Analysis of metabolic pathways and fluxes in a newly discovered thermophilic and ethanol-tolerant *Geobacillus* strain

Yinjie J. Tang[1,8], Rajat Sapra[2,4], Dominique Joyner[1,3] , Terry C. Hazen[1,3], Samuel Myers[5], David Reichmuth[4], Harvey Blanch[2,5,7], and Jay D. Keasling[1,2,5,6,7]

Running title: *flux analysis of an ethanol tolerant thermophile*

(1) Virtual Institute for Microbial Stress and Survival

(2) Joint Bio-Energy Institute, Emeryville, CA 94608

(3) Ecology Department, Lawrence Berkeley National Lab, Berkeley, 94720,

(4) Sandia National Laboratories, PO Box 969, Livermore, CA 94551-9951,

(5) Department of Chemical Engineering, University of California, Berkeley, CA 94720

(6) Department of Bioengineering, University of California, Berkeley, CA 94720

(7) Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

(8) Department of Energy, Environmental and Chemical Engineering, Washington University, St Louis, MO63130

Abstract

A recently discovered thermophilic bacterium, *Geobacillus thermoglucosidasius* M10EXG, ferments a range of C5 (e.g., xylose) and C6 sugars (e.g., glucose) and is tolerant to high ethanol concentrations (10% v/v). We have investigated the central metabolism of this bacterium using both *in vitro* enzyme assays and ^{13}C-based flux analysis to provide insights into the physiological properties of this extremophile and explore its metabolism for bioethanol or other bioprocess applications. Our findings show that glucose metabolism in *G. thermoglucosidasius* M10EXG proceeds via glycolysis, the pentose phosphate pathway, and the TCA cycle; the Entner-Doudoroff pathway and transhydrogenase activity were not detected. Anaplerotic reactions (including the glyoxylate shunt, pyruvate carboxylase and phospho*enol*pyruvate carboxykinase) were active, but fluxes through those pathways could not be accurately determined using amino acid labelling. When growth conditions were switched from aerobic to micro-aerobic conditions, fluxes (based on a normalized glucose uptake rate of 100 units (gm DCW)$^{-1}$·hr^{-1}) through the TCA cycle and oxidative pentose phosphate pathway were reduced from 64±3 to 25±2 and from 30±2 to 19±2, respectively. The carbon flux under microaerobic growth was directed to ethanol, L-lactate (>99% optical purity), acetate, and formate. Under fully anaerobic conditions, *G. thermoglucosidasius* M10EXG used a mixed acid fermentation process and exhibited a maximum ethanol yield of 0.38±0.07 mol mol^{-1} glucose. *In silico* flux balance modelling demonstrates that lactate and acetate production from *G. thermoglucosidasius* M10EXG reduces the maximum ethanol yield by approximately three folds, thus indicating that both pathways should be modified to maximize ethanol production.

Key words: C5 sugar, micro-aerobic, TCA cycle, anaplerotic pathway, flux balance model

Introduction

A recently discovered thermophilic ethanologen, *Geobacillus thermoglucosidasius* M10EXG (M10EXG), is a facultative anaerobe that has an optimal growth temperature of 60°C (Fong et al. 2006). It can ferment a range of C5 and C6 sugars and tolerate ethanol concentrations of up to 10% (v/v) (Fong et al. 2006), which makes it an ideal microbe for improved bio-ethanol production. Moreover, *Geobacillus* species have many other potential industrial applications for production of various thermostable enzymes, exopolysaccharides and bacteriocins; they have also been found to metabolize hydrocarbons in high temperature oil fields (Nazina et al. 2005) as well as degrading herbicides (such as organophosphonates) (McMullan et al. 2004). However, the genome sequence of this newly discovered *Geobacillus* strain is not yet available, and as such there is no functional genomics data. In order to engineer the metabolic pathways of the bacterium for optimizing ethanol production from C5 and C6 sugars, an understanding of the carbon fluxes and the maximum potential for ethanol production is required. In this study, we used *in vitro* enzyme activity assays and a ^{13}C-based isotopomer flux model to investigate central metabolic pathways of this thermophilic organism as a function of oxygen availability (Stephanopoulos et al. 1998; Tang et al. 2007a; Tang et al. 2007b; Wiechert et al. 2001). To accomplish this, cells were grown in minimal medium containing either [1-^{13}C] or [2-^{13}C]-labelled glucose as the sole carbon source, and the ^{13}C-labeling patterns of derivatized, intracellular amino acids were determined using gas chromatography-mass spectrometry (GC-MS). An isotopomer model was then constructed to simulate all of the atom transitions in the assumed biochemical network (based on both enzyme activity assay and metabolic pathways of its closest sequenced species *G. kaustophilus* (Takami et al. 2004)) and the label distribution in all central metabolites. We then searched for a set of active intracellular metabolic pathways and flux distributions that predicted the inferred isotopomer distribution of key metabolites resulting from the isotopomer pattern of the derivatized amino acids. We show that in the absence of genome information our approach provides an effective way to map the central pathways of a new fermentative organism (Tang et al. 2007e) and directly observe the functional output (i.e., metabolic fluxes) of the transcriptome, proteome, and metabolic changes under different growth conditions (Sauer 2004).

Materials and Methods

Culture conditions. M10EXG was obtained from the *Bacillus* Genetic Stock Center at the Ohio State University (Cat # W9A44). A complete minimal medium was used (Fong et al. 2006) for the cell culture under defined conditions. Since singly labeled carbon substrate can well resolve the Entner-Doudoroff pathway and pentose phosphate pathway, [1-^{13}C] D-glucose (10 g L^{-1}; >98%; Cambridge Isotope, Andover, MA) was used as the sole carbon source (Fischer et al. 2004). For anaerobic or micro-aerobic experiments, the cultures were incubated in sealed glass bottles with septum caps, and the headspace was filled with argon (anaerobic conditions) or air (micro-aerobic conditions, air liquid volume ratio 1:1). For aerobic cell cultures, cells were incubated in shake flasks at 200 rpm. All cultures with labeled medium were started with a ~3% inoculation volume from cells that had been first grown in Tryptic Soy Broth (BD Biosciences, San Jose, CA) to stationary phase and then sub-cultured into minimal medium with ~3% inoculation volume to remove the effect of naturally labelled carbon sources from the initial inocula. All cultures (aerobic, microaerobic, and anaerobic) were incubated at 60°C with shaking at 200 rpm. Total biomass growth was monitored by measuring the optical density at a wavelength of 600 nm (OD_{600}).

Enzyme Assays. Exponentially growing cells were centrifuged and the resulting cell pellets were resuspended in 1 ml 100 mM Tris-HCl pH 7.4 and lysed by sonication for all enzyme assays. Total protein concentration for cell lysates was determined using the Bradford method (Bio-Rad, Hercules, CA) with bovine serum albumin as the standard. All chemicals and coupling enzymes were purchased from Sigma Chemical (St. Louis, MO). All enzyme assays were performed at 55°C and monitored spectroscopically at their respective wavelengths. All enzyme assays were performed as previously reported (McKinlay et al. 2007; Sauer et al. 2004; Terada et al. 1991; Van der Werf et al. 1997). In brief, an isocitrate lyase assay contained 25 mM imidazole pH 6.8, 5 mM $MgCl_2$, 1 mM EDTA, 4 mM phenylhydrazine, and 1 mM D-L-isocitrate; the absorbance at 324 nm was used to monitor forming osazone derivatives. The oxaloacetate decarboxylase assay contained 41 mM triethanolamine-HCl (TEA) pH 8.0, 460 μM $MnCl_2$, 300 μM β-NADH, 11 units mL^{-1} of lactate dehydrogenase, and 2.3 mM oxaloacetate; the absorbance at 340 nm was used to monitor β-NADH oxidation. The α-ketoglutarate dehydrogenase assay contained 50 mM MOPS pH 7.4, 4 mM $MgCl_2$, 200 μM $CaCl_2$, 6 mM thiamine pyrophosphate, 6.7 mM β-NAD^+, 5.2 mM cysteine-HCl, and 25 mM α-ketoglutarate; the absorbance at 340 nm was used to monitor the increase in β-NADH. The phosphoenolpyruvate (PEP) carboxylase assay contained 100 mM Tris-acetate pH 8.5, 2 mM potassium PEP, 10 mM $KHCO_3$, 10 mM magnesium acetate, and 1.17 M dioxane; the absorbance at 340 nm was used to monitor β-NADH oxidation. The transhydrogenase assay contained 50 mM Tris-HCl, pH 7.6, 2 mM $MgCl_2$, 500 μM β-NADH and 1 mM 3-acetylpyridine adenine dinucleotide ($APAD^+$); the absorbance at 375 nm was used to monitor the loss of $APAD^+$. The PEP carboxykinase assay contained 25 mM HEPES, pH 7.1, 50 mM KCl, 2 mM $MgCl_2$, 50 mM $NaHCO_3$, 500 μM dithiothreitol, 20 μM β-NADH, 100 μM ADP-Mg, 5 mM glucose, 4 units mL^{-1} of malate dehydrogenase, 4 units mL^{-1} of hexokinase and 1 mM PEP; the absorbance at 340 nm was used to monitor β-NADH oxidation. The malic enzyme assay contained 67 mM TEA pH 7.4, 3.5 mM malic acid, 333 μM $NAD(P)^+$, and 5 mM $MnCl_2$; the absorbance at 340 nm was used to monitor NAD(P)H oxidation. The pyruvate carboxylase assay contained 95 mM TEA pH 8.0, 6.3 mM pyruvate, 0.11% BSA, 26 units mL^{-1} malate dehydrogenase, 50 μM acetyl CoA, 240 μM β-NADH, 15 mM $KHCO_3$, and 1 mM ATP; the absorbance at 340 nm was used to monitor β-NADH oxidation. All activity calculations had the basal reaction rate subtracts and were normalized for amount of protein added to the assay (Table 1).

