

Identification of the [FeFe]-Hydrogenase Responsible for Hydrogen Generation in *Thermoanaerobacterium saccharolyticum* and Demonstration of Increased Ethanol Yield via Hydrogenase Knockout^{∇†}

A. Joe Shaw,^{1,2} David A. Hogsett,² and Lee R. Lynd^{1,2,3*}

Thayer School of Engineering, Dartmouth College, Hanover, New Hampshire¹; Mascoma Corporation, Lebanon, New Hampshire²; and Department of Biological Sciences, Dartmouth College, Hanover, New Hampshire³

Received 9 April 2009/Accepted 25 July 2009

Three putative hydrogenase enzyme systems in *Thermoanaerobacterium saccharolyticum* were investigated at the genetic, mRNA, enzymatic, and phenotypic levels. A four-gene operon containing two [FeFe]-hydrogenase genes, provisionally termed *hfs* (hydrogenase-Fe-S), was found to be the main enzymatic catalyst of hydrogen production. *hfsB*, perhaps the most interesting gene of the operon, contains an [FeFe]-hydrogenase and a PAS sensory domain and has several conserved homologues among clostridial saccharolytic, cellulolytic, and pathogenic bacteria. A second hydrogenase gene cluster, *hyd*, exhibited methyl viologen-linked hydrogenase enzymatic activity, but *hyd* gene knockouts did not influence the hydrogen yield of cultures grown in closed-system batch fermentations. This result, combined with the observation that *hydB* contains NAD(P)⁺ and FMN binding sites, suggests that the *hyd* genes are specific to the transfer of electrons from NAD(P)H to hydrogen ions. A third gene cluster, a putative [NiFe]-hydrogenase with homology to the *ech* genes, did not exhibit hydrogenase activity under any of the conditions tested. Deletion of the *hfs* and *hydA* genes resulted in a loss of detectable methyl viologen-linked hydrogenase activity. Strains with a deletion of the *hfs* genes exhibited a 95% reduction in hydrogen and acetic acid production. A strain with *hfs* and *ldh* deletions exhibited an increased ethanol yield from consumed carbohydrates and represents a new strategy for engineering increased ethanol yields in *T. saccharolyticum*.

Thermophilic anaerobic bacteria have long been of interest for studies of cellulosic biomass conversion due to their native hydrolytic and fermentative abilities (5, 33). However, all thermophilic anaerobes isolated to date have branched fermentation pathways which produce organic acids in addition to solvents such as ethanol (12). For cellulosic fuel production, significant yield loss is likely to be economically unfeasible (11).

In their natural environments, saccharolytic fermentative bacteria participate in interspecies hydrogen transfer, producing hydrogen from carbohydrates that is rapidly consumed by methanogens and sulfate-reducing bacteria (30). As a result, the hydrogen partial pressure remains exceedingly low, allowing hydrogen (E_0' , -414 mV) to be produced not only from ferredoxin (E_0' , ~ -400 mV) but also from the less favorable electron source NAD(P)H (E_0' , -320 mV). Fermentative bacteria benefit from hydrogen production, because they are able to coproduce acetic acid and an additional ATP via acetate kinase (23). When grown in pure culture in a closed fermentation vessel, hydrogen is generated primarily from reduced ferredoxin, since generation from NAD(P)H becomes less favorable as the concentration of hydrogen increases (7).

We have recently demonstrated high-yield ethanol production in the thermophilic anaerobe *Thermoanaerobacterium saccharolyticum* JW/SL-YS485 through deletion of the L-lactate dehydrogenase (*ldh*), phosphate acetyltransferase (*pta*), and acetate kinase (*ack*) genes (20). In addition to producing ethanol at high yield, this strain produced significantly less hydrogen, as is required to balance end product electron stoichiometry, although hydrogenase activity in cell extracts remained high. In this study, we used gene knockout to identify gene clusters that are implicated in hydrogenase activity in *T. saccharolyticum* and to identify the *hfs* gene operon, which is required for hydrogen production. The *hfs* operon contains a protein with [FeFe]-hydrogenase and PAS sensory domains that is conserved among a few members of the genera *Clostridium* and *Thermoanaerobacter*. Strains with *hfs* deletions showed decreased acetic acid production, and a strain with *hfs* and *ldh* deletions produced ethanol at an increased yield.

MATERIALS AND METHODS

Media and growth conditions. *T. saccharolyticum* JW/SL-YS485 was grown in modified DSMZ 122 medium, containing (per liter) 1.3 g of $(\text{NH}_4)_2\text{SO}_4$, 2.6 g of $\text{MgCl}_2 \cdot 6 \text{H}_2\text{O}$, 1.43 g of KH_2PO_4 , 1.8 g of K_2HPO_4 , 0.13 g of $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 6 g of Na- β -glycerophosphate, 0.00013 g of $\text{FeSO}_4 \cdot 7 \text{H}_2\text{O}$, 4.5 g of yeast extract, 0.002 g of resazurin, 0.5 g of L-cysteine-HCl, and 10 g of agarose (for solid media only). Biochemicals were from Sigma-Aldrich, and yeast extract was from BD Difco. The cellobiose concentration and pH were adjusted for different purposes in this study. A cellobiose concentration of 5 g/liter was used for *T. saccharolyticum* routine growth, transformation recovery, and fermentation with knockout strains, while a cellobiose concentration of 1.8 g/liter was used for fermentation of different strains under conditions that included an N_2 or H_2 atmosphere. After *T. saccharolyticum* transformation, the medium was adjusted to pH 6.7 for se-

* Corresponding author. Mailing address: Dartmouth College, Thayer School of Engineering, 8000 Cummings Hall, Hanover, NH 03755. Phone: (603) 646-2231. Fax: (603) 646-2277. E-mail: lee.r.lynd@dartmouth.edu.

