Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC-PapersOnLine 49-26 (2016) 364–370

Maps for when the living gets tough: Maneuvering through a hostile energy landscape

Thierry D.G.A Mondeel*, Samrina Rehman, Yanfei Zhang*, Malkhey Verma***, Peter Dürre****, Matteo Barberis* and Hans V. Westerhoff*,**,*****

*Synthetic Systems Biology and Nuclear Organization, SILS, University of Amsterdam, The Netherlands
**Manchester Centre for Integrative Systems Biology, University of Manchester, United Kingdom
*** Centre for Biochemistry & Microbial Sciences, Central University of Punjab, Bathinda – 151001, India
**** Institute of Microbiology and Biotechnology, University of Ulm, Germany
*****Molecular Cell Physiology, VU University Amsterdam, Amsterdam, The Netherlands, (e-mail: hans.westerhoff@manchester.ac.uk)

Abstract: With genome sequencing of thousands of organisms, a scaffold has become available for data integration: molecular information can now be organized by attaching it to the genes and their gene-expression products. It is however, the genome that is selfish not the gene, making it necessary to organize the information into maps that enable functional interpretation of the fitness of the genome. Using flux balance analysis one can calculate the theoretical capabilities of the living organism. Here we examine whether according to this genome organized information, organisms such as the ones present when life on Earth began, are able to assimilate the Gibbs energy and carbon that life needs for its reproduction and maintenance, from a relatively poor Gibbs-energy environment. We shall address how *Clostridium ljungdahlii* may use at least two special features and one special pathway to this end: gear-shifting, electron bifurcation and the Wood-Ljungdahl pathway. Additionally, we examined whether the *C. ljungdahlii* map can also help solve the problem of waste management. We find that there is a definite effect of the choices of redox equivalents in the Wood-Ljungdahl pathway and the hydrogenase on the yield of interesting products like hydroxybutyrate. We provide a drawing of a subset of the metabolic network that may be utilized to project flux distributions onto by the community in future works. Furthermore, we make all the code leading to the results discussed here publicly available for the benefit of future work.

Keywords: Wood-Ljungdahl pathway, *Clostridium ljungdahlii*, genome-wide metabolic map, flux-balance analysis, electron bifurcation, gear-shifting, syngas, biodegradable plastics

1. INTRODUCTION

The question of how life got started in the early periods of our planet is still very much open. In the early soup of chemicals, some billion years ago, no organic carbon compounds may have been abundant, there was no oxygen and the question of how life got going is therefore intriguing. A critical question in this regard is how the early organisms could produce their ATP (Schuchmann and Müller, 2014) and grow autotrophically on the available mixtures of CO, CO_2 and H_2.

Schuchmann and Müller (2014) highlighted the Wood-Ljungdahl pathway (WLP) as a network feature enabling to overcome the difficulties faced by early colonizers of Earth. Since the WLP is ATP-neutral, an additional 'trick' is also required so as to be able to harvest the limited amount of Gibbs free energy made available by the WLP. This has been identified as electron bifurcation. Electron bifurcating enzymes consist of multiple subunits coupling an endergonic to an exergonic redox reaction, thereby achieving what would otherwise be impossible without ATP input or chemiosmotic coupling of the endergonic reaction.

Schuchmann and Müller (2014) described how WLP and electron bifurcation could lead to the production of ATP from ADP and phosphate provided the use of two more transmembrane enzyme complexes is made, i.e. the H^+-ATPase and the Rnf complex. They did not show whether other solutions to the ATP synthesis problem may be possible, and whether what they proposed was in immediate concordance with the knowledge integrated through the genome-wide metabolic maps of the organisms (Nagarajan, 2013).

In the present paper, we wish to compare the analysis by Schuchmann and Müller with the predictions of maximal ATP synthesis emanating from the genome-wide metabolic reconstruction of the model acetogen *Clostridium ljungdahlii* through flux-balance analysis (FBA) (Orth, 2010). Some analysis on the effect of redox equivalents on growth and product synthesis was already present in (Nagarajan, 2013). Specifically, it was shown that the genome-wide map predicts the possibility of growth on CO_2/H_2 and CO and the effect of various options in redox equivalents were analyzed under the knockout of acetate kinase. However, we hope to extend that analysis here by including various alternative reactions that were considered in the treatment by Schuchmann and Müller. Specifically, we will investigate alternatives in the electron

10.1016/j.ifacol.2017.03.002

donors/acceptors for various enzymes and their effect on ATP yield coupled to acetogenesis. Additionally, we will focus on the importance of the Wood-Ljungdahl *pathway* as opposed to single enzymes, the need for electron bifurcation and the Nfn complex, the concept of gear-shifting, the requirement of low gear and advantages of high gear operation, and how much product yield might be attained when engineering *C. Ljungdahlii* with two additional genes for producing poly-hydroxybutyrate (PHB) under various redox alterations.