Analytical methods for metabolite concentrations, biomass composition, and isotopomer labelling. The concentrations of glucose, formate, lactate, acetate, succinate, and ethanol in the medium were measured using enzyme linked assay kits (r-Biopharm, Darmstadt, Germany). Biomass constituents (proteinogenic amino acid composition) were measured by the Molecular Structure Center, University of California, Davis; the fatty acids of M10EXG were measured by Microbial ID (Newark, DE) (Supplementary Table S1). Most fatty acids were saturated in M10EXG (a fact that would lead to decreased membrane fluidity and allow cell growth at high temperatures and high ethanol concentrations (Daron 1970; Sullivan et al. 1979)), and the 16- and 17-carbon fatty acids (including branched-chain iso- and anteiso-) accounted for ~80% of total fatty acids. The weight fractions of the various macromolecules were assumed to be same as a typical bacterium: protein (52%), RNA (16%), DNA (3%), lipids (9%), and total carbohydrate (17%) (Stephanopoulos et al. 1998). The biomass constitute information was used to give the estimation of range for searching the optimal fluxes to biomass pools.

The GC-MS protocol for isotopomer measurement has been reported previously (Tang et al. 2007d). In brief, protein in cell pellets (from 50 ml culture) was hydrolyzed in 6 M HCl at 100°C for 24 hours. The resulting amino acid mixture was derivatized in 100 µL tetrahydrofuran (THF) and 100 µl N-(tert-butyldimethylsilyl)-N-methyl-trifluoroacetamide (Sigma-Aldrich, St. Louis, MO) at 70°C for 1 hour and analyzed using a gas chromatograph (Model 6890, Agilent, Wilmington, DE) equipped with a DB5 column and a mass spectrometer (Model 5973 Network, Agilent, Wilmington, Delaware). Two types of positively charged species were used in the model simulation: unfragmented amino acids, $[M\text{-}57]^+$, and fragmented amino acids that have lost their α carboxyl group, $[M\text{-}159]^+$. The natural abundance of isotopes was corrected using a published algorithm before using the data for calculating the label distribution (Hellerstein and Neese 1999).

Assumptions employed in isotopomer modelling. The development of the isotopomer model was based on six assumptions. (i) A quasi-steady state is assumed to be achieved using batch culture as a convenient and less expensive approach (Sauer et al. 1999; Tang et al. 2007b). This assumption is based on the fact that isotopic patterns of 14 proteinogenic amino acids did not change (less than 1% difference) during the exponential growth phase (OD_{600} 0.4-0.9). (ii) The central metabolic network in M10EXG was inferred from the pathways in closely-related *Geobacillus kaustophilus* (only sequenced *Geobacillus* species) and *Bacillus subtilis* (Christiansen et al. 2002; Sauer et al. 1997). (iii) The direction of flux was based on reaction thermodynamics, as suggested by a previous flux study on *Bacillus subtilis* (Sauer et al. 1997); considerations of potential reversibility of each reaction would make the model system highly underdetermined (Zhao and Shimizu 2003), thus only reactions between PEP and oxaloacetate were assumed reversible. (iv) Amino acids provide isotopomer information unique to their precursors in the central metabolic pathways. To avoid possible inaccuracies resulting from alternative amino acid biosynthesis routes, seven amino acids were used to determine ^{13}C fluxes in central metabolism (Supplementary Table S2). These corresponding metabolites and their amino acid precursors were pyruvate (alanine), acetyl-CoA (leucine), oxaloacetate (aspartic acid), 2-oxo-glutarate (glutamate), phospho*enol*pyruvate & erythrose-4-phosphate (phenylalanine), and 3-phosphoglycerate (serine and glycine) (Sauer et al. 1997). (v) The pathways included in the model were the tricarboxylic acid (TCA) cycle, pentose phosphate (PP) pathway, the Entner-Doudoroff pathway, and anaplerotic reactions. Oxalacetate decarboxylase and malic enzyme were not included in the model because of no activity was observed from *in vitro* enzyme assays (Table 1). (vi) The fluxes PEP←→OAA and Pyr←→OAA could not be clearly distinguished via isotopomer labelling; therefore, PEP and pyruvate were treated as a single metabolite pool and the reactions between PEP/Pyr pool and OAA were assumed reversible.

Algorithm for ^{13}C based flux calculation. The extra-cellular fluxes (production of formate, lactate, acetate and ethanol) were measured directly, and fluxes to biomass pools were calculated based on the biomass composition (Supplementary Table S1). These fluxes were used as inputs to the isotopomer model and tightly constrained within measurement noise. The remaining unknown fluxes were determined from isotopomer fractions, to identify the operative intracellular metabolic reactions as described before (Tang et al. 2007b). In brief, the complete fluxes were solved using the reaction stoichiometry and atom / isotopomer mapping matrices in an iterative scheme to obtain the steady-state isotopomer distributions in the intracellular metabolites pools. To avoid getting trapped in a local minimum, the model applied a grid search strategy (Antoniewicz et al. 2006): with the glucose uptake rate under three oxygen conditions normalized to a value of 100 units (gm DCW)$^{-1}$ hr^{-1}, the model exhaustively searched all combinations of independent variables (metabolite fluxes). Since the pyruvate shunt reaction (pyruvate+$CO_2 \rightarrow$oxaloacetate via pyruvate carboxylase) consumes CO_2 from the medium, the fraction of $^{13}CO_2$ in the medium was also estimated. The step size of the grid search algorithm was 1 (normalized to the glucose uptake rate of 100 units (gm DCW)$^{-1}$hr^{-1}) for unknown fluxes and 0.01 for the $^{13}CO_2$ fraction. All possible flux combinations were searched to determine the global minima of the objective function (Tang et al. 2007b):

$$\varepsilon(v_n) = \sum_{i=1}^{a} \left(\frac{M_i - N_i(v_n)}{\delta_i} \right)^2$$

where v_n are the unknown fluxes to be optimized in the program, M_i are the measured MS data, N_i are the corresponding model-simulated MS data, and δ_i are the corresponding standard deviations in the GC-MS data (1~2%). The unknown metabolic fluxes were calculated to minimize ε. To estimate the confidence interval for calculated fluxes, a Monte Carlo approach was employed (Zhao and Shimizu 2003). In brief, 20 isotopomer concentration data sets were generated by addition of normally distributed measurement noise to actual measurement data. The same optimization routine was used to estimate the best-fit flux distribution from these data sets. Confidence limits for each flux value were obtained from the probability distribution of calculated flux resulting from the simulated data sets. The model program was developed using MATLAB 7.0 (The Mathworks, Natick, MA). The calculations were carried out using a Quad-Process Server (Finetec, San Jose, CA) at the Lawrence Berkeley National Laboratory.