† Supplemental material for this article may be found at <http://jbb.asm.org/>.

[∇] Published ahead of print on 31 July 2009.

TABLE 1. Plasmids and strains used in this study

Plasmid or strain	Description ^a	Source or reference
Plasmids		
pMQ87	Cloning vector for yeast homologous recombination, Gen ^r , <i>ura3</i>	Presque Isle culture collection
pBluescriptSK II(+)	General-purpose cloning vector, Amp ^r	Stratagene
pSGD8-erm	<i>ldh</i> knockout, Ery ^r , Amp ^r	20
pIKM1	<i>T. saccharolyticum</i> - <i>E. coli</i> shuttle vector	6
pHKO2	<i>hydA</i> knockout, Ery ^r , Amp ^r	This study
pHKO3	<i>hyd</i> cluster II knockout, Kan ^r , Amp ^r	This study
pHKO7	<i>hyd</i> knockout, Kan ^r , Gen ^r , <i>ura3</i>	This study
pHKO8	<i>ech</i> knockout, Kan ^r , Gen ^r , <i>ura3</i>	This study
pHKO9	<i>ech</i> knockout, Ery ^r , Gen ^r , <i>ura3</i>	This study
pHKO10	<i>hfs</i> knockout Kan ^r , Gen ^r , <i>ura3</i>	This study
Strains		
NEB5 α	<i>E. coli</i> cloning strain	New England Biolabs
JW/SL-YS485	<i>T. saccharolyticum</i> wildtype strain	Juergen Wiegell
ALK2	<i>T. saccharolyticum</i> lacking <i>ldh</i> , <i>pta</i> , and <i>ack</i> , Kan ^r , Ery ^r	20
HKO1	<i>T. saccharolyticum</i> lacking <i>hyd</i> , Kan ^r	This study
HKO2	<i>T. saccharolyticum</i> lacking <i>ech</i> , Ery ^r	This study
HKO3	<i>T. saccharolyticum</i> lacking <i>hfs</i> , Kan ^r	This study
HKO4	<i>T. saccharolyticum</i> lacking <i>hyd</i> and <i>ech</i> , Kan ^r , Ery ^r	This study
HKO5	<i>T. saccharolyticum</i> lacking <i>hfs</i> and <i>ech</i> , Kan ^r , Ery ^r	This study
HKO6	<i>T. saccharolyticum</i> lacking <i>hfs</i> and <i>hydA</i> , Kan ^r , Ery ^r	This study
HKO7	<i>T. saccharolyticum</i> lacking <i>ldh</i> and <i>hfs</i> , Ery ^r , Kan ^r	This study
ATCC 27405	<i>Clostridium thermocellum</i>	Lynd laboratory culture
H10	<i>Clostridium cellulolyticum</i>	ATCC
DSM 8903	<i>Caldicellulosiruptor saccharolyticus</i>	DSMZ
39E	<i>Thermoanaerobacter ethanolicus</i> ATCC 33223	ATCC
ISDg	<i>Clostridium phytofermentans</i>	ATCC

^a Gen^r, gentamicin resistant; Amp^r, ampicillin resistant; Ery^r, erythromycin resistant; Kan^r, kanamycin resistant.

lection with kanamycin (200 μ g/ml) or to pH 6.1 for selection with erythromycin (10 μ g/ml). The pH was adjusted to 6.8 for fermentation with different strains under conditions that included an N₂ or H₂ atmosphere and to 6.4 for fermentation with *T. saccharolyticum* knockout strains and for routine growth. Batch cultures used for fermentation analysis were inoculated with 6% (vol/vol) of an overnight culture in the same medium and incubated without shaking at 55°C in Hungate tubes (8 ml of liquid volume and 8 ml of headspace). Determination of growth curves of knockout strains was performed using a Powerwave XS microplate reader (Biotek, Winooski, VT) specially outfitted for incubation at 55°C and a 96-well plate with 100 μ l of media per well sealed with microplate cover tape.

Strains and plasmids. Strains and plasmids used in this study are listed in Table 1. *T. saccharolyticum* was a gift from the laboratory of Juergen Wiegell. *Clostridium thermocellum* 27405 was from a stock maintained by the Lynd laboratory, originally from the laboratory of Arnold Demain. All other strains were from culture collections or commercial sources.

Plasmid construction. Knockout plasmids for *T. saccharolyticum* were generally designed as previously described (6) with either kanamycin or erythromycin resistance cassettes from pIKM1 or pSGD8-erm flanked by 1.0- to 0.5-kb regions homologous to areas 5' and 3' of the deletion target of interest. Plasmids pHKO2 and pHKO3 were constructed using traditional cloning techniques and pBluescript KS II(+) (16), and plasmids pHKO7 through pHKO10 were constructed by cloning with *Saccharomyces cerevisiae* based on in vivo recombination (18) using the *S. cerevisiae*-*E. coli* pMQ87 shuttle plasmid. Plasmid maps and primers used in this study are shown in Fig. S1 in the supplemental material.

Transformation of *T. saccharolyticum*. *T. saccharolyticum* was transformed (6, 24) and grown on the appropriate antibiotic and solid media in an anaerobic chamber (COY Labs, Grass Lake, MI). Posttransformation growth with kanamycin selection was performed at 55°C and erythromycin selection at 50°C. Transformation efficiencies of 10³ to 10⁴ transformants/ μ g of DNA within 48 h were routine for kanamycin-based constructs and of 10² to 10³ transformants/ μ g of DNA within 96 h for erythromycin-based constructs. Deletions with chromosomal integration of both flanking regions were confirmed by PCR with primers external to the areas of homologous recombination.

