Indian J. Microbiol. (June 2008) 48:252–266

ORIGINAL ARTICLE

Pyruvate catabolism and hydrogen synthesis pathway genes of *Clostridium thermocellum* ATCC 27405

Carlo R. Carere · Vipin Kalia · Richard Sparling · Nazim Cicek · David B. Levin

Received: 21 January 2008 / Accepted: 12 June 2008

Abstract *Clostridium thermocellum* is a gram-positive, acetogenic, thermophilic, anaerobic bacterium that degrades cellulose and carries out mixed product fermentation, catabolising cellulose to acetate, lactate, and ethanol under various growth conditions, with the concomitant release of H_2 and CO_2. Very little is known about the factors that determine metabolic fluxes influencing H_2 synthesis in anaerobic, cellulolytic bacteria like *C. thermocellum*. We have begun to investigate the relationships between genome content, gene expression, and end-product synthesis in *C. thermocellum* cultured under different conditions. Using bioinformatics tools and the complete *C. thermocellum* 27405 genome sequence, we identified genes encoding key enzymes in pyruvate catabolism and H_2-synthesis pathways, and have confirmed transcription of these genes throughout growth on α-cellulose by reverse transcriptase polymerase chain reaction. Bioinformatic analyses revealed two putative lactate dehydrogenases, one pyruvate formate lyase, four pyruvate:formate lyase activating enzymes, and at least three putative pyruvate:ferredoxin oxidoreductase (POR) or POR-like enzymes. Our data suggests that hydrogen may be generated through the action of either a Ferredoxin (Fd)-dependent NiFe hydrogenase, often referred to as "Energy-converting Hydrogenases", or via NAD(P)H-dependent Fe-only hydrogenases which would permit H_2 production from NADH generated during the glyceraldehyde-3-phosphate dehydrogenase reaction. Furthermore, our findings show the presence of a gene cluster putatively encoding a membrane integral NADH:Fd oxidoreductase, suggesting a possible mechanism in which electrons could be transferred between NADH and ferredoxin. The elucidation of pyruvate catabolism pathways and mechanisms of H_2 synthesis is the first step in developing strategies to increase hydrogen yields from biomass. Our studies have outlined the likely pathways leading to hydrogen synthesis in *C. thermocellum* strain 27405, but the actual functional roles of these gene products during pyruvate catabolism and in H_2 synthesis remain to be elucidated, and will need to be confirmed using both expression analysis and protein characterization.

Keywords *Clostridium thermocellum* · Fermentation · Cellulose · Hydrogen · Pyruvate catabolism

C. R. Carere[1] · V. Kalia[2] · R. Sparling[3] · N. Cicek[1] · D. B. Levin[1] (✉)
[1]Department of Biosystems Engineering,
University of Manitoba, Winnipeg MB, Canada, R3T 5V6
e-mail: umcarerc@cc.umanitoba.ca, nazim_cicek@cc.umantiboa.ca, levindb@cc.umanitoba.ca
[2]Microbial Biotechnology and Genomics,
Institute of Genomics and Integrative Biology (IGIB);
Council of Scientific and Industrial Research (CSIR);
Delhi University Campus; Mall Road,
Delhi - 110 007, India
e-mail: vckalia@igib.res.in, vc_kalia@yahoo.co.in
[3]Department of Microbiology, University of Manitoba,
Winnipeg MB, Canada, R3T 2N2;
e-mail: sparlng@cc.umanitoba.ca

Introduction

Clostridium thermocellum is a gram-positive, acetogenic, thermophilic, anaerobic bacterium, that degrades cellulose and carries out mixed product fermentation, catabolising cellulose to various amounts of acetate, lactate and ethanol, with the concomitant release of H_2 and CO_2, under different growth conditions [1–6]. Formate has also been reported

as a significant end-product in *C. thermocellum* ATCC 27405 [7]. *C. thermocellum* expresses a suite of cellulolytic enzymes that are assembled into a complex structure on the cell surface called the cellulosome [8, 9]. The bacteria attach to cellulose particles via the cellulosome, which efficiently degrades cellulose chains to cellobiose and other soluble cellulodextrans. *C. thermocellum* displays the highest rate of cellulose degradation of all known cellulose degrading microorganisms [2, 9, 10].

We have investigated hydrogen (H_2)-production by *C. thermocellum* strain 27405. *C. thermocellum* produces greater amounts of H_2 when cultured on cellulosic substrates compared with the soluble cellulodextran cellobiose, with an average yield of 1.55 mol H_2/mol glucose equivalent [11, 12]. We have observed the production of formate, ethanol, and acetate along with H_2 and CO_2 during exponential growth of the cells, with lactate being produced, as the cells entered stationary phase [7]. In order to develop strategies to enhance H_2 production by *C. thermocellum*, a greater understanding of the metabolic and genetic mechanisms by which H_2 is synthesized is required.

In many fermentative organisms, H_2 synthesis associated with synthesis of acetyl-CoA may occur via pathways mediated by either 1) Pyruvate:ferredoxin oxidoreductase (POR), which catalyzes $Pyruvate + CoA + 2\ Fd_{ox} \rightarrow Acetyl\text{-}CoA + CO_2 + Fd_{red}$ or 2) Pyruvate:formate lyase (PFL), which catalyzes $Pyruvate + CoA \rightarrow Acetyl\text{-}CoA + Formate$. POR (EC 1.2.7.1) mediated oxidation of pyruvate is typically observed in obligate anaerobic Eukarya, Archaea and Bacteria, including the Clostridiales [13]. Oxidation of pyruvate by POR generates acetyl-CoA plus Fd_{red} and drives H_2 synthesis via Ferredoxin-dependent hydrogenase (EC 1.12.7.2). Formate production directs reducing equivalents away from H_2 synthesis enzymes associated with the POR mediated pathway. In some organisms, formate can be further oxidized to CO_2 plus H_2 ($\Delta G = +1.3$ kJ/mol) by Formate:hydrogen lyase (FHL) yielding one mole of H_2 per mole of formate [14, 15]. The ΔG for FHL, however, is near neutrality, so the reaction is very much dependent on the concentrations of H_2 and formate in the cell and environment [16].

NADH generated during the oxidation of glyceraldehyde-3-phosphate can donate its electrons for the generation of H_2 through the use of a NADH-dependent hydrogenase. Thus, when acetic acid is the sole organic end-product, a theoretical maximum of 4 moles H_2 per mole of glucose catabolized is obtained: $C_6H_{12}O_6 + 2\ H_20 \rightarrow 2\ CH_3COOH + 4\ H_2 + 2\ CO_2$ [14, 15]. However, NADH is at a higher redox potential than H_2, making this reaction thermodynamically unfavourable when there is significant H_2 within the cell or in the environment of the cell [16].

A critical step in understanding the metabolic and genetic mechanisms by which H_2 is synthesized is to identify the genes encoding enzymes in H_2-synthesizing pathways. Pyruvate catabolism plays a pivotal role, since the relative flux through the various pyruvate metabolizing enzymes will determine, in part, the cells capacity to synthesize H_2. Cell-free extract activity has been observed for all three enzymes [7]. The objective of this study was to investigate the relationships between genome content, gene expression, and end-product synthesis with respect to H_2 production in *C. thermocellum* while cultured on cellulosic substrates.

Methods and materials

In Silico analyses

All sequence data for *C. thermocellum* was produced by the US Department of Energy Joint Genome Institute (http://www.jgi.doe.gov/). Gene numbers presented in this paper reflect the numbering of the final annotation. Nucleotide sequences encoding enzymes known to be involved in pyruvate catabolism in *Escherichia coli* or Firmicules such as *Clostridum perfringens*, *Clostridium acetobutylicum*, *Clostridum tetani*, and/or *Thermoanaerobacter tengcongensis* were used as probes to screen the *C. thermocellum* genome data base. Basic Local Alignment Search Tool Nucleotide-nucleotide (BLASTn) analysis, in conjunction with conserved amino acid domain searches (rpsBLAST) were used to identity genes of interest, as indicated by the genome database annotation. For BLASTn analyses, only E-score values of less than 2e-26 were accepted as positive identification of a gene and its putative function. In some cases, multiple genes with similar annotations were identified because of the presence of conserved amino acid sequence domains or high levels of sequence homology. ClustalW multiple alignments were performed to subsequently screen for conserved regions between these genes. In cases where a gene expected to be present in the genome was not found, a consensus sequence was constructed from *C. perfringens*, *C. acetobutylicum*, *C. tetani*, and/or *T. tengcongensis,* all of which are available at the National Center for Biotechnology Information (NCBI) website (www.ncbi.nlm.nih.gov/). BLASTn searches with these consensus sequences against the *C. thermocellum* genome were then performed to screen for their presence.

To identify genes encoding NiFe- or Fe-only hydrogenases in the *C. thermocellum* genome, and to determine if these genes encode Fd- or NAD-dependent hydrogenases, we took a two-step approach. First, BLAST searches using conserved amino acid sequences encoding catalytic

domains or subunits corresponding to either NiFe- or Fe-only hydrogenases (corresponding to NuoBCD vs. NuoG, respectively) were conducted. Second, BLAST analyses using the amino acid sequence of a flavin containing NAD(P)-binding subunit (for example related to NuoF), which is known to mediate the transfer of electrons between NADH and ferredoxin within the enzyme for the synthesis of H_2, were conducted within the same gene or operon.

