Elucidating Central Metabolic Redox Obstacles Hindering Ethanol

Production in *Clostridium thermocellum*

R. Adam Thompson^{1,2}, Donovan S. Layton^{2,3}, Adam M. Guss^{1,2,4}, Dan G. Olson^{2,5}, Lee R.
Lynd^{2,5}, and Cong T. Trinh^{1,2,3*}

¹Bredesen Center for Interdisciplinary Research and Graduate Education, the University of
Tennessee, Knoxville and Oak Ridge National Laboratory, Oak Ridge, TN, USA

²BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, TN, USA

³Department of Chemical and Biomolecular Engineering, the University of Tennessee,
Knoxville, TN, USA

⁴Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA

⁵Thayer School of Engineering, Dartmouth College, Hanover, NH, USA

*Corresponding author: ctrinh@utk.edu

Key Words: *Clostridium thermocellum*; redox metabolism; energy metabolism; elementary mode
analysis; minimal metabolic functionality; ethanol

ABSTRACT

Clostridium thermocellum is an anaerobic, Gram-positive, thermophilic bacterium that has generated great interest due to its ability to ferment lignocellulosic biomass to ethanol. However, ethanol production is low due to the complex and poorly understood branched metabolism of *C. thermocellum*, and in some cases overflow metabolism as well. In this work, we developed a predictive stoichiometric metabolic model for *C. thermocellum* which incorporates the current state of understanding, with particular attention to cofactor specificity in the atypical glycolytic enzymes and the complex energy, redox, and fermentative pathways with the goal of aiding metabolic engineering efforts. We validated the model's capability to encompass experimentally observed phenotypes for the parent strain and derived mutants designed for significant perturbation of redox and energy pathways. Metabolic flux distributions revealed significant alterations in key metabolic branch points (e.g., phosphoenol pyruvate, pyruvate, acetyl-CoA, and cofactor nodes) in engineered strains for channeling electron and carbon fluxes for enhanced ethanol synthesis, with the best performing strain doubling ethanol yield and titer compared to the parent strain. *In silico* predictions of a redox-imbalanced genotype incapable of growth were confirmed *in vivo*, and a mutant strain was used as a platform to probe redox bottlenecks in the central metabolism that hinder efficient ethanol production. The results highlight the robustness of the redox metabolism of *C. thermocellum* and the necessity of streamlined electron flux from reduced ferredoxin to NAD(P)H for high ethanol production. The model was further used to design a metabolic engineering strategy to phenotypically constrain *C. thermocellum* to achieve high ethanol yields while requiring minimal genetic manipulations. The model can be applied to design *C. thermocellum* as a platform microbe for consolidated bioprocessing to produce ethanol and other reduced metabolites.

1. INTRODUCTION

Sustainable energy production is a defining challenge for modern society, and achieving this goal will depend on the ability to produce fuels and chemicals from renewable sources. Consolidated bioprocessing (CBP) (Demain 2009, Guss et al., 2012) has the appeal of combining the steps of hydrolytic enzyme production, feedstock solubilization, and fermentation of multiple sugars into a single step (van Zyl et al., 2007), with the goal of maximizing cost-efficiency of lignocellulose conversion (Lynd et al., 2008).

C. thermocellum is an anaerobic, thermophilic, gram-positive bacterium which has been considered for CBP due to its strong capability to ferment lignocellulosic feedstocks while producing ethanol and hydrogen (Lynd et al., 2005, Ragauskas et al., 2014). However, wild-type *C. thermocellum* produces a wide range of fermentative products (e.g. acetate, formate, lactate) as well as amino acids, which limits the yield of ethanol (Ellis et al., 2012). Also, the wild-type cannot grow at the high concentrations of ethanol that are industrially necessary without extensive directed evolution (Herrero and Gomez 1980, Brown et al., 2011, Shao et al., 2011). Progress in engineering *C. thermocellum* for increased ethanol production has been made in several ways: isolating a high ethanol-tolerant mutant via adaptive evolution (Williams et al., 2007), elimination of competing carbon fermentative pathways (e.g., acetate (Tripathi et al., 2010), lactate (Biswas et al., 2014), and both acetate and lactate (van der Veen et al., 2013)), expressing an exogenous pyruvate kinase from *Thermoanaerobacter saccharolyticum* and deleting malic enzyme (Deng et al., 2013), as well as eliminating a major competitor for electrons, H₂ production (Biswas et al., 2015). None of these strategies, however, has been successful in reaching the goal of at least 90% conversion of fermentable sugars to ethanol as

75 seen in non-cellulolytic mesophilic organisms such as *Saccharomyces cerevisiae* (Kuyper et al.,
76 2005) and *Escherichia coli* (Yomano et al., 1998, Trinh et al., 2008).

77 An important obstacle to metabolic engineering of *C. thermocellum* for enhanced biofuel
78 production is an incomplete knowledge of its metabolism. Recent studies suggest that *C.*
79 *thermocellum* possesses an atypical central metabolism. For instance, the glycolysis does not
80 have an annotated pyruvate kinase (Deng et al., 2013). Fructose-6-phosphate kinase uses
81 pyrophosphate (PPi) instead of ATP, and GTP can substitute for ATP in several reactions which
82 might have important implications for metabolism (Zhou et al., 2013). The pentose phosphate
83 pathway lacks several enzymes in the oxidative branch which implies the need for alternative
84 pathways for NADPH synthesis (Rydzak et al., 2012). The Krebs cycle is also incomplete with
85 the lack of succinate dehydrogenase (Roberts et al., 2010). In addition, *C. thermocellum* has
86 many transhydrogenase reactions involving ferredoxins and both NADH and NADPH.

87 In this work, we consolidated the current understanding of *C. thermocellum* metabolism
88 into a predictive metabolic network model used for metabolic flux quantification and rational
89 strain design. We characterized engineered *C. thermocellum* strains designed and constructed
90 with significant perturbations of energy and redox pathways to validate the model. Redox
91 perturbation involved partial and full disruption of H₂ production while energy perturbation
92 disrupted acetate production. Their unique phenotypes were analyzed by quantitative metabolic
93 flux distributions. Next, we used the model to elucidate the key bottleneck in the redox
94 metabolism of *C. thermocellum* hindering high-yield ethanol production. Finally, the validated
95 model was used as an input for elementary mode analysis (Trinh et al., 2009) to design an
96 optimal genotype for efficient ethanol production in *C. thermocellum*.

2. MATERIALS AND METHODS

2.1. Strains and plasmids

C. thermocellum strains employed in this study (Table 1) were derived from the genetically tractable strain DSM 1313 with a Δhpt background, allowing for a two-stage selection counter-selection method using 8-azahypoxanthine (Argyros et al., 2011). *C. thermocellum* $\Delta hydG$ and $\Delta hydG \Delta ech$ strains were previously described (Biswas et al., 2015). The mutant $\Delta hydG \Delta pta-ack$ was constructed from the $\Delta hydG$ strain using the plasmid pAMG498, similar to previous reports (Argyros et al., 2011, van der Veen et al., 2013), but also deleting acetate kinase. Gene loci of *hpt*, *hydG*, *ech*, and *pta-ack* are Clo1313_2927, Clo1313_1571, Clo1313_0564-0575, and Clo1313_1185-1186, respectively. To eliminate pyruvate-formate lyase (PFL) activity and create an effective Δpfl strain, the parent strain or $\Delta hydG \Delta ech$ strain was cultured with 6 mM sodium hypophosphite (HPP), an analogue of formate and known PFL inhibitor (Unkrig et al., 1989, Rydzak et al., 2014).

2.2. Growth conditions

C. thermocellum strains were cultured in a defined MTC medium containing (per liter): 5 g cellobiose, 2 g urea, 1.5 g ammonium chloride, 2 g sodium citrate tribasic dihydrate, 1.5 g citric acid monohydrate, 1 g sodium sulfate, 1 g potassium phosphate monobasic, 2.5 g sodium bicarbonate, 1 g magnesium chloride hexahydrate, 0.2 g calcium chloride dihydrate, 0.1 g iron (II) chloride tetrahydrate, 1 g L-cysteine HCl, 20 mg pyridoxamine dihydrochloride, 4 mg p-aminobenzoic acid, 2 mg biotin, 2 mg vitamin B12, 1.25 mg manganese chloride tetrahydrate, 0.5 mg cobalt chloride hexahydrate, 0.125 mg zinc chloride, 0.125 mg copper chloride dihydrate,

0.125 mg boric acid, 0.125 mg sodium molybdate dihydrate, and 0.125 mg nickel chloride hexahydrate (Ozkan et al., 2001, Zhang and Lynd 2003, Kridelbaugh et al., 2013). Balch tube studies were conducted with 10 mL MTC plus the appropriate additional chemicals where noted. For bottle studies and inoculum preparation, 120 mL serum bottles were prepared with 90 mL MTC. Balch tube and serum bottle cultures also contained 5 g L⁻¹ MOPS for buffering, starting at a pH adjusted to 7.4 with 5M KOH. Bottles and tubes were sealed with butyl rubber stoppers and aluminum seals, while the head space contained a N₂:CO₂ (90:10) mixture. Cells were maintained at -80°C prior to initial inoculation. Inocula were incubated in water baths set at 55°C.