2. METHODS

2.1 Genome-scale reconstruction of Clostridium ljungdahlii

Our starting point was the previously published genome-scale metabolic reconstruction of *Clostridium ljungdahlii* (Nagarajan, 2013). We obtained the SBML file of the reconstruction from the BiGG database (King et al., 2016) through http://bigg.ucsd.edu/models/iHN637 on August 25th 2016. This reconstruction covers 698 metabolites, in 785 reactions, encoded by 637 genes. The map was shown to recover experimentally measured growth rates on media of various compositions (Nagarajan, 2013).

2.2 Extending the Genome-wide metabolic model with alternative reactions considered in Schuchmann et al.

In this work we are concerned with reproducing and understanding in more detail the analysis provided by Schuchmann and Müller (2014) for *C. ljungdahlii* with its genome-wide metabolic reconstruction. To that end we made sure all reactions considered in that study also came to exist in a (slightly) enhanced version of the the genome-wide metabolic map (GeMM) (which we shall call GeMM*), adding reactions where necessary, and we set the simulation conditions appropriately for the acetogenesis problem.

We adjusted the *in-silico* medium to be a mixture of CO_2 and H_2 in a 2:4 ratio by setting the lower bounds on the CO_2 and H_2 exchange reactions to -2 and -4 respectively, where the negative direction indicates uptake of the metabolites. We set the ATP maintenance reaction (ATP -> ADP + phosphate) as the objective function with a lower bound of zero, thereby asking with which flux distribution the network could make the ATP synthesis reaction as high as possible. Additionally, we forced the flux through the exchange reaction of acetate to be equal to 1 flux unit in the outward direction. Note that this does force the organism to make full use of both CO_2 and H_2 in the medium. If simulation conditions deviate from those described here, we will explicitly highlight the new conditions.

Several reactions described in Schuchmann and Müller (2014) are essential and were added to the metabolic reconstruction in order to enable the simulations in the scenarios considered, see Table 2. In traditional biochemistry considering a single compartment, it is of no importance to keep clear track of protons across reactions since the pH-buffer of the medium is large enough to assuage any problems. When chemiosmotic coupling plays a role however, one needs to keep track of protons that move across the membrane and because membrane potential is often the more important component of the proton motive force, contribute to the membrane potential. Thereby other charged species that move across that membrane should also be taken into account. However, because flux-balance analysis requires all species, i.e. also the protons, to be balanced, even one wrongly annotated proton can lead to problems including inaccurate bioenergetics/ATP synthesis. The modeler must account for each of these protons i.e. perform accurate bookkeeping of protons and do so while taking into account the protonation of the metabolites already existing in the reconstruction. Alternatively, one should become explicit in transmembrane charge movement, which is not customary in existing flux balance analysis. We listed the reactions (Table 1) that were manually added to the reconstruction but that are slightly different in terms of protons from those considered in Schuchmann and Müller (2014).

In the reconstruction downloaded from the BiGG database, the ACACT1r (Acetyl-CoA C-acetyltransferase) and HACD1 (3-hydroxyacyl-CoA dehydrogenase) reactions had been blocked, i.e. both the lower and upper bound had been set to zero, for unknown reasons. We unblocked these so as to allow flux into the beta-hydroxybutyrate synthesis pathway. Finally, we added a demand reaction for removing (S)-3-hydroxybutanoyl-CoA (while recycling the CoA) from the cell, so that we may predict its maximal production flux for various network perturbations.

Reaction ID	Formula
FDHH2	$CO_2 + H_2 \rightarrow Formate^- + H^+$
FDHFDNADPH	$2\,CO_2 + Fd^{2-} + NADPH + H^+ \rightarrow Fd + 2\,Formate^- + NADP^+$
MTHFD_alt	$Methenyl-THF^{2-} + NADH + Methylene-THF^{3-} + NAD^+$
MTHFR5_alt	$Methylene-THF^{3-} + NADH + 2\,H^+ \rightarrow Methyl-THF^{2-} + NAD^+$

Table 1. Reactions that were added to the genome scale reconstruction ('GeMM') from (Nagarajan, 2013). The genome-wide metabolic reconstruction with these reactions added will be referred to as GeMM*. Electric charges of compounds are important and subject to standards set by the existing GeMM. They present an issue to be elaborated upon elsewhere.