Microorganism and media

Clostridium thermocellum 27405 was obtained from the American Type Culture Collection (ATCC) and was employed for all growth experiments. Media preparation and culturing of *C. thermocellum* was performed as outlined by Sparling *et al*. [7]. All culturing experiments were performed at a final volume of 10 ml, at 1.1 g/L α-cellulose and incubated at 60°C.

Protein and cellulose determination

Two millilitre (2 ml) samples of fresh culture were centrifuged for 10 min at 14000 × g after which the pellet was washed with 0.9% NaCl and then resuspended in 0.2N NaOH. Total protein was used to follow growth of *C. thermocellum* on 1.1 g/L α-cellulose using the Bradford method of protein determination [17]. Protein standards were prepared using bovine serum albumin and all absorbance readings were performed at 595 nm using a PowerWave-XS single channel spectrophotometer using KCjunior software (BIO-TEK Instruments Inc., Winooski, Vermont USA)

Cellulose consumption was measured using a modified Anthrone assay for the determination of total carbohydrates [18]. Anthrone reagent was prepared 0.1% (w/v) in 95% H_2SO_4 at least 4 hours prior to each assay and was discarded if unused after 1 week. Supernatant samples (250 µl) were diluted such that the final concentration of total carbohydrates did not exceed 200 µg/ml, after which they were incubated at room temperature for 20 min in 2.0 ml 0.1% (w/v) Anthrone reagent. Samples were then incubated in a boiling water bath for 15 min and then allowed to cool for 45 min. Absorbencies were read at 595 nm against α-cellulose standards of known concentrations.

Determination of fermentation products and gas production

One milliliter (1 ml) samples of culture supernatant were transferred to 1.5 ml micro-centrifuge tubes and centrifuged at 10000 × g for ten min. Supernatants were then transferred into fresh tubes and stored at –20C until required. Formate, acetate and lactate production were measured by high performance liquid chromatography using an IonPac AS11-HC anion exchange column (Dionex Corporation, Sunnyvale, California, USA). The production of ethanol was measured using an Ethanol assay kit (Cat. No. 10 176 290 035) purchased from R-Biopharm (Marshall, Michigan, USA). The pH of each sample was measured directly from the culture supernatant with a model AP62 pH/mV meter (Fisher Scientific, Ottawa, Ontario, Canada) equipped with a needle probe. Product gas composition (H_2 and CO_2) was measured using a Multiple Gas Analyzer #1 Gas Chromatograph System Model 8610-0070 (SRI Instruments, Torrance, California, USA) using a 2-meter Molecular Sieve 13X column for the separation of H_2 and a 2-meter Silica Gel column for the separation of CO_2. A thermal conductivity (TCD) detector with detection limits between 200-500 ppm was used. All gas measurements were calculated taking into account atmospheric pressure, solubility in water, and for CO_2, bicarbonate equilibrium at the pH of the culture sample used.

RNA extraction and Reverse Transcriptase-PCR (RT-PCR)

Total RNA was isolated from freshly collected 10 ml cultures of C. *thermocellum* grown on 1.1 g/L α-cellulose using the Invitrogen TRIzol reagent kit (Invitrogen, Carlsbad , CA). Cells were pelleted by centrifugation at 10000 x g for 10 min and then total RNA was isolated following the manufacturers instructions for RNA extraction for cells grown in suspension. The obtained RNA pellet was dissolved in RNase free H_2O containing RNase inhibitor (Invitrogen) and dithiothreitol (DTT) at final concentrations of 0.5 U/µl and 1mM respectively. RNA was treated with DNase (Invitrogen) for 15 min at 20 °C prior to cDNA synthesis. The final concentrations for the DNase treatment reaction were 2 mM $MgCl_2$, 20 mM Tris (pH 8.4), 50 mM KCl, 0.1U/µl DNase. Ethylenediamine Tetra-acetic acid (EDTA) was added to a final concentration of 2.5 mM and the reaction mixture was incubated at 65 °C for 10 min to stop the reaction. First-strand cDNA synthesis using Invitrogen SuperScript II Reverse Transcriptase was performed following the manufacturers recommended protocol using random hexamer primers. Each reaction was performed under the following conditions: between 1-5 µg of total RNA, 2.5 ng/µl random hexamer primers, 0.5 mM dNTP's, 5 mM DTT, 1.25 mM $MgCl_2$, 0.5X RT buffer, 2.5 U/µl SuperScript II Reverse Transcriptase (Invitrogen) and 0.2 U/µl RNaseH (Invitrogen).

An Eppendorf Mastercycler thermocycler was used for all RT-PCR reactions with PCR products run on 1% Tris-Boric acid EDTA agarose gels, stained with Ethidium

bromide at 0.25 µg/ml and visualized with an EpiChem [3] Darkroom system (UVP, USA). Amplification consisted of an initial incubation for 2 min at 94 C, followed by 30 three-step cycles at 94 °C for 45 seconds (melting), 55 °C for 45 seconds (annealing), and 72 °C for 1 min (extension). Reactions were held at 72 °C for 10 min after the 30 cycles and then kept at 4 °C. Each reaction contained the following (final volume 25µl): 10 mM KCl, 10 mM $(NH_4)SO_4$, 20 mM Tris-HCl (pH 7.5), 1% Dimethylsulfoxide 100 µg/ml Bovine Serum Albumin, 2 mM $MgSO_4$, 1 mM forward primer, 1 mM reverse primer, 0.2 mM dATP, 0.2 mM dTTP, 0.2 mM dCTP, 0.2 mM dGTP, 0.04 U/µl Taq polymerase. Finally, 10 ng of cDNA or 10 ng of *C. thermocellum* ATCC 27405 genomic DNA was added to each reaction to a final volume of 25 µl.

Primer pairs were designed using Oligo software such that T_m values for each set of primers were within 1.5 °C of their complement and fell between 60 °C and 63 °C. Primers were selected in regions internal to the open reading frame (ORF) of the gene being investigated (Table 1). All PCR reaction products were cloned into pGEM-T (Promega, WI, USA) following the manufacturers suggested protocol and then sequenced in order to confirm amplicon identity.

Results

Cell Growth, substrate consumption, and end-product synthesis during fermentation

Growth of *C. thermocellum* ATCC 27405 on 1.1g/L -cellulose was followed for 49 hours. No significant lag was observed following inoculation, and cell mass (measured as total protein) increased exponentially until approx. 29 hrs, when the culture began to enter stationary phase. Initial α-cellulose concentrations decreased rapidly during the first 10 hours of growth (31.7–4.89 µmoles glucose equivalent/culture), and a final concentration of 1.25 µmoles glucose equivalent/culture was observed at 49 hours (Fig. 1a). Acetate, ethanol, and formate production followed cell growth closely. Lactate concentrations, however, remained low (0.1 µmoles/10 ml culture) until 24 hours post- inoculation (hrs pi), when a sharp increase to 3.54 µmoles/10 ml culture was observed. Lactate levels continued to increase up to 49 hrs and reached a concentration of 23.74 µmoles/10ml culture (Fig. 1b). H_2 and CO_2 gas production also followed cell growth closely, with concentrations reaching 18.33 and 17.22 µmoles/10ml culture for H_2 and CO_2 respectively, at 49 hrs pi (Fig. 1c).