Bioreactor studies were conducted with a Biostat B+ (Sartorius Stedim US, Bohemia, NY) dual 1.5 L fermentation system at a working volume of 1 L MTC media with MOPS omitted. Prior to inoculation, freshly prepared media were autoclaved and sparged overnight with a N₂:CO₂ (90:10) gas mixture at 500 mL/min at 55°C. Inoculation used ~10% v/v mid-log phase cell culture to give a starting optical density of 0.05 in bioreactors. At inoculation, the gas flow was slowed to 150 mL/min, and the gas input was altered to flush the headspace only. The culture was maintained at pH 7.0 with 5 M KOH and 40% H₃PO₄, an agitation rate of 200 rpm, and a temperature of 55°C. Batch bioreactor studies were performed at least in duplicate. For each run, triplicate samples were collected and analyzed at each time point. The fluxes, titers, and yields were reported as means ± standard deviation for all samples.

Optical density (OD) was measured via a spectrophotometer at 600 nm (Spectronic 200+, Thermo Fisher Scientific, Inc.) To establish a calibration curve between dry cell weight (DCW) and OD (i.e., 0.502 g DCW = 1 OD₆₀₀, R²>0.95) 80 mL culture samples at known optical densities were centrifuged and washed twice with 0.9% (w/v) NaCl in water. Washed cell pellets

were then oven dried at 70°C for at least three days, until DCW was stable. The calibration curve was constructed from at least two different cell samples for several different OD measurements. An elemental DCW composition of C₄H₇O₂N (101 g/mol) was used to calculate DCW in moles (Rydzak et al., 2009).

2.3. Analytical methods

For all studies, substrate and product concentrations were quantified by using a Shimadzu high pressure liquid chromatography (HPLC) system. One mL aliquots of culture were centrifuged, then the supernatants filtered at 0.2 micron and run on an Aminex 87H (Biorad Inc., Hercules, CA) column using a refractive index detector (RID) to monitor glucose consumption with lactate, formate, acetate, pyruvate, and ethanol production by an established method (Trinh et al., 2008).

Gas chromatography coupled with mass spectrometry (GC/MS) was also used to quantify extracellular amino acids. At each measured time point during fermentation, 100 µL aliquots of cell culture were centrifuged at 25,200 rcf for 5 minutes and cell supernatants collected for secreted amino acid analysis. Amino acids were then derivatized using the EZ-faast™ Free/Physiological Amino Acid GC/MS kit (Model # KG0-7166, Phenomenex Inc., Torrence, CA, USA). The manufacturer's protocols for derivatization and GC/MS separation/analysis were followed, and the concentrations of secreted amino acids were calculated from calibration curves with standard mixtures provided by the kit as well as control mixtures made in-house.

2.4. Real-time gas measurement from bioreactor runs

H₂ monitoring was accomplished in real-time using a universal gas analyzer (UGA-300) equipped with an RGA probe (SRS Inc., Sunnyvale, CA, USA) connected to the gas effluent stream from the bioreactor system. The UGA-300 was equipped with a multiple inlet valve for the concurrent sampling from both bioreactor vessels. Exhaust gases (mainly N₂, H₂, and CO₂) from the bioreactor vessels were monitored by in-line mass flow meters (Cole Palmer, Vernon Hills, IL) calibrated for a N₂:CO₂ mixture to allow for accurate calculation of H₂ concentration. The pressure within the reactor was maintained at 1.5 atm with a ball-type pressure relief valve in series with the effluent stream, downstream of the mass flow meter and before reaching the UGA-300. Accumulative H₂ (n_{H2} in mole) were determined by the following formula:

$$n_{H_2} = \int_0^t \frac{P \cdot v}{R \cdot T} \cdot X_{H_2} \cdot dt \quad [1]$$

where t (hr) is the fermentation time, P (Pa) is the pressure of the reactor, v (m³/hr) is the volumetric flow rate, R (J/mol/K) is the ideal gas constant, T (K) is temperature, and X_{H₂} is the molar fraction of H₂.

Carbon dioxide production was not measured directly due to the presence of CO₂ in the influent gas causing fluctuations in the CO₂ signal from the UGA-300. The fluctuations were sufficient to disrupt an accurate measurement of CO₂ concentration. In this study, the CO₂ production was estimated from the concentrations of ethanol, acetate, formate, and cell mass at specific time points during the fermentation by the formula:

$$[CO_2] \triangleq [ethanol] + [acetate] + 0.5461 [cell\ mass] - [formate] \quad [2]$$

where square brackets represent concentrations of respective species. The scalar operator for cell mass represents the sum of the requirement of acetyl-CoA and α-ketoglutarate minus the requirement of oxaloacetate for cell mass synthesis (Deng et al., 2013) (Supplementary Table

S1). It should be noted that DCW and cell mass are used interchangeably throughout the manuscript.

2.5. Determination of experimental fluxes

Concentration profiles of 11 potentially prominent components of the fermentation broth (i.e., cellobiose, glucose, succinate, lactate, pyruvate, acetate, ethanol, formate, CO₂, H₂, and cell mass) were determined at each time point. Concentrations of all 20 major amino acids were measured during exponential growth and at the final time point, and the prominent species were included in flux analysis and carbon balance. Their fluxes were determined as follows:

$$r_P = \mu \cdot Y_{P/X} = \mu \cdot \frac{d[P]/dt}{d[X]/dt} \quad [3]$$

where r_P (mmol/ g DCW/hr) is the specific consumption/production rate (or flux) of metabolite P; μ (hr⁻¹) is the specific growth rate; $Y_{X/P}$ is the yield of metabolite P per unit cell mass X; and [P] (mmol/L) and [X] (mmol/L) are the concentrations of metabolites P and X, respectively.

2.6. Construction of *C. thermocellum* model for metabolic pathway analysis

A metabolic network was constructed for *C. thermocellum* which utilized cellobiose, a representative substrate for cellulosic biomass, as the sole carbon source (Figure 1, Supplementary Table S1). The network reconstruction was based on the publically available curation of the DSM 1313 genome sequence (Feinberg et al., 2011) as well as an extensive literature review. The model accounts for the strict anaerobic metabolism of *C. thermocellum* which does not use the canonical forms of Embden-Meyerhof glycolysis (Zhou et al., 2013), the pentose phosphate pathway (Rydzak et al., 2012), or the Krebs (tricarboxylic acid, TCA) cycle (Roberts et al., 2010). In addition, the export of major fermentative products (i.e., ethanol,

acetate, lactate, formate, and H₂), CO₂, valine, and DCW were included in the model. The biomass term was constructed following a composition described previously (Gowen and Fong 2010, Roberts et al., 2010), adjusting amino acid and G+C content for *C. thermocellum* DSM 1313. All cellular components are derived from 12 precursor metabolites (Neidhardt et al., 1990). Total ATP requirement for DCW synthesis (i.e., $Y_{ATP}^{max} = 42.7$ mmol ATP/g DCW) was estimated following the well-established protocol (Stouthamer 1973). The maintenance energy requirement was modeled as a separate stoichiometric reaction.

Briefly, our model proceeds from cellobiose uptake via an ATP-binding cassette transporter and cleavage via cellobiose phosphorylase (Ng and Zeikus 1982, Nochur et al., 1992). Glucose subunits are then processed through glycolysis until the phosphoenolpyruvate (PEP) step, where PEP can be converted to pyruvate through two branches: i) inter-conversion to oxaloacetate and malate through the "malate shunt," effectively transferring electrons from NADH to NADPH in a transhydrogenase type of mechanism, ii) decarboxylation of oxaloacetate to pyruvate via oxaloacetate decarboxylase (ODC), and iii) direct conversion to pyruvate coupled with ATP production via pyruvate phosphate dikinase (PPDK). Enzymes in these pathways are transcribed (Riederer et al., 2011) and translated (Rydzak et al., 2012) according to recent reports, despite PPDK enzymatic activity not being confirmed in cell-free extracts (Zhou et al., 2013).

Redox metabolism of *C. thermocellum* is complex, with multiple potential pathways available for redox cofactor recycling (Carere et al., 2012). Our model incorporated detailed electron shuttling reactions, consistent with the most recent understanding (Rydzak et al., 2014). Most prominently represented in *C. thermocellum* is a complicated system which shuttles electrons between reduced ferredoxin, NADH, and NADPH using several different enzyme

complexes including reduced ferredoxin:NAD(P) oxidoreductases such as RNF or NFN, an Fe-Ni energy conservation hydrogenase (ECH), and several Fe-Fe hydrogenases which are activated by HydG (Figure 1). A full list of reactions with stoichiometry is included in the Supplementary Materials.

2.7. Elementary mode analysis

From the construction of the *C. thermocellum* metabolic network, the mass balance for intracellular metabolites can be written as:

$$\frac{dc}{dt} = \mathbf{S} \cdot \mathbf{r} - \mu \cdot \mathbf{c} \quad [4]$$

where $\mathbf{c}$ is the concentration vector of intracellular metabolites; $\mathbf{S}$ is the stoichiometric matrix whose rows and columns correspond to metabolites and reactions, respectively; $\mathbf{r}$ is the intracellular flux vectors ($r_i \geq 0$ for irreversible reactions i); and μ is the dilution effect. At quasi-steady state, equation 4 can be simplified to:

$$\mathbf{S} \cdot \mathbf{r} = 0 \quad [5]$$

$$r_i \geq 0 \quad [6]$$

Elementary mode analysis (EMA) was applied to calculate all possible solutions $\mathbf{r}$, called elementary modes (EMs) that satisfy equation 5, inequality 6, and an additional non-decomposability constraint. The constraint states that for any $\mathbf{EM}_1$ and $\mathbf{EM}_2$, $S(\mathbf{EM}_1)$ cannot be a subset of $S(\mathbf{EM}_2)$ and vice versa where $S(\mathbf{EM}_1)$ and $S(\mathbf{EM}_2)$ are indices of non-zero fluxes of $\mathbf{EM}_1$ and $\mathbf{EM}_2$, respectively (Trinh et al., 2009). EMs were calculated for the metabolic network of *C. thermocellum* DSM 1313 by using METATOOL (Pfeiffer et al., 1999, Kamp and Schuster 2006).

