Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/he

Genome based metabolic flux analysis of *Ethanoligenens harbinense* for enhanced hydrogen production

*J.F. Castro, V. Razmilic, Z.P. Gerdtzen**

Centre for Biochemical Engineering and Biotechnology, Department of Chemical Engineering and Biotechnology, University of Chile, Santiago, Chile

ARTICLE INFO

Article history:
Received 20 August 2012
Received in revised form
19 October 2012
Accepted 1 November 2012
Available online 5 December 2012

Keywords:
Ethanoligenens harbinense
Hydrogen production
Metabolic flux analysis

ABSTRACT

Ethanoligenens harbinense is a promising hydrogen producing microorganism due to its high inherent hydrogen production rate. Even though the effect of media optimization and inhibitory metabolites has been studied in order to improve the hydrogen productivity of these cultures, the identification of the underlying causes of the observed changes in productivity has not been targeted to date. In this work we present a genome based metabolic flux analysis (MFA) framework, for the comprehensive study of *E. harbinense* in culture, and the effect of inhibitory metabolites and media composition on its metabolic state. A metabolic model was constructed for *E. harbinense* based on its annotated genome sequence and proteomic evidence. This model was employed to perform MFA and obtain the intracellular flux distribution under different culture conditions. These results allow us to identify key elements in the metabolism that can be associated to the observed production phenotypes, and that can be potential targets for metabolic engineering in order to enhanced hydrogen production in *E. harbinense*.

1. Introduction

In recent years, research in the field of energy production has been directed toward the development of renewable energy sources which are cleaner, safer, and more efficient than currently used fossil fuels. Hydrogen has emerged as one of the most promising alternative energy carrier due to its high energy content (61,000 BTU/lb) in comparison to other alternative fuels such as ethanol (13,000 BTU/lb). Although hydrogen is produced conventionally by steam reforming of natural gas, its production by microorganisms has become a major area of research as this process has significant advantages, e.g. no distillation step is required for its recovery and its production process is usually operated at mesophilic temperatures and atmospheric pressure [1–4].

In fermentation based hydrogen production processes, *Ethanoligenens harbinense* is one of the most promising producer organisms due to its capability to generate hydrogen at high rates and efficiency [5]. *E. harbinense* is a gram-positive, mesophilic, strict anaerobic bacteria, phylogenetically related to the clostridia class [6]. These features make it an interesting target for physiological and genetic studies aiming to improve its metabolic properties and increase hydrogen productivity.

The metabolic pathway for hydrogen production in strict anaerobes begins with glucose breakdown through the Embden–Meyerhof–Parnas (EMP) pathway to pyruvate. Pyruvate

* *Corresponding author*. Tel.: +56 29784712; fax: +56 26991084.
E-mail addresses: jcastrofigueroa@ing.uchile.cl (J.F. Castro), vrazmilic@ing.uchile.cl (V. Razmilic), zgerdtze@ing.uchile.cl (Z.P. Gerdtzen).

0360-3199/$ – see front matter
http://dx.doi.org/10.1016/j.ijhydene.2012.11.007

is then degraded into acetyl-CoA and CO_2, in a reaction coupled to the reduction of ferredoxin by pyruvate:ferredoxin oxidoreductase (PFOR). Ferredoxin acts as an electron carrier and the electrons transferred are employed by hydrogenases to produce hydrogen. Acetyl-CoA is further converted to a variety of secondary metabolites, but in *E. harbinense*, acetate, ethanol and hydrogen have been reported as the most abundant products resulting from glucose degradation. The reaction pathway associated to the PFOR enzyme is present only in this type of anaerobic microorganisms [3,6,7].

Experimental results for *E. harbinense* in culture show that the concentration of end metabolites affects hydrogen production [8]. Acetate and ethanol addition to culture media reduce hydrogen production in approximately 71% and 54%, respectively [9]. Changes in the concentration of nutrients, liquid and gaseous metabolites, both in the growth media and in the reactor head space, can trigger a metabolic shift in *E. harbinense* that leads to alterations in the hydrogen productivity of the culture. One of the highest hydrogen yields reported, 2.26 mol H_2/mol glucose, has been achieved by using media with an optimized nutrient concentration [10].

Previous studies do not address the effect of culture conditions and nutrient concentrations in the metabolic state of *E. harbinense*. Characterizing the use of nutritional resources and product formation under given culture conditions allows determining the metabolic pathways involved in hydrogen production and identifying potential targets for modifications that could increase the culture's productivity. Metabolic flux analysis (MFA) is a powerful tool that allows addressing the study of production systems in a comprehensive manner, considering their specific metabolic capabilities, requirements and culture conditions [11]. This tool provides information regarding flux distribution among the different metabolic pathways based on experimental measurement for extracellular metabolites. Results obtained from MFA can help interpret current results and guide future experiments leading to an enhanced hydrogen yield [12]. For instance, MFA has been employed to study the effect of variations of initial glucose concentrations and operational pH on *Clostridium butyricum* [13]. More specifically, this methodology has been used to describe carbon distribution in *Escherichia coli* when lactate dehydrogenase is knocked-down [14] and to identify pathways that allow maximal theoretical hydrogen production in *Citrobacter amalonaticus* [15].

The motivation behind this research is to provide a framework for the study of *E. harbinense* from a systems point of view, in order to increase knowledge on its metabolism during hydrogen production. To do this, a genome based MFA methodology was implemented. A genome scale model was constructed, based on *E. harbinense* genomic annotation and available bibliographic information. Recently published experimental data was employed along with the model developed, to extract information regarding the metabolic behavior of *E. harbinense* under different culture conditions. The genomic model generated here can help guide metabolic engineering strategies to improve the metabolic capabilities of *E. harbinense* by identifying key elements of its metabolism that could be targeted for enhanced hydrogen production.

2. Materials and methods

2.1. Evaluating metabolic states through MFA

2.1.1. Theory behind MFA

MFA is a methodology based on mass balances of all metabolites considered to be involved in the metabolic network under study. The mass balances must include the definition of the metabolic network's structure and stoichiometry. To construct the model, stoichiometric coefficients for each reaction involved in the network are arranged in a matrix, where each column corresponds to a reaction and each row is associated to a single metabolite. Production and consumption rates for extracellular metabolites are experimentally determined and used to calculate the intracellular fluxes using a stoichiometric model under a pseudo-steady-state assumption. The outcome of this method is the net flux distribution map of the network, including estimated rates for each reaction. The results obtained depend strongly on the reactions considered in the network and their stoichiometry [16].