***In silico* flux balance analysis (FBA).** FBA was used to estimate ethanol production potential by M10EXG and prioritize the pathways for genetic engineering. Since the genome sequence and functional metabolic pathway information of M10EXG were not available, *in silico* modelling was constructed using the Simpheny platform (Genomatica, San Diego, CA) to coarsely predict the M10EXG metabolic network (Mahadevan et al. 2006), with the following modifications: 1) two unique reactions found in *Geobacillus* strains (L-lactate dehydrogenase and pyruvate carboxylase) were added to the model; 2) The biomass composition for the M10EXG model is given in Supplementary Table S1. The model included ~1075 reactions and ~760 constraints, and the flux calculation algorithm relied on implementing a series of physicochemical constraints, including thermodynamic directionality, enzymatic capacity constraints, and reaction stoichiometry constraints (Edwards and Palsson 2000a). Since the number of reactions is much greater than the number of metabolites, the system requires the assumption of an objective function for the *in silico* flux balance analysis, i.e, maximizing cell growth or ethanol production (Stephanopoulos et al. 1998).

Results and Discussion

Growth kinetics and cellular metabolites under various oxygen conditions. Since *Geobacillus thermoglucosidasius* M10EXG grows in a minimal glucose medium between 55-65ºC (its optimal growth temperature is 60ºC) with glucose as the sole carbon source, the bacterium contains complete biosynthesis pathways for all amino acids and other essential metabolites. The average doubling time in the exponential phase in minimal glucose medium was two hours under aerobic growth conditions and 3.5 hours under micro-aerobic conditions (Figure 1), i.e., the doubling time at high temperature was not faster than that of mesophilic bacteria (such as *E. coli*) under similar batch conditions (Shaikh et al. 2008). Under aerobic conditions, M10EXG produced ~0.64 mol acetate mol glucose $^{-1}$ (Table 2). Acetate accumulation in the medium can inhibit cell growth, especially for mesophilic bacteria (Lynd 1989), due to change in the intracellular pH or inhibition of activities of key enzymes in central metabolism (Luli and Strohl 1990; Tang et al. 2007c). The production of acetate by aerobically growing bacteria is often observed when the carbon source is in excess, so bacteria can regulate acetyl-CoA consumption rate and quickly generate ATP when the activity of key TCA cycle enzymes (e.g., citrate synthase) are inhibited (Majewski and Domach 1990). Under microaerobic or anaerobic conditions, the cells secreted lactate, ethanol, and formate, in addition to acetate (Table 2). Under completely anaerobic conditions, the L-lactate (>99% optical purity) production increased to 0.89 mol L-lactate mol^{-1} glucose, and the formate yield was ~1 mol mol^{-1} glucose (i.e., pyruvate-formate lyase replaced pyruvate dehydrogenase for acetyl-CoA production), while the molar yield of ethanol was ~ 0.38 mol mol^{-1} glucose. The large amounts of acids produced reduced the pH in the medium from 7.6 to ~5, and the cells entered the death phase (lysed) under micro-aerobic condition after 20 hrs (i.e., the OD_{600} dropped after 20 hrs). Finally, M10EXG can also grow in xylose minimal medium and exhibits similar growth kinetics under different oxygen conditions (Supplementary Table S3), which makes this microorganism an ideal candidate for bio-ethanol production from lignocellulosic biomass (contain up to 40% C5 sugars). However, M10EXG mainly utilized glucose as carbon source when both glucose (C6) and xylose (C5) were available (ratio 1:1) (Supplementary Table S4). This result indicates that the presence of glucose may strongly inhibit xylose metabolism.

^{13}C-based flux analysis of intra-cellular pathways under aerobic and micro-aerobic conditions. Isotopomer flux models were developed based on assumed central metabolic pathways to optimally fit all isotopomer data. The optimal flux distributions (based on a normalized glucose uptake rate of 100 units (gm DCW)$^{-1}$ hr^{-1}) and the confidence intervals of seven key intracellular fluxes (including glycolysis, PP pathway, TCA cycle and anapleurotic pathways) under both aerobic and micro-aerobic conditions are shown in Figure 2. The flux distribution results indicate that oxygen concentrations strongly affected the metabolic fluxes through the central pathways. Under aerobic conditions, approximately two-thirds of the glucose flowed through glycolysis (relative flux = 69) and the remainder through the pentose phosphate pathway (flux = 30), while the flux through citrate synthase (into the TCA cycle) was 64. Under micro-aerobic conditions, growth was slower (0.20 hr^{-1}) (Table 2), and the fluxes through TCA cycle and PP pathway (G6P→6PG) were reduced to 25 and 19, respectively. *In vitro* assays showed no evidence for transhydrogenase activity under our experimental conditions (Table 1). Flux through the PP pathway was sensitive to growth rate, most likely because NADPH to support biomass synthesis is mainly from PP pathway (Christiansen et al. 2002). On the other hand, *B. subtilis* showed much higher PP pathway flux (~70) that varied less with specific growth rate (Sauer et al. 1997). This is because *B. subtilis* can convert excess NADPH from the PP pathway to NADH via the transhydrogenase reaction and as such has greater flexibility in balancing redox (Dauner et al. 2001; Sauer et al. 1997).

The flux results indicate that there is no Entner-Doudneroff (ED) pathway activity under either experimental condition, which is consistent with the lack of a phosphogluconate dehydratase (a key ED pathway enzyme) in the annotated *Geobacillus kaustophilus* genome (Alm et al. 2005) and with the fact that *Bacillus* species do not use the ED pathway (Goldman and Blumenthal 1963). Several anapleurotic reactions in M10EXG (as inferred from the *Geobacillus kaustophilus* genome annotation) were present based on *in vitro* enzyme assays (Table 1): Pyruvate→OAA (pyruvate shunt via pyruvate carboxylase), OAA→PEP (via PEP carboxykinase), and PEP→OAA (via PEP carboxylase). The flux results indicate that these anapleurotic reactions were down-regulated under micro-aerobic conditions: the OAA→PEP / Pyruvate flux declined from 54 to 37, and the combined flux from PEP and pyruvate to OAA declined from 44 to 24. Those non-biomass-related anapleurotic reactions may provide M10EXG central metabolism with flexibility to cope with various growth conditions (Tang et al. 2007b). The glyoxylate shunt, which reduces carbon flow through the oxidative branch of the TCA cycle (coupled with other anapleurotic pathways) and provides an alternative route for acetyl-CoA metabolism, was also measurable under aerobic and micro-aerobic conditions in M10EXG (Fischer and Sauer 2003).

Analysis of anaerobic pathways of M10EXG. Under anaerobic conditions, M10EXG disposed part of the glucose through lactic acid and formic acid production. The formic acid yield (~1 mol mole^{-1} of glucose consumed) was approximately equal to the sum of the acetate and ethanol yields, indicating that acetyl-CoA is a precursor to ethanol (via alcohol dehydrogenase) and acetate under mixed acid fermentation (Figure 3). Based on the metabolite measurements and the reported mixed acid fermentation pathway of *B. subtilis* (Cruz Ramos et al. 2000), a simplified anaerobic pathway is proposed (Figure 3). Under anaerobic conditions, carbon flux between glycolysis and the PP pathway can be directly calculated based on the labelling information, because the flux ratio between the two pathways is reflected in the labelling pattern of 3-phosphoglycerate (inferred from serine) and pyruvate (inferred from alanine) (Sauer et al. 1997). Meanwhile, the *in vitro* activity of α-ketoglutarate dehydrogenase was one order magnitude lower than that measured when oxygen was available (Table 1). The data showed that the enzymes of the TCA cycle were significantly repressed under anaerobic conditions (Table 1) and that the TCA cycle was mainly used for biosynthesis. The major carbon fluxes were directed towards mixed acids and ethanol production, which could be directly measured. Eight key fluxes were shown in Figure 3: the flux was through glucose-6-phosphate dehydrogenase (entrance to the oxidative branch of the PP pathway) under anaerobic conditions (flux=15); pyruvate was converted to lactate via L-lactate dehydrogenase (flux=89) or to acetyl-

CoA and formic acid via pyruvate formate lyase (flux=103); acetyl-CoA was mainly used for ethanol (flux=38) and acetate (flux=61) production. Compared to micro-aerobic conditions, formic acid production was eight-fold higher under anaerobic conditions, suggesting that pyruvate formate-lyase (PFL) was induced under anaerobic conditions (note: FNR, a transcriptional regulator to mediate PFL gene (Sawers and Suppmann 1992) is annotated in *Geobacillus kaustophilus*). Under anaerobic conditions, M10EXG generated NADH primarily from glycolysis (glyceraldehyde-3-phosphate dehydrogenase) (flux = 190 ± 4) and consumed NADH mainly for lactate, ethanol, and acetate production (flux = 165 ± 7).