2.8. Metabolic flux analysis

We calculated the metabolic flux distribution for *C. thermocellum* based on the measured fluxes of extracellular metabolites under a given condition. The flux distribution was determined by a non-negative, linear combination of EMs using Poolman et al.'s algorithm (Poolman et al., 2004), and briefly described below:

$$\mathbf{r} = \mathbf{EM} \cdot \mathbf{w} \quad [7]$$

where $\mathbf{r}$ is a $nx1$ flux vector; $\mathbf{EM}$ is a nxk EM matrix; and $\mathbf{w}$ ($\in \mathfrak{R}^+$) is a $kx1$ weighing factor vector. By using the measured flux vector $\mathbf{r}_m$ ($\mathbf{r}_m \subseteq \mathbf{r}$), $\mathbf{w}$ can be calculated as follows:

$$\mathbf{r}_m = \mathbf{EM}_m \cdot \mathbf{w} \quad [8]$$

$$\mathbf{w} = \text{pinv}(\mathbf{EM}_m) \cdot \mathbf{r}_m \quad [9]$$

where $\text{pinv}(\mathbf{EM}_m)$ is the Moore-Penrose pseudo inverse matrix of $\mathbf{EM}_m$ (mxk). The flux calculation errors were determined as follows:

$$\mathbf{error} = \frac{\|\mathbf{r} - \mathbf{r}_m\|}{\|\mathbf{r}\|} \quad [10]$$

One-at-a-time sensitivity (OATS) analysis was performed by systematically removing an individual reaction from the EM space for a given genotype, calculating the flux distribution with the resulting EM space, and comparing the new calculated error to the error of the original flux distribution vector.

2.9. Metabolic flux ratio analysis

To quantitatively differentiate flux distributions at key metabolite nodes in the central metabolic network upon perturbations, we employed the metabolic flux ratio (METAFor) analysis (Sauer et al., 1999, Szyperski et al., 1999). The metabolic flux ratio for an incoming flux(es) at node i ($MFR_{i\leftarrow}$) is defined as

$$MFR_{i\leftarrow} = \frac{r_{j,i\leftarrow}}{\sum_k r_{k,i\leftarrow}} \quad [11]$$

where $r_{j,i\leftarrow}$ is the incoming flux(es) of node i through reaction j . $MFR_{i\rightarrow}$ can be defined similarly where the outgoing flux(es) of node i ($r_{j,i\rightarrow}$) is represented with the opposite arrow. Given the significant overflow metabolism in *C. thermocellum*, we are interested in carbon and electron fluxes directly to ethanol as well as upstream of ethanol synthesis. The key branch points in focus here are i) conversion of PEP to pyruvate ($\rightarrow$ PYR) via pyruvate phosphate dikinase (PPDK), oxaloacetate decarboxylase (ODC), and malic enzyme (MAE), with ODC and MAE jointly considered, ii) recycling reduced ferredoxin ($Fd_{rd}\rightarrow$) through the Ni-Fe energy conserving hydrogenase (ECH), the Rnf-type NADH:Fd oxidoreductase (RNF), the NADH- Fd_{rd} :NADP⁺ oxidoreductase (NFN), and the NADH- Fd_{rd} bifurcating hydrogenase (BIF), iii) production of hydrogen ($\rightarrow$ H₂) through ECH and BIF as well as the Fe-Fe NADPH-dependent hydrogenase (Fe-H₂), iv) consumption of pyruvate (PYR $\rightarrow$) through pyruvate formate lyase (PFL), pyruvate:ferredoxin oxidoreductase (PFOR), and lactate dehydrogenase (LDH), and v) acetyl-CoA flux (ACoA $\rightarrow$) to alcohol dehydrogenase (AdhE), phosphotransacetylase (PTA), and citrate synthase (TCA1) for ethanol, acetate, and cell mass production, respectively.

The workflow for integrating the experimental flux measurements to metabolic network analysis to elucidate reaction bottlenecks is shown in the Supplementary Figure S1.

3. RESULTS

299

300 **3.1. Quantitative analysis of cellular phenotypic states of *C. thermocellum***

301 We first characterized the *C. thermocellum* strain DSM1313 Δhpt (henceforth denoted
302 “parent strain”). Similar to literature reports, the parent strain produced a mixture of fermentative
303 products and amino acids (with valine as the most dominant) (Figure 2A) with a good carbon
304 recovery (Figure 2B, Table 2). Due to the mixed acid fermentation, the ethanol yield (0.16 ± 0.01
305 g ethanol/g cellobiose) was low, as expected.

306 From the time-dependent concentration data, we extracted experimental fluxes for
307 measured metabolites (Table 2), and calculated the metabolic flux distribution of the parent
308 strain (Supplementary Table S2). Linear regression between the experimental and calculated
309 fluxes gave an excellent fit ($R^2 = 0.96$, Supplementary Figure S2A), indicating that the calculated
310 flux distribution vector is a reasonable assessment of the central metabolism of *C. thermocellum*.
311 In the subsequent sections, flux distributions of various *C. thermocellum* strains were calculated
312 for quantitative analysis of their cellular physiological states, and were used to identify reaction
313 bottlenecks hindering efficient ethanol production.

314 We used METAFoR analysis to examine the resulting flux distribution at key nodes more
315 closely. Flux from pyruvate is determined directly by monitoring formate and lactate production.
316 Additionally, the fate of acetyl-CoA in our model is determined by fermentation products and
317 cell mass. However, the flux to pyruvate from PEP and flux through the redox pathways are only
318 constrained by H_2 production which might not be effectively constrained and require additional
319 experimental evidence (e.g., enzyme activities and proteomics). The best fit of experimental
320 fluxes onto the EM space predicts that over 99% of flux to pyruvate proceeds through ODC and
321 MAE combined, with the small remainder proceeding through PPDK (Figure 3B). One-at-a-time

sensitivity (OATS) analysis shows no effect when the malate shunt is deleted (Table 3), but this calculated flux distribution is consistent with high enzymatic activity in crude cell extracts (Zhou et al., 2013).

For H₂ flux, the best fit predicts over 90% proceeds through the ECH hydrogenase. This observation makes sense because, under high gas flow rate that was used in the experiments, the soluble H₂ was low and hence ECH was probably the major hydrogenase used in a thermodynamically favorable reaction. Since the network has many local redundancies around H₂ to allow for meaningful stoichiometric analysis for the parent strain, we also performed the OATS analysis to investigate the network structure's ability to give equivalent solutions. The results show that individual deletions of ECH, RNF, NFN, Fe-H₂, or BIF have no effect on the error of the experimental data fitting (Table 3). This suggests that for the wildtype the redundancies in the redox pathway can compensate for individual gene deletions, and that genetically perturbing the system is necessary to obtain additional insights.

3.2. Metabolic model illustrates robust redox metabolism of *C. thermocellum*

3.2.1 Effect of partial deletion of hydrogen production

We characterized previously constructed strains with disrupted H₂ production in order to elucidate the effect of redox perturbation on *C. thermocellum* metabolism and ethanol production (Biswas et al., 2015). The first characterized mutant was the $\Delta hydG$ strain containing a disruption of the maturase enzyme complex, which inactivates the three Fe-Fe hydrogenases in *C. thermocellum* (Feinberg et al., 2011, Mulder et al., 2011) (Figure 1). This strain also acquired a spontaneous mutation in the bifunctional alcohol/aldehyde dehydrogenase *adhE*, which

allowed it to not only utilize NADH as an electron donor for acetaldehyde reduction to ethanol, but also NADPH. Deletion of *hydG* caused a 40% drop in the specific growth rate compared to the parent strain, and a lower yield of cell mass on cellobiose (Table 2). While the cellobiose uptake flux was minimally affected, the ethanol flux and yield on cellobiose increased 18% and 24%, respectively, and this slight increase in ethanol production was accompanied by a corresponding drop in acetate production (Figure 3A), presumably due to greater electron availability for ethanol synthesis while lowering the acetyl-CoA pool used for acetate production.

Most interestingly, H₂ yield increased ~17% compared to the parent strain (Figure 3A), which resulted in a very limited increase in ethanol production. This phenotype was unexpected and conflicts with previous reports of serum bottle cultures of the $\Delta hydG$ strain having a significant drop in H₂ yield (Biswas et al., 2015). This discrepancy can be explained by culture conditions, as we used a venting mechanism to measure gas concentrations in real-time and controlled pH, which prevents accumulation of H₂ and likely increases flux through the Ech hydrogenase.

To gain insights into the redox metabolism of *C. thermocellum*, we examined the flux distributions of the $\Delta hydG$ strain (Supplementary Table S2), which matched well with the experimental data ($R^2 = 0.99$, Supplementary Figure S2B). A parametric plot of normalized flux distributions with respect to the cellobiose uptake fluxes shows no significant difference between the parent and $\Delta hydG$ strains ($R^2 = 0.98$, Figure 4A). Outliers of this plot demonstrate the discrepancies of fluxes between the parent strain and the mutants. The only significant outlier was the NADPH-dependent AdhE* reaction, which is known not to be active in the parent strain (Lamed and Zeikus 1980). Additionally, METAFoR analysis suggests that the malate shunt was

still the major pathway to channel the carbon flux from PEP to pyruvate, and the fate of acetyl-CoA was minimally affected by the deletion of *hydG* (Figure 3B). METAFoR analysis also shows that H₂ production can be compensated by the major ECH hydrogenase activity for effective electron shuttling in the absence of HydG-related enzyme activity. OATS analysis shows that removal of ECH from the $\Delta hydG$ network leads to a dramatic 170% change in error (Table 3).