Preparation of cleared cell lysates. Cells were grown to mid-exponential phase in 40 ml of medium, collected by centrifugation at 5,000 \times g for 20 min at 4°C, washed twice in degassed anaerobic 50 mM Tris-HCl (pH 7.6) buffer supple-

mented with 5 mM dithiothreitol, and resuspended in 1 ml of the same buffer. Sonication (XL-2000 sonicator; Misonix, Farmingdale, NY) was carried out inside an anaerobic chamber (COY Labs, Grass Lake, MI). The cell suspension was placed in a 2-ml microcentrifuge tube and incubated in an ice-water bath as six 30-s pulses at half-intensity were applied at 30-s intervals. Cell lysates were centrifuged at 5,000 \times g for 5 min to remove unbroken cells. The protein concentration of cleared cell lysates was measured by the Bradford method (Bio-Rad Laboratories, Hercules, CA) with bovine serum albumin as the standard.

Hydrogenase assay. Assays were performed anaerobically at 60°C in semi-microcuvettes with rubber stoppers (Starna Cells, Atascadero, CA) sealed in an anaerobic chamber (COY Labs, Grass Lake, MI) with an atmosphere of ~89% N₂, 10% CO₂, 1% H₂. Initial rates of methyl viologen (MV) reduction were measured with a spectrophotometer (Shimadzu, Columbia, MD) at 578 nm (ϵ = 9.7 mM⁻¹ cm⁻¹) (13). The reaction mixture contained 50 mM Tris-HCl (pH 7.6)–1.0 mM MV–sodium dithionite to reduce the MV to a faint blue color (A_{578} , 0.05 to 0.2; approximately 0.25 mM) and 15 to 20 μ g of cell extract. The reaction was begun with the addition of 0.02 mmol of hydrogen to the cuvette headspace (3).

Transcript analysis. RNA was purified using exponentially growing *T. saccharolyticum* cell culture (optical density [OD], 0.6 to 0.8) with an RNAeasy purification kit (Qiagen, Valencia CA), RNA-protect reagent, on-column DNase treatment, and a 10-min lysozyme incubation. cDNA was prepared with Thermoscript avian reverse transcriptase (Invitrogen, Carlsbad, CA) and used as a template for transcriptional structure analysis of the *hfs* genes with *Taq* polymerase (NEB, Ipswich, MA) and for quantitative reverse transcriptase PCR of genes homologous to hydrogenases with SYBR green detection using an iQ SYBR green Supermix kit (Bio-Rad, Hercules, CA), and a Bio-Rad iQ5 Q-PCR thermocycler. Primers (see Fig. S1 in the supplemental material) were designed to bridge open reading frames or to amplify 100- to 120-bp internal fragments and were tested to ensure specificity by agarose gel and melting curve analysis.

Analytical techniques. Fermentation metabolites were analyzed by high-performance liquid chromatography using an Aminex HPX-87H column (Bio-Rad Laboratories, Hercules, CA). Hydrogen was analyzed by gas chromatography on a silica gel column with nitrogen as the carrier gas with a TCD detector (SRI Instruments, Torrance, CA). Cell dry weight was correlated to OD measurements via a conversion of 0.4 g of dry cell weight/liter per OD unit. Carbon

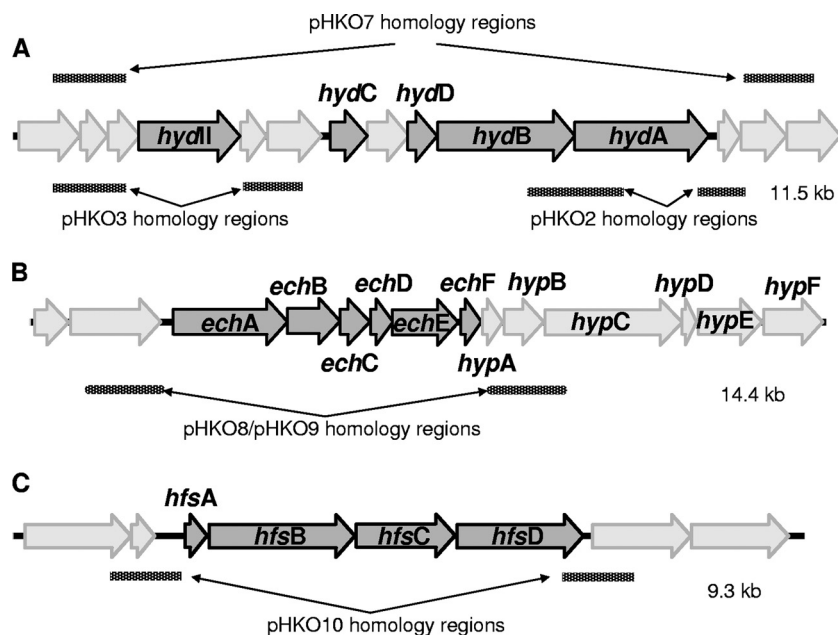


FIG. 1. Genetic maps of the *hyd* (A), *ech* (B), and *hfs* (C) genes. The phrase “homology regions” refers to areas of homologous recombination used to generate knockout strains; the deleted area lies between the upstream and downstream homology for each plasmid. The *ech* map includes the *hyp* gene cluster, which facilitates [Ni-Fe] packaging and *ech* activation.