Verification of the isotopomer flux model. The isotopomer-based flux analysis used herein is based on the labelling pattern of [^{13}C]-amino acids to infer the [^{13}C] labelling pattern of key metabolic intermediates. By tracing the path of ^{13}C from singly-labelled carbon substrate to those metabolites in the pathway network, an isotopomer model can predict the carbon flux distribution through central metabolism. To check the reliability of the flux analysis results, a Monte Carlo method was used to calculate confidence intervals of key intracellular fluxes to estimate uncertainty from measurement noise and experimental variation (as illustrated in Materials and Method section). The obtained confidence intervals for seven key intracellular pathways in Figure 2 showed that reaction G6P→6PG was best determined (confidence intervals for both aerobic and anaerobic conditions are within ±2), since 1-[^{13}C]-glucose was good for differentiating the reactions of the PP pathway from glycolysis (Fischer et al. 2004). However, fluxes of anapleurotic reactions between the pyruvate/PEP pool and OAA pool had the highest errors; for example, the confidence interval of OAA→PEP flux under aerobic conditions was ±38. This result indicated that the isotopomer data were not sufficient to constrain these two fluxes accurately. To further validate the calculated intracellular flux distribution, [2-^{13}C] glucose was used as the carbon source and the labeling of resulting key amino acids was used to estimate the flux distribution under both aerobic and micro-aerobic conditions. The results from [2-^{13}C] glucose experiments were qualitatively consistent with the results from [1-^{13}C] glucose experiments (Supplementary Figure S1). Fluxes through reactions of the PP pathway, glycolysis, and the TCA cycle (via citrate synthase) were very similar (difference <5). In contrast, the measured fluxes for PEP carboxylase, pyruvate shunt, and the glyoxylate shunt from the experiments using [2-^{13}C] glucose had larger differences (up to 10) compared to the measured fluxes from experiments using [1-^{13}C] glucose. Errors in calculated fluxes may arise from several sources: 1) the isotopomer information may be insufficient to constrain certain anapleurotic reactions very accurately; 2) measurement uncertainty of extracellular metabolites in batch cultures; 3) protein degradation and reincorporation of metabolites from catabolized amino acids into metabolic intermediates.

***In silico* analyses of metabolic network for ethanol production.** Since genetic engineering of thermophilic bacteria is very difficult, it will, therefore, be beneficial to know which reactions are the most important targets for genetic manipulation to improve ethanol production. As such, an *in silico* flux balance analysis (FBA) was performed to coarsely predict the optimal cellular metabolism for ethanol production via Simpheny Software from Genomatica (San Diego, CA) (Mahadevan et al. 2006). The FBA model did not require isotopomer information or detailed kinetic parameters for individual metabolic reactions (Edwards and Palsson 2000a; Edwards and Palsson 2000b). Although these models are underdetermined and may not reflect the actual metabolic flux distribution if typical objective function is assumed without any additional constraints (e.g., maximum biomass production, Supplementary Figure S2), they have proven to be a useful tool to provide important guidelines to explore the target pathways for genetic engineering (Stephanopoulos et al. 1998). The theoretical maximum yields of acetate, ethanol, lactate, formate, and biomass, as well as the theoretical maximum growth rate (assuming an average carbon substrate uptake rate of 5 mM glucose (gm DCW)$^{-1}$ hr^{-1}), were estimated using the Simpheny model. The predicted maximum yields of the metabolites and biomass were much higher than the corresponding measured yields (Table 2). A plot of the theoretical maximum ethanol production rate as the function of lactate and acetate production rates for two growth rates (0.1 hr^{-1} and 0.2 hr^{-1}), with or without formate production, indicates that mixed acid production or a high growth rate significantly reduce the ethanol production rate (Figure 4), because the mixed acid fermentation and biomass growth compete for the precursors and reducing power (NADH) with ethanol production. Lactate production (by L-lactate dehydrogenase) has the largest impact on ethanol yield followed by acetate production (acetate kinase and phosphotransacetylase), while formate production (by pyruvate-formate lyase) has the least impact on ethanol yield.

When growth rate maximization was used as the objective function, model results (Figure 5, the three dotted arrows linking the measured ethanol fluxes with the corresponding measured growth rates) indicated that ethanol production by M10EXG was much lower than the theoretical value. Meanwhile, the growth rate and all fluxes through reactions of the PP pathway and TCA cycle declined when more ethanol production was specified (Figure 5). The TCA cycle and oxidative PP pathway appeared to be the most sensitive to ethanol production, indicated by the slopes of the fluxes through each pathway as a function of ethanol production. Those pathways must be sufficiently down-regulated in order to produce high levels of ethanol. On the other hand, the non-oxidative PP pathway (G3P+S7P→E4P+F6P via transaldolase) appeared to be relatively insensitive to ethanol production.

Summary

The rising cost and use of fossil fuels has renewed focus on lignocellulosic ethanol (Lin and Tanaka 2006) production via simultaneous saccharification and fermentation process (SSF) (Lin and Tanaka 2006). However, enzymes employed to hydrolyze lignocellulosic biomass to simpler sugars for fermentation generally have temperature optimum of around 55°C, whereas the industrial organisms used to ferment the sugars to ethanol or other products (e.g., *Saccharomyces cerevisiae* (Antoni et al. 2007) and *E. coli* (Dien et al. 2003)) have a lower operating temperature. Secondly, yeast cannot typically ferment C5 sugars (Sonderegger et al. 2004) such as xylose, a major component of lignocellulosic biomass. Furthermore, common thermophilic ethanologens, e.g., *Clostridium thermosaccharolyticum*, are strict anaerobes (Lin and Tanaka 2006) and cannot tolerate high ethanol concentrations (>4% w/v) (Fong et al. 2006). *Geobacillus thermoglucosidasius* M10EXG overcomes some of these aforementioned limitations and has many potential advantages for ethanol or other bio-product production: it tolerates high ethanol concentrations (> 10% v/v); it can utilize a wide range of substrates (particularly pentoses and insoluble substrates), which makes it an attractive organism for simultaneous saccharification and fermentation of lignocellulosic biomass; there are lower risks of contamination by other microorganisms (Akao et al. 2007) due to growth at high temperatures; the growth medium will have desirable properties at high temperatures (reduced viscosity, increased diffusion rates and substrate solubility, reduced energy requirements for mixing, and the possibility of combining the fermentation and distillation processes to continuously extract ethanol) (Lin and Tanaka 2006; Lynd 1989). This study investigates this species' metabolic network via *in vitro* enzyme assays and ^{13}C based flux analysis. The obtained information

provides guidelines for engineering the metabolic pathways for bioethanol production as well as other environmental and industrial applications.

Acknowledgements

We thank Dr. Steve Van Dien (Genomatica) for helping with the Simpheny model and Jeannie Chu for helping with metabolite measurement. Financial support for this research was provided by the Sandia National Laboratories Laboratory Directed Research and Development Program. Sandia is a multi-program laboratory operated by the Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000. D.J. and T.C. H. acknowledges support by the Virtual Institute for Microbial Stress and Survival (http://VIMSS.lbl.gov) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomics: GTL Program through contract DE-AC02-05CH11231 between the Lawrence Berkeley National Laboratory and the US Department of Energy. This work is also a part of the Joint BioEnergy Institute supported by the U.S. Department of Energy.

Financial Interest

Jay D. Keasling has a consulting relationship with and a financial interest in Amyris and a financial interest in LS9, two biofuel companies.

References

Akao S, Tsuno H, Cheon J. 2007. Semi-continuous L-lactate fermentation of garbage without sterile condition and analysis of the microbial structure. Water Research 41:1774-1780.