The general form of the resulting system is:

$$S \cdot \nu(x) = \begin{bmatrix} r \\ 0 \end{bmatrix} \tag{1}$$

where S is the stoichiometric matrix, $\nu(x)$ is a vector of intracellular metabolic fluxes in steady-state, r is a vector that contains the experimentally measured production and consumption rates for extracellular metabolites, and a 0 vector is assigned to intracellular metabolites which are not produced or consumed.

Before solving the system, mass and charge balances must be checked for each reaction as follows:

$$E \cdot S = 0 \tag{2}$$

where E corresponds to the elemental matrix containing the number of carbon, hydrogen, oxygen, phosphorous and sulfur that are present in every compound in the network considered. The charge is balanced by adding protons to one side of the reaction.

The minimum number of specific rates that need to be experimentally determined in order to solve the system is given by the degrees of freedom (dof) of the system, which correspond to the difference between number of reactions and intracellular fluxes. If the system is over-determined (number of measurements $>$ dof) it can be solved by the least square method. It is important to verify that S is full rank for it to be invertible. If the system is under-determined (number of measurements $<$ dof) the system must be solved using linear programming [16].

The solution for an over-determined system is as follows:

$$\nu(x) = \left(\left(S^T \cdot S\right)^{-1} \cdot S^T \right) \cdot \begin{bmatrix} r \\ 0 \end{bmatrix} \tag{3}$$

where $\left((S^T \cdot S)^{-1} \cdot S^T\right)$ is the pseudo-inverse of S. With the expression in Eq. (3), intracellular metabolic fluxes are determined and a measure of the degree of engagement of various pathways in overall cellular functions and metabolic processes is obtained. This approach can help specify nutrient requirements, and identify metabolic steps that limit cell growth and metabolite production.

2.2. Reconstruction of central metabolism of E. harbinense

The genome of *E. harbinense* YUAN-3 was recently sequenced (NCBI, accession number NC_014828.1) [17] and its automated annotation is available through BioCyc [18]. Based on this annotation we developed a genome based metabolic reconstruction and a stoichiometric model for *E. harbinense* which considers the EMP pathway, tricarboxylic acid (TCA) cycle, pentose phosphate pathway (PPP), transport mechanisms, biosynthesis of building blocks, secondary metabolites, and H_2 production. The information acquired from BioCyc was manually curated and cross checked with bibliographic information in order to accurately define each reaction and metabolites involved in the network. Bioinformatic analyses were performed with the Basic Local Alignment Search Tool (BLAST) server (http://blast.ncbi.nlm.nih.gov/). The model developed was used to assess the metabolic flux distribution of *E. harbinense* under different culture conditions.

2.3. Experimental data

Two key experimental settings were considered for this MFA study: growth in the presence of the inhibitory metabolites sodium acetate and ethanol [9], and growth in an optimized media (OM) [10]. In the first case, batch experiments were performed at 37 °C in 50 mL of basal medium consisting of (g/L): glucose 10, polypepton 4, beef extract 2, yeast extract 1, NaCl 4, $MgCl_2$ 0.1, $FeSO_4 \cdot 7H_2O$ 0.1, K_2HPO_4 1.5, L-cysteine 0.5, trace elements and resazurine. Different concentrations of sodium acetate (10, 20, 50, 100, 150 mM) were added to the basal media for the sodium acetate effect experiments. For studies regarding the effect of ethanol in culture, different concentrations of ethanol were added to basal media (20, 40, 80, 100, 200 mM). These experiments were analyzed using the genome based stoichiometric model proposed in order to evaluate the inhibitory effect of these compounds on hydrogen yield. In the second case, *E. harbinense* was cultured in a batch reactor for 40 h at 35 °C in 50 mL of culture media. The culture medium consisted of 0.3–0.5 g/L yeast extract, 100–150 mM phosphate and 10 g/L glucose. All experimental data were normalized to mmol of metabolite per gram of dry cell weight (DCW) per unit time (h). Normalized experimental data are shown in (Table 1).

Table 1 – Specific consumption and production rates measured for *Ethanoligenens harbinense* under different culture conditions: sodium acetate (NaAc) and ethanol (ETH) presence in the media, and growth on optimized media (OM).

		Specific rates (mmol/g DCW/h)·10^{-3}						
		GLC	ACE	BUT	ETH	CO_2	H_2	BIO
NaAc (mM)	0	−1.36	0.96	0.00	1.49	2.74	2.99	0.61
	10	−1.35	0.92	0.00	1.45	2.39	2.32	0.65
	20	−1.35	0.87	0.00	1.28	2.65	2.07	0.65
	50	−1.40	0.54	0.46	1.07	2.37	1.73	0.65
	100	−1.79	−0.61	1.40	0.80	2.36	1.59	0.71
	150	−1.87	−1.44	1.87	0.67	2.07	1.20	0.64
ETH (mM)	0	−1.36	0.96	0.00	1.49	2.74	3.02	0.61
	20	−1.47	1.04	0.00	1.50	2.70	2.95	0.69
	40	−1.42	0.96	0.00	1.37	3.43	2.39	0.68
	80	−1.53	1.08	0.15	1.38	3.15	2.16	0.67
	100	−1.46	1.04	0.24	0.46	2.77	1.74	0.68
	200	−1.14	0.87	0.90	−1.17	1.86	1.16	0.67

	GLC	ACE	LAC	ETH	CO_2	H_2	BIO
OM	−1.84	1.29	0.07	2.03	3.72	4.17	0.70

3. Results and discussion

3.1. Proposed metabolic model

The reconstructed metabolic model of *E. harbinense* comprises the EMP pathway, TCA cycle, PPP, biosynthesis of secondary metabolites and building blocks for biomass assembly, transport reactions and H_2 production (shown in Fig. 1 and detailed in Table A1). Relevant aspects related to the model's construction process are discussed below.