balances were determined according to the following equations, with accounting of carbon dioxide through the stoichiometric relationship of its production to levels of acetic acid and ethanol. The carbon contained in the cell mass was estimated with a general empirical formula for cell composition ($\text{CH}_2\text{N}_{0.25}\text{O}_{0.5}$). The overall carbon balance is as follows:

$$C_t = \frac{144}{342}\text{CB} + \frac{72}{180}\text{G} + \frac{36}{90}\text{L} + \frac{36}{60}\text{A} + \frac{36}{46}\text{E} + \frac{12}{25.5}\text{CDW} \quad (1)$$

where C_t = total carbon, CB = cellobiose, G = glucose, L = lactic acid, A = acetic acid, E = ethanol, and CDW = cell dry weight (with all units in grams/liter) and

$$C_R = \frac{C_{tf}}{C_{t0}} \times 100\% \quad (2)$$

where C_R = carbon recovery, C_{t0} = total carbon at the initial time, and C_{tf} = total carbon at the final time. Electron recoveries were calculated in a similar manner, with the following numbers of available electrons per mole of compound: per mole 48 for cellobiose, 24 for glucose, 8 for acetic acid, 12 for lactic acid, 12 for ethanol, 2 for hydrogen, and 4.75 for cell mass.

$$E_t = \frac{48}{342}\text{CB} + \frac{24}{180}\text{G} + \frac{12}{90}\text{L} + \frac{8}{60}\text{A} + \frac{12}{46}\text{E} + \frac{4.75}{25.5}\text{CDW} \quad (3)$$

$$E_R = \frac{E_{tf}}{E_{t0}} \times 100\% \quad (4)$$

Protein sequence analysis. Protein sequences homologous to putative hydrogenases of *T. saccharolyticum* were identified using the BLASTP protein-protein algorithm and the National Center for Biotechnology Information (NCBI) database of nonredundant protein sequences. Alignments were performed with VectorNTI software (Invitrogen, Carlsbad, CA) using a ClustalW algorithm.

Nucleotide sequence accession numbers. The *T. saccharolyticum* gene sequences have been deposited with GenBank under the following accession numbers: for *hyd* cluster II, GQ354411; for the *hyd* operon, EU313771; for the *ech* operon, EU313772; and for the *hfs* operon, GQ354412.

RESULTS

Knockout of *hyd* and *ech* hydrogenases. A draft genome of *T. saccharolyticum* allowed a search of genes homologous to

known hydrogenase enzymes (19). Two gene clusters, one composed of five genes with high-level homology to the NAD-dependent [FeFe]-hydrogenase (*hyd*, also referred to as *hnd*) (15) found in *Thermoanaerobacter tengcongensis* and another composed of six genes with high-level homology to the membrane-bound [Ni-Fe] hydrogenase (*ech*) of *T. tengcongensis*, were identified as targets potentially involved in hydrogen production (Fig. 1) (19). Soboh et al. demonstrated that these two enzymes are responsible for hydrogen production in *T. tengcongensis* (21), and they are in fact the only two enzymes exhibiting hydrogenase activity in *T. tengcongensis* cell extracts (R. Hedderich, personal communication). The *ech* hydrogenase is interesting, as it appears to be of archaeal origin, having been acquired through horizontal gene transfer (4). In addition, it has been shown to transport H^+/Na^+ ions across the cell membrane during hydrogen production (16) and shares homology with subunits of the NADH-quinone oxidoreductase (complex I) in respiring organisms (8).

However, strains of *T. saccharolyticum* lacking either or both of the complete *hyd* or *ech* gene clusters do not show any decrease in hydrogen or acetic acid production in batch fermentation (Table 2). A gene that exhibits partial homology to an [FeFe]-hydrogenase 2 kb upstream of the *hyd* gene cluster that has been labeled *hyd* cluster II for this study also had no effect on hydrogen production when deleted (data not shown). The *hyd* cluster II location and amino acid sequence are conserved in many thermophilic bacteria containing the *hyd* genes; however, the *hyd* cluster II gene lacks a well-defined catalytic H cluster, as determined on the basis of sequence alignments.

Knockout of the *hfs* gene cluster. A further investigation of the draft genome identified a gene cluster of four open reading frames containing putative Fe-S clusters. Two of these open reading frames code for proteins homologous to [FeFe]-hydro-

TABLE 2. Fermentation profiles of *T. saccharolyticum* knockout strains^a

Strain	Fermentation profile ^b											
	Cellobiose (mM)	Lactic acid (mM)	Acetic acid (mM)	Ethanol (mM)	Hydrogen (mM)	CDW (g/liter)	Carbon recovery (%)	Electron recovery (%)	Y_{lactate} (M)	Y_{acetate} (M)	Y_{ethanol} (M)	Y_{hydrogen} (M)
YS485 WT	0	13.5	13.3	25.4	23.4	0.6	110	114	0.50	0.49	0.94	0.87
HKO1 <i>hyd</i>	0	11.1	14.3	28.1	23	0.6	112	114	0.41	0.53	1.04	0.85
HKO2 <i>ech</i>	0	14.5	12.2	25.1	19.6	0.6	108	111	0.54	0.45	0.93	0.73
HKO3 <i>hfs</i>	1.6	22.4	0.7	19.7	0.9	0.4	98	102	0.83	0.03	0.83	0.04
HKO4 <i>ech hyd</i>	0	8.7	19.1	32.5	24.9	0.6	124	125	0.32	0.71	1.20	0.92
HKO5 <i>ech hfs</i>	2	22.7	0	18	0.7	0.3	97	100	0.84	0.00	0.78	0.03
HKO6 <i>hfs hydA</i>	5	22.8	2.6	11	2.4	0.6	110	121	0.84	0.15	0.65	0.14
HKO7 <i>hfs ldh</i>	7	0	1.2	17.5	1.7	0.4	93	97	0.00	0.09	1.35	0.13
ALK2 <i>pta ack ldh</i>	0	0	0	54.7	0.4	0.3	107	111	0.00	0.00	2.03	0.01

^a Concentrations are reported in millimoles except for cell dry weight (CDW) values, which are reported in grams per liter. Initial cellobiose concentrations were 13.5 mM for all fermentations. Cultures were incubated for 24 h at 55°C in medium 122 at an initial pH of 6.4. WT, wild type.