Alm EJ, Huang KH, Price MN, Koche RP, Keller K, Dubchak IL, Arkin AP. 2005. The MicrobesOnline Web site for comparative genomics. Genome Res. **15**:1015-1022.

Antoni D, Zverlov VV, Schwarz WH. 2007. Biofuels from microbes. Appl Microbiol Biotechnol. 77(1):23-35.

Antoniewicz MR, Kelleher JK, Stephanopoulos G. 2006. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. Metabolic Engineering 8(4):324-37.

Christiansen T, Christensen B, Nielsen J. 2002. Metabolic network analysis of *Bacillus claussi* on minimal and semirich medium using ^{13}C-Labeled glucose. Metabolic Engineering(4):159-169.

Cruz Ramos H, Hoffmann T, Marino M, Nedjari H, Presecan-Siedel E, Dreesen O, Glaser P, Jahn D. 2000. Fermentative Metabolism of Bacillus subtilis: Physiology and Regulation of Gene Expression. Journal of Bacteriology 182(11):3072-80.

Daron HH. 1970. Fatty Acid Composition of Lipid Extracts of a Thermophilic Bacillus Species. Journal of Bacteriology 101(1):145-151.

Dauner M, Bailey JE, Sauer U. 2001. Metabolic flux analysis with a comprehensive isotopomer model in *Bacillus subtilis*. Biotechnology and Bioengineering 76(2):144-156.

Dien BS, Cotta MA, Jeffries TW. 2003. Bacteria engineered for fuel ethanol production: current status. Appl Microbiol Biotechnol. 63(3):258-66.

Edwards JS, Palsson BO. 2000a. The *Escherichia coli* MG1655 *in silico* metabolic genotype: Its definition, characteristics, and capabilities. PNAS 97(10):5528–5533.

Edwards JS, Palsson BO. 2000b. Robustness Analysis of the *Escherichia coli* Metabolic Network. Biotechnology Progress 16:927-939.

Fischer E, Sauer U. 2003. A novel metabolic cycle catalyzes glucose oxidation and anaplerosis in hungry *Escherichia coli*. J Biol Chem. 278(47):46446-51.

Fischer E, Zamboni N, Sauer U. 2004. High-throughput metabolic flux analysis based on gas chromatography-mass spectrometry derived 13C constraints. Analytical Biochemistry 325:308-316.

Fong JC, Svenson CJ, Nakasugi K, Leong CT, Bowman JP, Chen B, Glenn DR, Neilan BA, Rogers PL. 2006. Isolation and characterization of two novel ethanol-tolerant facultative-anaerobic thermophilic bacteria strains from waste compost. Extremophiles 10(5):363-72.

Goldman M, Blumenthal HJ. 1963. Pathways of glucose catabolism in *Bacillus subtilis*. Journal of Bacteriology 86:303-11.

Hellerstein MK, Neese RA. 1999. Mass isotopomer distribution analysis at eight years: theoretical, analytic, and experimental considerations. American Journal of Physiology-Endocrinology and Metabolism 276(6):E1146-E1170.

Lin Y, Tanaka S. 2006. Ethanol fermentation from biomass resources: current state and prospects. Applied Microbiology and Biotechnology 69:627-642.

Luli GW, Strohl WR. 1990. Comparison of growth, acetate production, and acetate inhibition of Escherichia coli strains in batch and fed-batch fermentations. Applied and Environmental Microbiology 56(4):1004-1011.

Lynd LR. 1989. Production of Ethanol from Lignocellulosic Materials Using Thermophilic Bacteria: Critical Evaluation of Potential and Review. Fiechter A, editor. Heidelberg: Springer-Verlag. 1-52 p.

Mahadevan R, Bond DR, Butler JE, Esteve-Nunez A, Coppi MV, Palsson BO, Schilling CH, Lovley DR. 2006. Characterization of metabolism in the Fe(III)-reducing organism *Geobacter sulfurreducens* by constraint-based modeling. Applied and Environmental Microbiology 72(2):1558-68.

Majewski RA, Domach MM. 1990. Simple Constrained-Optimization View of Acetate Overflow in E. coli. Biotechnology and Bioengineering 35(7):732-38.

McKinlay JB, Shachar-Hill Y, Zeikus JG, Vieille C. 2007. Determining *Actinobacillus succinogenes* metabolic pathways and fluxes by NMR and GC-MS analyses of ^{13}C-labeled metabolic product isotopomers. Metabolic Engineering 9(2):177-192.

McMullan G, Christie JM, Rahman TJ, Banat IM, Ternan NG, Marchant R. 2004. Habitat, applications and genomics of the aerobic, thermophilic genus *Geobacillus*. Biochemical Society Transactions 32(2):214-7.

Nazina TN, Sokolova DSh, Grigoryan AA, Shestakova NM, Mikhailova EM, Poltaraus AB, Tourova TP, Lysenko AM, Osipov GA, Belyaev SS. 2005. Geobacillus jurassicus sp. nov., a new thermophilic bacterium isolated from a high-temperature petroleum reservoir, and the validation of the Geobacillus species. Syst. Appl. Microbiol. 28(1):43-53.

Sauer U. 2004. High-throughput phenomics: experimental methods for mapping fluxomes. Current Opinion in Biotechnology 15:58-63.

Sauer U, Canonaco F, Heri S, Perrenoud A, Fischer E. 2004. The soluble and membrane-bound transhydrogenases UdhA and PntAB have divergent functions in NADPH metabolism of *Escherichia coli*. Journal of Biological Chemistry 279(8):6613-6619.

Sauer U, Hatzimanikatis V, Bailey JE, Hochuli M, Szyperski T, Wuthrich K. 1997. Metabolic fluxes in riboflavin-producing *Bacillus subtilis*. Nature Biotechnology 15(5):448-52.

Sauer U, Lasko DR, Fiaux J, Hochuli M, Glaser R, Szyperski T, Wuthrich K, Bailey JE. 1999. Metabolic flux ratio analysis of genetic and environmental modulations of Escherichia coli central carbon metabolism. Journal of Bacteriology 181(21):6679-6688.

Sawers G, Suppmann B. 1992. Anaerobic Induction of Pyruvate Formate-Lyase Gene Expression Is Mediated by the ArcA and FNR Proteins. Journal of Bacteriology 174(11):3474-3478.

Shaikh AS, Tang YJ, Mukhopadhyay A, Keasling JD. 2008. Isotopomer Distributions in Amino Acids from a Highly Expressed Protein as a Proxy for Those from Total Protein. Anal Chem. 80 (3):886-890.

Sonderegger M, Jeppsson M, Larsson C, Gorwa-Grauslund MF, Boles E, Olsson L, Spencer-Martins I, Hahn-Hägerdal B, Sauer U. 2004. Fermentation performance of engineered and evolved xylose-fermenting *Saccharomyces cerevisiae* strains. Biotechnol Bioeng. 87(1):90-8.

Stephanopoulos GN, Aristidou AA, Nielsen J. 1998. Metabolic Engineering Principles and Methodologies. San Diego: Academic Press. 75, 120-130 p.

Sullivan KH, Hegeman GD, Cordes EH. 1979. Alteration of the Fatty Acid Composition of *Escherichia coli* by Growth in the Presence of Normal Alcohols. Journal of Bacteriology 138(1):133-138.

Takami H, Takaki Y, Chee GJ, Nishi S, Shimamura S, Suzuki H, Matsui S, Uchiyama I. 2004. Thermoadaptation trait revealed by the genome sequence of thermophilic *Geobacillus kaustophilus*. Nucleic Acids Res. 32(21):6292-6303.

Tang YJ, Chakraborty R, Martin HG, Chu J, Hazen TC, Keasling JD. 2007a. Flux analysis of central metabolic pathways in *Geobacter metallireducens* during reduction of soluble Fe(III)-NTA. Applied and Environmental Microbiology 73(12):3859-3864.

Tang YJ, Hwang JS, Wemmer D, Keasling JD. 2007b. The *Shewanella oneidensis* MR-1 fluxome under various oxygen conditions. Applied and Environmental Microbiology 73(3):718-29.

Tang YJ, Meadows AL, Keasling JD. 2007c. A kinetic model describing *Shewanella oneidensis* MR-1 growth, substrate consumption, and product secretion. Biotechnology and Bioengineering 189(3):894-901.