2.3 Flux balance analysis

In all simulations in this paper we apply the computational technique of FBA (Orth, 2010). Briefly, this technique concerns the following linear programming problem:

$$\text{maximize or minimize } Z = c^T v, \text{such that for all k}$$
$$Sv = 0$$
$$\alpha_k \leq v_k \leq \beta_k.$$

where S is the stoichiometric matrix for the metabolites, v is the vector of fluxes through all reactions including exchange reactions with the environment of the system considered, c is a vector of weights generating the linear combination of fluxes that make up the objective function Z and α and β are the vectors of lower and upper bounds on these fluxes. A flux distribution returned by FBA is therefore such that all metabolites are produced and consumed in equal amounts, the flux boundaries are accommodated and the flux distribution maximizes (or minimizes) a linear combination of fluxes in the model.

2.4 Model check: ATP from nothing in all considered scenarios

In the main text we perform FBA simulations with various combinations of enzyme alternatives. Ensuring our added reactions did not introduce errors in the energetics, we checked in all considered scenarios of Table 2 and 3 that the model could not generate Gibbs free energy from nothing. This was accomplished by preventing all medium components from being taken up and then asking for flux through the ATP maintenance reaction. In all cases this returned a maximal yield of zero, meaning the model is not capable of generating energy from nothing.

2.4 Visualization of flux distributions

Using the Escher software package (King et al., 2015) we produced a static map of a subset of the reactions encoded in our (updated) genome-scale metabolic model of *C. ljungdahlii*. This static map can be used, with help of the COBRApy FBA package (Ebrahim, 2013), to visualize flux distributions for any number of situations the modeler may wish to explore. Hence it becomes a fluid map that may generate new visualizations on the fly. Various such images featuring in this work are provided as supplementary files in SVG format. We hope this map may be utilized and extended by others to aid in future work making use of the ljungdahlii genome-wide metabolic map.

2.5 Reproducibility: python code for model construction and analysis

All changes to the published reconstruction were performed through a python script that loads the original model, performs the changes with aid of the COBRApy package (Ebrahim, 2013) and exports the updated model in SBML version 3 with the FBC package (Olivier and Bergmann, 2015; Chaouiya *et al.*, 2015). The python script, the original and modified model are available as supplementary files. Furthermore, we provide, as a supplementary file, a Jupyter notebook with Python code that reproduces all the analyses discussed in this work. All the discussed models, model analysis code and the visualizations are available on a publicly available Github repository at: https://github.com/ThierryMondeel/FOSBE_2016/.

3. RESULTS

C. ljungdahlii is one of the few organisms that are able to grow under the highly challenging conditions that may characterize the early *billenia* of Planet Earth, in which there was neither organic carbon abundant, nor any other complex source of Gibbs free energy, nor molecular oxygen, or other suitable electron acceptors: essentially, *C. ljungdahlii* is able to grow on carbon dioxide in the dark, using molecular hydrogen and elements required for its biomass. It shares this property with other acetogens, organisms that are able to produce acetate from CO_2 and molecular hydrogen:

$$2\,CO_2 + 4\,H_2 \rightarrow CH_3COO^- + H^+ + 2\,H_2O$$

It has been proposed that acetogens have this fairly unique capability not by possessing a unique protein (gene), but rather by having a unique pathway, *i.e.* the Wood-Ljungdahl pathway (Schuchmann and Müller, 2014). We examined whether the information comprised in the genome sequence of *C. ljungdahlii* and the annotation thereof, confirms this possibility under alteration of certain enzymes in the pathway. To do this, we used the genome-wide metabolic map (GeMM) already published in (Nagarajan, 2013). We changed medium conditions and the objective function to reflect the acetogenesis problem. We also added any missing reactions that were considered in Schuchmann and Müller (2014), see the Methods section.

