into pMU1162. 2. Loss of *pyrF* containing plasmid. Primers used for PCR are indicated by the one-sided arrows and the size of the expected amplicon are shown to the left. (B) DNA image showing the results of colony PCR using primers X02051 and X02099 on wild type (lane 2) and a putative Δ *pta*::*gapDH*p-**cat** strain (lane 3). The numbers on the left indicate the band size in Kb for the NEB 1 KB ladder used as a marker (lane 1). (C) Growth analysis of the wt, Δ*pyrF* and Δ*pta*::*gapDH*p-*cat* strains in rich medium with or without supplementation of uracil. The graphs indicate the optical density (OD) at 600 nm of each culture over time. The μmax for wt, Δ*pyrF* and Δ*pta*::*gapDH*p-*cat* strains were 0.55 h-1, 0.58 h-1and 0.38 h-1, respectively and did not change significantly upon addition of uracil.

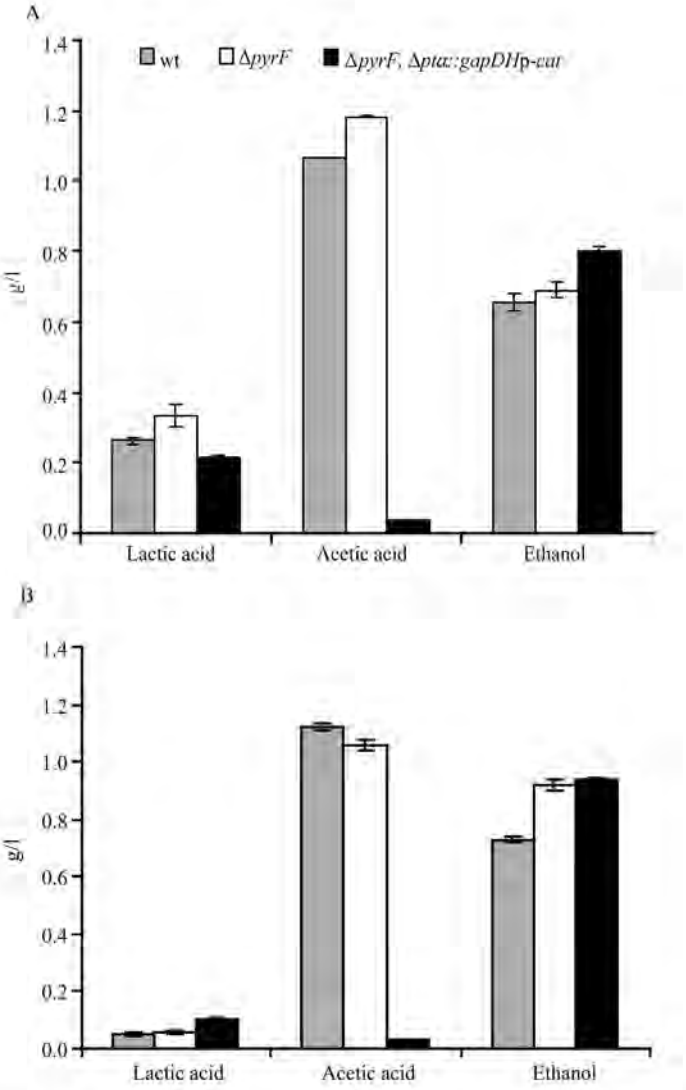

Fig. 4. Batch fermentation analysis of the Δ*pta*::*gapDH*p-*cat* strain at 55ºC in rich medium supplemented with uracil. (A) Measurement of lactic acid, acetic acid and ethanol resulting from the consumption of 5 g/l cellobiose in 48 hours by the wt, Δ*pyrF*, and Δ*pta*::*gapDH*p-*cat* strains. Error bars represent data from triplicate fermentations. (B) Measurement of lactic acid, acetic acid and ethanol resulting from the consumption of of 5 g/l Avicel in 48 hours by the wildtype, Δ*pyrF*, and Δ*pta*::*gapDH*p-*cat* strains. Error bars represent data from triplicate fermentations.