3.2.2. Effect of complete deletion of hydrogen production

To characterize a dramatic push of electron flux towards enhanced ethanol production, we investigated the dual deletion $\Delta hydG \Delta ech$ strain. The results show this mutant produced little H₂ and enhanced ethanol yield by 2.1 fold, while having similar growth characteristics as the $\Delta hydG$ strain (Table 2). Similar to the phenotype of the $\Delta hydG$ strain, the increase in ethanol production of the $\Delta hydG \Delta ech$ strain was accompanied by decreased acetate production. However, one striking phenotype was that the formate yield, not the lactate yield, increased two-fold (Figure 3A).

To elucidate how the $\Delta hydG \Delta ech$ strain changed its metabolism to balance redox reactions, we calculated and analyzed the metabolic flux distribution (Supplementary Table S2). The distribution matched well with the experimental data ($R^2 = 0.99$, Supplementary Figure S2C). Figure 4B shows the significant discrepancies between the parent and the $\Delta hydG \Delta ech$ strain ($R^2=0.82$). The outliers in this plot can be classified into two distinct groups, one associated with both carbon and electron fluxes (i.e., NADPH-dependent AdhE*, total ethanol production, PTA, and PFL) and the other associated with only electron fluxes (i.e., RNF, NFN, ECH, and total H₂ production). METAFoR analysis on the flux distribution of the $\Delta hydG \Delta ech$

strain further revealed the dramatic restructuring of redox metabolism. The calculated share of reduced ferredoxin through the RNF reaction increased four-fold compared to the parent strain, whereas the flux through the NFN reaction, which was practically inactive in the parent strain, increased to account for 20% of the reduced ferredoxin recycling (Figure 3B). This increase in NAD(P)H availability allowed for ethanol production to increase to 80% of the total flux through acetyl-CoA, and was matched by a drop in acetate production. Interestingly, ethanol yield only reached 59% of the theoretical limit while hydrogenase activities were diminished.

In summary, the complete disruption of H₂ production provided an effective means to increase flux of reducing equivalents to ethanol. This significant perturbation might have caused accumulation of reducing equivalents and hence triggered a regulatory mechanism for redox balancing where cells diverted pyruvate to acetyl-CoA through the redox neutral PFL reaction instead of the reduced ferredoxin-generating PFOR reaction. This undesirable loss of reducing equivalents lessened the potential ethanol yield. The effect of this bifurcation at the pyruvate node on redox balance and ethanol production is further investigated in Section 3.4.

3.3. The metabolic model encompasses several phenotypes

As an additional assessment of our model's predictive capability, we constructed and characterized the $\Delta hydG \Delta pta-ack$ strain which has perturbations designed to route carbon and electrons toward ethanol production with reduced production of H₂ and acetate. The results show that the $\Delta hydG \Delta pta-ack$ strain achieved a 61% increase in the ethanol yield (Figure 3A) while its growth rate decreased compared to the parent strain and the aforementioned $\Delta hydG$ strain by 50% and 30%, respectively (Table 1). This decreasing trend carried over to the cellobiose uptake rate, acetate production rate, and H₂ production rate (Table 2). The decreased growth rate and

cell mass yield, with respect to the parent strain, could be attributed to the lack of ATP synthesis via substrate level phosphorylation through the *pta-ack* reaction pathway. In the $\Delta hydG \Delta pta-ack$ strain, the lack of acetate production helped increase carbon flux from the acetyl-CoA node towards ethanol production, while also increasing flux of reducing equivalents (Figure 3A). The need to balance carbon and electron fluxes to ethanol also translated to a 72% decrease in formate production, compared to the parent strain.

Like the aforementioned mutants, the calculated flux distribution matched experimental data well ($R^2 = 0.96$, Supplementary Figure S2D), and when parametrically compared to the parent strain, one can quickly see a similar pattern of outlier reactions as seen in the $\Delta hydG \Delta ech$ strain (Figure 4C). The most prominent outliers focused again on carbon fluxes (i.e., NADPH-dependent AdhE*, total ethanol production, and PTA) and the electron fluxes (e.g., RNF, NFN, ECH, and total H₂ production). To elaborate on this point, METAFoR analysis (Figure 3B) shows how the $\Delta hydG \Delta pta-ack$ strain recycled reduced ferredoxin equivalently through the ECH and RNF reactions, as opposed to the parent strain predominately utilizing ECH. Furthermore, the NFN flux increased to about 14% of the total reduced ferredoxin recycling in the $\Delta hydG \Delta pta-ack$ strain. The increase in RNF and NFN fluxes provided the NAD(P)H necessary for ethanol production, and as expected the share of acetyl-CoA going to ethanol production increased over 60% compared to the parent strain (Figure 3B).

Taken all together, the model elucidates how *C. thermocellum* adjusts electron and carbon fluxes when its redox metabolism is perturbed towards ethanol production. These results give confidence in the model structure and any further phenotypic predictions.

3.4. Elucidating the redox bottlenecks hindering ethanol production

While the energy and redox perturbations enhanced ethanol flux, the ethanol yield was still low, only reaching 60% of the theoretical limit. This result indicates that the capacity for ethanol flux might still be limiting ethanol production. Since *C. thermocellum* employs multiple metabolic control valves, such as flux bifurcation at the pyruvate node via PFL and PFOR and/or distribution of electron fluxes via NFN, RNF, BIF, and the Fe-H₂ reactions, we used EMA to simulate additional redox-disrupted genotypes to identify key bottlenecks hindering high-yielding ethanol production.

3.4.1. PFL is vital for redox balancing when hydrogenases are inactive

Relative to the parent strain, a 70% increase in the flux to formate in the $\Delta hydG \Delta ech$ mutant suggests that accumulation of reduced ferredoxin results in a decrease in PFOR flux and increase in PFL flux. The importance of formate production is also apparent in the OATS analysis for the $\Delta hydG \Delta ech$ network (Table 3). To further examine this observation, we investigated whether additional deletion of *pfl* in the $\Delta hydG \Delta ech$ mutant would cause a redox imbalance which inhibits cell growth.

In silico results, via EMA, show that the $\Delta hydG \Delta ech \Delta pfl$ strain had zero EMs that supported cell growth (Supplementary Figure S3A). To date, this strain has not been constructed. However, hypophosphite (HPP) is a known PFL inhibitor (Rydzak et al., 2014) and can be used to experimentally eliminate PFL activity. Therefore, we tested this genotype experimentally by growing the $\Delta hydG \Delta ech$ strain in the presence of 6 mM HPP. While HPP had no effect on the growth of the parent strain, we were able to confirm the model's prediction of growth inhibition when HydG, Ech, and Pfl were inactivated (Figure 5A). These results suggest that without the

redox neutral relief valve of PFL, the lack of H₂ production causes a redox imbalance within the metabolism of *C. thermocellum* that eliminates cell growth.

3.4.2. Probing redox bottlenecks via external electron acceptors

To investigate if growth could be recovered when Pfl is inhibited in the $\Delta hydG \Delta ech$ strain, we simulated the addition of alternative electron consumption pathways into the model (see Table S1 for pathway description.) The electron sinks were chosen to test the accumulation of the major electron carrying species: 1) fumarate was tested for NADH recycling due to the presence of a fumarate reductase domain (Clo1313_3018) and a hypothetical protein with similarity to succinate dehydrogenase (Clo1313_2640); 2) 2-keto-isovalerate (KIV) was tested for NAD(P)H recycling due to experimental evidence of isobutanol production in *C. thermocellum* (Holwerda et al., 2014); and 3) sulfate was tested for ferredoxin recycling as at least one of the steps to sulfate reduction is ferredoxin dependent (Feinberg et al., 2011). Reactions are included in Table S1. *In silico* results demonstrate that addition of any single electron sink pathway into the $\Delta hydG \Delta ech \Delta pfl$ metabolic network restored the phenotypic space to the same area and yield ranges as the parent or $\Delta hydG$ strain (Supplementary Figure S3). To test these phenotypes experimentally, we investigated the growth behavior of the parent versus $\Delta hydG \Delta ech$ strains in the presence of various electron sinks, with and without inhibition of PFL by HPP.

First, the parent strain (positive control) could grow with or without the HPP inhibition regardless of supply of external electron acceptors (Figure 5A), as expected. For the $\Delta hydG \Delta ech$ strain, exogenous addition of 40 mM fumarate was unable to restore growth with PFL inhibition (Figure 5B). This result indicates that i) there is not an accumulation of NADH and/or ii) the

annotated fumarate reductase was not active. It should be noted that the level of fumarate addition was non-toxic because the parent strain and $\Delta hydG \Delta ech$ strains grew under conditions without HPP. Furthermore, addition of 40 mM KIV led to slight growth (Figure 5C) and isobutanol production (data not shown) after two days. This result indicates that accumulation of NAD(P)H could be occurring.