3.1 A network feature: The Wood-Ljungdahl pathway and the hydrogenase

In line with the analysis in Schuchmann and Müller (2014) we asked the genome-wide metabolic map to produce acetate rather than biomass at zero maintenance choosing the variant of the formate dehydrogenase that oxidizes ferredoxin. This indeed yields a flux distribution with positive flux through the Wood-Ljungdahl pathway plus the hydrogenase, Nfn and Rnf complexes and the H^+-ATPase. We next deleted every enzyme of this pathway one by one and confirmed that in all cases except the formate dehydrogenase this eliminated the production of acetate. The formate dehydrogenase is not essential because the formate may be alternatively synthesized through the pyruvate formate lyase. This shows that the total genomic information on *C. ljungdahlii* contained in the annotated map, confirms the notion that almost the entire Wood-Ljungdahl *pathway* is essential for autotrophic growth of *C. ljungdahlii* in the presence of hydrogen gas. It is the *pathway* that matters, not *just* a single enzymatic step.

Furthermore, we investigated the essentiality of the hydrogenase. When deleting both variants of the hydrogenase (i.e. the one with Fd + NAD acceptors and the one with Fd + NADP as electron acceptors) the organism is predicted to be unable to make acetate. With either of these hydrogenases the organism is predicted to be able to produce the acid. This shows that the network property allowing for acetate production autotrophically is the presence of the full Wood-Ljungdahl pathway and a hydrogenase.

In *these* computations, we configured the Wood-Ljungdahl pathway so as to use reduced ferredoxin both to reduce carbon dioxide to formate and to reduce a second such molecule to carbon monoxide. In the former, ‘methyl’ branch

of the pathway, two molecules of NADH were used for the further reduction of the formate to the methyl group. The hydrogenase reduced two oxidized ferredoxin molecules and two NAD molecules by using molecular hydrogen as the electron donor. In these processes one ATP is used in the formyl-THF synthetase reaction, whereas a second one is produced by the reverse operation of acetate kinase. The process merely dissipates the 40 kJ/mol of Gibbs energy that the 4 hydrogen molecules plus 2 CO_2 molecules have in excess compared to the acetate molecule (Schuchmann and Müller, 2014); it does

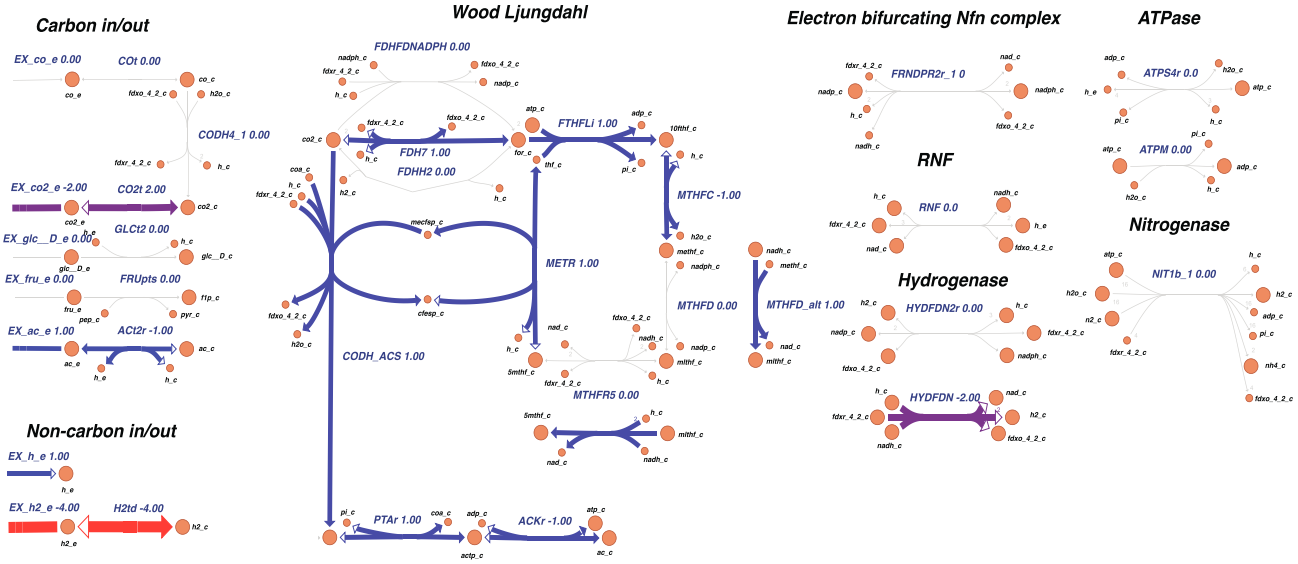

Fig. 1. Visualization of the flux distribution when asking for ATP and acetate on a CO_2 and H_2 mixture. CO_2 and H_2 are taken up in a 4:2 ratio and together reduce to 1 mole of acetate. To reproduce one of the scenarios in the analysis of Schuchmann and Müller, we allowed only the formate dehydrogenase alternative oxidizing ferredoxin, the hydrogenase reducing NAD^+ and the methylene-THF dehydrogenase and reductase as described in (Schuchmann and Müller, 2014) to carry flux. The combination of enzyme alternatives shown here does not have any further ability to couple this process to ATP synthesis.