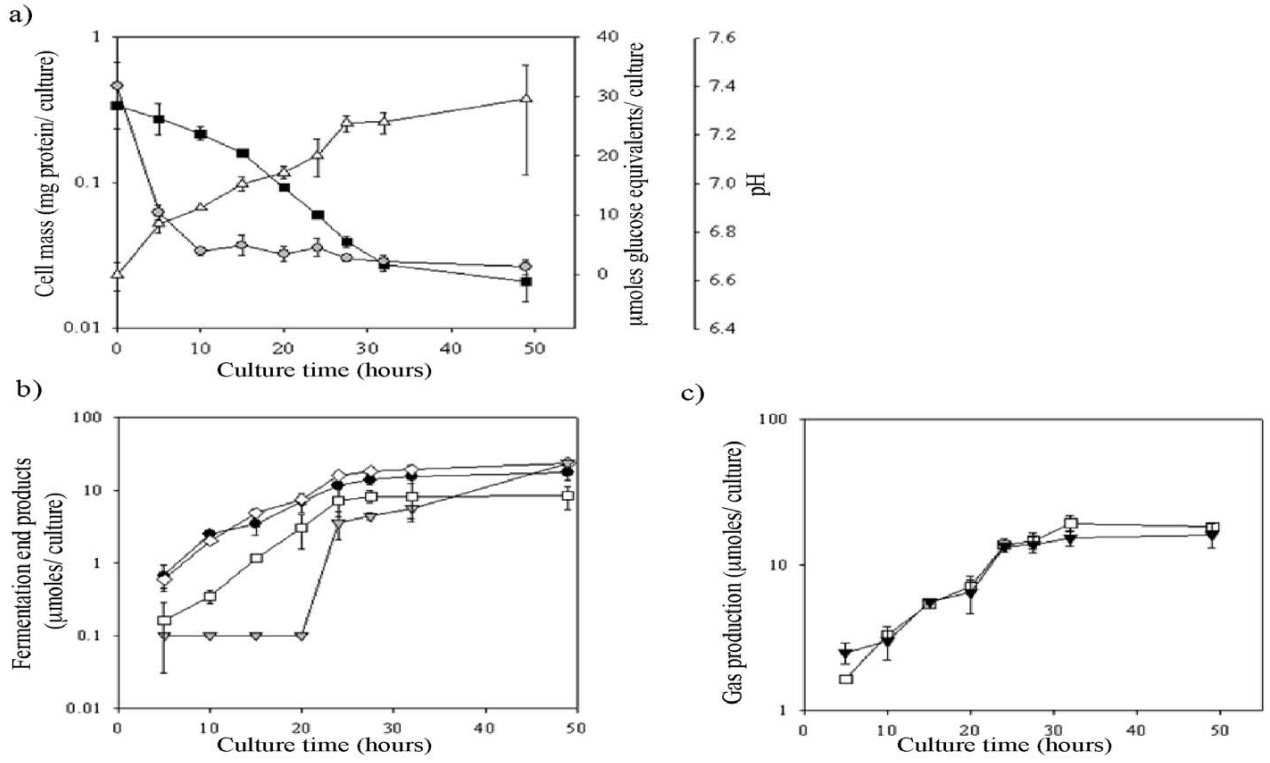

Fig. 1 a) Growth of *Clostridium thermocellum* on 1191 media at 1.1g/L α-cellulose incubated at 60°C as determined by the Bradford method of total protein determination (Δ) per 10 ml batch culture, cellulose consumption (○) and pH (■). b) Fermentation products synthesized during growth including formate (□), acetate (●), lactate (▽) and ethanol (◇), c) Hydrogen (□) and CO_2 (▼) synthesis.

Pyruvate catabolism genes

Using bioinformatic analyses, we identified genes encoding key enzymes in pyruvate catabolism pathways including 2 putative LDHs, 1 putative PFL, at least 3 putative PFL-AE and 3 putative POR or POR-like enzymes. RT-PCR confirmed that these genes are transcribed throughout growth of *C. thermocellum* on α-cellulose (Fig. 2a). BLAST analysis comparing the two putative LDHs revealed that they share 48 % amino acid sequence identity (Fig. 3a,b). Gene

Table 1 Amino acid sequences used for ClustalW analysis of Clostridium thermocellum ATCC 27405 pyruvate catabolism pathways

Organism	Enzyme	Accession number	Reference*
Clostridium thermocellum ATCC 27405	PFL	ZP_00510887.1	
Alkaliphilus metalliredigenes QYMF	PFL	ZP_00799914.1	
Thermosynechococcus elongatus BP-1	PFL	NP_681780.1	[39]
Bacillus cereus subsp. cytotoxis NVH	PFL	ZP_01181225.1	
Bacillus thuringiensis serovar konkukian	PFL	YP_034774.1	
Clostridium thermocellum ATCC 27405	LDH (gene 345)	ZP_00510243.1	
Thermoanaerobacterium saccharolyticum	LDH	AAP34686.1	[40]
Clostridium tetani E88	LDH	NP_782567.1	[26]
Clostridium acetobutylicum ATCC 824	LDH	NP_346908.1	[41]
Clostridium perfringens SM101	LDH	YP_697439.1	[42]
Clostridium thermocellum ATCC 27405	LDH (gene 1053)	Q8KQC4	[19]

* No entry in the reference column refers to sequences that have been submitted to the NCBI database but have not yet been published.

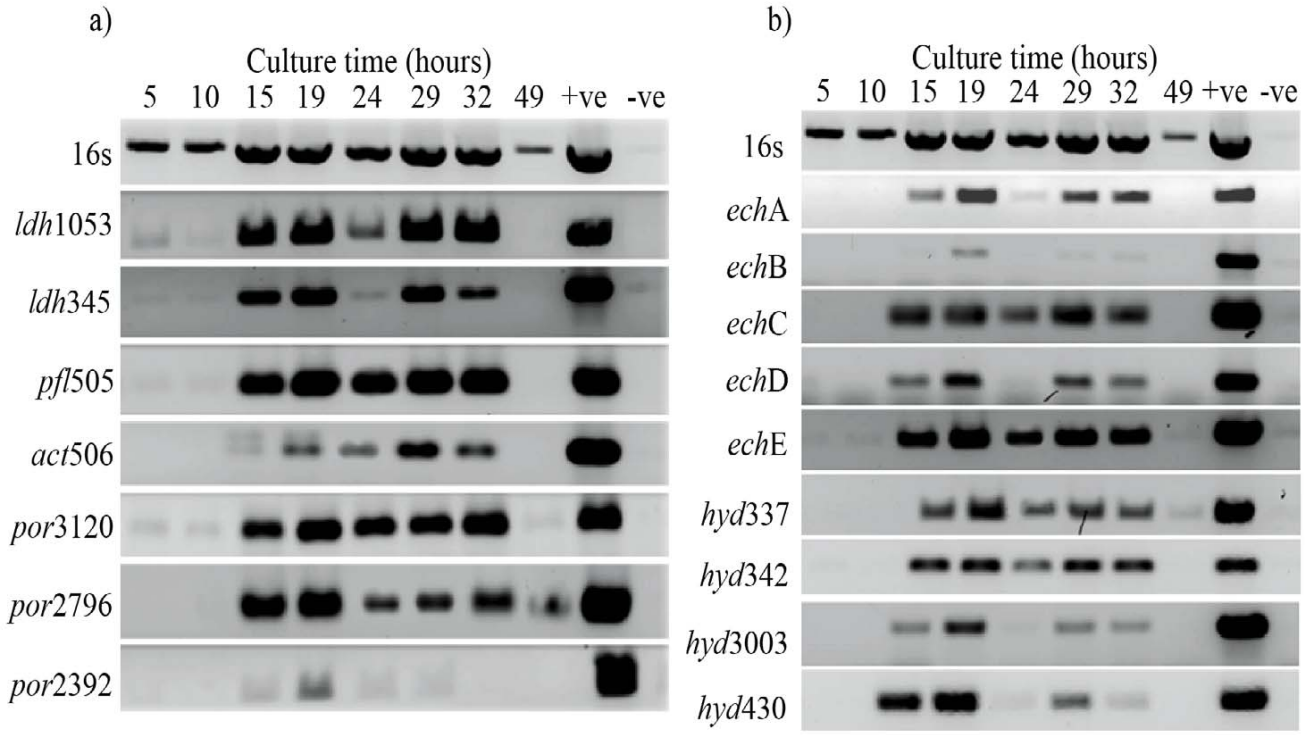

Fig. 2 Reverse transcriptase Polymerase Chain Reaction from total RNA of *C. thermocellum* ATCC 27405. Batch cultures (10 ml) were grown in 1191 media at 60°C on 1.1g/L α-cellulose. a) Lactate dehydrogenase genes (*ldh*345, *ldh*1053), pyruvate:formate lyase (*pfl*505), PFL-activating enzyme (*act*506) and pyruvate:ferredoxin oxidoreductase (*por*3120, *por*2796, *por*2392) were all probed for transcription during 49 hours of growth. b) Identified NiFe hydrogenases (*ech*ABCDE) and identified Fe-only hydrogenases (*hyd*337, *hyd*342, *hyd*3003 *hyd*430) were also probed for transcription.