Table 1. Oligonucleotides used in this study

Primer	Sequence 5'-3'
X00105	actccgcggccgggaagccgatctcggcttgaacgaattgggatcctctagagtcgacct
X00106	ggtcgatgtttgatgttatggagcagcaacggaaagcagagcaatgtggaattgggaacg
X00109	cctgataaatgcttcaataatattgaaaaaggaagagtatccagtattctgacatgggtg
X00110	cttttctcttccaattgtctaaatcaattttattaaagttcattaatatcgcctcctattg
X00967	ggcggaattcggaggcttacttgtctgctt
X00968	ggcgagatctgttgtgtggaattgtgagcg
X00969	ggcggaattcggtttcttagacgtcaggtg
X00970	ggcgagatcttgacgctcagtggaacgaaa
X00973	aggtcgactctagaggatccccgggtaccgagctcgaatttaatatcgcctcctattg
X00974	aaaggatctcaagaagatcctttgatcttttctacggggatcgcttgcctgtaacttac
X01840	atcagttattacccactttcggggaaatgtgcgcggaacccattgcgaagatctggacc
X01841	gcactgttaatatcgtccctagggtctaatcctacaacgg
X01842	ccgttgtaggattagaccctagggacgatattaacagtgc
X01843	gcagtgagcgcaacgcaattaatgtgagttagctcaccgtcgttcatgtttctgtgg
X01893	ggcgggtacccgtcgttcatgtttctgtgg
X02004	catgagtgctgcaatgggag
X02005	tatccttgtaccttacgcc
X02051	ctaataaccctcatgtcaag
X01905	ggcgtctagaccattgcgaagatctggacc
X03061	aagaaaccattattatcatgacattaacctataaaaaaggtctagagagtcgtgactaagaacgtc
X03062	cgtttcacttctgagttcgcgaattcaagaacgcaggattgaatagag
X03063	ctctattcaatcctgcgttcttgaattcgcgaactcagaagtgaaacg
X03064	gaattgtgagcggataacaatttcacacaggaaacagcttctagaacaagggttaccggaatttac
X03079	caaataaacgaaaatttttaaggaggacgaaagatgtttattgatacattaattgaaaagatta
X03080	acccggataacttcttttttgtcataccttgctttattacttcctgtctcgcaacgcactg
X03852	cagtgaggcacctatctcagcgatctgtcaattcgcggccgcgaaaagtgggtaataactg
X03853	ctgcaggtcgactctagaggatccccgggtaccgcggccgcgatcgcttgcctgtaacttac
X03886	attacgccaagcttgcatgccctgcaggtcgactctagagcgtcgttcatgtttctgtgg
X03889	tgtaagttacaggcaagcgatcgcggccgcggtacccgggccattgcgaagatctggacc
X03896	aaggataatccttccgttgtaggattagaccctactagtagcccttcaaacttcccaaag
X03897	cactagggctcgcctttgggaagtttgaagggctactagtagggtctaatcctacaacgg
X03898	ctttttttaaaagtcaatcccgtttgttgaactactgtacaagggacgatattaacagtgc
X03899	tgtctcgcaacgcactgttaatatcgtcccttgtacagtagttcaacaaacgggattgac
X05109	tgtaagttacaggcaagcgatcgcggccgcggtacccgggcaggcaaagtccaactatgc
X05110	ctttgggaagtttgaagggctgcggctttaataaccctcag
X05111	ctgagggttattaaagccgcagcccttcaaacttcccaaag
X05112	gacatattgagcctgaaccggtagttcaacaaacgggattgac
X05113	gtcaatcccgtttgttgaactaccggttcaggctcaatatgtc
X05114	attacgccaagcttgcatgccctgcaggtcgactctagagtgtcatctccctttctgcg
X05850	gcttcttaagacaggaaagc