Tang YJ, Meadows AL, Kirby J, Keasling JD. 2007d. Anaerobic central metabolic pathways in *Shewanella oneidensis* MR-1 reinterpreted in the light of isotopic metabolite labeling. Journal of Bacteriology 189(3):894-901.

Tang YJ, Pingitore F, Mukhopadhyay A, Phan R, Hazen TC, Keasling JD. 2007e. Pathway confirmation and flux analysis of central metabolic pathways in *Desulfovibrio vulgaris* Hildenborough using GC-MS and FT-ICR mass spectrometry. Journal of Bacteriology 189(3):940-949.

Terada K, Murata T, Izui K. 1991. Site-Directed Mutagenesis of Phosphoenolpyruvate Carboxylase from *E. coli*: The Role of His^{579} in the Catalytic and Regulatory Functions. J. Biochem. 109:49 - 54.

Van der Werf MJ, Guettler MV, Jain MK, Zeikus JG. 1997. Environmental and physiological factors affecting the succinate product ratio during carbohydrate fermentation by *Actinobacillus* sp. 130Z. Arch Microbiol. 167(6):332-342.

Wiechert W, Mollney M, Petersen S, de Graaf AA. 2001. A Universal Framework for ^{13}C Metabolic Flux Analysis. Metabolic Engineering 3:265-283.

Zhao J, Shimizu K. 2003. Metabolic flux analysis of *Escherichia coli* K12 grown on 13C-labeled acetate and glucose using GC-MS and powerful flux calculation method. Journal of Biotechnology 101:101-117.

Figure and Table Captions

Figure 1: Growth kinetics of M10EXG under three oxygen conditions: □ aerobic, ◊ micro-aerobic, ○ anaerobic.

Figure 2. Pathways and flux distributions of glucose metabolism under aerobic (top) and micro-aerobic (bottom) conditions. The amino acids used for isotopomer models were shown in parentheses. The glucose uptake rates were normalized to a value of 100. Dotted lines indicate that the pathways are not active. Abbreviations: Acetyl-CoA, acetyl-coenzyme A; CIT, citrate; E4P, erythrose-4-phosphate; C1, 5,10-Me-THF; F6P, fructose-6-phosphate; G6P, glucose-6-phosphate; 6PG, 6-phosphogluconate; ICT, isocitrate; MAL, malate; OAA, oxaloacetate; OXO, 2-oxoglutarate; PEP, phosphoenolpyruvate; PGA, 3-phosphoglycerate; C5P, ribose-5-phosphate (or ribulose-5-phosphate or xylulose-5-phosphate); S7P, sedoheptulose-7-phosphate; SUC, succinate; T3P, triose-3-phosphate.

Figure 3. M10EXG mixed acid fermentation. The abbreviations were the same as those in Figure 2. Key reactions (and their corresponding relative fluxes): 1. glucose-6-phosphate isomerase; 2. glucose-6-phosphate dehydrogenase; 3. T3P dehydrogenase; 4. L-lactate dehydrogenase; 5. pyruvate-formate lyase; 6. acetaldehyde dehydrogenase; 7. alcohol dehydrogenase; 8. phosphate acetyltransferase/acetate kinase. The arrows were drawn in proportion to the fluxes. Fluxes below 10% of the glucose uptake rate were represented by non-scaled hairlines.

Figure 4. Effect of mixed acids production and biomass growth rate on ethanol production as calculated by the FBA model. The glucose uptake rate was set to 5 mmol hr^{-1} g^{-1} biomass. The units for ethanol and acids production rates are mmol hr^{-1} g^{-1} biomass. (a) Growth rate = 0.1 hr^{-1}, formate production = 0. (b) Growth rate = 0.2 hr^{-1}, formate production = 0. (c) Growth rate = 0.1 hr^{-1}, formate production was assumed to equal the sum of the ethanol and acetate production rates; (d) Growth rate = 0.2 hr^{-1}, formate production was assumed to equal the sum of the ethanol and acetate production rates.

Figure 5. Change in central metabolism as a function of ethanol production as predicted by the *in silico* flux balance model (Simpheny). The objective function used for the calculations was the maximal biomass production. Symbols: growth rate (○); flux into the TCA cycle via citrate synthase (▲); flux into the pentose phosphate pathway via glucose 6-phosphate dehydrogenase (♦) and via transaldolase (GAP+S7P→E4P+F6P) (●); flux through the pyruvate shunt (■). The three dashed arrows linked the measured ethanol flux values with their corresponding measured growth rates for the three growth conditions (aerobic, micro-aerobic, and anaerobic). The fact that the lines (skewed dashed arrows) were not vertical indicates a difference between *in silico* model predicted flux (optimal metabolism) and experimentally measured flux (actual metabolism).

Table 1. Enzyme activities in cell extracts of *Geobacillus thermoglucosidasius* M10EXG under three oxygen conditions (n=3).

Table 2. Growth kinetics and yields of ethanol and organic acids under the three oxygen conditions: aerobic growth (G + O_2), micro-aerobic growth (G + μO_2), anaerobic growth (G - O_2).

Figures

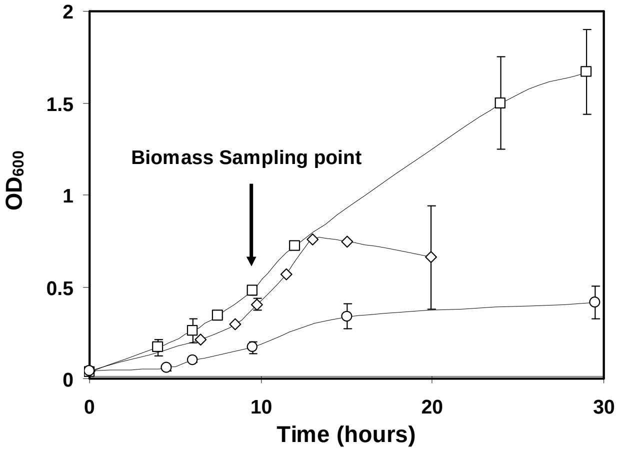

Figure 1. Growth kinetics of M10EXG under three oxygen conditions: □ aerobic, ◊ micro-aerobic, ○ anaerobic.