Two enzymes involved in hydrogen production have been identified in the *E. harbinense* genome: one is a [FeFe]-hydrogenase (E.C. 1.6.5.3, gene Ethha_2293) and the other one is a bifurcative hydrogenase (E.C. 1.12.1.4, genes Ethha_2614-15-16). [FeFe]-hydrogenase has been characterized and identified as the main enzyme responsible for hydrogen production in *E. harbinense* (protein accession number DQ_177326) [7]. Zhao et al. in 2010 demonstrated that the gene sequence Ethha_2293 (*hydA*) cloned into *E. coli* effectively leads to hydrogen production in this microorganism. Another study shows that a gene with 95% identity to *hydA* from *E. harbinense* is in fact expressed *in vivo* during hydrogen production in fermentation by *E. harbinense*-like microorganisms [19]. Ferredoxin hydrogenase activity has been assigned to this enzyme based on the presence of a domain inferred by InterPro. Considering this information, a hydrogen production reaction with this specific characteristic was explicitly included in the model (reaction 22, Table A1). A bifurcative hydrogenase was identified using bioinformatic tools and comprises three genes (Ethha_2614-15-16) which share high homology to its counterpart in *Clostridium thermocellum* (Cthe_0428-29-30). This heterotrimeric hydrogenase has been characterized in *Thermotoga maritima* and it has been proposed to employ the exergonic oxidation of ferredoxin to drive the oxidation of NADH to produce hydrogen [20]. The bifurcative activity of this enzyme has been taken into account in our model. The electrons required for hydrogen production by both, [FeFe]- and bifurcative hydrogenases, are provided by ferredoxin which acts as a multi-functional electron carrier in diverse hydrogen production systems [21]. Reduced ferredoxin can be supplied by two reaction pathways: pyruvate:ferredoxin oxidoreductase (PFOR) enzyme (E.C. 1.2.7.1, gene Ethha_2733), which catalyzes the decarboxylation of pyruvate to acetyl-CoA, and 2-oxoglutarate:ferredoxin oxidoreductase (KGOR) complex (E.C. 1.2.7.3, genes Ethha_0231-32-33-34), which oxidatively decarboxylates α-ketoglutarate to form succinyl-CoA. Both steps are coupled to ferredoxin reduction and CO_2 production.

Formate can also be used as a substrate for hydrogen production by the formate-hydrogen lyase complex (FHL) which produces hydrogen and CO_2. However, FHL is characteristic of facultative anaerobes microorganisms and is not

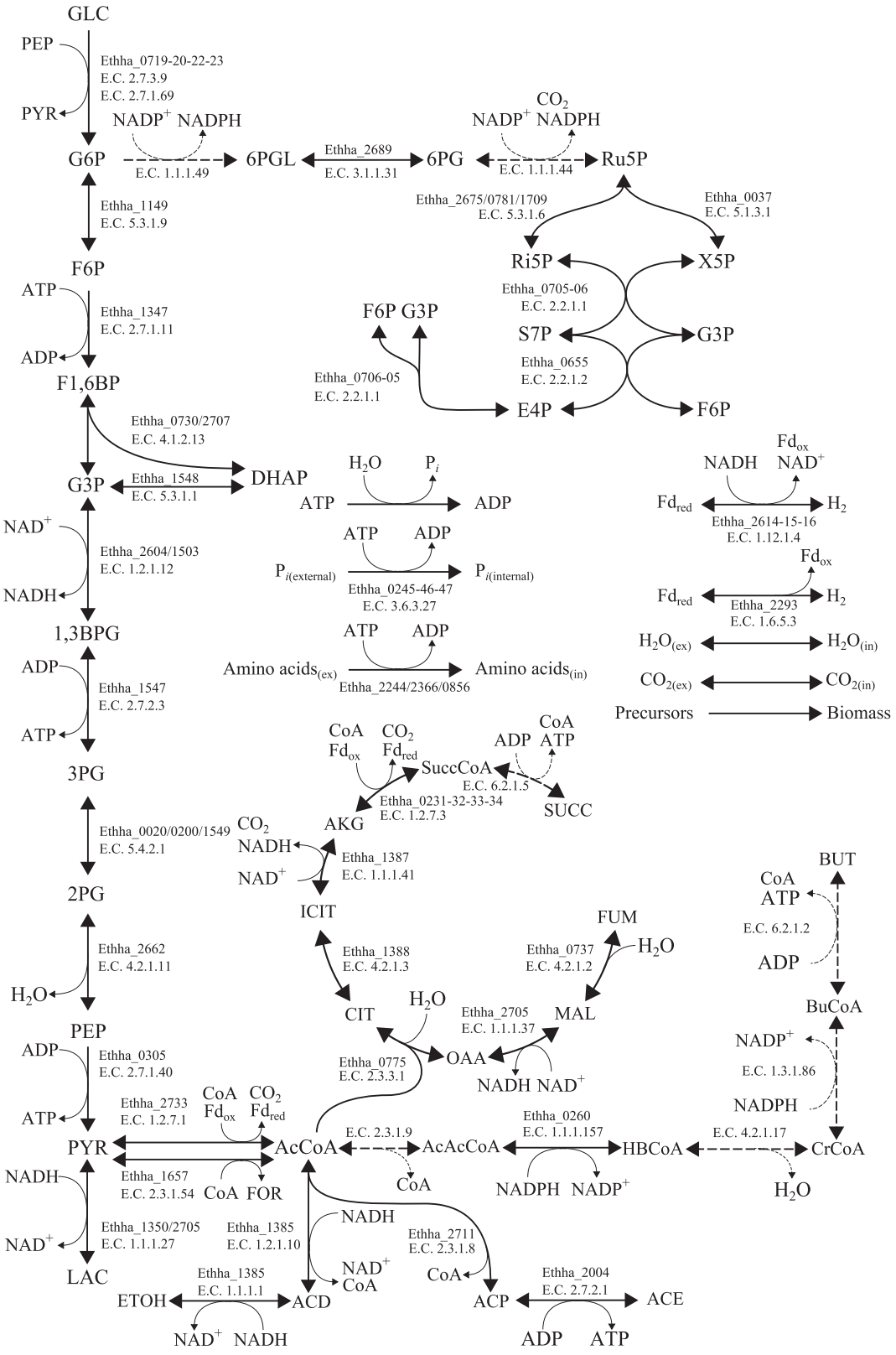

Fig. 1 – Genomic model for *Ethanoligenens harbinense*. Enzyme commission (E.C.) numbers and gene indices are shown. Dashed lines indicate enzyme added by gap filling. Metabolites abbreviations are listed in Table B1.

usually present on strict anaerobes such as *E. harbinense* [3]. Based on this, FHL was not considered in the model presented in this work. Formate is produced by the pyruvate formate lyase (PFL) enzyme (E.C. 2.3.1.54, gene Ethha_1657), which catalyzes the conversion of pyruvate to acetyl-CoA with concomitant production of formate. PFL was included in the model since the PFL gene is present in the *E. harbinense* genome.

