AEM Accepts, published online ahead of print on 30 September 2011
Appl. Environ. Microbiol. doi:10.1128/AEM.00646-11

High Ethanol Titers from Cellulose using Metabolically Engineered Thermophilic, Anaerobic Microbes

Running Title: High Titer Cellulosic Ethanol from Thermophiles

D. Aaron Argyros[a,e], Shital A. Tripathi[a,c,e], Trisha F. Barrett[a] , Stephen R. Rogers[a] , Lawrence F. Feinberg[a], Daniel G. Olson[a,b], Justine M. Foden[a], Bethany B. Miller[a], Lee R. Lynd[a,b], David A. Hogsett[a], and Nicky C. Caiazza[a,d*]

Author affiliations:

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

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

Current Address:

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

[d]Synthetic Genomics, 11149 North Torrey Pines Rd, LaJolla, CA 92037

[e]These authors contributed equally to this work

* To whom correspondence should be addressed. Nicky C. Caiazza. Synthetic Genomics., 11149 North Torrey Pines Rd, LaJolla, CA 92037. Phone: 858-754-2983. Fax: 858-754-2988. E-mail: ncaiazza@syntheticgenomics.com

Abstract

This work describes novel genetic tools for use in *Clostridium thermocellum* that allow creation of unmarked mutations while using a replicating plasmid. The strategy employed counter selections developed from the native *C. thermocellum hpt* gene and the *Thermoanaerobacterium saccharolyticum tdk* gene, and was used to delete the genes for both lactate dehydrogenase (Ldh) and phosphotransacetylase (Pta). The Δ*ldh* Δ*pta* mutant was evolved for 2,000 hrs, resulting in a stable strain with 40:1 ethanol selectivity and 4.2 fold increase in ethanol yield over the wild type strain. Ethanol production from cellulose was investigated with an engineered co-culture of organic acid deficient engineered strains of both *C. thermocellum* and *T. saccharolyticum*. Fermentation of 92 g/L Avicel by this co-culture resulted in 38 g/L ethanol with acetic and lactic acids below detection limits in 146 hours. These results demonstrate that ethanol production by thermophilic, cellulolytic microbes is amenable to substantial improvement by metabolic engineering.

Introduction

Lignocellulosic biomass is the most abundant feedstock for biofuel production and at $50/metric tonne it is one of the cheapest, equivalent on an energy basis to oil at $17/barrel (16). However, a significant cost is attributed to the saccharolytic enzymes required to release soluble sugars from lignocellulose (16). In order to match the current economics of starch hydrolysis from corn kernels, the cost of cellulase enzymes would need to decrease nearly an order of magnitude (1). A promising solution to reduce enzyme cost is consolidated bioprocessing (CBP), which requires an organism or collection of organisms capable of both releasing and fermenting cellulosic sugars without added enzymes and at high ethanol yield (17).

Expression of saccharolytic enzymes has been demonstrated in microorganisms with established genetic systems such as *S. cerevisiae* (7, 15, 29, 32) but the overall enzymatic activity has not rivaled that of natively cellulolytic organisms. Alternatively, metabolic engineering has produced high yield ethanol fermentation in model organisms that can be genetically manipulated (24, 27, 33). Therefore, metabolic engineering of a naturally cellulolytic microbe should make it possible to create a strain that can both hydrolyze and ferment the products of cellulose hydrolysis into ethanol as the major fermentation end product.

An attractive candidate for metabolic engineering is the thermophillic anaerobe *Clostridium thermocellum*, which is capable of hydrolyzing cellulose at rates approaching 2.5 g/l h (10). This high rate of cellulose hydrolysis results from the activity of a multi-protein complex called the cellulosome. This cell associated, organelle-like appendage has been extensively reviewed and contains an assemblage of structural proteins and enzymes designed to access, bind, and hydrolyze cellulose (2, 6). It is widely recognized that development of *C. thermocellum* as a CBP organism has been limited by the genetic tools required to create a stable strain with high ethanol to organic acid ratios (1, 18). Recent reports have described development of basic genetic tools for *C. thermocellum*, including transformation of plasmid DNA and the ability to make marked gene deletions (19, 28, 30). However, these tools do not provide the versatility associated with creating unmarked genetic mutations and while they could be used to eliminate organic acid production they would not be ideal. In particular, these approaches rely on replacement of gene targets with antibiotic resistance cassettes limiting the number of gene deletions to the available number of antibiotic resistance markers. Furthermore, this strategy can only be performed using the *C. thermocellum* Δ*pyrF* strain which has a growth defect and is not an optimal background for future genetic manipulation (28). The development of additional genetic tools, in particular counter selectable markers, is thus desirable.

Nucleic acid metabolism has been the basis for numerous counter selectable markers. The cellular activity of hypoxanthine phosphoribosyl transferase (Hpt) re-assimilates purines such as hypoxanthine, xanthine, and guanine for the purpose of DNA and RNA synthesis (25), but can lead to cellular toxicity in the presence of purine antimetabolites such as 8-azahypoxanthine (AZH). Recently *hpt* was developed into a useful genetic marker for counter selection in *Archaea* (20). The cellular toxicity of fluoro-deoxyuracil (FUDR) is dependent on the presence of two enzymes involved in pyrimidine metabolism: thymidine kinase (Tdk) and thymidilate synthetase (ThyA). Tdk converts FUDR to fluoro-dUMP (F-dUMP) which is a covalent inhibitor of ThyA and the basis for counter selection in a variety of eukaryotic organisms (5, 9, 12, 26).

C. thermocellum uses the Embden-Myerhof pathway to ferment sugars to pyruvate, with branched end product metabolism resulting in formation of ethanol, acetic acid and lactic acid (13). As reviewed elsewhere (17) a substantial effort has been devoted to strain isolation, optimization of culture conditions (14), and strain development via mutation and selection for improved thermophilic fermentation of cellulose and/or xylose to ethanol at high yield. However, these efforts did not result in robust strains that consistently produce ethanol at high yields under a broad range of conditions and in the hands of different investigators. Metabolic engineering using molecular techniques has been applied to manipulate end-product metabolism towards increased ethanol yield in the non-cellulolytic, thermophillic bacterium *Thermoanerobacterium saccharolyticum* (23-24) but not in the highly cellulolytic *C. thermocellum*. The benefit of a co-culture of these two organisms, would result from a combination of cellulose hydrolysis by *C. thermocellum* with broad sugar utilization and high ethanol yield of *T. saccharolyticum.* Prior efforts along these lines has been recently reviewed (6), but until now organic acid production by *C. thermocellum* has inhibited growth and decreased ethanol yield. In this study, we show that ethanol production by *C. thermocellum* was amenable to substantial improvement by metabolic engineering and demonstrate the long standing potential of a co-culture between ethanologenic versions of these two organisms.

Material and Methods

Molecular techniques. All plasmids constructions in this work were created using yeast gap repair cloning. A detailed description of the particular techniques, and yeast strains used for plasmid construction is described in detail by Shanks et al (22). A complete list of strains, plasmids, and primers used in this study to generate knockouts are listed in Tables 1 and 2. Primer design was based on the *C. thermocellum* 27405 genome (http://www.ncbi.nlm.nih.gov). Transformation was performed as in Tripathi *et al* (28).