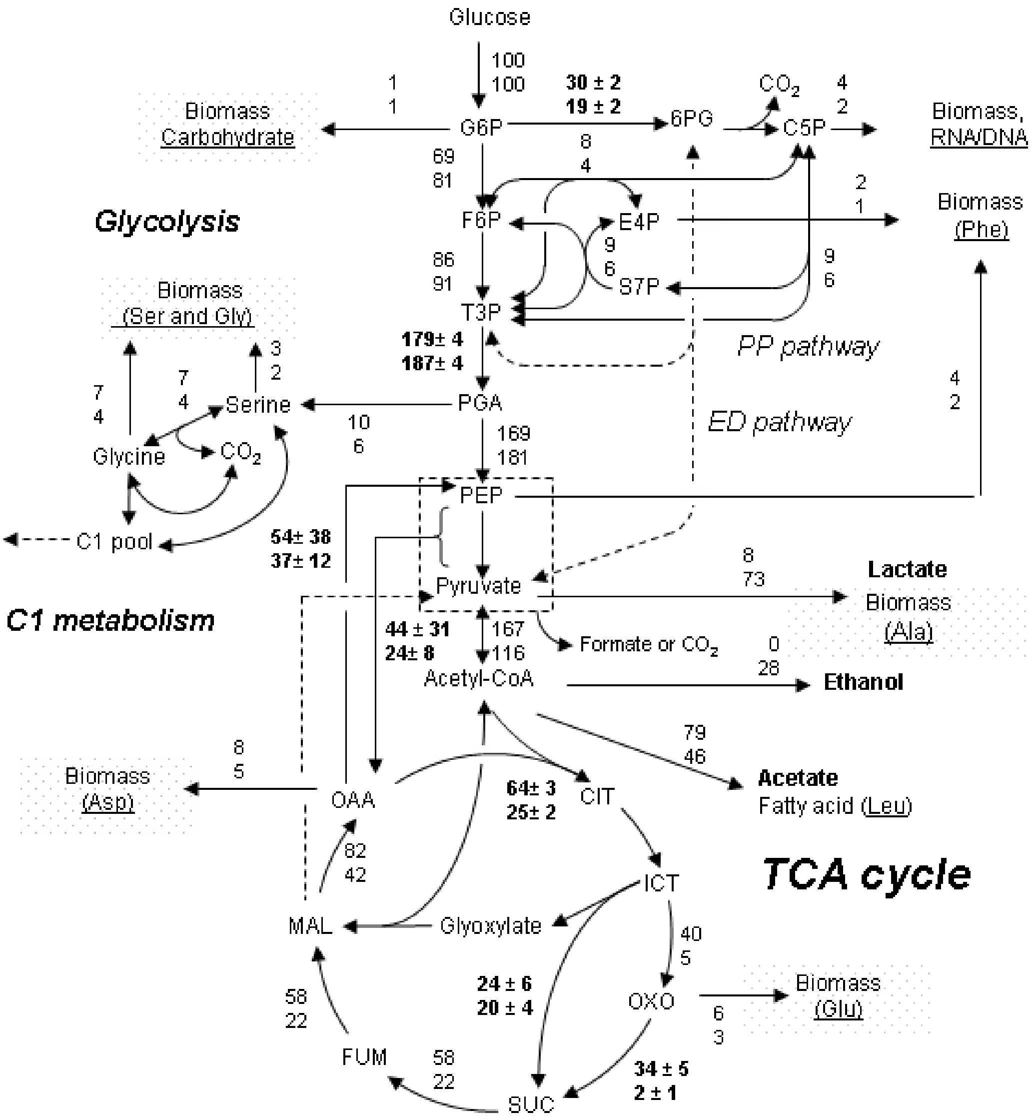

Figure 2. Pathways and flux distributions of glucose metabolism under aerobic (top) and micro-aerobic (bottom) conditions. The amino acids used for isotopomer models were shown in parentheses. The glucose uptake rates were normalized to a value of 100. Dotted lines indicate that the pathways are not active. Abbreviations: Acetyl-CoA, acetyl-coenzyme A; CIT, citrate; E4P, erythrose-4-phosphate; C1, 5,10-Me-THF; F6P, fructose-6-phosphate; G6P, glucose-6-phosphate; 6PG, 6-phosphogluconate; ICT, isocitrate; MAL, malate; OAA, oxaloacetate; OXO, 2-oxoglutarate; PEP, phosphoenolpyruvate; PGA, 3-phosphoglycerate; C5P, ribose-5-phosphate (or ribulose-5-phosphate or xylulose-5-phosphate); S7P, sedoheptulose-7-phosphate; SUC, succinate; T3P, triose-3-phosphate.

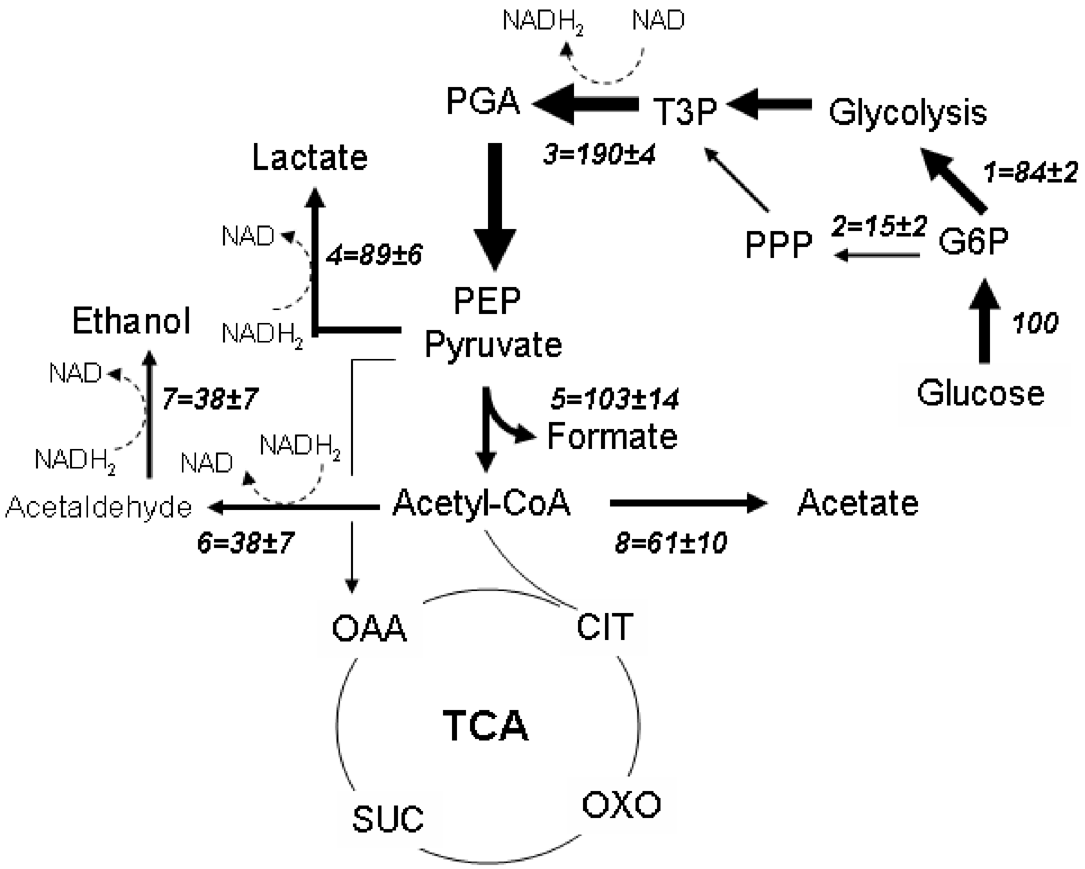

Figure 3. M10EXG mixed acid fermentation. The abbreviations were the same as those in Figure 2. Key reactions (and their corresponding relative fluxes): 1. glucose-6-phosphate isomerase; 2. glucose-6-phosphate dehydrogenase; 3. T3P dehydrogenase; 4. L-lactate dehydrogenase; 5. pyruvate-formate lyase; 6. acetaldehyde dehydrogenase; 7. alcohol dehydrogenase; 8. phosphate acetyltransferase/acetate kinase. The arrows were drawn in proportion to the fluxes. Fluxes below 10% of the glucose uptake rate were represented by non-scaled hairlines.