Bioinformatic analysis indicates that Ethha_1350 and Ethha_2705 are associated to the lactate dehydrogenase (LDH) family. This includes both LDH and malate dehydrogenase (MDH) enzymes, which are structurally similar and use the same coenzyme NADH. It is common for these genes to be misannotated when analyzed solely on the basis of sequence homology [22]. In our model we associate the Ethha_1350 gene to LDH, based on sequence homology to *Clostridium acetobutylicum* (55% identity to LDH, 98% coverage, *E*-value $1e^{-123}$) and *C. thermocellum* (39% identity, 96% coverage, *E*-value $6e^{-78}$). The Ethha_2705 gene is associated to both LDH and MDH, as it has 72% homology to LDH in *Clostridium leptum* with 95% coverage, E-value $6e^{-160}$ and 39% homology to MDH in *C. thermocellum* with 98% coverage, *E*-value $3e^{-83}$.

When building a genome based model, some genes encoding for enzymes within a specific pathway may be absent in the annotated genome sequence (gaps) even though there is experimental evidence of their activity. For instance, only one enzyme involved in butyrate metabolism was found in the genome of *E. harbinense* (gene Ethha_0260). However, previous studies report butyrate synthesis [9], indicating activity of the pathway. Based on these findings, all the reactions involved in butyrate biosynthesis were added to the model. The same criteria was applied for including enzymes in PPP (E.C. 1.1.1.49 and E.C. 1.1.1.44) and TCA cycle (E.C. 6.2.1.5).

A partial TCA cycle was considered. Fumarate reductase (E.C. 1.3.5.4), which catalyzes the fumarate to succinate conversion step was not found in the *E. harbinense* genome. Bioinformatic analysis suggests that the gene Ethha_0039 (annotated as E.C. 1.3.5.4) is more similar to aspartate oxidases rather than fumarate reductases, inferred by the presence of an aspartate oxidase multi-domain. Therefore this enzyme was not included in the proposed model.

The genomic model proposed for *E. harbinense*, depicted in Fig. 1, consists of 46 reactions, 34 of which are associated to intracellular metabolites and 10 correspond to transport reactions. A biomass production reaction was included based on precursor requirements [23]. In addition, an ATP hydrolysis reaction which represents energy requirements not associated to cell growth was added. Among intracellular reactions, 28 can be associated to a gene present in the sequenced genome and 6 result from the gap filling process described above.

To reduce the number of reactions without affecting the final flux distribution results, sequential reactions were lumped together and represented by a single reaction step, therefore eliminating intermediate metabolites. As a result a simplified model is obtained, consisting of a total of 31 reactions (Table A1).

This is the first genome based model reported for *E. harbinense*, and incorporates most of the relevant information available to date regarding its metabolism. The model has been manually curated, with a careful selection of reactions and genes based on the current annotation of the *E. harbinense* genome, proteomic evidence, and bioinformatic analysis. This is a framework for the comparative analysis of experimental data under different culture conditions, as illustrated in the results that follow. This framework could be improved as further research in the field provides us with new information regarding *E. harbinense*'s genome and metabolism, which may contribute to improve the proposed model.

3.2. Flux distribution under different culture conditions of *E. harbinense* and their effect on hydrogen yield

Genome based MFA was used to study two experimental settings: the effect of inhibitory metabolites on hydrogen production and an optimized media scenario. The intracellular flux distribution obtained for each experimental setting is shown in Fig. 2. In *E. harbinense* hydrogen is produced by the action of both, the [FeFe]- and bifurcative hydrogenases. The first requires reduced ferredoxin (Fd_{red}) only and the second employs Fd_{red} and NADH (Fig. 2, $\nu 22$ and $\nu 23$, respectively). Fd_{red} is supplied by pyruvate decarboxylation in $\nu 7.1$. Lactate is also synthesized from pyruvate ($\nu 20$) in a reaction that requires NADH. This NADH can be supplied either by the EMP pathway ($\nu 4$) or by the consumption of ethanol ($\nu 19$). Acetyl-CoA is an entry point for the TCA cycle, but also a precursor for acetate ($\nu 18$), butyrate ($\nu 21$) and ethanol ($\nu 19$) synthesis. In addition, for butyrate biosynthesis, NADPH and ADP are required. NADPH is supplied by PPP ($\nu 12$). ADP is produced in $\nu 25$, which represents a summarized expression for the physiological energy requirements of the cell.

3.2.1. Fermentation experiments in the presence of sodium acetate and ethanol

In order to achieve good hydrogen production rates it is necessary to maintain the concentration of inhibitors, such as acetate and ethanol, below critical levels. In *E. harbinense* the concentrations of ethanol and acetate required to achieve a half-maximal degree of inhibition of hydrogen production ($C_{I,\ 50}$) are 154 and 62 mM, respectively [9].

3.2.1.1. Sodium acetate. Fermentation experiments under increasing concentrations of sodium acetate showed a decline in hydrogen and acetate production; even acetate consumption was reported at high concentration of sodium acetate (100 and 150 mM). At these concentrations, ethanol production was very low and was accompanied by butyrate production [9]. Tang et al. suggested that the high levels of sodium acetate present in the medium, negatively altered the metabolic state of *E. harbinense*, leading to the observed phenotype.

The intracellular flux distribution obtained for *E. harbinense* growing in the presence of sodium acetate at different concentrations ranging from 0 to 150 mM is shown in Fig. 2A. The distribution is based on experimental data shown in Table 1 [9]. As acetate concentration is increased, a reduction in acetate production or even its consumption is observed ($\nu 18$). Consumption of acetate leads to higher levels of acetyl-CoA, which results in a reduction of the pyruvate to acetyl-CoA flux ($\nu 7.1$) and an increase in butyrate production ($\nu 21$). The flux from phosphoenolpyruvate to pyruvate ($\nu 6$) remains unchanged as acetate is added to the media. Reduction in $\nu 7.1$ with increasing concentrations of sodium acetate, leads to a higher predicted lactate production flux ($\nu 20$). An increased fraction of the glucose flux that is taken up by the cells is directed toward PPP to supply NADPH for butyrate synthesis. In the first reaction of PPP, NADPH is generated as Ru5P is produced ($\nu 12$). As sodium acetate concentration increases there is a raise in the $\nu 12$ flux, which in turn leads to an increment in NADPH production.