Interestingly, the normal levels of sulfate in MTC media ($1 \text{ g L}^{-1} \text{ Na}_2\text{SO}_4$) were not sufficient to allow growth of $\Delta hydG \Delta ech \Delta pfl$. However, addition of extra sulfate ($2 \text{ g L}^{-1} \text{ Na}_2\text{SO}_4$ total) into the culture medium was able to bring the maximum OD_{600} to levels comparable to the control (Figure 5D) after a lag phase. This result suggests that reduced ferredoxin is the species most likely accumulating, causing the lethal electron imbalance.

Collectively, perturbation of redox metabolism was effective for pushing electron and carbon fluxes towards the production of reduced metabolites e.g. ethanol. Accumulation of reduced ferredoxin triggered the bifurcation at the pyruvate node to balance redox potentials. Thus, either transfer of electrons from ferredoxin to NAD(P)H or electron pulling by the native ethanol pathway was not sufficient to maximize ethanol production. This mechanism was supported by the observed high formate flux in the $\Delta hydG \Delta ech$ strain as well as *in silico* and *in vivo* probing of redox bottlenecks in *C. thermocellum* using different external electron acceptors.

3.5. Design of optimal genotype of *C. thermocellum* for high-yield ethanol production

With the trained metabolic model, we next applied EMA to analyze the network structure and give insights for rational strain design (Trinh et al., 2009).

3.5.1. Additional modifications are necessary to constrain *C. thermocellum* phenotype

To explain why the above metabolic engineering strategies were not sufficient to achieve high-ethanol yields, we used EMA to qualitatively assess phenotypic spaces of the parent, $\Delta hydG$, $\Delta hydG \Delta ech$, and $\Delta hydG \Delta pta-ack$ strains, solely based on the metabolic network structure. Table 4 summarizes the EM properties of each genotype while Figure 6 compares the phenotypic spaces of the parent strain with the three mutant strains focusing on the 2-D projection of the ethanol versus cell mass yields on cellobiose. This representation is commonly used to qualitatively compare what phenotypes a metabolic network can feasibly attain given its stoichiometric structure. By definition, the experimental phenotypes of the wildtype or any knock-out mutant must fall within the phenotypic space predicted by the model unless new reactions are added to the network; otherwise, there is at least one error present in the model structure.

When examining the presented strains, an immediately noticeable result is the similarity of the phenotypic spaces. The $\Delta hydG$ strain had the same space boundaries as the parent strain (the area ‘abcd’, Figure 6), while the $\Delta hydG \Delta ech$ strain shows a slightly contracted space (the area ‘abefd’) with regards to the cell mass yield. Since cellobiose and dry cell weight have equivalent oxidation states, the contracted space can be attributed to the lack of carbon-free electron disposal in the form of hydrogen production requiring the synthesis of reduced carbon compounds such as ethanol and lactate to balance redox. This flux of carbon to other reduced products limits the yield of dry cell weight. The $\Delta hydG \Delta pta-ack$ strain exhibits a truncated boundary on the cell mass yield axis (the area ‘abgd’) which can be attributed to a lower ATP availability for cell synthesis due to a lack of substrate-level phosphorylation by *pta-ack*.

Most importantly, for each genotype, the admissible range of ethanol still stretched from 0.00 to 0.67 (Cmol/Cmol). This indicates that there are accessible phenotypic states for these

strains that do not produce any ethanol; the existence of the undesired EMs in these strains helps explain why they did not reach high ethanol yields experimentally (Table 2). It should be noted that each phenotypic space represent the metabolic capabilities of each genotype at quasi steady state. Under a given operating condition, each genotype will occupy a point within its own space, i.e., symbols in Figure 6 designate the phenotypic states of the parent strain and its mutants under our characterized conditions.

3.5.2. Metabolic model predicts feasibility of high-ethanol yielding phenotype

We investigated the ability of the central metabolic model to describe a high ethanol yielding ($> 75\%$ theoretical maximum) phenotype by performing the Minimal Metabolic Functionality (MMF) algorithm (Trinh et al., 2008). The MMF method constrains the phenotypic space by iteratively performing gene deletions *in silico*, aiming to reduce the total number of elementary modes and the accessible range of the desired product. When running the MMF algorithm, we opted to start with the $\Delta hydG \Delta ech$ genotype, since an increase in ethanol yield was observed in that strain, the dual NADH/NADPH activity has been shown for ethanol synthesis, and genetic manipulation in *C. thermocellum* is a bottleneck in strain design (Tripathi et al., 2010, Mohr et al., 2013). Figure 7 presents the stepwise progress of the MMF algorithm on the number of EMs describing the genotype, as well as the range of ethanol yields. While simulating a $\Delta hydG \Delta ech \Delta pta-ack$ strain reduced the total number of elementary modes by about half, and simulating $\Delta hydG \Delta ech \Delta pta-ack \Delta ldh$ reduced the total EMs by another 40%, the range of ethanol yield was not constrained (Figure 7). However, simulating $\Delta hydG \Delta ech \Delta pta-ack \Delta ldh \Delta val_{out}$ brought the ethanol yield range to 0.50-0.67 (Cmol ETOH/Cmol CB) for non-growth associated production and 0.50-0.62 (Cmol ETOH/Cmol CB) for growth-associated

ethanol production, while reducing the total number of EMs again by over half. This result invites the hypothesis that central metabolism can be constrained to a high-yielding phenotypic space with only three additional gene deletions. The ability to delete or overexpress genes in *C. thermocellum* is a significant challenge in its own right (Guss et al., 2012, Mohr et al., 2013), and as such the construction of the optimal strain predicted here is beyond the scope of this report.

4. Discussion

We developed and validated a predictive, stoichiometric, central metabolic model of *C. thermocellum* for flux quantification to elucidate cellular phenotypes and rational strain design based on EMA. The model was able to illuminate how *C. thermocellum* dramatically shifts its cofactor recycling processes when exposed to significant energy and redox perturbations. In each of the mutants presented, a change from the parent strain central metabolism led to an increase in ethanol production. The proposed mechanism is an increase in conversion of reduced ferredoxin to NAD(P)H through NFN and RNF.

Our study reveals the major driving force in ethanol production is a need to balance redox metabolism, as opposed to the push of carbon flux. At the same time, the diverse routes for regeneration of oxidized electron carriers offer many different avenues for producing reduced products other than ethanol. For instance, in the $\Delta hydG$ strain, the results suggest that the reduced ferredoxin positively affects ECH flux since specific H_2 production increased without a significant change in the production of ethanol. Additionally, the $\Delta hydG \Delta ech$ and $\Delta hydG \Delta pta-ack$ strains had similar acetate yields, but the $\Delta hydG \Delta ech$ strain had a 28% increase in ethanol yield compared to the $\Delta hydG \Delta pta-ack$ strain (Figure 3A).

The need to balance redox metabolism as a driving force is also intuitive when considering acetyl-CoA is made available by both PFOR and PFL reactions as well as the fact that the sum of PFOR and PFL fluxes is not significantly altered by genotype (Table S2). This suggests AdhE* activity is not acetyl-CoA limited. The experimental data suggests that regenerating reduced ferredoxin solely through PFOR is not sufficient to match the rate of cellobiose uptake and glycolysis. To compensate, pyruvate consumption is diverted from the stalling PFOR reaction towards the redox neutral PFL reaction (Figure 3B), which is consistent with literature reports of conditions with disrupted H₂ synthesis (Rydzak et al., 2014, Biswas et al., 2015). The decrease in flux through PFOR can be attributed to rate-limitation at either i) reduced ferredoxin recycling via NFN or RNF, or ii) the push of NAD(P)H to ethanol synthesis (i.e. AdhE* activity).

It is worth noting the difference in H₂ production in the $\Delta hydG$ strain in serum bottles (Biswas et al., 2015) and the bioreactor studies presented here could be linked to inhibition of ECH activity and an increase in the activities of NFN and RNF. The calculated flux distribution for the $\Delta hydG \Delta ech$ strain shows that NFN and RNF become highly active to compensate for diminished hydrogenase activity. Therefore, it is feasible to propose that the higher partial pressure of H₂ in serum bottle cultures increases the availability of NAD(P)H in the $\Delta hydG$ strain. Indeed, it has been shown that artificially increasing H₂ in the culture atmosphere leads to increased ethanol production (Rydzak et al., 2011). The hypothesis that H₂ inhibition is lower for ECH than the bifurcating hydrogenase or Fe-Fe hydrogenases could explain the higher H₂ yield in the parent strain serum bottle studies, since all hydrogenases should be actively present. In contrast, the only way that the $\Delta hydG \Delta ech$ strain can recycle reduced ferredoxin is to produce NAD(P)H via RNF or NFN.

To further highlight the importance of redox balance, inhibiting PFL activity in the $\Delta hydG \Delta ech$ strain causes growth to cease. This phenotype was predicted by our model and then confirmed in batch cultures. This effective genotype serves as a promising experimental condition to probe bottlenecks in redox cofactor recycling by use of exogenous electron sinks. We anticipated that the addition of fumarate would serve as a NADH probe and restore redox balance through succinate production. However, the fumarate study was inconclusive as the inability to restore growth could be due to a lack of fumarate reductase expression, lack of a fumarate transport protein, or an incorrect annotation of the fumarate reductase protein.

The inconclusive results from the fumarate study led to the use of KIV as a NAD(P)H probe. This gave more promising results, as growth was slowly able to recover, cellobiose was consumed in the process, and isobutanol was detected. The slow growth recovery clearly shows additional reducing equivalents can be recycled via KIV conversion to isobutanol. The slow rate of growth likely reflects a kinetic bottleneck in the conversion of reduced ferredoxin to NAD(P)H and/or low activity of KIV reductase, although isobutylaldehyde reduction kinetics or isobutanol toxicity cannot be ruled out.