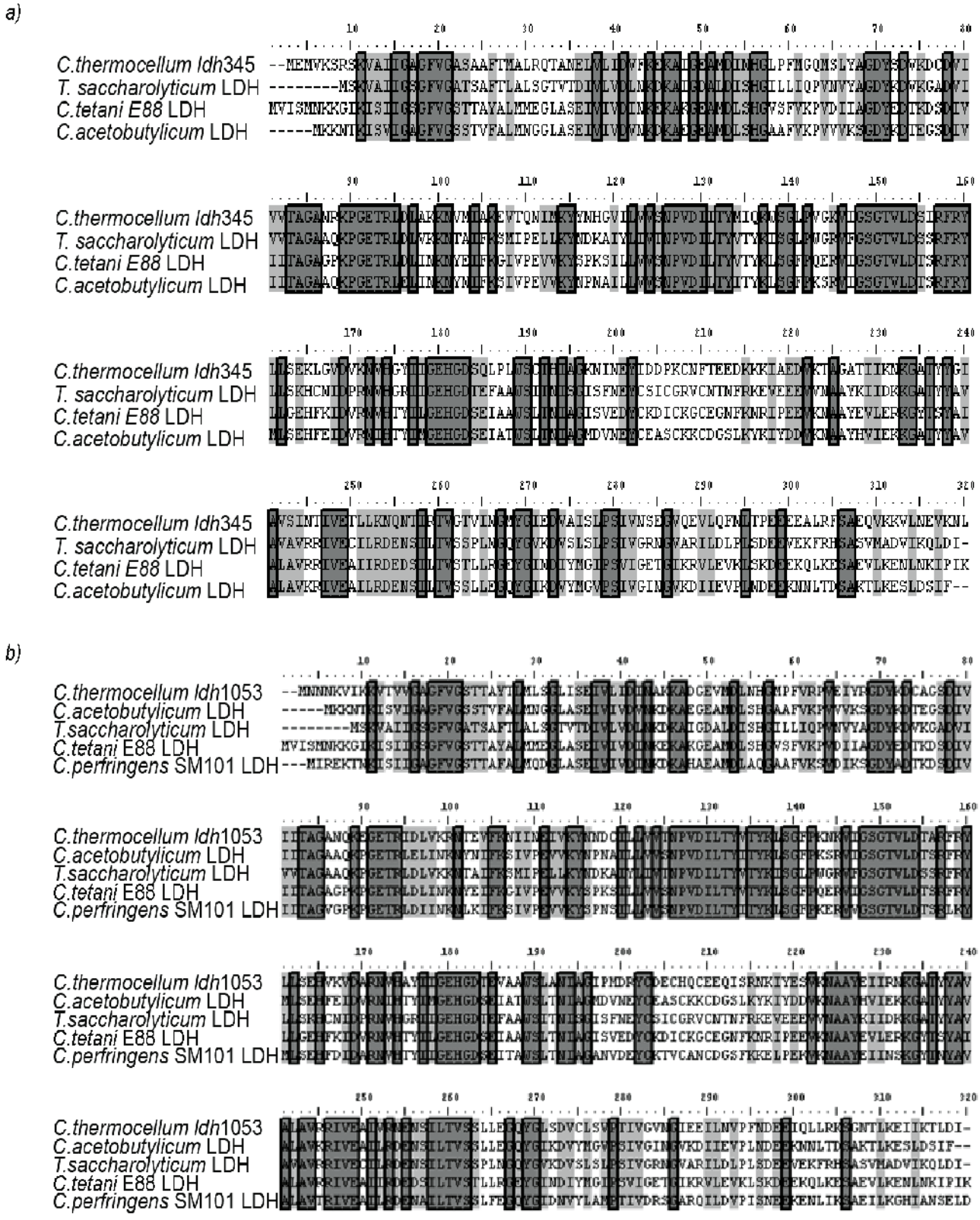

Fig. 3 C. *thermocellum* ATCC 27405 amino acid sequence alignment of the putative Lactate dehydrogenases encoded by gene 345 with the a) LDH encoded by *T. saccharolyticum*, *C. tetani* and *C. acetobutyliticum* and b) LDH encoded by gene 1053 aligned with LDH from *C. acetobutylicum*, *T. saccharolyticum*, *C. tetani* and *C. perfringens*. Identical residues are shaded grey and outlined while similar residues are shaded light grey.

345 putatively encodes a 318 amino acid polypeptide that shares 52 % amino acid sequence identity with an LDH from *Thermoanaerobacterium saccharolyticum* (E-value = 2e-86), and 46% amino acid sequence identity with LDHs from *C. tetani* and *C. acetobutylicum* (E-value = 6e-82 and 8e-82, respectively; Table 1). A putative malate dehydrogenase (EC 1.1.1.38), gene 344, is located immediately upstream suggesting an apparent malo-lactic fermentation operon. The malate dehydrogenase (MDH) encodes a polypeptide with 65 % amino acid similarity to the MDH of *Carboxydothermus hydrogenoformans* and 64 % amino acid sequence identity with the same enzyme in *Thermoanaerobacter tengcongensis* MB4 (E-values of 6e-136 and 4e-134 respectively).

Gene 1053 encodes a putative 317 amino acid polypeptide that shares 64 % identity with an LDH encoded by *Clostridium acetobutylicum* (E-value = 1.0e-111), 61 % amino acid sequence identity with LDHs encoded by *C. tetani* and *C. perfringens* (E-values = 1.0e-106 and 2e-97, respectively), and 59 % amino acid sequence identity with a second LDH encoded by *T. saccharolyticum* (E-value = 1.0e-104; Fig. 4b). This putative LDH has previously been cloned and expressed within *E. coli* strain FMJ39 and functionally characterized by Özkan *et al*. [19]. An analysis of genes 1053 and 345 for conserved amino acid motifs (rpsBLAST) revealed conserved MDH, LDH-like_MDH, Mdh and LDH_MDH domains, in addition to the LDH domain. RT-PCR products of the putative *C. thermocellum ldh* (gene 345) were detected throughout growth on α-cellulose. Weak amplification was observed between 5 and 10 hrs pi, while strong amplification was evident from samples taken 15-32 hrs pi. No amplification was observed during late stationary phase at 49 hrs (Fig. 2a). A similar pattern of amplification was observed for the RT-PCR products of the putative *C. thermocellum ldh* (gene 1053).

Our analyses have identified genes encoding both PFL and POR, which mediate oxidation of pyruvate to acetyl-CoA. *C. thermocellum pfl* (gene 505; Table 2) encodes a 742 aa polypeptide with a predicted molecular mass of 84,405 Da (TrEMBL accession number Q4CBR1). The identified PFL shares 72 % amino acid sequence identity with *Alkaliphilus metalliredigenes* QYMF, *Thermosynechococcus elongatus* BP-1 and *Bacillus cereus* subsp. cytotoxis NVH and 70 % identity with *Bacillus thuringiensis* serovar *konkukian* (E values =0.0; Fig. 4). Weak amplification of *pfl* was observed between 5 and 10 hrs pi, while strong amplification was observed during log phase, between 15 and 32 hrs pi. No amplification was observed during late stationary phase at 49 hrs (Fig. 2a).

As observed in other genomes, *C. thermocellum pfl* is adjacent to a gene encoding a PFL-activating enzyme (*act*, gene 506; Table 2), which is required for activation of PFL, but is transcribed independently in most organisms studied including *C. pasteurianum* [20]. This enzyme shares 53 % amino acid similarity with the PFL-AE encoded by *Bacillus cereus* G9241 (E score = 2e-73). A survey of the *C. thermocellum* genome revealed an additional three genes encoding putative PFL-AE like enzymes. Gene 1578 (229 amino acids), gene 1167 (289 amino acids) and gene 647 (280 amino acids) share 28%, 26% and 29% amino acid sequence homology, respectively, with the PFL-AE previously identified (gene 506). No RT-PCR products of the putative *C. thermocellum act* gene (506) were observed at 5 and 10 hrs pi. While weak RT-PCR products were observed at 15 hrs pi, the intensity of RT-PCR bands increased from 19 to 32 hrs pi, and no RT-PCR products of *pfl* were observed at 49 hrs pi (Fig. 2a).

Bioinformatic analyses using conserved amino acid sequence domains revealed the presence of two putative multi-subunit POR enzymes (genes 2794-2797 and genes 2390-2393) and a large open reading frame (ORF) encoding a single putative POR polypeptide (gene 3120) within the *C. thermocellum* genome (Table 2).

Genes 2794-2797 form an operon expressing putative POR γ (192 aa), δ (101 aa), α (394 aa), and β (311 aa) subunits respectively. The 192 aa ORF encoded by gene 2721 has a predicted molecular mass of 21,203 Da (TrEMBL accession # Q4CHL1) and rpsBLAST analysis has revealed the presence of the Por_G conserved domain which is associated with catalytic or regulatory function. The 311 aa ORF encoded by gene 2797 has a predicted molecular mass of 34,649 Da (TrEMBL accession # Q4CHL4) and contains both thiamine pyrophosphate and divalent cation binding domains while the ORF of 101 amino acids encoded by gene 2795 possesses putative domains for a 4Fe-4S center. The amino acid sequences of these gene products share varying levels of sequence identity with POR γ (70 %), POR δ (54 %), POR α (59 %), and POR β (65 %) subunits encoded by *C. tetani* (E-values = 4.0e-72, 2.0e-26, 1.0e-131, and 1.0e-114, respectively).

Genes 2390-2393 putatively encode a second multi-subunit POR enzyme. These gene products have amino acid sequence identity with POR enzymes encoded by thermophilic Archaea (43 % identity with *Methanopyrus kandleri* POR γ, 4.0e-34; 47 % identity with *Pyrococcus abyssi* POR δ, 4.0E-19; 50 % identity with *Methanobacterium thermoautotrophicum* POR α, 3.0e-95; and 49 % identity with *M. kandleri* POR β, 1.0E-77). BLAST analysis revealed no significant similarity between the γ (gene 2390), δ (gene 2391), α (gene 2392) and β (gene 2393) subunits with the corresponding POR encoding subunits (genes 2794-2797). rpsBLAST analysis of genes 2390-2393 revealed conserved

Table 2 Key enzymes identified within Clostridium thermocellum ATCC 27405 involved in pyruvate catabolism and hydrogen synthesis