Evidence from a related organism, *C. thermocellum*, shows that PFL specific activity is 18 times lower than PFOR activity

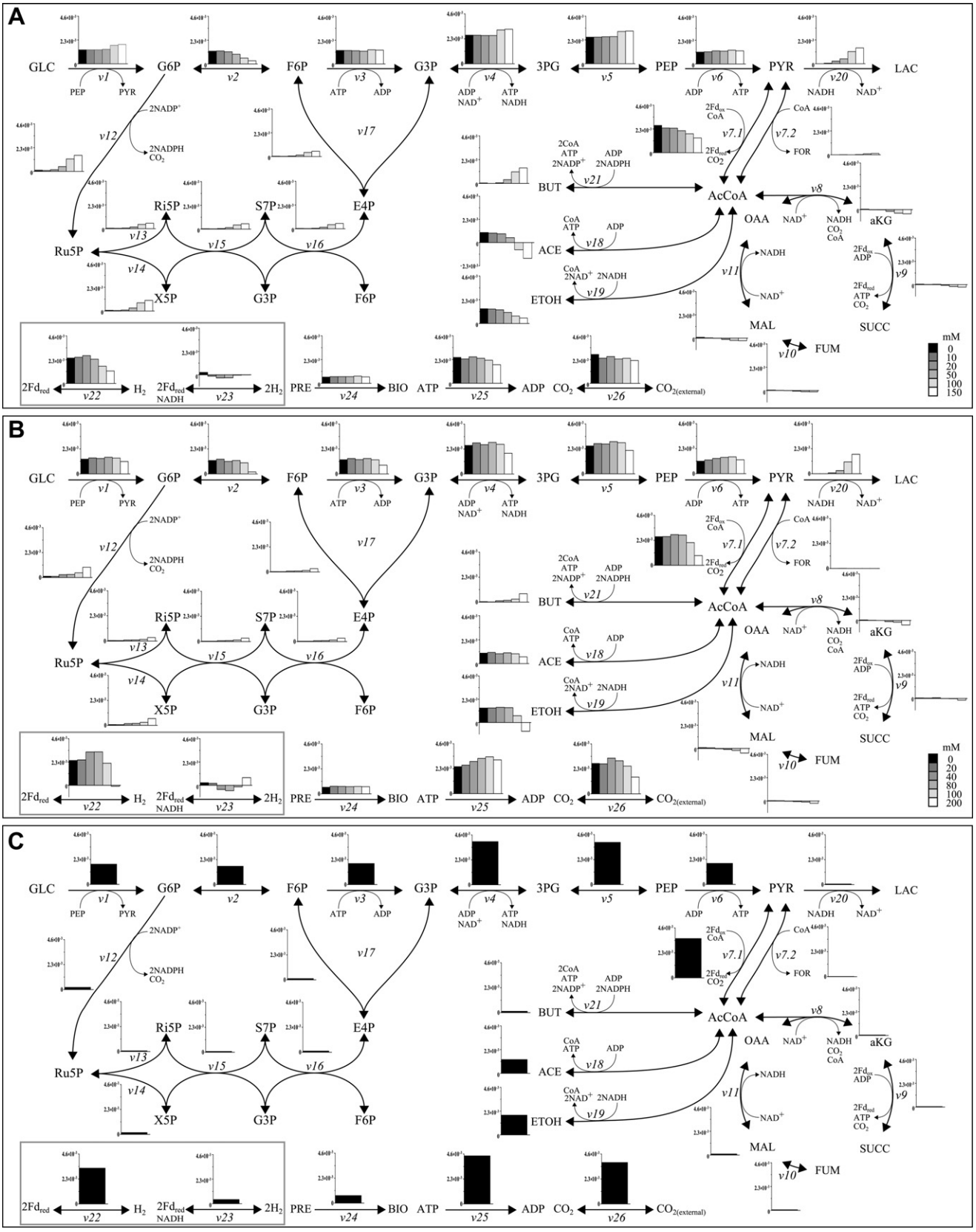

Fig. 2 – Intracellular flux distribution (mmol/g DCW/h) resulting from genome based MFA for *Ethanoligenens harbinense*, during its cultivation under increasing concentration of sodium acetate (A) and ethanol (B) and growing in optimized media (C). All fluxes are represented by bars and are shown in the same scale. Hydrogenase fluxes are highlighted in a gray box.

under anaerobic growth conditions [24]. This results in low formate production levels. This is in agreement with our simulation results as flux through reaction *v7.2* (PFL) is one of the lowest obtained upon increase of sodium acetate concentrations (see Fig. 2A).

Reduced ferredoxin is the main electron carrier for hydrogen production by both, [FeFe]- and bifurcative hydrogenases (*v22*) [21]. The contribution of TCA cycle (*v9*) to supplying Fd_{red} is negligible. Therefore, the phenotype observed experimentally under these culture conditions, can be explained by the fact that a reduction of the pyruvate to acetyl-CoA flux (*v7.1*) triggers a reduction in hydrogen synthesis due to limited availability of Fd_{red}.

A similar study was carried out in *C. thermoaceticum*. In this study a greater production of end metabolites was correlated to a lower hydrogen yield. Wang et al., attributed the decreased productivity to cell lysis caused by a higher ionic force in the media generated by increased end metabolites concentration [25]. In *Caldicellulosiruptor saccharolyticus*, reduced hydrogen production has also been associated to cell lysis due to autolysins activation (gene Ethha_1410) in the presence of high concentration of end metabolites [26].

Simulation results suggest a complementary explanation for the experimental results of Tang et al. [9]. The presence of sodium acetate triggers a redistribution of glucose flux toward PPP and a reduced flux toward acetyl-CoA from pyruvate. This results in an increment of NADPH production and a decrease in Fd_{red} levels. At 100 mM sodium acetate concentration there is a 50% reduction in hydrogen production flux (*v22*). The limited availability of Fd_{red} would be the underlying cause of the reduced hydrogen productivity observed.

[FeFe]-hydrogenase is mainly responsible for hydrogen evolution (*v22*), since flux through this pathway is 10 times greater than flux through bifurcating hydrogenase (*v23*) in this case. As sodium acetate concentration is increased, the flux for both hydrogenases is reduced. For bifurcating hydrogenase (*v23*) it even reverses, producing low levels of Fd_{red} and NADH. However, [FeFe]-hydrogenase (*v22*) remains the main hydrogen production pathway (see Fig. 2A).

3.2.1.2. *Ethanol*. In ethanol addition experiments, Tang et al. reported that the production of acetate remains almost unchanged up to 100 mM of ethanol. In addition, ethanol consumption was reported at 200 mM of ethanol. Tang et al. explained these results in terms of a metabolic shift cause by feedback inhibition and self-regulation of bacteria, above critical concentrations of ethanol.