Hpt deletion vector (pMU1657). A sequence including 65 bp of coding sequence and 1121 bp upstream of *hpt* was amplified using primers X107384 and X07385 and a sequence including the 10 bp of coding sequence and 1025 bp downstream of *hpt* was amplified using primers X07503 and X07383. These fragments were gap repair cloned into AatII digested pMU749 (28) along with a cassette containing the *hpt* gene expressed from the *C. thermocellum* cellobiose phosphorylase promoter and an antibiotic resistance cassette consisting of the *cat* gene from pNW33N fused to the *C. thermocellum* glyceraldehyde 3-phosphate dehydrogenase promoter.

Deletion vector backbone (pMU1647). The entire annotated sequence of plasmid pMU1647 is deposited in Genbank as accession #HQ264098 and is described pictorially in appendix figure A1. An operon was created to link the positive (*cat*) and negative (*hpt*) selectable markers. A 15 bp segment located between the first two genes of the *C. thermocellum* 1313 urease operon was used as a spacer and ribosome binding site for the *hpt* gene. To facilitate subsequent cloning of homologous flanks used for targeted deletions, the *gapDHp-cat-hpt* operon was flanked by FspI restriction sites. As a counter selection against the plasmid and selection for chromosomal integration of the operon, the *tdk* gene was amplified from *T. saccharolyticum* and fused to the *C. thermocellum* cellobiose phosphorylase promoter.

LDH deletion vector (pMU1758). To create plasmid pMU1758, plasmid pMU1647 was digest with FspI which cuts immediately upstream and downstream of the *gapDH*p-*cat*-*hpt* cassette. An 840 bp fragment located inside the lactate dehydrogenase (*ldh)* ORF was amplified from *C. thermocellum* genomic DNA using primers X08205 and X08206 with tails homologous to vector sequence surrounding the upstream FspI site in pMU1647. An 824 bp region upstream of the *ldh* ORF was amplified using primers X08207 and X08208 and a 952 bp region downstream of the *ldh* ORF was amplified using primers X08209 and X08210. These two fragments had homologous tails for each other and the sequence surrounding the downstream FspI site. The yeast origins and selectable markers were removed by digestion of pMU1758 with NotI, gel purification and ligation of remaining plasmid to create pMU1777.

PTA deletion vector. To create plasmid pMU1817, plasmid pMU1647 was digested with FspI. A 632 bp fragment located inside the *pta* ORF was amplified from *C. thermocellum* 1313 genomic DNA using primers X09013 and X09014 which contain tails homologous to vector sequence surrounding the upstream FspI. An 843 bp region upstream of the *pta* ORF was amplified using primers X09015 and X0915 and a 1035 bp region downstream of the *pta* ORF was amplified using primers X09017 and X09018 with tails for each other and the sequence surrounding the downstream FspI site.

Medium and culture conditions. Routine growth of *C. thermocellum* for genetic manipulations was done at 55º C using modified DSMZ medium 122 containing 5 g/l cellobiose. The modifications included a reduction in the K_2HPO_4 concentration to 1.8 g/l and the glutathione was replaced with 0.5 g/l Cysteine-HCl. Analysis of mutant strains was done in CM3 medium (3) containing Avicel at indicated concentrations. The inoculums for all fermentations were taken from an overnight culture in CM3 medium containing 10 g/l Avicel to induce cellulosome production with exception to those used in the batch reactor which where pre-grown in chemically defined MTC medium (34-36).

Evolution of Δhpt Δldh Δpta strain. The initial *Δhpt Δpta Δldh* strain was evolved for faster growth by serial transfer in CM3 medium containing 10 g/l Avicel for 10 transfers, 20 g/l for the second twenty transfers, 40 g/l for the third twenty transfers and 60 g/l for a final ten transfers. Transfers were by done by sub-culturing at a final dilution of 1:100 every 24-48 hrs.

Anti-metabolite selection media. Plating and selection against the *hpt* gene was done using a final concentration of 500 µg/ml 8-azahypoxanthine (Acros Organics, 202590050) in defined DSMZ medium122 which contained a previously described vitamin concentrations as a substitute for yeast extract (11). Selection against *tdk* was done in modified DSMZ medium 122 containing 10 µg/ml 5-fluoro-2'-deoxy-uridine (Sigma, F0503). All selections were done using cultures grown to mid-log phase.

Co-cultures *C. thermocellum* and *T. saccharolyticum* strains were initially grown for 18-20 hours at 55°C in mono-culture on CM3 medium containing 25 g/L cellobiose at pH 6.8. The cultures were transferred twice in this manner prior to inoculation. Co-culture experiments were conducted using duplicate 50 ml final volumes of CM3 containing ~17.2 g/L Avicel and 5 g/L calcium carbonate at pH 6.8 in 160 ml serum bottles purged with 5% carbon dioxide, 95% nitrogen gas mix prior to autoclaving. A 5% inoculum of each strain was introduced at the outset. Cultures were incubated at 55°C and shaken at 300 rpm. Well mixed samples were taken aseptically using a 3 ml syringe fitted with a 21G needle.

For reactor batch cultures, inocula for *C. thermocellum* and *T. saccharolyticum* strains were initially grown for 18-24 hours at 55°C in mono-culture on MTC medium containing respectively 10 g/L Avicel or cellobiose. The co-culture reactor was initiated using a 10% inoculum of each strain into a 2L Sartorius BIOSTAT® Bplus bioreactor (Sartorius BBI Systems, Inc., Bethlehem, PA, USA) with a 1L working volume containing MTC medium supplemented with 5 g/L calcium carbonate and 100 g/L (wet weight) Avicel. Incubations were conducted anaerobically at 55°C at 300 rpm stirring with a starting pH 6.3. Well mixed samples were taken periodically from the sampling port using a syringe.

Analytical techniques. Residual cellulose from Avicel was estimated gravimetrically and used to compute carbon balance as follows. Mixed 1.5 ml culture samples were centrifuged (10,000 x g) in 2 ml tubes for 3 minutes and the supernatants discarded. To solubilize residual carbonate, one ml of a 10% w/v citric acid solution was added and allowed to stand for 5 minutes after thorough mixing. The samples were then centrifuged and washed three times in 1 ml distilled water as described above prior to transfer to aluminum weighing tins. Drying overnight at 55-60°C was followed by 45 minutes at 105°C prior to weighing. The substrate concentrations were then computed based on the processed sample volumes. Yield (carbon balance) was calculated as the ratio of products to residual substrate (g/L glucose equivalents) in the fermentation samples according to the equation,

$$\text{Yield} = [(1.053 * \text{Cb}) + \text{Glu} + \text{LA} + (\text{AA}/0.667) + (\text{EtOH}/0.511)] / [1.111*(S_0 - S_f)],$$

where Cb, Glu, LA, AA and EtOH are the concentrations (g/L) of cellobiose, glucose, lactic acid, acetic acid and ethanol in the fermentation broth determined by HPLC and S_0 and S_f are the initial and final substrate concentrations measured as above. This equation includes carbon contributions from carbon dioxide, accounted for by stoichiometric correlation to ethanol and acetic acid formation, and cells, included with residual substrate. The maximum difference of computing converted substrate under the latter assumption is ±2% which is well within the sampling error (~5%) at low substrate concentrations. Fermentation metabolites were analyzed by HPLC methods described in Shaw *et al* (24).