Enzyme name	Gene	Subunits	Primers
Lactate dehydrogenase	1053	LDH	F:CAAAGACTGTGCCGGATCCGA
			R:GGCTGTGTCCAAAACCGTTCC
	345	LDH	F:GCCGGAGCCAACAGAAAACCT
			R:CGTCAACGCCCAATTTTTCGC
Pyruvate formate lyase	505	PFL	F:CCGAAGCTTATGGCCACAGTG
			R:GCAATCAGCCTGTCAACGCCA
Pyruvate Formate Lyase - Activating Enzyme	506		F: TTGGGACACTGGACGGACCG
			R:ATCCATTGGTATCCAGCGCCG
Pyruvate:ferredoxin Oxidoreductase*	2390-2393	γ, δ, α, β	F:GGGAAATGAAGCAGTGGCGGA
			R:AGCCCCTGGGATGAAGTTGCC
	2794-2797	γ, δ, α, β	F:ACCGATGATGCCGATGTTGCC
			R:AACAAAGGTCCTCCGGCTGCA
	3120	γ, δ, α, β	F:GCAGGGGCATTGACGACCACT
			R:TAATGGCCGAAAGATGCGAA
16S RNA			F:TGACGGGCGGTGTGTACAAGG
			R:GGTGGGGACGACGTCAAATCA
Fe-only hydrogenase	342		F:ATAATGGCCTGTCCCGGTGGT
			R:CGTGAGCTTTATGACTGCCCG
	430		F:TTCGAAAAGCGGGCATCAAGC
			R:CCCGACAGTGATTGGCAAGCA
	3003		F:TACAGCTGCAGCCGTGGTTCC
			R:CAAATCGCAGGTGAAACGGGC
NiFe hydrogenase	3024	echA	F:CCCAGATGCCCTTTTCCTCCT
			R:TGCGTCGCTTTTGGAGATTGC
	3023	echB	F:TTACGGCCCGCATTCCTTCAA
			R:GGCTGAACGGTGACCTGAAGG
	3022	echC	F:TTACGAACAAATGGCGGACCC
			R:TGCCTTTCCTCCAATATGCCG
	3021	echD	F:TTCCTTTGACAGCGGCAGTGA
			R:CCCGTTATATTCACCCCAAAA
	3020	echE	F:ATCCCCTTCGGCCCTCAACAT
			R:TATGCGTGACAGCTCTGCCCA

* RT-PCR primers designed to amplify the α subunit

amino acid domains consistent with those described within genes 2794-2797.

BLAST analyses also detected an ORF (gene 3120) containing the same functional domains observed within the two multi-subunit POR operons. This single large ORF (3,527 bp) is expected to encode one polypeptide similar to the single subunit POR isolated from *C. acetobutylicum* [21]. This predicted ORF of 1175 amino acids likewise contains domains for a 4Fe-4S center in addition to a Thiamine pyrophosphate binding domain and a catalytic or regulatory domain. BLASTp analysis revealed 75% amino acid sequence identity to the POR of *Thermoanaerobacter ethanolicus* ATCC 33223 (E score = 0.0).

RT-PCR products of gene 3120 were detected throughout growth on α-cellulose. Weak amplification was observed at 5 and 10 hrs pi, while strong amplification was apparent from samples taken 15-32 hrs pi. No amplification was observed during late stationary phase (Fig. 2a). Conversely, RT-PCR products of the putative *por* operon (genes 2794-2797) were detected during log phase growth on α-cellulose, between 15 and 32 hrs post-inoculation with some weak amplification observed at 49 hrs pi (Fig. 3).

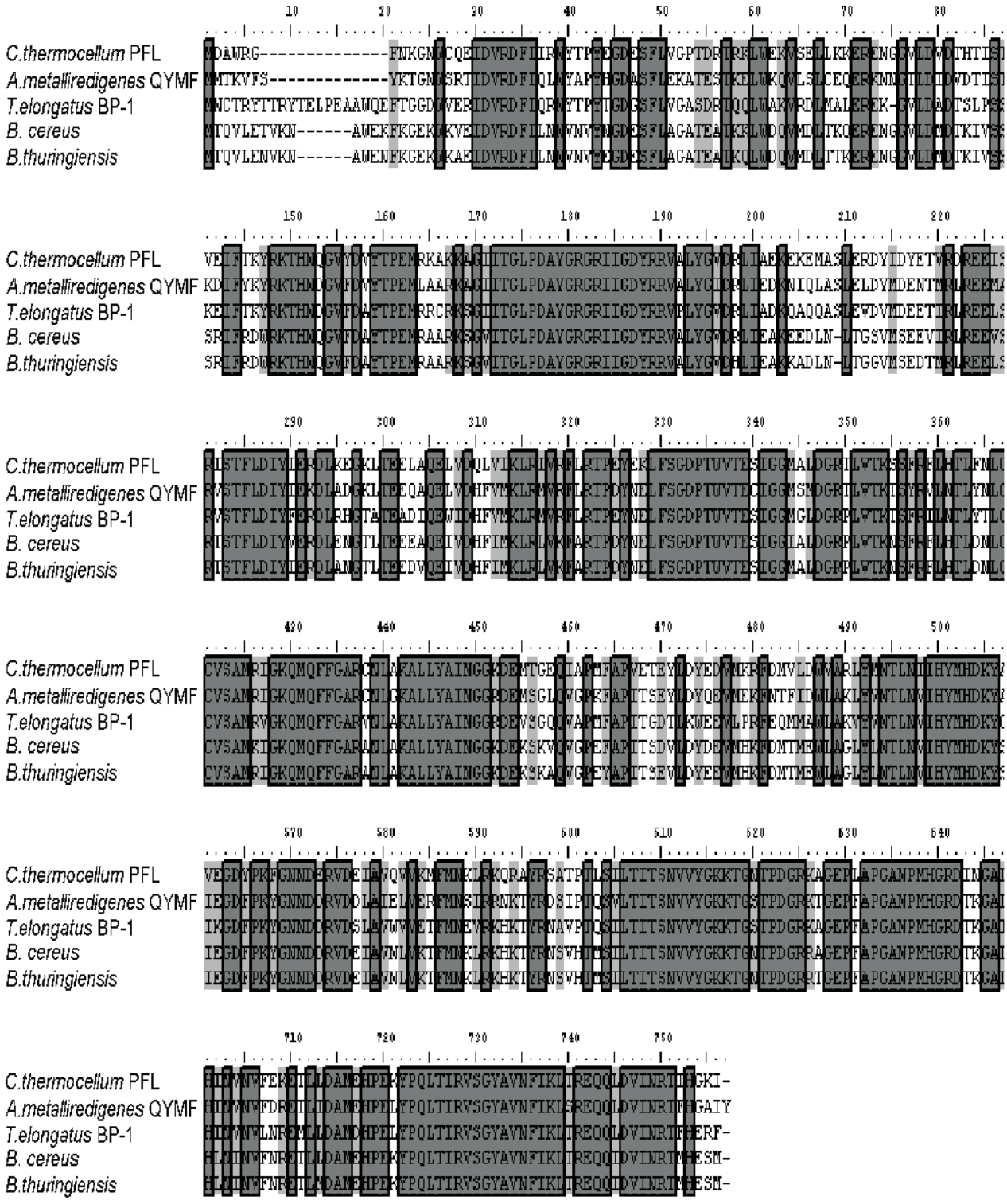

Fig. 4 Amino acid sequence alignment of the putative Pyruvate formate lyase encoded by *C. thermocellum* ATCC 27405 with Pyruvate formate lyases encoded by *A. metalliredigenes, T. elongatus*, *B. cereus* and *B. thuringiensis*. Identical residues are shaded grey and outlined while similar residues are shaded light grey.

Weak amplification of the *por* operon (genes 2390-2393) were observed only between 15 and 24 hrs pi.

Amplification of *C. thermocellum* 16S *rrna* was used as a positive control for RNA quality and RT-PCR amplification. Low levels of 16S *rrna* RT-PCR products were observable at 5 and 10 hrs pi, strong 16S *rrna* RT-PCR products were observed between 15 and 32 hrs pi, and weak 16S *rrna* RT-PCR products were observed at 49 hrs pi (Fig. 2a). The absence of detectable levels of RT-PCR products for *act* (gene 506), *por* (gene 2796), and *por* (gene 2392) and the presence of very weak bands for *ldh* (gene 345), *ldh* (gene 1053), *pfl* (gene 505), and *por* (gene 3120) at 5 and 10 hrs pi strongly contrast the low, but clearly detectable levels of 16S *rrna* RT-PCR products at these time points. While no significant lag in cell growth was observed following inoculation, and cell mass (measured as total protein) increased exponentially until approx. 29 hrs pi, we observed a lag in transcription levels across the genes probed.

Hydrogenases

Our analyses revealed the presence of at least three genes that encode putative subunits related to the NuoG component of Fe-only hydrogenase, and one NiFe hydrogenase on the basis of putative subunits encoding for sequences related to NuoBCD in *C. thermocellum* (Table 2).

Gene 342 encodes a 582 aa ORF that shares 99 % aa sequence identity (E-value = 1.0e-108) with a *C. thermocellum* hydrogenase 1 gene that was previously sequenced [22]. Gene 3003 encodes a 644 amino acid ORF with a predicted molecular mass of 71,780 Da (TrEMBL accession # Q4CD10). The amino acid sequence of this gene product shares 44 % sequence identity with the 75 kD subunit NADH dehydrogenase/NADH:ubiquinone oxidoreductase (NuoG) from *T. tengcongensis*. This is a putative Fe-only hydrogenase catalytic subunit (E-value = 1.7e-135) that contains a 4Fe-4S ferredoxin, iron-sulfur binding domain. Finally, gene 439 encodes a 566 aa ORF that shares 45 % aa sequence identity (E-value = 1.0e-140) with the *C. thermocellum* hydrogenase 1 gene that was previously sequenced [22]. As expected for the catalytic subunit of the Fe-hydrogenase, all three of these sequences share homology with NuoG (Fig. 5).