^b Data represent averages of the results of three replicate fermentation experiments; standard deviations were less than 2 mM for cellobiose, lactic acid, acetic acid, and ethanol and less than 5 mM for hydrogen. Y_{lactate} , Y_{acetate} , Y_{ethanol} , and Y_{hydrogen} values represent molar yields (Y) based on consumed glucose equivalents.

genes. Deletion of this gene cluster resulted in strain HKO3 (Δhfs), which produced >95% less hydrogen than the wild-type strain (Table 2) and produced significantly less acetic acid, as expected due to the requirement to balance electron flux (12, 23). Strains with this deletion also displayed a growth defect, reaching a lower final cell density than the wild-type strain (Fig. 2).

***hfs* transcript analysis.** Based on genomic sequence data, the *hfs* genes compose a likely operon, as the open reading frames of *hfsA*, *hfsB*, and *hfsC* all overlap each other by a single base pair and as there is a 45-bp gap between the open reading frames of *hfsC* and *hfsD*. To test this possibility, wild-type *T. saccharolyticum* cDNA was used as a template for PCR using primers amplifying the gaps between *hfs* genes and genes outside of the putative operon. Figure 3 shows that a transcriptional operon was detected for the *hfs* genes and that transcription did not extend further in the 5' or 3' direction.

Enzymatic assay and quantitative reverse transcriptase PCR of hydrogenase genes. As seen in Table 3, MV hydrogenase activity was assayed with cleared lysate cell extracts of

strains with *hyd*, *ech*, and *hfs* deletions as well as the *hfs hydA* double mutant. MV acts as a universal electron acceptor-donor and is able to interact with hydrogenases that have either NAD(P)H or ferredoxin as a natural substrate (27). It is expected that the majority of enzymatic activity present in whole cells is also present in the cleared lysates; however, is it possible that enzymatic activity present in the membrane fraction could be underrepresented. The cell extract from the Δhyd strain showed a decrease of more than 50% in MV hydrogenase activity, although strain HKO1 produced hydrogen yields nearly identical to those seen with the wild-type strain. The Δhfs and Δech strains had MV hydrogenase activities comparable to those of the wild-type strain. However, cell extract from a $\Delta hfs \Delta hydA$ strain had no detectable hydrogenase activity, suggesting that these two enzymes are responsible for the observed MV hydrogenase activity in the wild-type strain.

In addition, transcript levels were tested in these strains by quantitative reverse transcriptase PCR to detect genes with homology to hydrogenase catalytic subunits and are reported

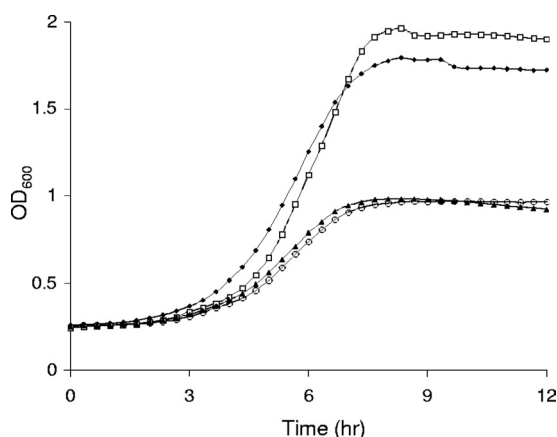


FIG. 2. Growth curves of the wild-type strain (\blacklozenge) and of strains lacking *hyd* (\square), *hfs* (\blacktriangle), and *ldh* plus *hfs* (\circ) grown at 55°C in medium 122 in a microplate reader. Data represent the averages of the results of eight replicate experiments.

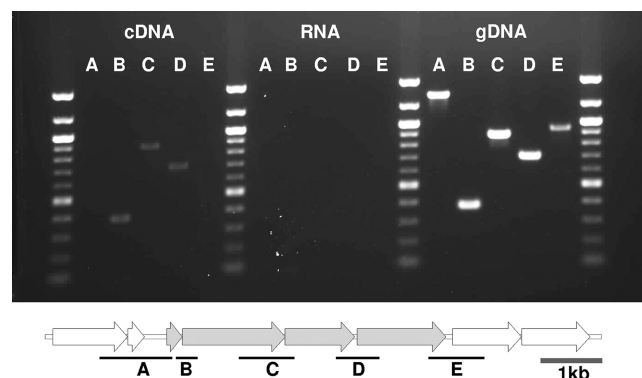


FIG. 3. Reverse transcriptase PCR of *hfs* and adjacent genes. Primers were designed to amplify the regions between the *hfs* genes (lanes B, C, and D) and between the *hfs* genes and adjacent genes (lanes A and E). cDNA from the wild-type strain (lanes 1 to 5) was prepared from RNA, which was tested for genomic DNA (gDNA) contamination (lanes 6 to 10). gDNA controls (lanes 11 to 15) were also included.