Results

Development of selectable markers for use in *C. thermocellum*

Plating experiments revealed that *C. thermocellum* is sensitive to AZH (appendix table A1) and subsequent genetic analysis confirmed that open reading frame (orf) Cthe2254 is an *hpt* homolog. Figure 1a illustrates deletion of *hpt* from *C. thermocellum* using AZH counter selection. PCR analysis (Fig. 1b) of the *hpt* locus indicated that AZH resistant colonies contained an in-frame deletion of *hpt* that was confirmed by DNA sequencing. A cassette was constructed on pMU1647 (appendix figure A1) that contained the *C. thermocellum hpt* gene cloned in an operon with *cat* expressed from the *C. thermocellum gapDH* promoter (*gapDH*p). The functionality of the cassette was verified by transforming the Δ*hpt* strain with pMU1647, selecting on Tm and demonstrating that the resulting transformants were sensitive to AZH while the parental Δ*hpt* strain was not (appendix table A1).

Analysis of the *C. thermocellum* genome indicated orf Cthe1227 is a *thyA* homolog, but a homolog of *tdk* could not be found. Consistent with the absence of *tdk*, *C. thermocellum* was insensitive to FUDR (appendix table A1). *T. saccharolyticum tdk* was expressed from the *C. thermocellum* cellobiose phosphorylase promoter (*cbp*p) and transformed into *C. thermocellum* using pMU1647. Counter selection in the presence of FUDR was evident in the form of a 96% reduction in colony forming units (appendix table A1).

Deletion of genes responsible for organic acid formation using a removable marker system

Deletion of *ldh* and *pta*, encoding lactate dehydrogenase (Ldh) and phosphotransacetylase (Pta) respectively, was undertaken pursuant to development of a strain of *C. thermocellum* with high ethanol yield. Figure 2a illustrates deletion of *ldh* using the removable marker system described above. After step 2, Tm and FUDR resistant colonies were screened by PCR and the dominant product was 5.7 kb consistent with insertion of the integration cassette onto the chromosome (Fig. 2b). After step 3, AZH resistant colonies were screened by PCR and the dominant product was 2.0 kb consistent with deletion of the integration cassette and *ldh* from the chromosome (Fig. 2c), an event that was confirmed by DNA sequencing. These selections were very stringent as positive results were observed with 95% and 100% of colonies screened (*n* = 20) after step 1 and step 2, respectively.

The strategy in figure 2a was used to delete *pta* in both the Δ*hpt* and Δ*hpt* Δ*ldh* strains. Using the same primer set, PCR amplification shows a 1.0 kb difference between the wild-type *pta* locus (3.0 kb) and that of the Δ*pta* strains (2.0 kb) (Fig. 2d). All deletions were confirmed by DNA sequencing.

Fermentation analysis of *C. thermocellum* organic acid mutants

Fermentation profiles support molecular evidence for the creation of Δ*hpt* Δ*ldh*, Δ*hpt* Δ*pta*, and Δ*hpt* Δ*ldh* Δ*pta* strains. Results in figure 3a show the product profile at the end of batch fermentations using 19.5 g/l Avicel, a model microcrystalline cellulosic substrate. The Δ*hpt* strain, the parent of all genetically engineered strains in this study, produced acetate, lactate, and ethanol in a ratio of 1.7 : 1.5 : 1.0, similar to the 2.1 : 1.9 : 1.0 ratio produced by wild-type. The Δ*hpt* Δ*ldh* strain did not produce significant levels of lactate and had a 1.4 : 1.0 ratio of acetate to ethanol. Similarly, the Δ*hpt* Δ*pta* strain did not produced acetate and had a 1.9 : 1.0 ratio of lactate to ethanol. The Δ*hpt* Δ*ldh* Δ*pta* strain achieved ethanol selectivity of 40 : 1 relative to organic acids.

Acetate titer was not effected in the Δ*hpt* Δ*ldh* fermentation indicating that flux through this node is not altered by deleting lactate dehydrogenase if there is capability to make ethanol, which increased by 23% (Fig. 3a). In contrast, lactate titer increased 62% (Fig. 3a) in the Δ*hpt* Δ*pta* fermentation while ethanol titer still only increased 23%, indicating a significant increase in flux from pyruvate to lactate if acetate production is prevented. Fermentation with the Δ*hpt* Δ*ldh* Δ*pta* strain, where cellular carbon flux is more constrained towards ethanol, showed the most dramatic effect with a 56% increase in ethanol titer (Fig. 3a).

Pyruvate was detected in the Δ*hpt* Δ*ldh* Δ*pta* fermentation and could be an indication of a metabolic imbalance within the strain (Fig. 3a). Since metabolism and growth are related, a growth-based evolutionary engineering approach was taken to increase the efficiency of metabolism. The Δ*hpt* Δ*ldh* Δ*pta* strain was transferred in batch culture for 2,000 hrs and resulted in a strain that no longer secreted pyruvate and increased ethanol titer ~4-fold relative to wild-type, achieving 5.61 g/l (Fig 3a) from consumption of 18.4 g/l Avicel, corresponding to a yield of 0.27 g ethanol/g substrate.

Fermentation analysis of an engineered *C. thermocellum* and *T. saccharolyticum* co-culture

C. thermocellum is classified as a mixed acid fermenter, however in laboratory fermentations a significant portion of the carbon balance remains open when only these end products are taken into account. Furthermore, additional fermentation products that close the carbon balance have not been routinely identified. For the Δ*hpt* strain 73% of the carbon could be accounted for during Avicel fermentation (measurements used to calculate Avicel hydrolysis and carbon recovery for each strain are provided in appendix table A2). The carbon recovery of the Δ*hpt* Δ*ldh* Δ*pta* strain dropped to 33% and through evolution was improved to 61%. To direct more cellulosic carbon to known end products we co-cultured *C. thermocellum* with *T. saccharolyticum*. a non-cellulolytic mixed acid fermenter which has two beneficial traits; it has previously been engineered for high ethanol yield and all carbon can be accounted for following fermentation (24). Batch co-culture fermentations containing 17.2 g/l Avicel were conducted using wild-type or engineered versions of these two strains to make ethanol as the sole fermentation product. The fermentation profiles with respect to acetate, lactate, and ethanol are shown in figure 3b and additional fermentation measurements addressing Avicel conversion and carbon recovery are found in appendix table A3. During the course of both fermentations 95%-96% of the Avicel was hydrolyzed and the wild-type strain co-cultures produced acetate, lactate, and ethanol with a 1.3 : 0.8 : 1.0 ratio, while the engineered strain co-cultures made predominantly ethanol resulting in ~2-fold increase in titer at 6.7 g/l (Fig. 3b). The wild-type co-culture had a carbon recovery of 79%, which is a moderate improvement over the 72% recovery observed for wild-type *C. thermocellum*. However, the engineered co-culture had a carbon recovery of 76%, a significant improvement over the 61% observed for the evolved Δ*hpt* Δ*ldh* Δ*pta* strain.