While these 3 sequences are clearly related, they are not identical, and are found each in a different context. Genes 342 and 430 are preceded by NuoF related NADH-binding subunits (genes 341 and 429), consistent with a function as NADH- dependent hydrogenase. Both putative hydrogenases share sequence homology with the purified NADH-dependent, Fe-only hydrogenase of *T. tengcongensis* [23]. In contrast, gene 3003 is not adjacent to a NuoF related sequence. Rather, it is adjacent to a NADPH-dependent glutamate synthase β-chain (gene 3004), the subunit that contains the NADPH-binding site [24]. These data suggest that the two genes may form a NADPH-dependent hydrogenase. If these genes encode gene products with completely separate functions, it is possible that gene 3003 encodes a protein that may function as a soluble, ferredoxin-dependent Fe-only hydrogenase. Nevertheless, NADP+ dependent hydrogenase activity has been detected in extracts of *C. thermocellum* (unpublished observations).

Genes 3024-3013 correspond to genes TTE0123-TTE0134 from *T. tengcongensis*, which encode a membrane bound, ferredoxin-dependent, energy converting (Ech) NiFe-hydrogenase (Table 2) and associated hydrogenase maturation proteins. Amino acid sequence analysis reveals the presence of transmembrane peptide regions in genes 3024 and 3023, consistent with EchA and EchB subunits (37 %, 1e-115 and 41 %, 3e-62 amino acid identity to *T. tengcongensis*). Both the N and C-terminal conserved NiFe binding motifs associated with a NiFe hydrogenase large subunit (N-terminal RXCXXCXXXH and C-terminal DPCXXCXX(H/R)) are present within gene 3020 and this corresponds to the EchE subunit of *T. tengcongensis* (63 %, 1e-135). Although gene 3022 corresponds to an *ech*C subunit (68 %, 8e-60 amino acid identity to *T. tengcongensis*); analysis of the putatively identified protein indicates that one of the four conserved cysteine residues required for binding a 4Fe-4S center has been replaced with a glutamic acid residue at position 24 of the predicted peptide. Two putative 4Fe-4S binding motifs ($CX_2CX_2CX_3CP$) consistent with other EchF subunits [25] have been identified within gene 3019. RT-PCR of the identified *ech* genes confirmed transcription throughout growth on –cellulose (Fig. 2b).

There is evidence in the *C. thermocellum* genome for an Rnf type membrane associated ferredoxin NAD oxidoreductase. On the basis of the *C. tetani* genome [26], which has a putative Na+-pumping NADH ferredoxin oxidoreductase, genes in *C. thermocellum* have been found corresponding to RnfCDGEAB. Genes 2430-2435 exhibit 46 % (1e-110), 54 % (3e-96), 41 % (1e-28), 63 % (7e-66), 61 % (2e-62) and 47 % (3e-72) amino acid identity to the corresponding RnfCDGEAB genes of *C. tetani*.

Discussion

We have detected several genes in the *C. thermocellum* genome that encode gene products with amino acid sequences that are consistent with key enzymes in pyruvate catabolic pathways principally mediated by the oxidation of pyruvate to lactate by LDH or to acetyl-CoA by PFL and POR. Our

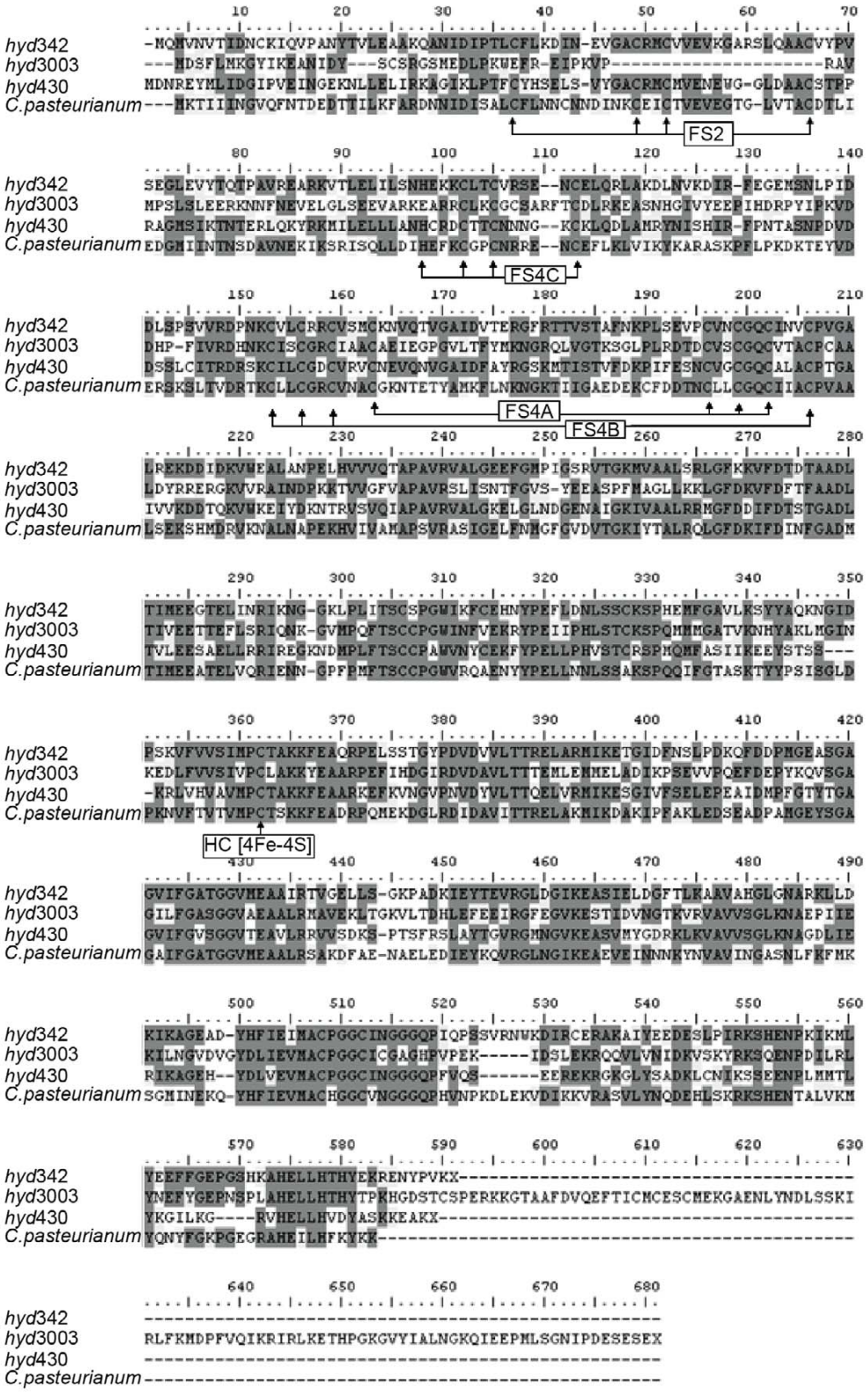

Fig. 5 ClustalW analysis of the predicted amino acid sequences from the identified Fe-only hydrogenases of *C. thermocellum* (*hyd*342, *hyd*3003 and *hyd*430) against the Fe-only hydrogenase of *C. pasteurianum* (gi|4139441). The predicted [4Fe-4S] binding domains FS4A, FS4B and FS4C, [2Fe-2S] binding domain (FS2) and active-site [4Fe-4S] subcluster binding site (HC [4Fe-4S]) are indicated

data suggest that conversion of pyruvate to acetyl-CoA occurs via both the PFL and POR mediated pathways in *C. thermocellum*, consistent with previous observations of end products [7, 11, 12].

Two ORF's putatively encoding *ldh* genes have been identified (genes 345 and 1053). Lactate production has previously been described within other clostridial species including *C. cellulolyticum* [27], *C. pasteurianum* [28], and *C. acetobutyliticum*. LDH gene 1053 has previously been cloned and expressed in *E. coli* and has been demonstrated to oxidize pyruvate to lactate [19]. The L-lactate dehydrogenase family includes both LDH and malate dehydrogenase (MDH; EC 1.1.1.37) enzymes, which are structurally similar and use the same coenzyme (NADH) to catalyze a redox interconversion by similar mechanisms. It has been demonstrated experimentally that the substitution of a single amino acid at position 86 can change the substrate binding specificity from lactate to malate [29], and it is therefore common for these genes to be misannotated when analyzed solely on the basis of sequence homology.