Flux distributions for *E. harbinense* growing in the presence of ethanol were obtained based on data on Table 1 [9]. Similarly to sodium acetate experiments, the presence of ethanol results in a redistribution of glucose flux toward PPP and a reduced flux toward acetyl-CoA from pyruvate (Fig. 2B). Reduction of this flux (*v7.1*) results in limited availability of Fd_{red}, which would explain the decreased hydrogen production observed. The contribution of PFL to acetyl-CoA production is negligible (*v7.2*). The flux from phosphoenolpyruvate to pyruvate (*v6*) also remains unchanged as ethanol is added to the media. Increasing lactate production fluxes are also predicted by the model upon ethanol addition. At 200 mM, ethanol hydrogen production through *v22* ceases and hydrogen is produced through *v23* as a result of NADH availability from ethanol consumption (*v19*). This reduction in hydrogen production at high ethanol concentrations predicted by our model is in agreement with experimental observations by Tang et al. [9].

The main differences that can be pointed out between ethanol and acetate experiments, are the shift toward ethanol consumption in *v19* at high concentration of ethanol (200 mM), while acetate flux remains almost unchanged (*v18*). At high acetate concentrations the preferred reaction for hydrogen production is [FeFe]-hydrogenase (*v22*) while at high ethanol concentrations hydrogen is produced by the bifurcative hydrogenase complex (*v23*). The acetyl-CoA generated from ethanol is used for butyrate production (*v21*), which requires ADP and NADPH. The amount of butyrate generated is less than in the acetate case, due to a lower uptake rate of the inhibitory metabolite. ADP is supplied through an increased ATP hydrolysis flux (*v25*). A reduced PPP flux (*v12*) is observed compared to the acetate case, associated to lower NADPH requirements for butyrate production. Ethanol has been reported to extend the lag phase and decrease the exponential growth rate of bacteria in culture [27]. In this context, the reduced growth rate associated to ethanol presence, could be associated to diminished ATP availability due to an increased ATP hydrolysis flux (*v25*).

3.2.2. *Optimized media*

Intracellular flux distribution for *E. harbinense* growing in an optimized media was obtained based on data in Table 1 [10]. Fig. 2C shows that pyruvate is produced from glucose through the EMP pathway, contributing to the generation of ATP and NADH. Pyruvate is also generated by the phosphotransferase system (PTS) (*v1*) during glucose uptake. The resulting pyruvate pool is used to generate acetyl-CoA, concomitant with ferredoxin reduction by the PFOR (*v7.1*). Reduced ferredoxin (Fd_{red}) is required by both hydrogenases ([FeFe]- and bifurcative hydrogenase) for hydrogen production (*v22* and *v23*, respectively).

Pyruvate is a precursor molecule for lactate, acetate, ethanol, butyrate and biomass synthesis. In this experiment a low flux from pyruvate to lactate (*v20*) was observed. This can be explained by pyruvate being employed to synthesize ethanol and acetate from acetyl-CoA (*v19* and *v18*, respectively). NAD^+ and ATP are generated as a result of ethanol and acetate synthesis. The large flux of ethanol production, requires high levels of NADH that can be supplied by the G3P to 3PG reaction *v4*. Although experimental measurements are not available for butyrate production, the model predicts a low butyrate production flux, which is consistent with experimental results obtained for *E. harbinense* growing without ethanol or acetate, shown in Fig. 2A and B (0 mM data points) [10].

TCA cycle is the final stage of glucose oxidation in which acetyl-CoA molecules are oxidated generating CO_2 and reducing power. For cells growing in optimized media, very low fluxes are observed for the TCA cycle. These low fluxes indicate that the reaction *v9* has a very small contribution in supplying Fd_{red}, supporting the idea that PFOR is a key enzyme for the production of this electron carrier [7]. In addition, since TCA fluxes are small, reducing power and ATP must be provided by G3P to 3PG reaction (*v4*).

The main purpose of PPP is to provide NADPH and precursors for biomass synthesis. Fluxes through the PPP are also very small. In fact, only 1/8th of the glucose taken up by the cells (*v1*) is

directed toward Ru5P ($v12$). Low PPP fluxes result in low NADPH and precursors levels, which lead to a low biomass synthesis rate ($v24$) and a very low butyrate production flux ($v21$).

The results shown in this section provide an alternative explanation for the reduced hydrogen productivity observed in the presence of inhibitory metabolites such as acetate and ethanol. This explanation is based on the metabolic capabilities of *E. harbinense*, summarized in a genome based metabolic model, and empirical evidence. In both cases, simulation results point to Fd_{red} as the key factor for hydrogen synthesis. Fd_{red} becomes limiting in the presence of secondary metabolites such as sodium acetate and ethanol, leading to decreased hydrogen production fluxes. The model also predicts changes in metabolites concentrations that were not measured in all experiments such as lactate and butyrate production fluxes, showing that genome based MFA can be a powerful tool for data analysis and experimental design.

3.2.3. *Targets for improving hydrogen production*

The analysis of experimental data using the mathematical model proposed for *E. harbinense* allows the identification of potential targets for metabolic cell engineering and changes in culture conditions that might enhance hydrogen production.

The first target to address for increasing hydrogen production is media supplementation. For *E. harbinense* growing in optimized media, an increase in the hydrogen production rate and biomass generation was observed as yeast extract was supplemented to the media [10]. This indicates that there is room for improvement in media design in order to increase biomass yield, and that hydrogen productivity can be correlated to biomass levels.

Previous studies have shown a relationship between lactate production and hydrogen yield. High levels of lactate are associated with low hydrogen yield in *C. thermocellum* [28]. Mutations in LDH have been shown to result in greater yields of hydrogen on *E. coli* [29]. Our simulation results predict lactate production in the presence of acetate and ethanol, and associate decreased hydrogen production to reduced availability of Fd_{red}. These results along with the experimental findings detailed above, strongly indicate that hydrogen production could be increased by eliminating or reducing the pyruvate–lactate flux ($v20$) via LDH (Ethha_1350/2705) knockout or knock-down. This modification would increase the available pyruvate pool for PFOR and increase the availability of Fd_{red}. The availability of reduced ferredoxin would lead to an increased hydrogen production through the action of [FeFe]-hydrogenase ($v22$).

Our results shows that [FeFe]-hydrogenase is the main hydrogen production enzyme, even in the presence of inhibitory metabolites. All the metabolic engineering strategies suggested above may lead to [FeFe]-hydrogenase becoming limiting for ferredoxin oxidation. Therefore overexpressing this enzyme could also help improve hydrogen productivity in modified *E. harbinense* cells.

4. Conclusions

A genomic model was developed for *E. harbinense* based on its genomic annotation, bioinformatic analyses, proteomic and experimental data. The model was used to perform genome based MFA under different culture conditions, and in the presence of hydrogen synthesis inhibitory metabolites. The MFA results led to a proposed mechanism whereby inhibition on hydrogen production observed in the presence of ethanol and acetate, was associated to a limited availability of Fd_{red}. Additionally, we identified potential targets for metabolic engineering: lactate dehydrogenase ($v20$) and biomass ($v24$), which would allow to enhance hydrogen yield in *E. harbinense*.