Figure 4 shows ethanol yield for all fermentations expressed as a percentage of the theoretical maximum based on Avicel hydrolysis (theoretical ethanol yields for each fermentation are reported in appendix tables A2 and A3). The Δ*hpt* strain is the reference for all engineered *C. thermocellum* strains and exhibited a small improvement over wild-type with respect to ethanol yield and achieved ~18% of theoretical maximum (Fig. 4). Compared to this strain, both the Δ*hpt* Δ*ldh* and Δ*hpt* Δ*pta* mutants did not drastically improve ethanol yield, while the Δ*hpt* Δ*ldh* Δ*pta* mutant improved yield by only ~33%, or ~27% of theoretical maximum. The evolved version of the Δ*hpt* Δ*ldh* Δ*pta* strain had a more significant improvement over the reference strain, increasing ethanol yield by ~70%, which is ~60% of theoretical maximum. Ethanol yield of the wild-type co-culture was 35% of theoretical maximum, an improvement of ~58% relative to wild-type *C. thermocellum*. This was more than doubled by using engineered strains in co-culture fermentation, achieving ~75% of theoretical maximum.

In our hands the highest ethanol titer that could be achieved using the evolved Δ*hpt* Δ*ldh* Δ*pta* strain was 14 g/l from 40 g/l Avicel. To confirm that the limitation is not the cellulolytic capability of *C. thermocellum* but rather the metabolism of the organism we tested higher substrate concentrations in the context of a co-culture. A batch fermentation performed in a reactor with constant stirring and no pH control was used to test the ability of an engineered *C. thermocellum* – *T. saccharolyticum* co-culture. Fermentation of 92.2 g/l Avicel was carried out for 146 h at which point pH became limiting and ~90% of the Avicel had been hydrolyzed (appendix table A3). Throughout the fermentation, ethanol was the only fermentation product detected reaching 38.1 g/l (Fig. 5), which was ~80% of theoretical maximum (Fig. 4).

Discussion

A removable marker system for *C. thermocellum* was developed and used to delete genes responsible for organic acid formation. This development required information about host metabolism which is available for a vast number of organisms in the form of annotated genome sequences that identify potential enzymatic reactions that might be targets for anti-metabolites. We used this information to create counter selections for use in *C. thermocellum* that are based on both endogenous (*hpt*) and exogenous (*tdk*) genes. Combining the markers allowed us to use replicating plasmids to insert and remove markers from the host chromosome, creating in-frame gene deletions in the process. Like *C. thermocellum,* many organisms have been isolated from nature because they exhibit interesting biology or industrial promise but cannot be studied through routine genetic manipulation. As we demonstrated with *C. thermocellum*, it is possible to overcome poor transformation efficiency and/or inadequate selections that hinder development of these organisms by exploiting information encoded within the host genome.

Low yield remains an issue for *C. thermocellum,* however this was ameliorated through co-culture with *T. saccharolyticum*. While the reason for the improved performance of the co-culture is unclear, it may be that *C. thermocellum* secretes undetected metabolites that can be consumed by the co-culture partner and turned into product. Since no soluble sugars were observed during co-culture fermentation, it may be that the co-culture partner rapidly consumes excess sugar and prevents overflow metabolism from occurring in *C. thermocellum*.

Evolutionary engineering was used to improve upon the fermentation performance of the Δ*hpt* Δ*ldh* Δ*pta* strain but there are indications that additional metabolic tuning is needed. An important node that determines if acetyl-CoA is directed to acetate or ethanol is the mechanism by which the pool of reduced ferredoxin, formed during the conversion of pyruvate to acetyl-CoA by pyruvate:ferredoxin oxidoreductase, is re-oxidized. Since acetate is the primary fermentation product, it appears the most efficient way to oxidize reduced ferredoxin in *C. thermocellum* is through H_2 evolution as opposed to the formation of NADH by ferredoxin:NADH oxidoreductase (FNOR) and subsequent redox balance by ethanol production. Thus, in the absence of sufficient FNOR activity, lack of Pta may cause redox imbalance and the accumulation of central metabolites leading to and including acetyl-CoA. In the Δ*hpt* Δ*pta* strain this bottleneck is partially alleviated by increasing the production of lactate. The Δ*hpt* Δ*pta* Δ*ldh* strain does not have the ability to make lactate in response to redox imbalance and this may be the reason pyruvate is secreted into the growth medium and the overall product yield is lower. Exploration of the POR node and the identification of possible deficiencies, such as an FNOR-like activity capable of transferring electrons from reduced ferredoxin to NAD+, may provide a means to enhance the performance of *C. thermocellum* ethanol fermentation.

The evolved Δ*hpt* Δ*ldh* Δ*pta* strain produced a 40:1 molar ratio of ethanol to organic acid while making 5.61 g/l ethanol from 19.6 g/l Avicel. This strain, in co-culture with an engineered *T. saccharolyticum* ethanologen, fermented 92.2 g/l Avicel into 38.1 g/l ethanol with high conversion and yield comparable to actively growing *S. cerevisiae*. This is the highest ethanol titer produced by a thermophilic, cellulolytic culture to date. The ability to improve fermentation of ethanol from cellulose by *C. thermocellum* in co-culture with non-cellulolytic thermophilic anaerobes has been established with numerous pentose utilizing organisms including *T. saccharolyticum* (4, 8, 13). Recent advances described here and by Shaw *et al* (24) have allowed us to eliminate organic acid production which has been a major limitation to realizing the potential of these co-cultures. With the development of the tools and strains described in this report there are no insurmountable obstacles preventing the development of *C. thermocellum* as the cornerstone of a robust cellulolytic platform that is the basis for a CBP solution for cellulosic ethanol.

Acknowledgments

This research was supported by Mascoma Corporation and a grant from the BioEnergy Science Center (BESC), 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. We thank A. Joe Shaw, Christopher D. Herring, and William R. Kenealy for useful discussion.

Authors declare a conflict of interest.

References

1. **Alper, H., and G. Stephanopoulos.** 2009. Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? Nat Rev Microbiol **7:**715-723.
2. **Bayer, E. A., R. Lamed, B. A. White, and H. J. Flint.** 2008. From cellulosomes to cellulosomics. Chem Rec **8:**364-377.
3. **Benoit, L., C. Cailliez, E. Petitdemange, and J. Gitton.** 1992. Isolation of cellulolytic mesophilic clostridia from a municipal solid waste digestor. Microb. Ecol. **23:**117-125.
4. **Bryant, F. O., J. Wiegel, and L. G. Ljungdahl.** 1988. Purification and Properties of Primary and Secondary Alcohol Dehydrogenases from Thermoanaerobacter ethanolicus. Appl Environ Microbiol **54:**460-465.