Conversion of pyruvate to formate via PFL occurs in enteric bacteria such as *E. coli* as well as in obligate anaerobes including members of the genus *Clostridium* [4]. Recently, we demonstrated that *C. thermocellum* synthesizes formate during exponential growth on cellobiose or cellulosic substrates [7]. Using RT-PCR, we demonstrated the presence of *pfl* transcripts during *C. thermocellum* growth on these substrates. We also found that the *C. thermocellum* genome contains the *act* gene, which encodes the *E. coli* PFL-activating enzyme, and *adh*E, which encodes an enzyme (ADH-E) known to negatively regulate formate synthesis in other Gram positive bacteria [5, 30] and demonstrated the presence of *act* and *adh*E transcripts during *C. thermocellum* growth on cellobiose. As within the genomes of *E. coli* and *C. pasteurianum*, we detected the presence of an *act* gene immediately downstream from *pfl* within the *C. thermocellum* genome. In *E. coli*, *pfl* and *act* are transcribed independently with *pfl* under the transcriptional control of several different promoter regions and *act* under the control of its own constitutive promoter [16]. Independent transcriptional regulation has also been observed within *C. pasteurianum* and it is likely, based on genetic organization of the two genes, that *C. thermocellum* exhibits a similar transcriptional uncoupling. In *Streptococcus mutans, pfl* and *act* are not closely linked in the genome [31]. With the low amino acid sequence identity among them, the roles of the putative *C. thermocellum act* gene products in activation of PFL and the pyruvate catabolism remain to be demonstrated.

The catabolism of pyruvate via the PFL mediated pathway in *C. thermocellum* likely does not serve as an overflow pathway at the pyruvate branch point since both POR and PFL appear to be expressed simultaneously under the growth conditions tested. Furthermore, oxidation via this pathway aids in recycling reducing equivalents through the conversion of acetyl-CoA to ethanol, while still reserving a portion of the Acetyl-CoA for ATP synthesis resulting in increased ATP production relative to lactate production only. The observed dramatic increase in lactate production at 24 hours, as the cells near stationary phase under cellulose excess, may illustrate an inability of both the *pfl* and *por* mediated pathways to support the carbon flow at the pyruvate branch point. Lactate production could therefore represent an overflow pathway as the catabolism of pyruvate to lactate does not appear to compete with the ATP generating pathways (*por* and *pfl* mediated) but does represent a source for recycling reducing equivalents.

The purpose for 3 distinct, simultaneously expressed POR is unclear. Pyruvate, 2-ketoisovalerate-, α-ketoglutarate-, and indolepyruvate:ferredoxin oxidoreductases are phylogenetically related [32]. However, since the latter 3 are primarily involved in amino acid catabolism, their presence at high level is unlikely. Another possibility would be that these POR might be differentially expressed at different growth temperatures. While the 4 subunit type of POR seems to be most common in hyperthermophiles [33], its presence has also been observed in mesophiles [34]. Conversely the single subunit type, also found in the mesophile *C. acetobutylicum* has an *in vitro* temperature optimum of 60°C [21].

We have detected several genes in the *C. thermocellum* genome that encode gene products with amino acid sequences consistent with key enzymes associated with H_2 metabolism. Ferredoxin dependent H_2 evolution has previously been described within other fermentative species including *C. pasteurianum* and *T. tengcongensis* [23, 35]. In these species, glucose is fermented via the Embden-Meyerhoff pathway, with pyruvate being oxidized to acetyl-CoA by pyruvate:ferredoxin oxidoreductase (POR). In *C. pasteurianum*, reduced ferredoxin functions as an electron donor for two soluble monomeric Fe-only hydrogenases [35]. The presence of the *echABCDEF* operon (genes 3013-3024), however, suggests that *C. thermocellum* employs a membrane-bound Fd-dependent NiFe hydrogenase similar to that described in *T. tengcongensis.* In both *C. thermocellum* and *T. tengcongensis*, the *echABCDEF* genes are predicted to encode two membrane-bound proteins (EchA and B), a conserved hydrogenase small subunit with one [4Fe-4S] cluster (EchC), an additional protein with two [4Fe-4S] cluster binding motifs, a NiFe hydrogenase large subunit (EchE), and a hydrophilic subunit with no predicted cofactor-binding site (Ech D) [23, 25].

In both *P. furiosus* and *T. tengcongensis*, membranes containing the partially purified Ech hydrogenase preparations were able to catalyze H_2 production with either reduced methylviologen or ferredoxin as the electron donor. Furthermore, the catalytic efficiency coefficient (K_{cat}/K_m) of the *T. tengcongensis* enzyme was found to be 7.3×10^7 M^{-1} s^{-1}, strongly suggesting that H_2 evolution is the physiological reaction catalyzed by this hydrogenase [23]. On the basis of sequence identity, it therefore seems likely that Fd-dependent H_2 production in *C. thermocellum* is likely accomplished through the action of the Ech hydrogenase identified. The use of an Ech type hydrogenase would be expected to allow the cells to take advantage of the difference in Eh between reduced ferredoxin and hydrogen by coupling this reaction with the generation of a transmembrane proton gradient [36]

Our analyses revealed the presence of three ORFs with significant homology with NuoG as expected for the catalytic subunit of Fe-only hydrogenases [36]. These show homology to the hydrogenase of *C. pasteurianum*. The Fe-only hydrogenase of *C. pasteurianum* is a soluble Fd-dependent enzyme that catalyzes H_2 evolution from reducing equivalents generated by the action of POR. As observed in *C. pasteurianum*, two [4Fe-4S] binding domains immediately adjacent to the active site were detected within the *C. thermocellum* Fe-only hydrogenases (genes 342, 3003 and 432). These domains, FS4A and FS4B, coordinate [4Fe-4S] binding by four cysteine residues. They are suspected to mediate electron transport from the initial electron acceptors located on the proteins surface, the FS2 and FS4C domains, to the active-site cluster [35]. It should be noted that gene 3003 is missing critical residues within the FS2 and FS4C domains suggesting an alternative method of initial electron acceptance from the physiological electron donor. The presence of ORFs putatively encoding NuoF homologous sequences (genes 341 and 429) is consistent with Fe-only NADH hydrogenases [36]. An NADPH (gene 3004) binding subunit adjacent to the identified Fe-only hydrogenase subunit (gene 3003) suggests that this enzyme may utilize NADPH as electron an donor.

Balancing the cells need for electrons as NADH and ferredoxin may also require a mechanism for shuttling electrons from one cofactor to the other. Indeed this type of activity has been described in the early literature of the clostridiales [37]. These have typically been described as small, independent, flavin containing proteins [38], and have been implicated in the transfer of electrons between NADH and ferredoxin for the synthesis of H_2. A putative 6 subunit operon consistent with a membrane integral Na+ pumping NADH:Fd oxidoreductase previously described in proteolytic clostridia [37], however, has been detected in *C. thermocellum*. In the NAD+ reduction direction it would be expected to take advantage of the Eh difference between Fd_r and NADH to generate a transmembrane ion motive force.

Conclusions

Pyruvate is a key intermediate during fermentation and represents a critical branch point with respect to managing bacterial energy requirements. The *C. thermocellum* genome contains genes that encode two putative lactate dehydrogenases, one putative pyruvate:formate lyase, four putative pyruvate:formate lyase activating enzymes, and at least three putative POR or POR-like enzymes. Hydrogen synthesis appears to occur only in the pyruvate:ferredoxin oxidoreductase-mediated pathway during pyruvate catabolism, as the *C. thermocellum* genome does not contain genes for formate dehydrogenase, a major component of the H_2 evolving complex formate:hydrogen lyase. *C. thermocellum* encodes hydrogenases that putatively enable NADH, NADPH and Fd-dependent hydrogen production through the action of both Fe-only and membrane-bound NiFe containing enzymes. As within other species, including *T. tengcongensis* and *C. pasteurianum*, these enzymes serve to release excess reducing equivalents generated during fermentation as H_2 gas. Although transcription of these genes has been confirmed, the actual functional roles of the gene products relating to H_2 synthesis remain to be elucidated and will need to be confirmed using both expression analysis and protein characterisation. This is especially true for families of enzymes in which slight alterations of structure or different accessory subunits may alter enzyme specificity and function.

A critical step in understanding the metabolic and genetic mechanisms by which H_2 is synthesized is to identify the genes encoding enzymes in H_2-synthesizing pathways. The emergence of bioinformatic tools has allowed the relationship between genome content, gene expression, and end-product synthesis to be investigated. With respect to H_2 production, the findings presented here will help cultivate future strategies aiming to influence metabolic flux towards increased production.

Acknowledgements This work was supported by funds provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), through a Strategic Programs grant (STPGP 306944-04), the BIOCAP Canada Foundation, and by the Manitoba Conservation Sustainable Development and Innovation Fund. VCK acknowledges Overseas Associate-ship, Department of Biotechnology and Director, IGIB, CSIR for Government of India for support.