TABLE 3. Hydrogenase enzymatic activity assayed with respect to MV reduction and RNA expression of putative catalytic hydrogenase subunits relative to 16S RNA expression

Strain	Genotype ^a	MV reduction by H ₂		Transcript expression level relative to 16S expression				
		Sp act (μmol/min/mg of protein)	SD	<i>hydA</i>	<i>hyd2</i>	<i>echE</i>	<i>hfsB</i>	<i>hfsD</i>
YS485	WT	2.1	0.27	0.6	0.04	0.09	0.02	0.09
HKO1	<i>hyd</i>	0.76	0.15	0	0	0.1	0.03	0.3
HKO2	<i>ech</i>	2.38	0.18	0.7	0.04	0	0.01	0.2
HKO3	<i>hfs</i>	1.92	0.36	0.003	0.002	0.003	0	0
HKO6	<i>hfs hydA</i>	0	0	0	0.02	0.03	0	0

^a WT, wild type.

relative to 16S RNA transcript levels (Table 3). In the wild-type strain, the *hydA*, *hyd2*, *echE*, *hfsB*, and *hfsD* genes all appear to be expressed. In the *hyd* and *ech* knockout strains, *hfsD* may be upregulated relative to the wild-type level, influencing cell extract hydrogenase activities. Surprisingly, transcript levels for hydrogenases in *hfs* knockout strains are significantly lower than in the wild-type strain. It is not yet clear why the other hydrogenase transcript levels are so reduced in strains HKO3 and HKO6, although it was noted that the *hfs* deletion imparts a growth defect that could affect transcript levels.

Fermentation products of knockout strains. The Δhfs strain produced primarily lactic acid in place of acetic acid under the conditions tested, resulting in an ethanol yield relatively the same as the wild-type strain yield. A small amount of hydrogen was produced by the Δhfs strain, and strains were constructed with *hydA* and *ech* deletions in the background of a Δhfs strain to see whether residual hydrogen would be abolished. This result was not seen, however, and a $\Delta hyd \Delta ech$ strain did not have any impact on hydrogen yields either. The ALK2 strain ($\Delta dh \Delta pta \Delta ack$) (20) also produced a small amount of hydrogen, a result not stoichiometrically predicted for catabolic metabolism, as all available reducing equivalents from carbohydrates are expected to be directed to ethanol via ferredoxin-NAD(P)H oxidoreductase, acetaldehyde dehydrogenase, and alcohol dehydrogenase. This fermentation study was done using rich media to facilitate cell growth of knockout strains, and it is possible that hydrogen was being produced from a pathway separate from carbohydrate metabolism. Also, levels of carbon and electron recovery from strains with *hfs* deletions were about 10% less than the levels seen with strains without this deletion. It is possible that some metabolic flux is redirected to an as-yet-unidentified compound in these strains.

In order to test the hypothesis that ethanol yields can be increased via a hydrogenase knockout, strain HKO7 ($\Delta hfs \Delta dh$) was created. This strain produced ethanol at a yield of 1.35 mM ethanol per mM consumed glucose equivalent, a 44% increase from the wild-type level (Table 2). The ethanol yield was lower than that seen with strain ALK2, due to the production of a small amount of acetic acid and a lower level of carbon recovery. It is not yet clear why this strain was unable to completely consume the available cellobiose.

DISCUSSION

Deletion of the *hfs* gene cluster was found to have a significant impact on hydrogen and acetic acid production in *T.*

saccharolyticum, suggesting that it is the physiological ferredoxin-linked hydrogenase active in this organism. Although the majority of MV hydrogenase activity in *T. saccharolyticum* is due to the *hyd* enzyme, deletion of these genes did not alter fermentation product yields when *T. saccharolyticum* was grown in pure culture, which is consistent with it being an NAD(P)H-linked enzyme operating in the direction of hydrogen generation. In addition, the *T. saccharolyticum hydB* gene aligns closely with the *hydB* genes from *Thermotoga maritima* and *T. tengcongensis*, which have been shown to contain NAD(P)⁺ and flavin mononucleotide binding sites (21, 26). The *hyd* enzyme purified from *T. tengcongensis* has a catalytic efficiency (k_{cat}/K_m) of $1.6 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ for hydrogen generation compared to an efficiency of $1.5 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ for hydrogen uptake, suggesting that the physiological activity of this enzyme is hydrogen production (21).

In contrast to that of *T. tengcongensis*, the *ech* enzyme of *T. saccharolyticum* does not appear to participate in hydrogen production or result in hydrogenase activity under the conditions observed in this study. In addition, growth in a minimal medium treated with an Ni²⁺ binding resin (Chelex-100) showed no difference in product yields compared to growth in an identical medium supplemented with 0.1 mM NiCl₂, further suggesting that the *ech* enzyme, which contains an active [Ni-Fe] site cluster, does not participate in end product formation (data not shown).

hfsB and *hfsD* both encode proteins with homology to [FeFe]-hydrogenases. In addition to multiple groups of conserved cysteine residues, an indication of the presence of Fe-S clusters (see Fig. S1 in the supplemental material), both genes have conserved amino acid residues in the three regions of the H-cluster active site: H1 ([140]TTSCPSVN), H2 ([181]IGPCXXKKXE), and H3 ([297]CDGGCINGP) in *hfsB* and H1 ([276]TSSCCPAIV), H2 ([327]IGPCXXKK), and H3 ([451]CIGGCIGGAGV) in *hfsD* (14, 26) (numbers in brackets are the starting amino acid positions of the sequences, underlining indicates highly conserved residues, and capital letters without underlining are less highly conserved residues). *hfsB* also has homology to a PAS sensor motif near its C terminus. PAS domains are typically part of two-component kinase signaling systems and are known to respond to redox potential and overall cellular energy levels, among other stimuli (22). Genes containing a regulatory or sensory domain in addition to an [FeFe]-hydrogenase domain have been noted during sequencing studies, for example, an [FeFe]-hydrogenase domain fusion with a sigma-54-dependent transcriptional regulator in *Carboxydotherrmus hydrogeniformans* (31) and three genes in *Halothermothrix orenii* with PAS and tran-