5. **Czako, M., and L. Marton.** 1994. The herpes simplex virus thymidine kinase gene as a conditional negative-selection marker gene in Arabidopsis thaliana. Plant Physiol **104:**1067-1071.

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

7. **du Plessis, L., S. H. Rose, and W. H. van Zyl.** 2010. Exploring improved endoglucanase expression in *Saccharomyces cerevisiae* strains. Appl Microbiol Biotechnol **86:**1503-1511.

8. **Duong C.T.V., E. A. J., A.L. Demain.** 1983. Thermophillic, anaerobic, cellulolytic bacteria. Topics in Enzyme and Fermentation biotechnology **7:**156-195.

9. **Gardiner, D. M., and B. J. Howlett.** 2004. Negative selection using thymidine kinase increases the efficiency of recovery of transformants with targeted genes in the filamentous fungus Leptosphaeria maculans. Curr Genet **45:**249-255.

10. **Hogsett, D. A. L.** 1995. Cellulose hydrolysis and fermentation by Clostridium thermocellum for the production of ethanol. Thesis (Ph. D.)--Dartmouth College.

11. **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.

12. **Khang, C. H., S. Y. Park, Y. H. Lee, and S. Kang.** 2005. A dual selection based, targeted gene replacement tool for Magnaporthe grisea and Fusarium oxysporum. Fungal Genet Biol **42:**483-492.

13. **Lamed, R., and J. G. Zeikus.** 1980. Ethanol production by thermophilic bacteria: relationship between fermentation product yields of and catabolic enzyme activities in Clostridium thermocellum and Thermoanaerobium brockii. J Bacteriol **144:**569-578.

14. **Lamed, R.J, J.H. Lobos, T.M. Su.** 1988. Effect of stirring and hydrogen on fermentation products of *Clostridium thermocellum.* Appl Environ Microbiol 54:1216-1221

15. **Lilly, M., H. P. Fierobe, W. H. van Zyl, and H. Volschenk.** 2009. Heterologous expression of a Clostridium minicellulosome in Saccharomyces cerevisiae. FEMS Yeast Res **9:**1236-1249.

16. **Lynd, L. R., M. S. Laser, D. Bransby, B. E. Dale, B. Davison, R. Hamilton, M. Himmel, M. Keller, J. D. McMillan, J. Sheehan, and C. E. Wyman.** 2008. How biotech can transform biofuels. Nat Biotechnol **26:**169-172.

17. **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

18. **Maki, M., K. T. Leung, and W. Qin.** 2009. The prospects of cellulase-producing bacteria for the bioconversion of lignocellulosic biomass. Int J Biol Sci **5:**500-516.

19. **Olson, D. G., S. A. Tripathi, R. J. Giannone, J. Lo, N. C. Caiazza, D. A. Hogsett, R. L. Hettich, A. M. Guss, G. Dubrovsky, and L. R. Lynd.** 2010. Deletion of the Cel48S cellulase from Clostridium thermocellum. Proc Natl Acad Sci U S A. 107(41):17727-17732

20. **Pritchett, M. A., J. K. Zhang, and W. W. Metcalf.** 2004. Development of a markerless genetic exchange method for Methanosarcina acetivorans C2A and its use in construction of new genetic tools for methanogenic archaea. Appl Environ Microbiol **70:**1425-1433.

21. **Sato, K., M. Tomita, S. Yonemura, S. Goto, K. Sekine, E. Okuma, Y. Takagi, K., and a. T. S. Hon-Nami.** 1993. Characterization and ethanol hyper-production by Clostridium thermocellum I-1-B. Biosci. Biotechnol. Biochem. **57:**2116-2121.

22. **Shanks, R. M., 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.

23. **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 **191:**6457-6464.

24. **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 U S A **105:**13769-13774.

25. **Stout, J. T., and C. T. Caskey.** 1985. HPRT: gene structure, expression, and mutation. Annu Rev Genet **19:**127-148.

26. **Szybalski, W.** 1992. Use of the HPRT gene and the HAT selection technique in DNA-mediated transformation of mammalian cells: first steps toward developing hybridoma techniques and gene therapy. Bioessays **14:**495-500.

27. **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.

28. **Tripathi, S. A., D. G. Olson, D. A. Argyros, B. B. Miller, T. F. Barrett, D. M. Murphy, J. D. McCool, A. K. Warner, V. B. Rajgarhia, L. R. Lynd, D. A. Hogsett, and N. C. Caiazza.** 2010. Development of pyrF-based genetic system for targeted gene deletion in Clostridium thermocellum and creation of a pta mutant. Appl Environ Microbiol. 2010. Oct; 76(19):6591-9.

29. **Tsai, S. L., J. Oh, S. Singh, R. Chen, and W. Chen.** 2009. Functional assembly of minicellulosomes on the Saccharomyces cerevisiae cell surface for cellulose hydrolysis and ethanol production. Appl Environ Microbiol **75:**6087-6093.

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

31. **Wang, D. I. C., G. C. Avgerinos, I. Biocic, S. D. Wang, and H. Y. Fang.** 1983. Ethanol from cellulosic biomass. Philos. Trans. R. Soc. Lond. Ser.**:**323-333.

32. **Yanase, S., R. Yamada, S. Kaneko, H. Noda, T. Hasunuma, T. Tanaka, C. Ogino, H. Fukuda, and A. Kondo.** 2010. Ethanol production from cellulosic materials using cellulase-expressing yeast. Biotechnol J. May;5(5):449-55

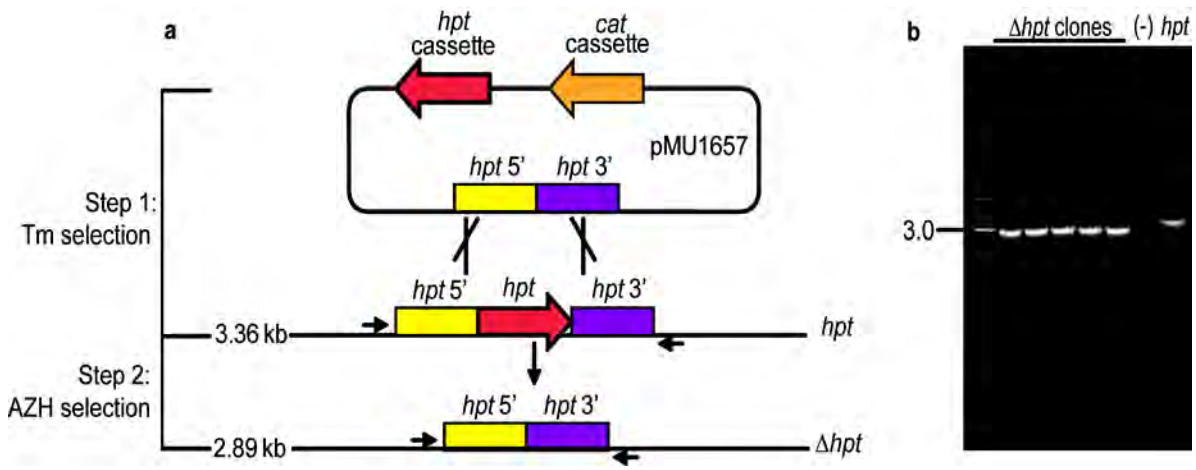
33. **Yomano, L. P., S. W. York, S. Zhou, K. T. Shanmugam, and L. O. Ingram.** 2008. Re-engineering Escherichia coli for ethanol production. Biotechnol Lett **30:**2097-2103.