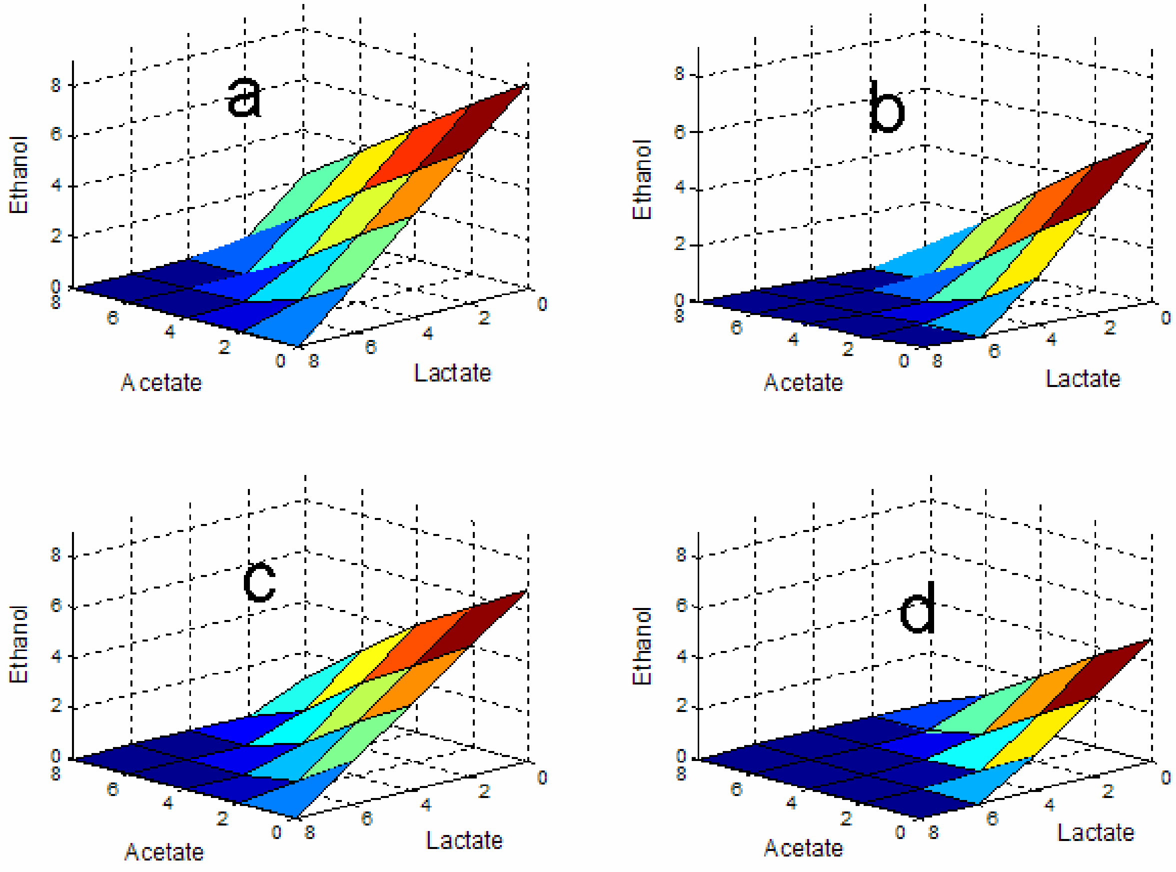

Figure 4. Effect of mixed acids production and biomass growth rate on ethanol production as calculated by the FBA model. The glucose uptake rate was set to 5 mmol hr^{-1} g^{-1} biomass. The units for ethanol and acids production rates are mmol hr^{-1} g^{-1} biomass. (a) Growth rate = 0.1 hr^{-1}, formate production = 0. (b) Growth rate = 0.2 hr^{-1}, formate production = 0. (c) Growth rate = 0.1 hr^{-1}, formate production was assumed to equal the sum of the ethanol and acetate production rates; (d) Growth rate = 0.2 hr^{-1}, formate production was assumed to equal the sum of the ethanol and acetate production rates.

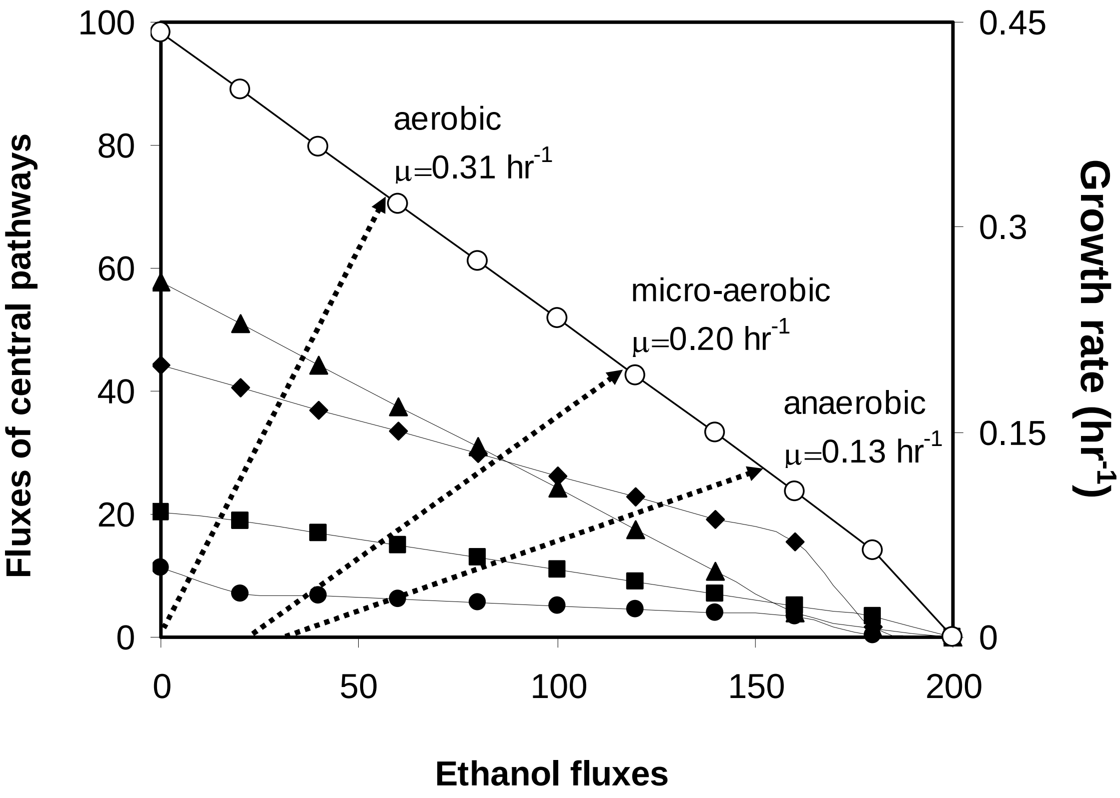

Figure 5. Change in central metabolism as a function of ethanol production as predicted by the *in silico* flux balance model (Simpheny). The objective function used for the calculations was the maximal biomass production. Symbols: growth rate (○); flux into the TCA cycle via citrate synthase (▲); flux into the pentose phosphate pathway via glucose 6-phosphate dehydrogenase (♦) and via transaldolase (GAP+S7P→E4P+F6P) (●); flux through the pyruvate shunt (■). The three dashed arrows linked the measured ethanol flux values with their corresponding measured growth rates for the three growth conditions (aerobic, micro-aerobic, and anaerobic). The fact that the dashed arrows were not vertical indicated a difference between *in silico* model predicted fluxes (optimal metabolism) and experimentally measured fluxes (actual metabolism).

Tables

Table 1. Enzyme activities in cell extracts of *Geobacillus thermoglucosidasius* M10EXG under three oxygen conditions (n=3).

Enzymes	**EC number**	**Specific activity (units g protein^{-1})**		
		Aerobic	Micro-aerobic	Anaerobic
Oxalacetate decarboxylase (- Na^+)	EC 4.1.1.3	0	3 ± 3	0
Oxalacetate decarboxylase (+ Na^+)	EC 4.1.1.3	0	2 ± 3	0
Malic Enzyme (NADP+)	EC 1.1.1.40	0	0	0
Malic Enzyme (NAD+)	EC 1.1.1.38	0	0	0
α-ketoglutarate dehydrogenase	EC 1.2.4.2	230 ± 52	92 ± 75	26 ± 18
Pyruvate carboxylase	EC 6.4.1.1	682 ± 385	660 ± 279	615 ± 246
PEP carboxykinase	EC 4.1.1.49	249 ± 50	373 ± 25	298 ± 22
Isocitrate lyase	EC 4.1.3.1	26 ± 2	22 ± 4	21 ± 5
PEP carboxylase	EC 4.1.1.31	89 ± 58	79 ± 70	49 ± 21
Transhydrogenase	EC 1.6.1.1.	0	0	0

Note: one unit catalyzes the formation of one µmol of substrate per minute.

Table 2. Growth kinetics and yields of ethanol and organic acids under the three oxygen conditions: aerobic growth (G + O_2), micro-aerobic growth (G + μO_2), anaerobic growth (G - O_2).

Yield[1]	G + O_2	G + μO_2	G-O_2[2]	Max[3]
$Y_{ace/s}$	0.64±0.12	0.40±0.05	0.61±0.10	2.6
$Y_{lact/s}$	0.02±0.01	0.67±0.07	0.89±0.06	2
$Y_{etho/s}$	0.01±0.01	0.28±0.04	0.38±0.07	2
$Y_{form/s}$	0	0.13±0.05	1.03±0.14	5.6
$Y_{biomass/s}$	0.27±0.05	0.19±0.04	0.08±0.03	0.34
Growth rate, hr^{-1}	0.31±0.04	0.20±0.04	0.13±0.03	0.44

[1] Metabolite yield unit, mol metabolites mol^{-1} glucose. Biomass yield unit, DCW g^{-1} glucose.

[2] A small amount of succinate was also detected.

[3] The maximum yield for each metabolite was predicted using Simpheny. The model assumed a glucose uptake rate equal to 5 mM hr^{-1} DCW $^{-1}$.