not capture part of this Gibbs free energy in the form of phosphorylated ADP. Since some ATP is necessary for the conversion of acetate to biomass according to the biomass equation, the pathway shown in Fig. 1 can produce acetate but cannot produce growth. It cannot even take place in realistic organisms that are not growing, since these require Gibbs free energy for maintenance.

3.2 Reproducing Schuchmann and Müller Figure 3 with the genome-wide metabolic map

In Figure 3 of Schuchmann and Müller (2014) the ATP yield per acetate is considered for 6 different combinations of 2 hydrogenases and 3 forms of the formate dehydrogenase. The chemical reaction in Eq. 1 may be coupled to the generation of up to roughly 1 ATP. The precise yield of ATP per acetate depends on the electron donors and acceptors used in various steps in the WPL and the hydrogenase and on the presence or absence of electron bifurcating steps. Figure 3 in Schuchmann and Müller (2014) therefore considered the question of finding the unknown coefficient X in

$$2\,CO_2 + 4\,H_2 \rightarrow CH_3COO^- + H^+ + 2\,H_2O + X\,ATP$$

It is not trivial that asking the genome-wide metabolic map GeMM* for maximum ATP yield given a medium of 2 CO_2 + 4 H_2 would yield the same results as the theoretical analysis in Schuchmann and Müller (2014). First of all, reactions may be encoded with different stoichiometries and co-factors. Secondly, the complete genome might contain reactions not considered by Schuchmann and Müller. We therefore analyzed the agreement between the GeMM and the analysis by Schuchmann and Müller and extended it with variants of the methylene-THF reductase and the methylene-THF dehydrogenase that Schuchmann and Müller did not consider in their calculations. As a preliminary check we made sure the model was not capable of generating ATP from nothing in any of the scenarios, see methods section. We then performed FBA with the maintenance reaction as the objective function given the CO_2/H_2 medium described before and forcing the organism to complete the reaction $2\,CO_2 + 4\,H_2 \rightarrow 1\,Acetate$. The corresponding maximal ATP yields are listed on in the first two columns of Table 2. These predictions agree perfectly with the analysis of Schuchmann and Müller.

We included the flux through the Nfn complex in parentheses (Table 2). With respect to the suggestion by Schumann and Müller that electron bifurcation is essential for ATP production, this shows that if so, this need not be electron bifurcation through the Nfn complex. There are multiple scenarios in which positive ATP yield is obtained whilst the Nfn is inactive. Also note that the direction of the flux through the Nfn complex flux differs between scenarios.

3.3 Additional scenarios: alternatives, electron bifurcation and breaking the second law of thermodynamics

We expanded the scenarios considered (Schuchmann and Müller, 2014) by adding two variations on the MTHFD and MTHFR reactions. This yields a total number of 3*2*2*2 = 24 scenarios (Table 2). Judging from the simulation results and from careful manual bookkeeping, the electron bifurcating MTHFR5 reaction structurally yields an additional 0.5 ATP. Rather than using a single molecule of NADH to reduce methylene-THF to methyl-THF, it couples this exergonic reaction to the endergonic reduction of ferredoxin by NADH. This electron-bifurcating reaction yields a mole of extra NAD^+ and a mole of reduced ferredoxin per mole of flux through the WLP. Through the Rnf these can together pump 2 additional protons across the membrane which yields 0.5 ATP through the H^+-ATPase. Similarly, the methylene-THF dehydrogenase oxidizing NADH depicted in the analysis by Schuchmann and Müller yields 0.25 ATP more than the NADPH variant which was encoded in the GeMM. In the latter the NADP needs to be re-reduced at the cost of oxidation of NADH costing one half mole of reduced ferredoxin and NADH, which are made available by half a turnover less of the Rnf complex. This results in one proton less being pumped over the membrane by the latter and therefore a 0.25 loss in ATP. The FBA computation can break the second law of thermodynamics, by predicting and ATP/acetate yields > 0.8, the ratio allowed by the energetics of 40 kJ in the overall process reduction of CO_2 by hydrogen to acetate relative to the Gibbs energy of ATP synthesis from ADP and phosphate, which may be close to 48 kJ per mole (Westerhoff and Van Dam, 1987).