34. **Zhang, Y., and L. R. Lynd.** 2003. Quantification of cell and cellulase mass concentrations during anaerobic cellulose fermentation: development of an

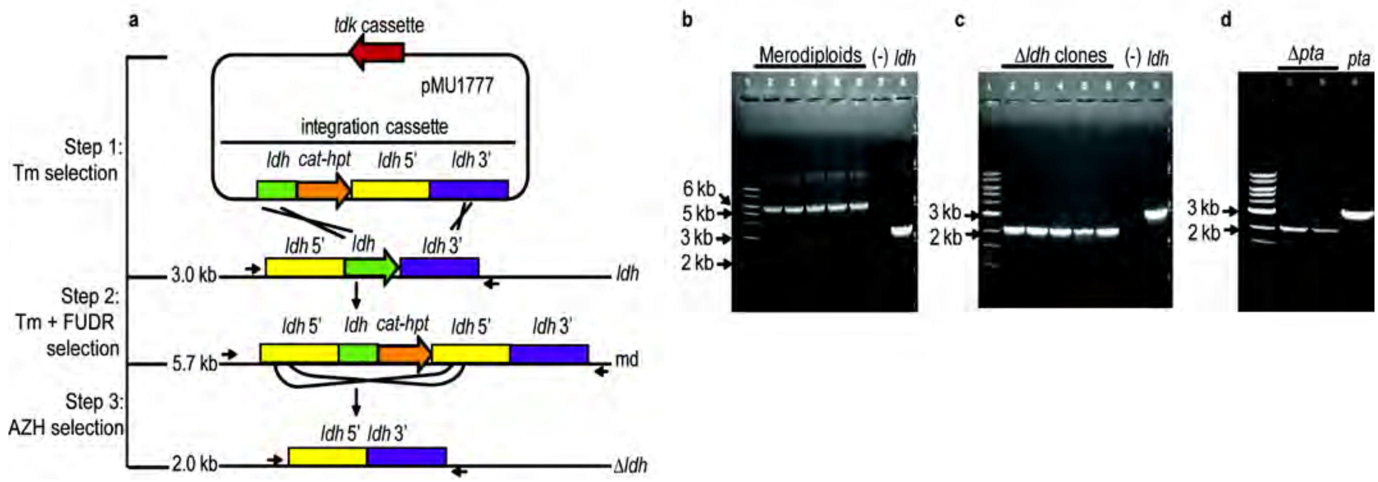
enzyme-linked immunosorbent assay-based method with application to Clostridium thermocellum batch cultures. Anal Chem **75:**219-227.

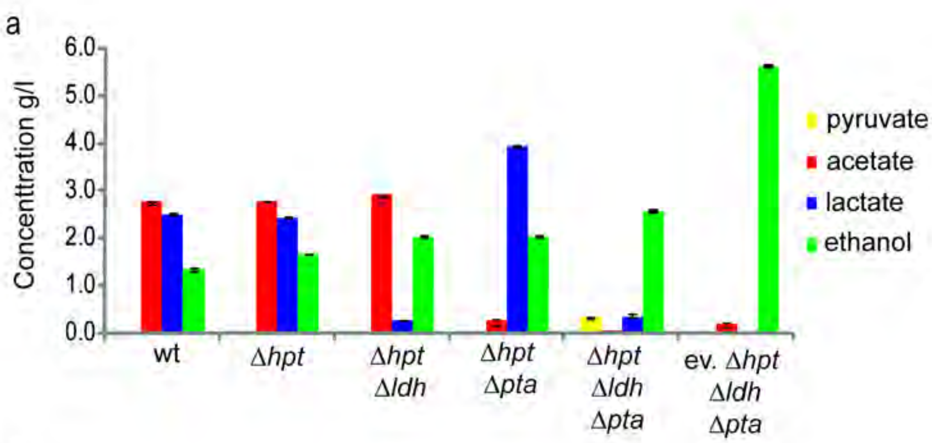
35. **Zhang, Y. H., and L. R. Lynd.** 2004. Kinetics and relative importance of phosphorolytic and hydrolytic cleavage of cellodextrins and cellobiose in cell extracts of Clostridium thermocellum. Appl Environ Microbiol **70:**1563-1569.

36. **Zhang, Y. H., and L. R. Lynd.** 2005. Regulation of cellulase synthesis in batch and continuous cultures of Clostridium thermocellum. J Bacteriol **187:**99-106.

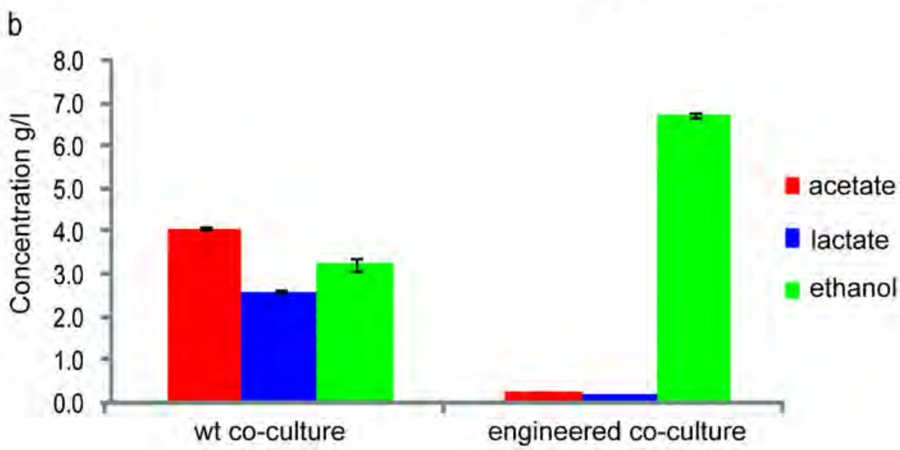


a
hpt cassette
cat cassette
pMU1657
hpt 5'
hpt 3'
Step 1:
Tm selection
hpt 5'
hpt
hpt 3'
3.36 kb
hpt
Step 2:
AZH selection
hpt 5'
hpt 3'
2.89 kb
Δhpt
b
Δhpt clones
(-)
hpt
3.0