Results obtained illustrate that genome based MFA is a powerful tool to obtain the intracellular flux distribution associated to a set of extracellular flux measurements under different environmental scenarios. It serves as a first reference for global data analysis and can provide a framework for the theoretical studies of new hypothesis before their experimental validation in *E. harbinense*.

Acknowledgment

J.F.C. and V.R. were supported by National Doctoral Scholarship, Conicyt, Chile. We acknowledge funding from Initiation Grant 11090268 from Fondecyt, Chile.

Appendix A. Metabolic reactions in the model

Table A1 – Metabolic reactions incorporated in the genomic model of *Ethanoligenens harbinense*.

v	Stoichiometric equation
Embden–Meyerhof–Parnas	
1	$GLU + PEP \rightarrow G6P + PYR$
2	$G6P \leftrightarrow F6P$
3	$ATP + F6P \rightarrow 2 \cdot G3P$
4	$G3P + P_i \leftrightarrow 3PG + ATP + NADH$
5	$3PG \leftrightarrow H_2O + PEP$
6	$PEP \rightarrow ATP + PYR$
TCA cycle	
7.1	$CoA + PYR \leftrightarrow AcCoA + CO_2 + 2 \cdot Fd_{red}$
7.2	$PYR + CoA \rightarrow FOR + AcCoA$
8	$AcCoA + H_2O + OAA \leftrightarrow AKG + CO_2 + CoA + NADH$
9	$AKG + P_i \leftrightarrow ATP + CO_2 + 2 \cdot Fd_{red} + SUCC$
10	$H_2O + FUM \leftrightarrow MAL$
11	$MAL \leftrightarrow NADH + OAA$
Pentose phosphate	
12	$G6P + H_2O \rightarrow CO_2 + 2 \cdot NADPH + Ru5P$
13	$Ru5P \leftrightarrow Ri5P$
14	$Ru5P \leftrightarrow X5P$
15	$Ri5P + X5P \leftrightarrow G3P + S7P$
16	$G3P + S7P \leftrightarrow E4P + F6P$
17	$E4P + X5P \leftrightarrow F6P + G3P$
Fermentation products	
18	$AcCoA + P_i \leftrightarrow ACE + ATP + CoA$
19	$AcCoA + 2 \cdot NADH \leftrightarrow CoA + ETH$
20	$PYR + NADH \leftrightarrow LAC$
21	$2 \cdot AcCoA + 2 \cdot NADPH + P_i \leftrightarrow ATP + BUT + 2 \cdot CoA + H_2O$
Hydrogen	
22	$2 \cdot Fd_{red} \leftrightarrow H_2$
23	$2 \cdot Fd_{red} + NADH \leftrightarrow 2 \cdot H_2$

Table A1 – (*continued*)

ν	Stoichiometric equation
Biomass	
24	$3.515\cdot10^{-2}\cdot3PG + 8.807\cdot10^{-2}\cdot AcCoA + 8.483\cdot10^{-3}\cdot E4P + 1.666\cdot10^{-3}\cdot F6P + 3.031\cdot10^{-3}\cdot G3P + 4.817\cdot10^{-3}\cdot G6P + 6.009\cdot10^{-3}\cdot GLN + 1.161\cdot10^{-1}\cdot GLU + 4.199\cdot10^{-2}\cdot OAA + 1.219\cdot10^{-2}\cdot PEP + 6.657\cdot10^{-2}\cdot PYR + 2.109\cdot10^{-2}\cdot Ri5P + 1.405\cdot ATP + 1.405\cdot H_2O + 3.061\cdot10^{-1}\cdot NADPH \rightarrow BIO + 9.677\cdot10^{-2}\cdot AKG + 8.807\cdot10^{-2}\cdot CoA + 8.335\cdot10^{-2}\cdot NADH + 1.405\cdot P_i$
Generalized energy utilization and others	
25	$ATP + H_2O \rightarrow ADP + P_i$
26	$CO_2 \leftrightarrow CO_{2(external)}$
27	$ATP + GLN_{(external)} + H_2O \leftrightarrow GLN + P_i$
28	$ATP + GLU_{(external)} + H_2O \leftrightarrow GLU + P_i$
29	$H_2O \leftrightarrow H_2O_{(external)}$
30	$P_{i(external)} + H_2O + ATP \rightarrow 2\cdot P_i$

Appendix B. Table B1: Nomenclature of metabolites

Abbreviation	Metabolite
1,3BPG	1,3-Bisphosphoglycerate
2PG	2-Phosphoglycerate
3PG	3-Phosphoglycerate
6PG	6-Phosphogluconate
6PGL	6-Phosphoglucono-1,5-lactone
AcAcCoA	Acetoacetyl-CoA
AcCoA	Acetyl-CoA
ACD	Acetaldehyde
ACE	Acetate
ACP	Acetylphosphate
ADP	Adenosine diphosphate
AKG	α-Ketoglutarate
ATP	Adenosine triphosphate
BIO	Biomass
BuCoA	Butyryl-CoA
BUT	Butyrate
CIT	Citrate
CO_2	Carbon dioxide
CoA	Coenzyme A
CrCoA	Crototonyl-CoA
DHAP	Dihydroxyacetone phosphate
E4P	Erythrose-4-phosphate
ETH	Ethanol
F1,6BP	Fructose-1,6-bisphosphate
F6P	Fructose-6-phosphate
Fd_{ox}	Oxidized ferredoxin
Fd_{red}	Reduced ferredoxin
FOR	Formate
FUM	Fumarate
G3P	Glyceraldehyde-3-phosphate
G6P	Glucose-6-phosphate
GLC	Glucose
GLN	Glutamine
GLU	Glutamate
H_2	Hydrogen
H_2O	Water
HBCoA	3-Hydroxybutyryl-CoA
ICIT	Isocitrate
LAC	Lactate
MAL	Malate
NAD^+	Oxidized nicotinamide adenine dinucleotide
NADH	Reduced nicotinamide adenine dinucleotide
$NADP^+$	Oxidized nicotinamide adenine dinucleotide phosphate
NADPH	Reduced nicotinamide adenine dinucleotide phosphate
OAA	Oxaloacetate
P_i	Orthophosphate
PEP	Phosphoenolpyruvate
PRE	Precursors
PYR	Pyruvate
Ri5P	Ribose-5-phosphate
Ru5P	Ribulose-5-phosphate
S7P	Sedoheptulose-7-phosphate
SCoA	Succinyl-CoA
SUCC	Succinate
X5P	Xylulose-5-phosphate

REFERENCES

[1] Das D. Advances in biohydrogen production processes: an approach towards commercialization. Int J Hydrogen Energy 2009;34(17):7349–57.