TABLE 4. Amino acid sequence homology to the *hfs* gene cluster of *T. saccharolyticum*

Organism and gene	GenBank accession no.	<i>T. saccharolyticum hfs</i> gene cluster ^a	
		% Homology	% Identity
<i>Clostridium botulinum</i> A ATCC 19397 ^b			
<i>hfsA</i>	YP_001384086	57.3	45.1
<i>hfsB</i>	YP_001384087	54.8	38.3
<i>hfsC</i>	YP_001384088	63.2	48.1
<i>Clostridium cellulolyticum</i> H10			
<i>hfsA</i>	ZP_01573820	53.5	45.3
<i>hfsB</i>	ZP_01573821	58.5	42.0
<i>hfsC</i> ^c	ZP_01573822	65.1	45.9
<i>Clostridium phytofermentans</i> ATCC 700394(T)			
<i>hfsA</i>	YP_001557222	65.4	44.4
<i>hfsB</i>	YP_001557221	56.0	41.3
<i>hfsC</i>	YP_001557220	59.7	41.5
<i>hfsD</i>	YP_001557219	56.4	43.8
<i>Clostridium thermocellum</i> ATCC 27405			
<i>hfsA</i>	YP_001036856	59.8	39.0
<i>hfsB</i>	YP_001036857	53.5	37.1
<i>hfsC</i>	YP_001036858	59.6	42.4
<i>hfsD</i> ^d	YP_001036861	33.5	22.5
<i>Thermoanaerobacter ethanolicus</i> 39E ATCC 33223			
<i>hfsB</i>	ZP_00778561	57.9	43.4
<i>hfsC</i>	ZP_00778562	61.5	45.5
<i>Thermoanaerobacter tengcongensis</i> MB4(T)			
<i>hfsB</i>	NP_622355	57.8	43.5
<i>hfsC</i>	NP_622356	67.1	45.2

^a Protein alignments are shown in Fig. S2 in the supplemental material.

^b The gene with highest homology (57.1%) and identity (43.8%) to *T. saccharolyticum hfsD* was that of YP_001385620; however, it is not adjacent to the other *hfs* genes in *C. botulinum*. The *C. botulinum hfsABC* genes are located near putative [NiFe]- and [FeFe]-hydrogenases but are transcribed in the reverse direction compared to these genes.

^c *C. cellulolyticum* has a putative transposase 16 bp downstream of *hfsC*.

^d *C. thermocellum* has three genes immediately downstream of *hfsC* that have closest homology to the *hydA*, *hydB*, and *hydC* genes. For this comparison, the label *hfsD* has been given to the gene containing a catalytic Fe-hydrogenase motif.

scriptional regulatory domains (15). The protein sequence of *hfsC* shares homology with predicted serine phosphatases and SpoEII, a sporulation regulatory phosphatase involved in σ^F factor activation in *Bacillus subtilis* (1), further suggesting that the *hfs* genes may have a regulatory as well as a metabolic role in *T. saccharolyticum*. *hfsA* does not align closely with proteins of known function but exhibits some homology to the large subunit of [FeFe]-hydrogenases.

Homologues to the *hfs* gene cluster are present in several clostridia that have sequenced genomes. Somewhat surprisingly, the closest identity match across the entire gene cluster is with *Clostridium phytofermentans*, a mesophilic cellulose degrader (28). The *hfsABC* genes are well conserved across sequenced strains of *C. botulinum*, *C. cellulolyticum*, *C. phytofermentans*, and *C. thermocellum* (Table 4). The *C. thermocellum hfsD* gene closely aligns with *hydA* homologs, and two other genes with high homology to *hydB* and *hydC* are located between *hfsC* and *hfsD*. *C. thermocellum* contains separate copies of *hydA*, *hydB*, and *hydC* in the standard *hyd* gene cluster on its genome. Both *T. ethanolicus* 39E and *T. tengcongensis* contain homologues for only *hfsB* and *hfsC*. The diverse genomic arrangements of these genes in other organisms raise questions about their function and could be an interesting area for further research.

Growth inhibition and end product shifts in the presence of hydrogen have been reported to various degrees for many of the thermophilic saccharolytic bacteria. At one extreme, *Caldicellulosiruptor saccharolyticus* and *T. tengcongensis* are reported to be completely inhibited by a headspace of 20- to 56-kPa hydrogen (21, 25, 32). In contrast, during the course of this study it was found that the growth and product profile of *T. saccharolyticum* is not influenced in the presence of up to 185-kPa hydrogen. *T. ethanolicus* 39E is reported to be inhibited by 100- to 200-kPa hydrogen, although an adapted strain showed higher hydrogen tolerance (10, 29), and *C. thermocellum* is reported to grow in the presence of >250-kPa hydrogen, albeit with a product profile shift (2, 9). To see whether the *hfs* genes are associated with hydrogen tolerance in these organisms, several strains were grown in the absence and presence of hydrogen (Table 5). Interestingly, strains lacking or containing a partial *hfs* gene cluster showed the greatest reduction in end product formation when grown under conditions that included a hydrogen atmosphere, although *C. phytofermentans*, with a complete *hfs* gene cluster, showed a product shift away from acetate formation. Use of a soluble ferredoxin-linked hydrogenase, such as that encoded by the *hfs* genes, rather than a membrane-bound, energy-conserving hydrogenase, such as the *ech* hydrogenase, would produce 0.25 mol less ATP per mole of