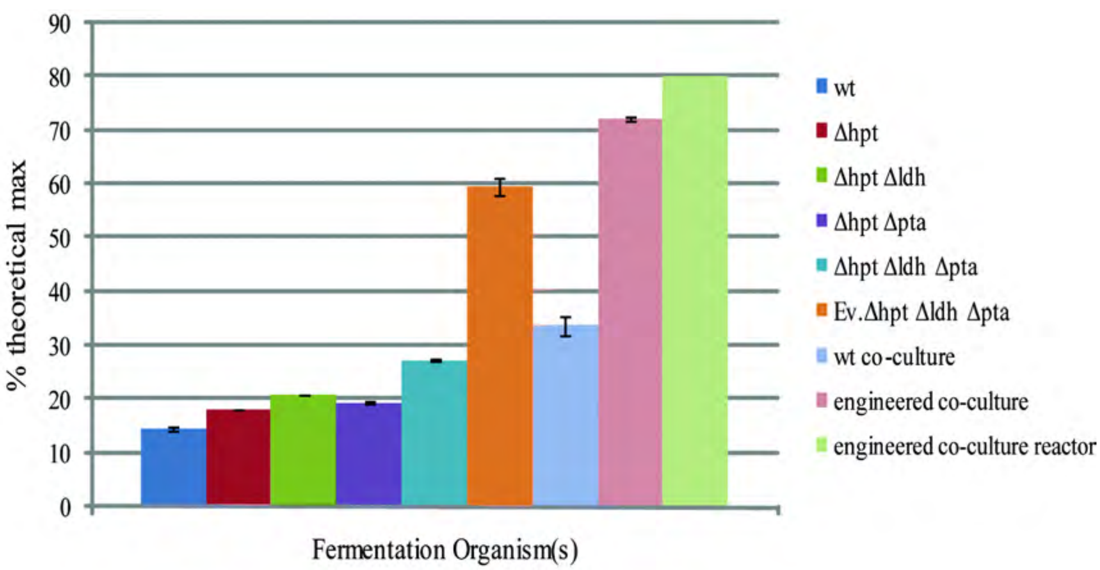

a
tdk cassette
pMU1777
integration cassette
Step 1:
Tm selection
ldh
cat-hpt
ldh 5'
ldh 3'
ldh 5'
ldh
ldh 3'
3.0 kb
ldh
Step 2:
Tm + FUDR
selection
ldh 5'
ldh
cat-hpt
ldh 5'
ldh 3'
5.7 kb
md
Step 3:
AZH selection
ldh 5'
ldh 3'
2.0 kb
Δldh
b
Merodiploids
(-)
ldh
6 kb
5 kb
3 kb
2 kb
c
Δldh clones
(-)
ldh
3 kb
2 kb
d
Δpta
pta
3 kb
2 kb



a
Concenttration g/l
6.0
5.0
4.0
3.0
2.0
1.0
0.0
pyruvate
acetate
lactate
ethanol
wt
Δhpt
Δhpt Δldh
Δhpt Δpta
Δhpt Δldh Δpta
ev. Δhpt Δldh Δpta



b
Concentration g/l
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
acetate
lactate
ethanol
wt co-culture
engineered co-culture


90
80
70
60
50
40
30
20
10
0
% theoretical max
Fermentation Organism(s)
wt
Δhpt
Δhpt Δldh
Δhpt Δpta
Δhpt Δldh Δpta
Ev.Δhpt Δldh Δpta
wt co-culture
engineered co-culture
engineered co-culture reactor

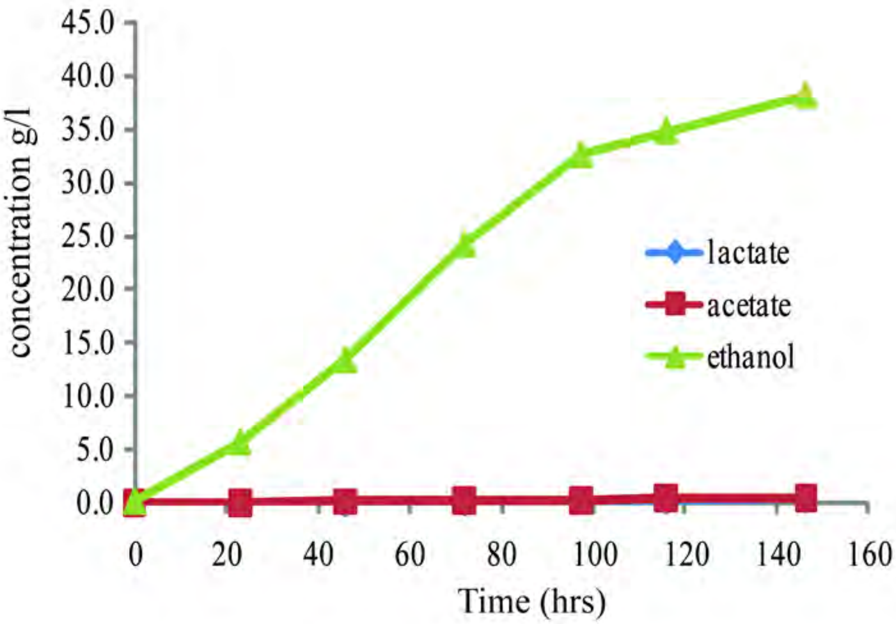

45.0
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
concentration g/l
0
20
40
60
80
100
120
140
160
Time (hrs)
lactate
acetate
ethanol

Figure Legends

Figure 1

Fig 1. Deletion of the *C. thermocellum hpt* gene. (a) Cartoon schematic illustrating the deletion of *hpt*. The plasmid used to delete *hpt*, pMU1657, contained: 1) a deletion cassette comprised of *hpt* 5' and 3' flanking DNA. 2) A *cat* cassette - *cat* driven by *C. thermocellum gapDH* promoter encoding Tm resistance. 3) An *hpt* cassette - *hpt* driven by the *C. thermocellum cbp* promoter encoding sensitivity to AZH. Step 1 - Plasmid transformation. *C. thermocellum* transformants selected using Tm resistance encoded by *cat* cassette. Step 2 – Plasmid curing and deletion of chromosomal copy of *hpt.* Plating on AZH selected for plasmid loss, mediated by the *hpt* cassette and deletion of chromosomal *hpt* by homologous recombination. Depiction of chromosomal regions, labels on right, indicate primer binding sites (→) used in diagnostic PCR. Expected amplicon size is shown to the left. (b) DNA gel showing results of diagnostic PCR at the *hpt* locus after step 2. Gel image contains 1 kb ladder (lane 1), diagnostic PCR of clones subjected to selection conditions (lanes 2-6), no DNA negative control (lane 7), *C .thermocellum* 1313 genomic DNA used as template to amplify the wild-type *hpt* locus as a positive control (lane 8).