[2] National Research Council. Microbial processes: promising technologies for developing countries. Books for Business; 1979.

[3] Hallenbeck PC, Ghosh D. Improvements in fermentative biological hydrogen production through metabolic engineering. J Environ Manage 2012;95(Suppl. 0):S360–4.

[4] Williams RH, Larson ED, Katofsky RE, Chen J. Methanol and hydrogen from biomass for transportation. Energy Sustain Dev 1995;1(5):18–34.

[5] Ren NQ, Lin HL, Zhang K, Zheng GX, Duan ZJ, Lin M. Cloning, expression, and characterization of an acetate kinase from a high rate of biohydrogen bacterial strain *Ethanoligenens* sp. hit b49. Curr Microbiol 2007;55:167–72.

[6] Xing D, Ren NQ, Li Q, Lin M, Wang A, Zhao L. *Ethanoligenens harbinense* gen. nov., sp. nov., isolated from molasses wastewater. Int J Syst Evol Microbiol 2006;56(4):755–60.

[7] Zhao X, Xing D, Zhang L, Ren NQ. Characterization and overexpression of a [FeFe]-hydrogenase gene of a novel hydrogen-producing bacterium *Ethanoligenens harbinense*. Int J Hydrogen Energy 2010;35(18):9598–602.

[8] Wang XJ, Ren NQ, Xiang WS, Guo WQ. Influence of gaseous end-products inhibition and nutrient limitations on the growth and hydrogen production by hydrogen-producing fermentative bacterial b49. Int J Hydrogen Energy 2007;32(6):748–54.

[9] Tang J, Yuan Y, Guo WQ, Ren NQ. Inhibitory effects of acetate and ethanol on biohydrogen production of *Ethanoligenens harbinese* B49. Int J Hydrogen Energy 2012;37(1):741–7.

[10] Xu L, Ren NQ, Wang X, Jia Y. Biohydrogen production by *Ethanoligenens harbinense* B49: nutrient optimization. Int J Hydrogen Energy 2008;33(23):6962–7.

[11] Koffas M, Roberge C, Lee K, Stephanopoulos G. Metabolic engineering. Annu Rev Biomed Eng 1999;1(1):535–57.

[12] Cai G, Jin B, Monis P, Saint C. Metabolic flux network and analysis of fermentative hydrogen production. Biotechnol Adv 2011;29(4):375–87.
[13] Cai G, Jin B, Saint C, Monis P. Metabolic flux analysis of hydrogen production network by *Clostridium butyricum* W5: effect of pH and glucose concentrations. Int J Hydrogen Energy 2010;35(13):6681–90.
[14] Kim S, Seol E, Oh YK, Wang G, Park S. Hydrogen production and metabolic flux analysis of metabolically engineered *Escherichia coli* strains. Int J Hydrogen Energy 2009;34(17): 7417–27.
[15] Oh YK, Kim HJ, Park S, Kim MS, Ryu DD. Metabolic-flux analysis of hydrogen production pathway in *Citrobacter amalonaticus* Y19. Int J Hydrogen Energy 2008;33(5): 1471–82.
[16] Stephanopoulos G, Aristidou A, Nielsen J. Metabolic engineering: methodology and principles. Academic Press; 1998.
[17] Hemme CL, Mouttaki H, Lee YJ, Zhang G, Goodwin L, Lucas S, et al. Sequencing of multiple clostridial genomes related to biomass conversion and biofuel production. J Bacteriol 2010; 192:6494–6.
[18] Karp PD, Ouzounis CA, Moore-Kochlacs C, Goldovsky L, Kaipa P, Ahren D, et al. Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. Nucleic Acids Res 2005;33(19):6083–9.
[19] Xing D, Ren NQ, Rittmann BE. Genetic diversity of hydrogen-producing bacteria in an acidophilic ethanol-H_2-coproducing system, analyzed using the [Fe]-hydrogenase gene. Appl Environ Microbiol 2008;74(4):1232–9.
[20] Schut GJ, Adams MWW. The iron-hydrogenase of *Thermotoga maritima* utilizes ferredoxin and NADH synergistically: a new perspective on anaerobic hydrogen production. J Bacteriol 2009;191(13):4451–7.
[21] Mathews J, Wang G. Metabolic pathway engineering for enhanced biohydrogen production. Int J Hydrogen Energy 2009;34(17):7404–16.
[22] Carere C, Kalia V, Sparling R, Cicek N, Levin D. Pyruvate catabolism and hydrogen synthesis pathway genes of *Clostridium thermocellum* ATCC 27405. Indian J Microbiol 2008; 48:252–66.
[23] Orth J, Fleming R, Palsson B. *Escherichia coli* and *Salmonella*: cellular and molecular biology; chap. 10.2.1: reconstruction and use of microbial metabolic networks: the core *Escherichia coli* metabolic model as an educational guide. ASM Press; 2010. p. 56–99.
[24] Rydzak T, Levin DB, Cicek N, Sparling R. Growth phase-dependant enzyme profile of pyruvate catabolism and end-product formation in *Clostridium thermocellum* ATCC 27405. J Biotechnol 2009;140:169–75.
[25] Wang Y, Zhao QB, Mu Y, Yu HQ, Harada H, Li YY. Biohydrogen production with mixed anaerobic cultures in the presence of high-concentration acetate. Int J Hydrogen Energy 2008;33(4):1164–71.
[26] van Niel E, Claassen P, Stams A. Substrate and product inhibition of hydrogen production by the extreme thermophile, *Caldicellulosiruptor saccharolyticus*. Biotechnol Bioeng 2003;81(3):255–62.
[27] Daifas D, Smith J, Blanchfield B, Cadieux B, Sanders G, Austin J. Effect of ethanol on the growth of *Clostridium botulinum*. J Food Prot 2003;66(4):610–7.
[28] Levin DB, Islam R, Cicek N, Sparling R. Hydrogen production by *Clostridium thermocellum* 27405 from cellulosic biomass substrates. Int J Hydrogen Energy 2006;31(11):1496–503.
[29] Yoshida A, Nishimura T, Kawaguchi H, Inui M, Yukawa H. Enhanced hydrogen production from glucose using *ldh* and *frd* inactivated *Escherichia coli* strains. Appl Microbiol Biotechnol 2006;73:67–72.