The final electron sink, sulfate, was designed to probe the accumulation of reduced ferredoxin. After a lag phase, growth was recovered to a rate comparable to controls without HPP. The lag phase could be due to an adjustment in gene expression levels for greater sulfate reduction. These results allude to reduced ferredoxin being the main accumulating redox cofactor, which implies that re-oxidation of ferredoxin is a rate limiting step in the electron recycling pathway of *C. thermocellum*. Future experiments overexpressing these potential bottleneck enzymes should enhance ethanol production.

Our application of the MMF method to *C. thermocellum* revealed that elimination of H₂ production coupled with the removal of acetate and lactate synthesis as well as valine secretion effectively constrains the network to a high yielding ethanol phenotype (Figure 7). It is worthwhile again to note that PFL is not included in this set of reaction deletions, as the model predicts that removing H₂ and formate production does not allow cell growth. While the MMF method seeks to find all reaction deletion sets that have the minimum size (i.e., cardinality) and exclude reactions downstream of a linear pathway (Trinh et al., 2009), it is possible to identify other unique sets with larger sizes by using the cMCS tool to achieve the same high yielding ethanol phenotype (Hädicke and Klamt 2011). Due to many possible candidate design strains generated by the cMCS method (>5000 for our network), great care must be taken to screen out those that are not biologically relevant. Given the redox bottlenecks in reduced ferredoxin recycling, as previously discussed, simply deleting the abovementioned genes might not immediately facilitate high ethanol productivity. Presumably, a directed evolutionary approach to select for faster growth or, more proactively, the intentional overexpression of NFN, RNF, and AdhE*, will be necessary for the designed strain to overcome metabolic obstacles and reach the desired, high ethanol yielding phenotype.

It should be emphasized that even though the model-driven approach presented in this study is powerful for strain design and optimization, validating the design is not a trivial task to implement in *C. thermocellum*. The main reasons are that genetic tools (e.g., gene deletion and expression) for *C. thermocellum* are not well established and native complex genetics are not well understood compared to those of the model organisms (e.g., *E. coli* and *S. cerevisiae*) despite significant progress made in very recent years (Guss et al., 2012, Olson and Lynd 2012, Olson et al., 2015). It is certain that the development of the powerful genome editing tools like

CRISPR/Cas9 (Wang et al., 2015) and siRNA (Tummala et al., 2003, Na et al., 2013) for *C. thermocellum*, if successful, will be beneficial for metabolic engineering of this organism for production of fuels and high-value chemicals from lignocellulosic biomass. Even though our metabolic model is constructed for the *C. thermocellum* DSM1313, different from the iSR432 model developed for *C. thermocellum* ATCC 27405 (Gowen and Fong 2010, Roberts et al., 2010), the important features (e.g. the atypical glycolysis and a complete model of redox regulatory mechanisms) should be applicable to ATCC 27405, and hence their incorporation should improve the prediction of the iSR432 model.

In summary, our study accents the hypothesis that constricting electron flux contributes more than carbon redirection to ethanol production in *C. thermocellum*, and identifies bottlenecks in the redox metabolism hindering high ethanol yields. Rational strain design based on the MMF method and *in vivo* characterizations predict increasing flux through NFN, RNF, and/or AdhE* while eliminating acetate, lactate, H₂, and valine production is the most promising strategy for optimizing the production of ethanol. We envision that our trained metabolic model will be used to engineer *C. thermocellum* as a CBP platform to produce other reduced metabolites.

ACKNOWLEDGEMENTS

This research was supported under the grant (DE-AC05-000R22725) by the BioEnergy Science Center (BESC), a U.S. Department of Energy Bioenergy Research Center funded by the Office of Biological and Environmental Research in the DOE Office of Science. The authors would like to thank Dr. Joe Bozell (the Center of Renewable Carbon, UTK) for use of the GC/MS instrument and Dr. Gary Sayler (the Center of Environmental Biotechnology, UTK) for use of

663 the SRS UGA-300 Universal Gas Analyzer. We also thank Prof. Hans Van Dijken for useful
664 discussion.

LIST OF TABLES

Table 1. List of strains and plasmids used in this study. Abbreviations: V, *in vivo*; S, *in silico*; and HPP, sodium hypophosphite.

Table 2. Experimentally determined growth rates, metabolic fluxes, and yields per cellobiose for selected fermentative products for the parent, $\Delta hydG$, $\Delta hydG \Delta ech$, and $\Delta hydG \Delta pta-ack$ strains. Experimental data were obtained through pH controlled bioreactors

Table 3. Sensitivity analysis for select reactions

Table 4. Elementary mode analysis for various *C. thermocellum* mutants

677 **Table 1.** List of strains and plasmids used in this study. Abbreviations: V, *in vivo*; S, *in silico*;
678 and HPP, sodium hypophosphite.

679

Strain Name	Genotype / Description	Characterization Method	Source
<i>C. thermocellum</i>			
DSM 1313	Wildtype		Leibniz-Institut DSMZ
M1354	DSM1313 Δhpt (Parent strain)	V,S	(Argyros et al., 2011)
$\Delta hydG$	M1354 $\Delta hydG adhE^*(D494G)$	V,S	(Biswas et al., 2015)
$\Delta hydG \Delta ech$	M1354 $\Delta hydG \Delta hypABFCDE-echABCDEF adhE^*(D494G)$	V,S	(Biswas et al., 2015)
$\Delta hydG \Delta pta-ack$	M1354 $\Delta hydG \Delta pta-ack adhE^*(D494G)$	V,S	This study
$\Delta hydG \Delta ech \Delta pfl$	M1354 $\Delta hydG \Delta hypABFCDE-echABCDEF adhE^*(D494G) + HPP$	V,S	This study

680

681

Table 2. Experimentally determined growth rates, metabolic fluxes, and yields per cellobiose for selected fermentative products for the parent, $\Delta hydG$, $\Delta hydG \Delta ech$, and $\Delta hydG \Delta pta-ack$ strains. Experimental data were obtained through pH controlled bioreactors. n.d. – Not detected

Fermentation parameters	Parent	$\Delta hydG$	$\Delta hydG \Delta ech$	$\Delta hydG \Delta pta-ack$
μ (hr ⁻¹)	0.33±0.01	0.24±0.02	0.22±0.02	0.16±0.02
Ethanol yield (g/g)	0.16±0.01	0.19±0.01	0.32±0.04	0.25±0.01
Maximum ethanol titer (g/L)	0.63±0.03	0.79±0.01	1.27±0.05	1.01±0.07
Carbon balance (% Recovery)	90.5±1.6	79.3±1.6	89.9±9.2	83.5±4.1
Experimental flux (mmol species/gDCW/hr)				
$r_{CB \text{ up}}$	3.58±0.16	3.43±0.25	3.33±0.62	2.42±0.05
r_{ETH}	4.19±0.10	4.96±0.24	7.95±1.37	4.65±0.25
r_{ACE}	2.63±0.86	2.24±0.57	0.67±0.17	0.53±0.21
r_{CO_2}	6.82±0.01	7.25±0.11	6.80±1.01	5.29±0.50
r_{FOR}	1.77±0.01	1.22±0.60	3.01±0.56	0.74±0.06
r_{H_2}	7.86±0.38	9.11±0.84	0.58±0.06	3.17±0.84
r_{VAL}	0.78±0.12	0.35±0.01	0.31±0.01	0.55±0.07
r_{LAC}	0.18±0.01	n.d.	0.05±0.01	0.03±0.01

Table 3. One-at-a-time sensitivity (OATS) analysis for select reactions. *n/a* designates reactions which were removed prior to sensitivity analysis due to genotype.

Nodes of interest	Reactions	% Change in fitting error upon reaction removal			
		Parent	Δ_{hydG}	$\Delta_{hydG} \Delta_{ech}$	$\Delta_{hydG} \Delta_{pta-ack}$
→ PYR	MAE+ODC	0	0	0	0
	PPDK	0	0	0	0
PYR →	PFOR	228	163	583	288
	PFL	28	4	245	10
Electrons →	ECH	0	170	n/a	113
	Fe-H ₂	0	n/a	n/a	n/a
	BIF	0	n/a	n/a	n/a
	RNF	0	0	32	0
	NFN	0	0	0	0
Products →	Cell Growth	16	11	31	13
	ETOH_out	114	71	764	246
	ACE_out	286	176	226	123
	CO ₂ _out	309	221	717	374
	FOR_out	28	4	245	10
	H ₂ _out	267	170	n/a	113

Table 4. Elementary mode analysis for various *C. thermocellum* mutants

Network characteristics	Parent	Δ_{hydG}	$\Delta_{hydG} \Delta_{ech}$	$\Delta_{hydG} \Delta_{pta-ack}$
Total EMs	5202	2258	1148	816
ETH EMs	1982	1357	807	499
BIO EMs	3858	1478	718	558
Both EMs	1143	842	470	334
Range ETH (Cmol/Cmol)	0.000-0.667	0.000-0.667	0.000-0.667	0.000-0.667
Range BIO (Cmol/Cmol)	0.000-0.394	0.000-0.394	0.000-0.330	0.000-0.247