MTHFR	**NADH**				**2 NADH - Fd^0**			
MTHFD	**NADH**		**NADPH**		**NADH**		**NADPH**	
FDH \ HYD	**Fd + NAD**	**Fd + NADP**	**Fd + NAD**	**Fd + NADP**	**Fd + NAD**	**Fd + NADP**	**Fd + NAD**	**Fd + NADP**
Fd^{2-}	0.0 (0.0)	0.5 (-1.0)	NP	0.25 (-0.5)	0.50 (0.0)	1.0 (-1.0)	0.25 (0.5)	0.75 (-0.5)
Fd^{2-} + NADPH	0.125 (0.25)	0.625 (-0.75)	NP	0.375 (-0.25)	0.625 (0.25)	1.125 (-0.75)	0.375 (0.75)	0.875 (-0.25)
H2	0.25 (0.0)	0.625 (-0.75)	0.0 (0.5)	0.375 (-0.25)	0.75 (0.0)	1.125 (-0.75)	0.5 (0.5)	0.875 (-0.25)

Table 2. Predicted moles of ATP per acetate when performing flux-balance analysis on the extended genome-wide reconstruction of *C. ljungdahlii* for various electron donors and acceptors of the formate dehydrogenase, the hydrogenase and the methylene-THF reductase (MTHFR) and dehydrogenase (MTHFD), as indicated by the naming around the fields with numbers. The electron donor to MTHFR being indicated as 2 NADH - Fd^0 refers to an enzyme complex such as the one in *Moorella thermoacetica* (Schuchmann and Müller, 2014) where 2 molecules of NADH serve as electron donor and oxidized ferredoxin and CH_2-THF as electron acceptors, which is then another electron bifurcating reaction. The – in front of Fd^0 refers to its use as electron acceptor, not donor. The superscript to its formal electric charge. The results in italics (the two left-most columns) are in perfect agreement with the analysis by Schuchmann and Müller and here extended to include the methylene-THF dehydrogenase and reductase alternatives. In parentheses we show the flux through the Nfn complex reaction (relative to the acetate production flux of 1), counted positive when in the direction of oxidizing Fd^{2-} and NADH. Cells colored red are 'NP' signifying that it is not feasible to produce acetate; in practice due to a negative ATP yield. Cells colored orange are produced by the FBA simulations but thermodynamically not feasible. Cells colored yellow correspond to situations in which acetate may be produced but not coupled to ATP production. This similarly cannot occur in living organisms due to their maintenance ATP requirement.

MTHFR	**NADH**				**2 NADH - Fd^0**			
MTHFD	**NADH**		**NADPH**		**NADH**		**NADPH**	
FDH \ HYD	**Fd + NAD**	**Fd + NADP**	**Fd + NAD**	**Fd + NADP**	**Fd + NAD**	**Fd + NADP**	**Fd + NAD**	**Fd + NADP**
Fd^{2-}	0.29 (0.0)	0.47 (-0.65)	0.222 (0.22)	0.35 (-0.24)	0.50 (0.25)	0.50 (-0.13)	0.43 (0.5)	0.50 (-0.13)
Fd^{2-} + NADPH	0.36 (0.25)	0.50 (-0.25)	0.25 (0.38)	0.41 (0.06)	0.50 (0.75)	0.50 (0.12)	0.50 (0.75)	0.50 (0.12)
H_2	0.43 (0.0)	0.50 (-0.25)	0.29 (0.29)	0.41 (0.06)	0.50 (0.5)	0.50 (0.13)	0.50 (0.5)	0.50 (0.13)

Table 3. Beta-hydroxybutanoyl yields under the same scenarios as Table 1 when *C. ljungdahlii* is grown *in silico* on a CO/H_2 mixture in the ratio 2:4. Again, in brackets we highlight the flux through the Nfn reaction.