TABLE 5. Effect of hydrogen on fermentation end products^a

Strain	Headspace gas ^b	End product (mM)				Presence of hydrogenase:		
		Lactic acid	Acetic acid	Ethanol	Total	<i>hyd</i>	<i>ech</i>	<i>hfs</i>
<i>T. saccharolyticum</i> JW/SL-YS485	N ₂	0.0	5.1	9.6	14.7	Yes	Yes	Yes
	H ₂	0.0	5.1	9.8	14.9			
<i>C. cellulolyticum</i> H10	N ₂	0.0	5.0	0.0	5.0	Yes	Yes	Partial
	H ₂	0.5	3.6	0.0	4.1			
<i>T. ethanolicus</i> 39E ATCC 33223	N ₂	5.6	2.3	11.6	19.5	Yes	Yes	Partial
	H ₂	5.3	0.7	8.0	14.0			
<i>C. saccharolyticus</i> DSM 8903	N ₂	4.2	8.7	0.0	12.9	Yes	Yes	No
	H ₂	1.2	1.8	0.0	3.0			
<i>C. thermocellum</i> ATCC 27405	N ₂	0.0	7.9	3.3	11.2	Yes	Yes	Yes
	H ₂	0.0	8.6	3.5	12.1			
<i>C. phytofermentans</i> ISDg	N ₂	1.4	4.6	7.2	13.2	Yes	Yes	Yes
	H ₂	2.3	2.0	9.0	13.3			

^a All strains were grown in medium 122 adjusted to an initial pH of 6.8 with 5 mM cellobiose. Cultures were incubated for 72 to 96 h at the following temperatures: for *T. saccharolyticum*, *T. ethanolicus*, and *C. thermocellum*, 55°C; for *C. cellulolyticum* and *C. phytofermentans*, 35°C; and for *C. saccharolyticus*, 70°C. Data represent averages of the results of three replicate culture experiments, with a standard deviation of less than 1 mM for end product measurements.

^b The nitrogen headspace absolute pressure was 170 kPa; the hydrogen headspace absolute pressure was 185 kPa.

hydrogen generated due to electron transport phosphorylation, assuming a requirement of four protons per molecule of ATP formed by ATP synthase (17). However, utilization of the *hfs* hydrogenase may facilitate growth in the presence of hydrogen, which would be beneficial for organisms in environments where interspecies hydrogen transfer has not been well established.

To increase the ethanol yield in *T. saccharolyticum*, metabolic steps consuming pyruvate are key branch points in the fermentation pathway. Unlike the lactic acid pathway, where electrons are added to pyruvate from reduced NADH generated during glycolysis, the pathways to acetic acid and ethanol begin with a further oxidation of pyruvate to acetyl coenzyme A and reduced ferredoxin. This additional reduction of electron carriers, which must ultimately be balanced in end product formation, allows two distinct control points that determine acetic acid and ethanol yields (12). Panel A of Fig. 4 shows the approach taken with strain ALK2, which can be considered “carbon” centered with respect to directing flux away from acetic acid and toward ethanol. Panel B shows an “electron”-

centered approach, where acetic acid production is limited by removing the ability to produce hydrogen, as demonstrated by the strain HKO7 results shown here. The electron-centered metabolic engineering approach provides a new route for raising solvent yields in the industrially relevant thermophilic and mesophilic anaerobes that share fermentation pathways similar to those of *T. saccharolyticum* (5, 12, 22).

Conclusion. This study began with the intent to demonstrate that ethanol yields of *T. saccharolyticum* could be increased by inactivation of the hydrogen production pathway. When it was found that deletion of the *hyd* and *ech* genes, which are responsible for hydrogen production in *T. tengcongensis*, did not affect the ethanol yield, the study took a somewhat new direction toward discovery of the genes responsible for hydrogen production. Identification of the *hfs* genes provides some insight into how *T. saccharolyticum* produces hydrogen when grown in pure culture but also raises more questions about how these genes are regulated in response to environmental or

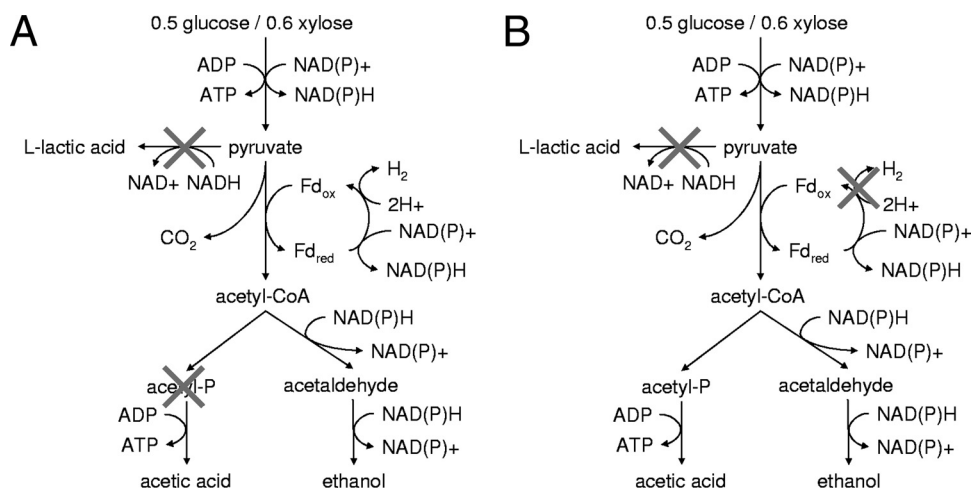


FIG. 4. Knockout strategies for increasing ethanol yield. (A) Carbon-centered approach, exemplified by strain ALK2 ($\Delta dh \Delta pta \Delta ack$). (B) Electron-centered approach, exemplified by strain HKO7 ($\Delta dh \Delta hfs$). CoA, coenzyme A; Fd, ferredoxin.