Figure 2

Fig 2. Deletion of the *C. thermocellum ldh* and *pta* genes. (**a**) Cartoon schematic illustrating the deletion of *ldh*. The plasmid used to delete *ldh*, pMU1777, contains two major features: 1) *tdk* cassette - *T. saccharolyticum tdk* driven by *C. thermocellum cbp* promoter 2) an integration cassette – internal fragment of the *ldh* gene, transcriptional fusion of *cat* and *hpt* driven by the *C. thermocellum gapDH* promoter, and 5' and 3' flanking DNA labeled as *ldh 5'* and *ldh 3'*. The first step consists of transforming the *C. thermocellum* Δ*hpt* strain with pMU1777 and selecting on Tm. The second step combines Tm and FUDR which selects for recombination of the integration cassette onto the chromosome and for loss of the plasmid backbone containing the *tdk* cassette. This creates a merodiploid (md) strain in which the *ldh* 5' flanking DNA is duplicated. The third step utilizes AZH counter selection to isolate cells that have lost the *cat-hpt* cassette through recombination between the duplicated *ldh* 5' flanking DNA, resulting in deletion of *ldh*. Chromosomal regions are labeled to the right and indicate primer binding sites (→) used in diagnostic PCR. Expected amplicon size is shown to the left. (**b**) DNA gel showing results of diagnostic PCR at the *ldh* locus after step 2. (**c**) DNA gel showing results of diagnostic PCR at the *ldh* locus after step 3. Both gel images contain a 1 kb DNA Ladder (lane 1), diagnostic PCR of clones subjected to selection conditions (lanes 2-6), no DNA negative control (lane 7), Δ*hpt* strain genomic DNA used as template to generate wild-type *ldh* locus as a positive control (lane 8). (**d**) Gel image showing molecular confirmation of Δ*hpt* Δ*pta* and Δ*hpt* Δ*ldh* Δ*pta* strains. The same selections as those shown in figure 2a were performed to delete *pta* in both the Δ*hpt* and Δ*hpt* Δ*ldh* strains. Lanes 2-4 show the results of diagnostic PCR at the *pta* locus for the Δ*hpt* Δ*pta* strain (lane 2), the Δ*hpt* Δ*ldh* Δ*pta* strain (lane 3), and the Δ*hpt* strain (lane 4). The expected size of the wild-type *pta* amplicon is 3.0 kb, while the Δ*pta* amplicon is 2.1 kb. The 1 kb DNA Ladder is loaded in lane 1.

Figure 3

Fig 3. Fermentation profiles of engineered *C. thermocellum* mono-cultures and co-culture using both wt or engineered strains of *C. thermocellum* and *T. saccharolyticum*. (**a**) Batch fermentations were maintained at 55°C with initial pH of 7.0 and 19.5 g/l Avicel. Data represent the averages and standard deviations of duplicate fermentations sampled at 72 h. The ev. Δ*hpt* Δ*ldh* Δ*pta* strain is an evolutionarily engineered version of the Δ*hpt* Δ*ldh* Δ*pta* strain that was transferred in batch culture for 2,000 h. (**b**) Batch fermentations were maintained at 55°C with initial pH of 6.75 and 17.2 g/l Avicel. Data represent the averages and standard deviations of duplicate fermentations sampled at 120 h. *C. thermocellum* ev. Δ*hpt* Δ*ldh* Δ*pta* and *T. saccharolyticum* ALK2 comprised the engineered co-culture.

Figure 4

Fig 4. Percent theoretical maximum ethanol of fermentations performed in this study. Maximum theoretical ethanol yield was based on Avicel conversion and is listed in Supplementary tables 2 and 3 online. For a given strain or co-culture, data represents averaged measured ethanol generated during fermentation divided by the theoretical maximum ethanol. The engineered co-culture run in the reactor represents a single experiment.

Figure 5

Fig 5. High ethanol titer using *C. thermocellum* ev. Δ*hpt* Δ*ldh* Δ*pta* and *T. saccharolyticum* ALK2 co-culture. A batch reactor with continuous stirring set at 300 rpm with initial pH of 6.3 and 92.2 g/l Avicel was maintained at 55°C. Data represent a single fermentation sampled over the course of 146 h.

Table 1. Plasmids and strains used in this study

Plasmids/ strains	Description and relevant characteristics	Source or references
Plasmids		
pMU1164	Cloning vector; contains *cpb*p-*tdk* and *gapD*p-*cat-hpt* operon	This study
pMU1657	HPT KO vector with *cpb*p-*hpt* and *gapDH-cat* outside deletion flank	This study
pMU1758	Lactate dehydrogenase deletion vector	This study
pMU1777	Lactate dehydrogenase deletion vector without yeast machinery	This study
pMU1817	phosphotransacetylase deletion vector	This study
Strains		
C. thermocellum		
M0003	Wild type *C. thermocellum* DSM 1313	DSMZ[1]
M1354	*C. thermocellum* DSM 1313 Δ*hpt*	This study
M1375	M1354, Δ*ldh*	This study
M1448	M1354, Δ*pta*	This study
M1434	M1375, Δ*pta*	This study
T. saccharolyticum		
ALK2	*T. saccharolyticum* Δ*ldh*, Δ*pta-ack*	[24]

[1] Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Germany

Table 2. Oligonucleotides used in this study

Primer	Sequence 5’-3’
XO7384	TATATTTTTTAGTCCATATCTTCTTTGTCCGTATAAACAAGTTCCTCTCTGGTAACCAA
XO7385	CGTGTAAGTTACAGGCAAGCGATCGCGGCCGCGGTACGAGGCACAGGTCTTGACGGACT
XO7383	TAAAGAAATTTTGGTTACCAGAGAGGAACTTGTTTATACGGACAAAGAAGATATGGACT
XO7503	ACTAGGGCTCGCCTTTGGGAAGTTTGAAGGGCTACCGTCCATTTCAACCAACAGCTGAT
XO8205	GAATCTTTTCCTCTCTTTCGGAAAAGAAATACAATTTCTTCATCGTTGAAAGGCACGTT
XO8206	GTGTAGAAAGTGCCATGAAGTCCCGCGGACTTAAGGTTCCACCACAGCTTATACATTGA
XO8207	CTTTTAGAGTGTTTCCGGACTTTCTGAGAAGCTGTACAAAGCCTGCACCAACTACGGTT
XO8208	TATATTGCTATAAAGAATGAGGAGGGAACTAGTTGAATGACACTATCCTGTATCCTGAT
XO8209	TACCCGGGGATCCTCTAGAGTCGACCTGCAGCCATGCCTGGGAGGCTCTGTATAGAGAA
XO8210	AAGTAACCGTAGTTGGTGCAGGCTTTGTACAGCTTCTCAGAAAGTCCGGAAACACT
XO9013	AGAATCTTTTCCTCTCTTTCGGAAAAGAAATACCAATGGCGGCATCCACCTGAAGTTCT
XO9014	AGGTGTAGAAAGTGCCATGAAGTCCCGCGGACTTAAGGGAATTGCAAAGGTTGTACTGA
XO9015	GTACCCGGGGATCCTCTAGAGTCGACCTGCAGTGTCATCTCCCTTTTCTGCGGCATCCT
XO9016	GAAAGTACGGATCTGAGGGTTATTAAAGCCGCCGGTTCAGGCTCAATATGTCAAGGCAT
XO9017	AATTTATGCCTTGACATATTGAGCCTGAACCGGCGGCTTTAATAACCCTCAGATCCGTA
XO9018	AGGTATCGTTATATGGATACTGATAATTATCGCCAGGCAAAGTCCAACTATGCATTGGT