References

1. Lamed R and Zeikus G (1980) Ethanol production by thermophilic bacteria: Relationship between fermentation product yields of and catabolic enzyme activities in *Clostridium thermocellum* and *Thermoanerobium brockii*. J Bacteriol 144:569–578
2. Lynd LR and Grethlein HG (1987) Hydrolysis of dilute acid pretreated hardwood and purified microcyrstalline cellulose by cell-free broth from *Clostridium thermocellum*. Biotechnol Bioeng 29:92–100
3. Ng TK, Weimer PJ and Zeikus JG (1977) Cellulolytic and physiological properties of *Clostridium thermocellum*. Arch Microbiol 114:1–7
4. Patni NJ and Alexander JK (1971a) Catabolism of fructose and mannitol by *Clostridium thermocellum*: Presence of phosphoenolpyruvate:fructose phosphotransferase, fructose-1-phosphate kinase, phosphoenol- pyruvate:mannitol phosphotransferase, and mannitol-1-phosphate dehydrogenase in cell extracts. J Bacteriol 105:226–231
5. Patni NJ and Alexander JK (1971b) Utilization of glucose by *Clostridium thermocellum*: Presence of glucokinase and other glycolytic enzymes in cell extracts. J Bacteriol 105:220–225
6. Thauer RK, Jungermann KA and Decker K (1977) Energy conservation in chemotrophic anaerobic bacteria. Bacteriol Rev 41:100–180
7. Sparling R, Islam R, Cicek N, Carere C, Chow H and Levin DB (2006) Formate synthesis by *Clostridium thermocellum* during anaerobic fermentation. Can J Microbiol 52: 681–688
8. Demain AL, Newcomb M and Wu JHD (2005) Cellulase, Clostridia, and ethanol. Microbiol Mol Biol Rev 69: 124–154
9. Lynd LR, Weimer PJ, van Zyl WH and Pretorius IS (2002) Microbial cellulose utilization: Fundamentals and biotechnology. Micro Mol Biol Rev 66:506–577
10. Lynd LR, Grethlein HG and Wolkin RH (1989) Fermentation of cellulose substrates in batch and continuous culture by *Clostridium thermocellum*. App Environ Microbiol 55: 3131–3139
11. Islam R, Cicek N, Sparling R and Levin DB (2006) Effect of substrate loading on hydrogen production during anaerobic fermentation by *Clostridium thermocellum* 27405. Appl Microbiol Biotechnol 72(3):576–583
12. Levin DB, Sparling R, Islam R and Cicek N (2006) Hydrogen production by *Clostridium thermocellum* 27405 from cellulosic biomass substrates. Int J Hydrogen Energy 31(11): 1496–1503
13. Charon MH, Volbeda A, Chabriére E, Pieulle L and Fontecilla-Camps JC (1999) Structure and electron transfer mechanism of pyruvate:ferredodin oxidoreductase. Curr Opin Struct Biol 9:663–669
14. Hallenbeck PC and Benemann JR (2002) Biological hydrogen production; fundamentals and limiting processes. Int J Hydrogen Energy 27:1185–1193
15. Hallenbeck PC (2005) Fundamentals of the fermentative production of hydrogen. Water Sci Technol 52:21–29
16. Sauter M and Sawers G (1990) Transcriptional analysis of the gene encoding Pyruvate formate lyase activating enzyme of *Escherichia coli.* Mol Microbiol 4:355–363
17. Bradford MM (1976) A rapid and sensitive method for the estimation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254
18. Sirko A, Zehelein E, Freundlich M and Sawers G (1993) Integration host factor is required for anaerobic pyruvate induction of *pfl* operon expression in *Escherichia coli*. J Bacteriol 175:5769–5777
19. Özkan M, Ylmaz E, Lynd LR and Özcengiz G (2004) Cloning and Expression of the *Clostridium thermocellum* L-lactate Dehydrogenase in *Escherichia coli* and Enzyme Characterization. Can J Microbiol 50:845–851
20. Weidner G and Sawers G (1996) Molecular characterization of the genes encoding pyruvate formate-lyase and its activating enzyme of *Clostridium pasteurianum*. J Bacteriol 178: 2440–2444
21. Meinecke B, Bertram J and Gottschalk G (1989) Purification and characterization of the pyruvate-ferredoxin oxidoreductase of *Clostridium acetobutylicum*. Arch Microbiol 152: 244–250
22. Desai SG, Steven DM, Prince HL, Guerinot ML, Lynd LH (1999) *Clostridium thermocellum* hydrogenase 1. GenBank accession # Q9XC55. Direct Submission
23. Soboh B, Linder D and Hedderich R (2004) A multisubunit membrane-bound [NiFe] hydrogenase and an NADH-dependent Fe-only hydrogenase in the fermenting bacterium *Thermoanaerobacter tengcongensis*. Microbiology 150: 2451–2463
24. Vanoni MA, Verzotti E, Zanetti G and Curti B (1996 Properties of the recombinant b subunit of glutamate synthase. European J Biochem 236:937–946
25. Forzi L, Koch J, Guss AM, Radosevich CG, Metcalf W and Hedderich R (2005) Assignment of the [4Fe-4S] clusters of Ech hydrogenase from *Methanosarcina barkeri* to individual subunits via the characterization of site-directed mutants. FEBS Journal 272:4741–4753
26. Bruggemann H, Baumer S, Fricke WF, Wiezer A, Liesegang H, Decker I, Herzberg C, Martinez-Arias R, Merkl R, Henne A and Gottschalk G (2003) The genome sequence of *Clostridium tetani*, the causative agent of tetanus disease. Proc Natl Acad Sci USA 100:1316–1321
27. Guedon E, Payot S, Desvaux M and Petitdemanger H (1999) Carbon and electron flow in *Clostridium cellulolyticum* grown in chemostat culture on synthetic medium. J Bacteriol 181:3262–3269
28. Dabrock B, Bahl H and Gottschalk G (1992) Parameters affecting solvent production in *Clostridium pasteurianum*. Appl Environ Microbiol 58:1233–1239
29. Viles F and Silverman L (1949) Determination of starch and cellulose. Anal Chem 21:950–953
30. Thauer RK, Kirchniawy FH and Jungermann KA (1972) Properties and function of the pyruvate-formate-lyase reaction in clostridiae. Eur J Biochem 23:282–290
31. Vasconcelos I, Girbal L and Soucaille P (1994) Regulation of carbon and electron flow in *Clostridium acetobutyliticum* grown in chemostat culture at neutral pH on mixtures of glucose and glycerol. J Bacteriol 176(5): 1443–1450
32. Kletzin A and Adams MWW (1996) Molecular and phylogenetic characterization of pyruvate and 2-ketoisovalerate ferredoxin oxidoreductases from *Pyrococcus furiosis* and

pyruvate ferredoxin oxidoreductase from *Thermotoga maritime*. J Bacteriol 178:248–257
33. Kunow J, Linder D and Thauer RK (1995) Pyruvate:ferredoxin oxidoreductase from sulfate reducing *Archaeoglubus fulgidis*: molecular composition, catalytic properties and sequence alignments. Arch Microbiol 63:21–28
34. Hughes NJ, Chalk PA, Clayton CL and Kelly DJ (1995) Identification of carboxylation enzymes and characterization of a novel four-subunit Pyruvate:Flavodoxin Oxidoreductase from *Helicobacter pylori*. J Bacteriol 177(14):3953–3959
35. Peters JW, Lanzilotta WN, Lemon BJ and Seefeldt LC (1998) X-ray crystal structure of the Fe-Only hydrogenase (CpI) from *Clostridium pasteurianum* to 1.8 Angstrom resolution. Science, 282:1853–1858
36. Vignais PM, Billoud B and Meyer J (2001) Classification and phylogeny of hydrogenases. FEMS Microbiol Reviews 25:455–501
37. Jungermann K, Thauer RK, Leimenstoll G and Decker K (1973) Function of reduced pyridine nucleotide-ferredoxin oxidoreductases in saccharolytic *Clostridia.* Biochimica et Biophysica Acta – Bioenergetics, 305:268–280
38. Chen YP and Yoch DC (1989) Isolation, characterization and biological activity of ferredoxin-NAD+ reductase from the methane oxidizer *Methylosinus trichosporium* OB3b. J Bacteriol 171:5012–5016
39. NakamuraY, Kaneko T, Sato S, Ikeuchi M, Katoh H, Sasamoto S, Watanabe A, Iriguchi M, Kawashima K, Kimura T, Kishida Y, Kiyokawa C, Kohara M, Matsumoto M, Matsuno A, Nakazaki N, Shimpo S, Sugimoto M, Takeuchi C, Yamada M and Tabata S (2002) Complete genome structure of the thermophilic cyanobacterium *Thermosynechococcus elongatus* BP-1. DNA Res 9(4):123–130
40. Desai SG, Guerinot ML and Lynd LR (2004) Cloning of L-lactate dehydrogenase and elimination of lactic acid production via gene knockout in *Thermoanaerobacterium saccharolyticum* JW/SL-YS485. Appl Microbiol Biotechnol 65(5):600–605
41. Nolling J, Breton G, Omelchenko MV, Markarova KS, Zeng Q, Gibson R, Lee HM, Dubois J, Qiu D, Hitti J, Wolf YI, Tatusov RL, Sabathe F, Doucette-Stamm L, Soucaille P, Daly MJ, Bennett GN, Koonin EV and Smith DR (2001) Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. J Bacteriol 183(16): 4823–4838
42. Myers GS, Rasko DA, Cheung JK, Ravel J, Seshadri R, DeBoy RT, Ren Q, Varga J, Awad MM, Brinkac LM, Daugherty SC, Haft DH, Dodson RJ, Madupu R, Nelson WC, Rosovitz MJ, Sullivan SA, Khouri H, Dimitrov GI, Watkins KL, Mulligan S, Benton J, Radune D, Fisher DJ, Atkins HS, Hiscox T, Jost BH, Billington SJ, Songer JG, McClane BA, Titball RW, Rood JI, Melville SB and Paulsen IT (2006) Skewed genomic variability in strains of the toxigenic bacterial pathogen, *Clostridium perfringens*. Genome Res 16(8): 1031–1040