3.4 Electron bifurcation might also promote beta-hydroxybutyrate yield

C. ljungdahlii is currently also investigated with the perspective of producing precursors for biodegradable plastics from waste, for example in the context of the SYNPOL project (http://www.synpol.org), by growing the organism on syngas ($CO/CO_2/H_2$) and having it produce beta-hydroxybutyrate, possibly after inserting one or two reaction capabilities into its genome. Genome-wide metabolic modeling, like shown here, may help strain-design for improved yield and growth rates. Table 3 shows a similar

scheme to Table 2 for the objective of beta-hydroxybutanoyl synthesis from a CO/H_2 mixture. We conclude that the variants among the WPL enzymes and the hydrogenases and formate dehydrogenases could impact the yield of the product. This suggests that model-assisted genome-editing or over/under-expressing the optimal variants with high yield in Table 3 could improve the real-world yield of desired products like beta-hydroxybutyrate and its product poly-hydroxybutyrate (PHB), a biodegradable plastic. For *C. Ljungdahlii* this would require the engineering in of an epimerase and a polymerase. There may be similar implications for the possible production of other products of possible interest such as butane-2,3 diol and for ethanol as shown in (Nagarajan, 2013).

3.5 Gear shifting: the concept

Given the results in Fig. 3 of Schuchmann and Müller (2014) and the analysis above, why would the organism contain alternatives in its genome for the various redox reactions other than the combination yielding optimal ATP/acetate yield? One could argue that the less optimal alternatives are left-overs from a distant past in which life was more difficult. However, over the billions of years of evolution these unused genes would not have any selective advantage and have obtained many mutations rendering them inactive. Perhaps they are inactive, encoding pseudo enzymes (Leslie, 2013). An alternative is that these genes and the proteins they code for still serve some catalytic function yielding a selectable advantage. What could this function be? We posit that these less optimal alternatives might serve as lower speed gears in what could effectively be considered a gear-shifting energy system. Many of the alternatives computed in Table 2 should thereby be at the disposition of acetogenic organisms, allowing them smartly to regulate their expression in order to generate appropriate ATP yields. Why would an organism not always go for the most ATP? Gibbs energy dissipation, which goes at the cost of thermodynamic efficiency, is the thermodynamic driver of processes; more than just serving as the arrow of time, flux tends to increase with the increase in Gibbs energy dissipation over a process (Van der Meer *et al*., 1980). Producing more ATP at the same ATP/ADP ratio, coupled to acetogenesis, reduces Gibbs free energy dissipation and as a consequence decreases reaction rates. Thereby living organisms face a rate/efficiency trade-off (Westerhoff & Van Dam, 1987).

To examine this in a simplified setting, we asked the genome wide metabolic map to produce acetate from CO_2 and hydrogen gas in the presence of a maintenance flux. By stepwise increases in the upper bound of the maintenance flux the gear-shifting phenomenon can be enforced and shows a sequence of ever more ATP yielding pathways among linear combinations of those highlighted above (Table 3). Although we in a sense artificially construct this sequence of pathways, it shows that if the organism was to maximize a different 'objective' (more on one side of the trade-off) the different pathways could pop up.

There are other cases in biology where such gear shifting, or at least the possibility to use low gear becomes important under some conditions. An example is again a case where an organism has to deal with extreme conditions, now of high temperature, the pathway being the lower half of glycolysis. At moderate temperatures the step from glyceraldehyde 3-phosphate to 1,3-bisphosphoglycerate is already problematic in terms of the standard Gibbs free energy difference. This problem is solved by *S. cerevisiae* and by the human, by a rapid subsequent reaction maintaining the concentration of 1,3-bisphoshoglycerate low and thereby the Gibbs energy change across the GAPDH reaction negative. However, at 80 ℃ this fails to work. Here, by using a dynamic computational model (Kouril *et al*., 2012) we found (data not shown) that the hyperthermophylic *Archaea S. solfataricus* then chooses a lower gear (in terms of the number of ATP molecules made per pyruvate produced) by enlisting the non-phosphorylating GAPN reaction for the flux.