LIST OF FIGURES

Figure 1. The central metabolic network of *C. thermocellum*. Genetic modifications investigated in this study are marked with a red X. Reactions which contribute to metabolite nodes of interest are highlighted. Pyruvate phosphate dikinase (PPDK), oxaloacetate decarboxylase (ODC), and malic enzyme (MAE) contribute to $\rightarrow$ PYR (green box). Pyruvate formate lyase (PFL), pyruvate ferredoxin oxidoreductase (PFOR), and lactate dehydrogenase (LDH) contribute to PYR $\rightarrow$ (purple box). ACoA $\rightarrow$ (brown box) consists of alcohol dehydrogenase (AdhE), phosphotransacetylase (PTA), and citrate synthase (TCA1). The redox pathway is highlighted in the blue box. Fd_{rd} $\rightarrow$ consists of the Ni-Fe energy conserving hydrogenase (ECH), the Rnf-type NADH:Fd oxidoreductase (RNF), the NADH-Fd_{rd}:NADP⁺ oxidoreductase (NFN), and the NADH-Fd_{rd} bifurcating hydrogenase (BIF). The $\rightarrow$ H₂ node includes ECH, BIF, as well as the Fe-Fe NADPH-dependent hydrogenase (Fe-H₂)

Figure 2. A) A concentration profile for the major metabolic end-products of the parent strain in a representative batch fermentation. **B)** A representative carbon-balance applied throughout a parent strain batch fermentation. “Removed” indicates carbon removed during sampling while “other” consists of trace amino acids, mostly alanine and isoleucine. Abbreviations: CDW: cell dried weight; CB: cellobiose; LAC: lactate; FOR: formate; ACE: acetate; ETOH: ethanol; CO₂: carbon dioxide; H₂: hydrogen; and VAL: valine

Figure 3. A) Experimental yields of major fermentative products on cellobiose for different genotypes. **B)** Fraction of contribution of fluxes through specific metabolic nodes via METAFoR analysis for different genotypes. The groups are based on the metabolic nodes to which they contribute (Methods). Abbreviations: M+O, malic enzyme and oxaloacetate decarboxylase are collectively considered; PPK, pyruvate phosphate dikinase; ECH, Ni-Fe Ech-type hydrogenase; RNF, NADH:Fd oxidoreductase; NFN, NADH-Fd_{rd}:NADP⁺ oxidoreductase. BIF, bifurcating hydrogenase; Fe-H₂, NADPH dependent Fe-Fe hydrogenase; PFL, pyruvate formate lyase; PFOR, Pyruvate:Fd oxidoreductase; LDH, lactate dehydrogenase; TCA, citrate synthase (towards cell mass); PTA, acetate synthesis; and ADHE, ethanol synthesis. In both panels, student t-test significance compared to the parent strain is reported by asterisks, i.e., ‘*’ for *p*-value < 0.01, and ‘**’ for *p*-value < 0.05

Figure 4. Parametric plots of flux distributions between **(A)** the parent and $\Delta hydG$ strains, **(B)** the parent and $\Delta hydG \Delta ech$ strains, and **(C)** the parent and $\Delta hydG \Delta pta-ack$ strains. Reaction outliers are labeled as (1) NADPH-dependent adhE* activity, (2) RNF, (3) ethanol export, (4) PFL, (5) acetate export, (6) ECH, (7) hydrogen export, and (8) NFN

Figure 5. Growth characteristics of parent strain (triangles) and $\Delta hydG \Delta ech$ (circles) in MTC media (filled symbols) or MTC with the PFL inhibitor hypophosphite (open symbols). To investigate redox bottlenecks, no additional electron sink **(A)**, 20 mM fumarate **(B)**, 20 mM 2-ketoisovalerate **(C)**, or 2 g/L total sulfate **(D)** were included in the medium to probe NADH, NAD(P)H, and Fd_{rd}, respectively

Figure 6. The 2-D phenotypic spaces of ethanol versus cell mass yields on cellobiose for various genotypes. The phenotypic spaces (or convex hull) are encompassed by the areas ‘abcd’ for the parent and $\Delta hydG$ strains, ‘abefd’ for $\Delta hydG \Delta ech$ strain, and ‘abgd’ for $\Delta hydG \Delta pta-ack$ strain. Symbols within these areas are experimental yields

Figure 7. Application of MMF algorithm to identify the optimal genotype of *C. thermocellum* for high-yield ethanol production. A) Effect of reaction deletions on the total EMs, ethanol-producing EMs, cell mass-producing EMs, and coupled ethanol- and cell mass-producing EMs. B) Effect of reaction deletions on minimum and maximum yields of ethanol and cell mass on cellobiose.

Figure 1

Figure 2

A

B

Figure 3

Figure 4.

771 **Figure 5.**

772

773

774

775 **Figure 6.**

Figure 7.

REFERENCES

- Argyros, D. A., S. A. Tripathi, T. F. Barrett, S. R. Rogers, L. F. Feinberg, D. G. Olson, J. M. Foden, B. B. Miller, L. R. Lynd, D. A. Hogsett and N. C. Caiazza 2011. "High Ethanol Titters from Cellulose by Using Metabolically Engineered Thermophilic, Anaerobic Microbes." *Applied and Environmental Microbiology* 77(23): 8288-8294.
- Biswas, R., S. Prabhu, L. R. Lynd and A. M. Guss 2014. "Increase in Ethanol Yield via Elimination of Lactate Production in an Ethanol-Tolerant Mutant of *Clostridium thermocellum*." *PLoS ONE* 9(2): e86389.
- Biswas, R., T. Zheng, D. Olson, L. Lynd and A. Guss 2015. "Elimination of hydrogenase active site assembly blocks H₂ production and increases ethanol yield in *Clostridium thermocellum*." *Biotechnology for Biofuels* 8(1): 20.
- Brown, S. D., A. M. Guss, T. V. Karpinets, J. M. Parks, N. Smolin, S. Yang, M. L. Land, D. M. Klingeman, A. Bhandiwad, M. Rodriguez, Jr., B. Raman, X. Shao, J. R. Mielenz, J. C. Smith, M. Keller and L. R. Lynd 2011. "Mutant alcohol dehydrogenase leads to improved ethanol tolerance in *Clostridium thermocellum*." *Proceedings of the National Academy of Sciences of the United States of America* 108(33): 13752-13757.
- Carere, C., T. Rydzak, T. Verbeke, N. Cicek, D. Levin and R. Sparling 2012. "Linking genome content to biofuel production yields: a meta-analysis of major catabolic pathways among select H₂ and ethanol-producing bacteria." *BMC Microbiology* 12(1): 295.
- Demain, A. L. 2009. "Biosolutions to the energy problem." *Journal of Industrial Microbiology and Biotechnology* 36(3): 319-332.
- Deng, Y., D. G. Olson, J. Zhou, C. D. Herring, A. J. Shaw and L. R. Lynd 2013. "Redirecting carbon flux through exogenous pyruvate kinase to achieve high ethanol yields in *Clostridium thermocellum*." *Metabolic Engineering* 15: 151-158.
- Ellis, L. D., E. K. Holwerda, D. Hogsett, S. Rogers, X. Shao, T. Tschaplinski, P. Thorne and L. R. Lynd 2012. "Closing the carbon balance for fermentation by *Clostridium thermocellum* (ATCC 27405)." *Bioresource Technology* 103(1): 293-299.
- Feinberg, L., J. Foden, T. Barrett, K. W. Davenport, D. Bruce, C. Detter, R. Tapia, C. Han, A. Lapidus, S. Lucas, J.-F. Cheng, S. Pitluck, T. Woyke, N. Ivanova, N. Mikhailova, M. Land, L. Hauser, D. A. Argyros, L. Goodwin, D. Hogsett and N. Caiazza 2011. "Complete Genome Sequence of the Cellulolytic Thermophile *Clostridium thermocellum* DSM1313." *Journal of Bacteriology* 193(11): 2906-2907.
- Gowen, C. M. and S. S. Fong 2010. "Genome-scale metabolic model integrated with RNAseq data to identify metabolic states of *Clostridium thermocellum*." *Biotechnology Journal* 5(7): 759-767.
- Guss, A. M., D. G. Olson, N. C. Caiazza and L. R. Lynd 2012. "Dcm methylation is detrimental to plasmid transformation in *Clostridium thermocellum*." *Biotechnology for Biofuels* 5.
- Hädicke, O. and S. Klamt 2011. "Computing complex metabolic intervention strategies using constrained minimal cut sets." *Metabolic Engineering* 13(2): 204-213.
- Herrero, A. A. and R. F. Gomez 1980. "Development of ethanol tolerance in *Clostridium thermocellum*: effect of growth temperature." *Applied and Environmental Microbiology* 40(3): 571-577.
- Holwerda, E., P. Thorne, D. Olson, D. Amador-Noguez, N. Engle, T. Tschaplinski, J. van Dijken and L. Lynd 2014. "The exometabolome of *Clostridium thermocellum* reveals overflow metabolism at high cellulose loading." *Biotechnology for Biofuels* 7(1): 155.