Table 2. Plasmids and strains used in this study

Plasmids/ strains	Description and relevant characteristics	Source or references
Plasmids		
pUC19	General purpose cloning vector, Ap	NEB[1]
pMU102	pMU104, region between the FokI and EcoRI sites has been deleted, Cm	This study
pMU104	pNW33N, *E.coli-C.thermocellum* shuttle vector, Cm	BGSC[2]
pMU110	pMQ87, *S.cerevisiae.-E.coli* shuttle vector, Ura+, Gm	This study
pMU111	pMU110 with *aac1* (Gm) replaced by *cat* from pMU104, Ura+, Cm	This study
pMU113	pMU111 with *C. the. gapDH*p driving *cat*, Ura+, Cm	This study
pMU245	*E.coli-C. thermoceelum* cloning vector, Ap	This study
pMU357	*S.cerevisiae-E.coli-C. thermocellum* shuttle vector for expressing genes in *C. the., C. the gapDHp, Ura+, Cm*	This study
pMU440	*S.cerevisiae-E.coli-C. thermocellum* shuttle vector, *C. therm* Δ*pyrF* cassette, Ura+, Cm	This study
pMU482	*E.coli-C. thermocellum* shuttle vector, *C. the* Δ*pyrF* cassette, Cm	This study
pMU582	pMU110, *C. the. cbp* promoter, *cbp* gene, T1T2 terminator , Ura+, Gm	This study
pMU597	pMU582, *C.the cbp* gene replaced by *C. the. pyrF* gene - creating cbpp-pyrF cassette, Ura+, Gm	This study
pMU612	pMU102 containing cbpp-pyrF cassette, Cm	This study
pMU749	pMU245, CEN6/*ARSH4, ura3 -S. cer.-E.coli-C. the.* vector, Ura+, Ap	This study
pMU769	pMU749 with ΔpyrF::gapDHp-cat cassette, Ura+, Ap, Cm	This study
pMU1016	pMU749 with Δ*pta*::*gapDH*p-*cat* cassette, Ura+, Ap, Cm	This study
pMU1162	pMU1016 with *cbp*p-*pyrF* cassette, Ura+, Ap, Cm	This study
Strains		
E. coli		
Top10	cloning strain	Invitrogen[3]
S.cereviciae		
InvSC1	Ura3$^-$ for gap repair cloning	Invitrogen[3]
C. thermocellum		
M0003	Wild type *C. thermocellum* DSM 1313	DSMZ[4]
M0970	*C. thermocellum* DSM 1313 Δ*pyrF*	This study
M0971	*C. thermocellum* DSM 1313 Δ*pyrF* Δ*pta*::*gapDH*p-*cat*	This study
M1061	*C. thermocellum* DSM 1313 Δ*pyrF* pMU612	This study
M1062	*C. thermocellum* DSM 1313 Δ*pyrF* pMU102	This study

[1]New England Biolabs, Ipswich, MA
[2]Bacillus Genetic Stock Center, http://www.bgsc.org/
[3]Invitrogen, Carlsbad, California
[4]Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Germany

Table 3. Phenotype and complementation of the Δ*pyrF* strain

Strain	Colony-forming units (cfu's)*			
	Rich Medium	Rich Medium + 5-FOA	Defined Medium	Defined Medium + Uracil
wt	1.0×10^7	3.0×10^4	1.8×10^7	1.2×10^7
Δ*pyrF*	6.9×10^7	5.7×10^7	1.0×10^3	2.7×10^7
Δ*pyrF* + pMU102	4.7×10^7	4.0×10^7	4.0×10^5	7.0×10^7
Δ*pyrF* + pMU612	2.3×10^7	1.0×10^7	3.7×10^7	5.1×10^7

* Each strain was dilution plated on the indicated medium and cfu's resulted after 48 hours are shown from a representative data set of three independent experiments. The phenotypes of the Δ*pyrF* strain were consistent with respect to uracil auxotrophy and 5-FOA resistance in all of the experiments, whereas the overall number of cfu's varied based on the plating efficiency.

TABLE 4. Dry weight and avicel conversion profile of the Δ*pta*::*gapDH*p-*cat* strain

Strain	Dry weight		% Avicel conversion	
	DW[a] g/l	SD[b]	% conversion[c]	% conversion[d]
WT	0.532	0.133	89.4	100.1
Δ*pyrF*	0.508	0.034	89.8	100.7
Δ*pta*::*gapDH*p-*cat*	0.537	0.033	89.3	95.9

[a] The average dry weight (DW) from triplicate samples of each strain at the end of 96 hours of fermentation on Avicel are indicated with the standard deviation (SD).

[b] Standard Deviation

[c] The % conversion of Avicel is calculated based on the starting Avicel concentration (5 g/l) and the final DW, which includes residual Avicel and dry cell weight.

[d] The % conversion of Avicel is calculated based on the starting Avicel concentration (5 g/l) and the estimated final Avicel concentration. We assumed that the dry cell mass for each strain after 96 hours of fermentation, is roughly equivalent to that obtained from cells that have consumed 5 g/l cellobiose (data reported in the text). Therefore, the estimated final Avicel concentration was calculated by subtracting dry cell mass from DW.