- Kamp, A. v. and S. Schuster 2006. "Metatool 5.0: fast and flexible elementary modes analysis." *Bioinformatics* 22(15): 1930-1931.
- Kridelbaugh, D. M., J. Nelson, N. L. Engle, T. J. Tschaplinski and D. E. Graham 2013. "Nitrogen and sulfur requirements for *Clostridium thermocellum* and *Caldicellulosiruptor bescii* on cellulosic substrates in minimal nutrient media." *Bioresource Technology* 130(0): 125-135.
- Kuyper, M., M. J. Toirkens, J. A. Diderich, A. A. Winkler, J. P. van Dijken and J. T. Pronk 2005. "Evolutionary engineering of mixed-sugar utilization by a xylose-fermenting *Saccharomyces cerevisiae* strain." *FEMS Yeast Research* 5(10): 925-934.
- 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*." *Journal of Bacteriology* 144(2): 569-578.
- 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(2): 169-172.
- Lynd, L. R., W. H. Van Zyl, J. E. McBride and M. Laser 2005. "Consolidated bioprocessing of cellulosic biomass: an update." *Curr Opin Biotechnol* 16(5): 577-583.
- Mohr, G., W. Hong, J. Zhang, G.-z. Cui, Y. Yang, Q. Cui, Y.-j. Liu and A. M. Lambowitz 2013. "A Targetron System for Gene Targeting in Thermophiles and Its Application in *Clostridium thermocellum*." *PLoS ONE* 8(7): e69032.
- Mulder, David W., Eric M. Shepard, Jonathan E. Meuser, N. Joshi, Paul W. King, Matthew C. Posewitz, Joan B. Broderick and John W. Peters 2011. "Insights into [FeFe]-Hydrogenase Structure, Mechanism, and Maturation." *Structure* 19(8): 1038-1052.
- Na, D., S. M. Yoo, H. Chung, H. Park, J. H. Park and S. Y. Lee 2013. "Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs." *Nat Biotech* 31(2): 170-174.
- Neidhardt, F. C., J. L. Ingraham and M. Schaechter (1990). *Physiology of the bacterial cell: a molecular approach*, Sinauer Associates.
- Ng, T. K. and J. G. Zeikus 1982. "Differential metabolism of cellobiose and glucose by *Clostridium thermocellum* and *Clostridium thermohydrosulfuricum*." *Journal of Bacteriology* 150(3): 1391-1399.
- Nochur, S. V., G. R. Jacobson, M. F. Roberts and A. L. Demain 1992. "Mode of sugar phosphorylation in *Clostridium thermocellum*." *Applied Biochemistry and Biotechnology* 33(1): 33-41.
- Olson, D. G. and L. R. Lynd 2012. "Transformation of *Clostridium thermocellum* by electroporation." *Method Enzymol.*: 317-330.
- Olson, D. G., M. Maloney, A. A. Lanahan, S. Hon, L. J. Hauser and L. R. Lynd 2015. "Identifying promoters for gene expression in *Clostridium thermocellum*." *Metabolic Engineering Communications*(0).
- Ozkan, M., S. G. Desai, Y. Zhang, D. M. Stevenson, J. Beane, E. A. White, M. L. Guerinot and L. R. Lynd 2001. "Characterization of 13 newly isolated strains of anaerobic, cellulolytic, thermophilic bacteria." *Journal of Industrial Microbiology & Biotechnology* 27(5): 275.
- Pfeiffer, T., I. Sanchez-Valdenebro, J. C. Nuno, F. Montero and S. Schuster 1999. "METATOOL: for studying metabolic networks." *Bioinformatics* 15(3): 251-257.
- Poolman, M. G., K. V. Venkatesh, M. K. Pidcock and D. A. Fell 2004. "A method for the determination of flux in elementary modes, and its application to *Lactobacillus rhamnosus*." *Biotechnology and bioengineering* 88(5): 601-612.

- Ragauskas, A., H. Akinoshio, K. Yee and D. Close 2014. "The emergence of *Clostridium thermocellum* as a high utility candidate for consolidated bioprocessing applications." *Frontiers in Chemistry* 2.
- Riederer, A., T. E. Takasuka, S.-i. Makino, D. M. Stevenson, Y. V. Bukhman, N. L. Elsen and B. G. Fox 2011. "Global Gene Expression Patterns in *Clostridium thermocellum* as Determined by Microarray Analysis of Chemostat Cultures on Cellulose or Cellobiose." *Applied and Environmental Microbiology* 77(4): 1243-1253.
- Roberts, S., C. Gowen, J. P. Brooks and S. Fong 2010. "Genome-scale metabolic analysis of *Clostridium thermocellum* for bioethanol production." *BMC Systems Biology* 4(1): 31.
- Rydzak, T., M. Grigoryan, Z. Cunningham, O. Krokhin, P. Ezzati, N. Cicek, D. Levin, J. Wilkins and R. Sparling 2014. "Insights into electron flux through manipulation of fermentation conditions and assessment of protein expression profiles in *Clostridium thermocellum*." *Applied Microbiology and Biotechnology*: 1-14.
- Rydzak, T., D. Levin, N. Cicek and R. Sparling 2009. "Growth phase-dependant enzyme profile of pyruvate catabolism and end-product formation in *Clostridium thermocellum* ATCC 27405." *J Biotechnol* 140(3-4): 169 - 175.
- Rydzak, T., D. B. Levin, N. Cicek and R. Sparling 2011. "End-product induced metabolic shifts in *Clostridium thermocellum* ATCC 27405." *Appl Microbiol Biotechnol* 92(1): 199-209.
- Rydzak, T., P. D. McQueen, O. V. Krokhin, V. Spicer, P. Ezzati, R. C. Dwivedi, D. Shamshurin, D. B. Levin, J. A. Wilkins and R. Sparling 2012. "Proteomic analysis of *Clostridium thermocellum* core metabolism: relative protein expression profiles and growth phase-dependent changes in protein expression." *BMC Microbiol* 12: 214.
- Sauer, U., D. R. Lasko, J. Fiaux, M. Hochuli, R. Glaser, T. Szyperski, K. Wüthrich and J. E. Bailey 1999. "Metabolic Flux Ratio Analysis of Genetic and Environmental Modulations of *Escherichia coli* Central Carbon Metabolism." *Journal of Bacteriology* 181(21): 6679-6688.
- Shao, X., B. Raman, M. Zhu, J. Mielenz, S. Brown, A. Guss and L. Lynd 2011. "Mutant selection and phenotypic and genetic characterization of ethanol-tolerant strains of *Clostridium thermocellum*." *Applied Microbiology and Biotechnology* 92(3): 641-652.
- Stouthamer, A. H. 1973. "A theoretical study on the amount of ATP required for synthesis of microbial cell material." *Antonie van Leeuwenhoek* 39(1): 545-565.
- Szyperski, T., R. W. Glaser, M. Hochuli, J. Fiaux, U. Sauer, J. E. Bailey and K. Wuthrich 1999. "Bioreaction Network Topology and Metabolic Flux Ratio Analysis by Biosynthetic Fractional ¹³C Labeling and Two-Dimensional NMR Spectroscopy." *Metabolic engineering* 1(3): 189-197.
- Trinh, C. T., P. Unrean and F. Sreenc 2008. "Minimal *Escherichia coli* cell for the most efficient production of ethanol from hexoses and pentoses." *Applied and Environmental Microbiology* 74: 3634-3643.
- Trinh, C. T., A. Wlaschin and F. Sreenc 2009. "Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism." *Applied Microbiology and Biotechnology* 81(5): 813-826.
- 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.* 76(19): 6591-6599.

- 921 Tummala, S. B., N. E. Welker and E. T. Papoutsakis 2003. "Design of antisense RNA constructs
922 for downregulation of the acetone formation pathway of *Clostridium acetobutylicum*."
923 Journal of Bacteriology 185(6): 1923-1934.
- 924 Unkrig, V., F. A. Neugebauer and J. Knappe 1989. "The free radical of pyruvate formate-lyase."
925 European Journal of Biochemistry 184(3): 723-728.
- 926 van der Veen, D., J. Lo, S. D. Brown, C. M. Johnson, T. J. Tschaplinski, M. Martin, N. L. Engle,
927 R. A. van den Berg, A. D. Argyros, N. C. Caiazza, A. M. Guss and L. R. Lynd 2013.
928 "Characterization of *Clostridium thermocellum* strains with disrupted fermentation end-
929 product pathways." Journal of Industrial Microbiology & Biotechnology 40(7): 725-734.
- 930 van Zyl, W. H., L. R. Lynd, R. den Haan and J. E. McBride 2007. "Consolidated bioprocessing
931 for bioethanol production using *Saccharomyces cerevisiae*." Advances in Biochemical
932 Engineering/Biotechnology 108(Journal Article): 205-235.
- 933 Wang, Y., Z.-T. Zhang, S.-O. Seo, K. Choi, T. Lu, Y.-S. Jin and H. P. Blaschek 2015.
934 "Markerless chromosomal gene deletion in *Clostridium beijerinckii* using CRISPR/Cas9
935 system." Journal of Biotechnology 200: 1-5.
- 936 Williams, T., J. Combs, B. Lynn and H. Strobel 2007. "Proteomic profile changes in membranes
937 of ethanol-tolerant *Clostridium thermocellum*." Applied Microbiology and Biotechnology
938 74(2): 422-432.
- 939 Yomano, L. P., S. W. York and L. O. Ingram 1998. "Isolation and characterization of ethanol-
940 tolerant mutants of *Escherichia coli* KO11 for fuel ethanol production." Journal of
941 industrial microbiology & biotechnology 20(2): 132-138.
- 942 Zhang, Y. and L. Lynd 2003. "Quantification of Cell and Cellulase Mass Concentrations during
943 Anaerobic Cellulose Fermentation: Development of an Enzyme-Linked Immunosorbent
944 Assay-Based Method with Application to *Clostridium thermocellum* Batch Cultures."
945 Anal Chem 75(2): 219 - 227.
- 946 Zhou, J., D. G. Olson, D. A. Argyros, Y. Deng, W. M. van Gulik, J. P. van Dijken and L. R.
947 Lynd 2013. "Atypical Glycolysis in *Clostridium thermocellum*." Applied and
948 Environmental Microbiology 79(9): 3000-3008.