4. CONCLUSIONS

In this text we have shown that the two propositions by (Schuchmann and Müller, 2014), i.e. that early life on this planet, if exemplified by the acetogen *C. ljungdahlii*, depends on the Wood-Ljungdahl pathway and on electron bifurcation through the Nfn complex, require ramifications if judged from the genome-wide knowledge as associated with the present genome wide model (Nagarajan, 2013): The ATP synthesis required can be facilitated by the electron bifurcation in the Nfn complex, but may also be attained by other reactions in the absence of Nfn activity or with reverse flow through Nfn. In some of these cases electron bifurcation then occurs in other reactions such as the methylene-THF reductase. The strict requirement of the Wood-Ljungdahl pathway can also do with some ramification: it is actually the network consisting of the Wood-Ljungdahl pathway, a hydrogenase, the Rnf complex and the H^+-ATPase that is required, where it seems that the formate dehydrogenase can actually be missed, when pyruvate formate lyase digs in (but then additional enzymes are needed as well in order to synthesize the pyruvate from acetylCoA using reduced ferredoxin as electron donor to drive the CO_2 fixation). The ATP yield per acetate through the WLP predicted in Schuchmann and Müller (2014) can be reproduced using the genome-wide metabolic map of *C. ljungdahlii*, but only after many of the redox enzymes have been bestowed with the substrate and product specificities assumed by Schuchmann and Müller. Furthermore, one should be very precise about proton stoichiometries.

The main take-away from this is that indeed acetogens such as *C. ljungdahlii* (and Archaea such as *S. solfataricus*) may be capable of life at the edge of what is energetically possible by appropriating the optimal gear settings and gear boxes (in terms of electron bifurcation) of the redox network. Therefore, the work shown here may help highlight differences between the current genome-wide reconstructions' predictions and those made by precise

microbial biochemistry, and show how differences may be resolved. Finally, the analysis here (and in Schuchmann and Müller (2014)) suggest multiple interesting targets for strain-design: singularly expressing only those enzyme variants that generate the highest yield of biomass, or of desired metabolic products (Table 3).

Acknowledgements

The authors wish to acknowledge Olga Revelles, Stéphanie Follonier and Tanja Narancic for key discussions during the early stages of this work. The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under Grant Agreement No. 311815 (SYNPOL project).

References

Chaouiya, C., Keating, S.M., Berenguier, D., Naldi, A., Thieffry, D., van Iersel, M.P., Le Novère, N., Helikar, T. (2015). The Systems Biology Markup Language (SBML) Level 3 Package: Qualitative Models, Version 1, Release 1. *Journal of Integrative Bioinformatics*, 12(2): 270.

Ebrahim, A., Lerman, J.A., Palsson B.O., Hyduke D.R. (2013). COBRApy: COnstraints-Based Reconstruction and Analysis for Python. *BMC Systems Biology*, 7(1):74.

King, Z.A., Dräger, A., Ebrahim, A., Sonnenschein, N., Lewis, N.E., Palsson, B.O. (2015). Escher: A Web Application for Building, Sharing, and Embedding Data-Rich Visualizations of Biological Pathways. P. P. Gardner, ed. *PLoS Computational Biology*, 11(8): e1004321.

King, Z.A., Lu, J., Dräger, A., Miller, P., Federowicz, S., Lerman, J.A., Ebrahim, A., Palsson, B.O., Lewis, N.E. (2016). BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Research*, 44(D1), D515-D522.

Kouril, T., Kolodkin, A., Zaparty, M., Steuer, R., Ruoff, P., Westerhoff, H.V., Snoep, J., Siebers, B., Sulfo, S Y S Consortium, (2012). Sulfolobus Systems Biology: Cool hot design for metabolic pathways. *Systems Microbiology: Current Topics and Applications*, p.151.

Leslie M. (2013). 'Dead' enzymes show signs of life. *Science*, 340(6128), 25-27. doi: 10.1126/science.340.6128.25.

Nagarajan, H., Sahin, M., Nogales, J., Latif, H., Lovley, D.R, Ebrahim, A., Zengler, K. (2013). Characterizing acetogenic metabolism using a genome-scale metabolic reconstruction of *Clostridium ljungdahlii*. *Microbial Cell Factories*, 12(1), 118.

Olivier, B.G. and Bergmann, F.T. (2015). The Systems Biology Markup Language (SBML) Level 3 Package: Flux Balance Constraints. *Journal of Integrative Bioinformatics*, 12(2): 269.

Orth, J.D., Thiele, I. and Palsson, B.Ø. (2010) What is flux balance analysis? *Nature Biotechnology*, 28(3), 245–248. doi: 10.1038/nbt.1614.

Schuchmann, K. and Müller, V. (2014). Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria. *Nature Reviews Microbiology*, 12(12), 809–821.

Van der Meer, R., Westerhoff, H. V and Van Dam, K. (1980). Linear relation between rate and thermodynamic force in enzyme-catalyzed reactions. *Biochimica et Biophysica Acta*, 591(2), 488–493.

Westerhoff, H.V. and Van Dam, K. (1987). Thermodynamics and control of biological free-energy transduction. Elsevier, Amsterdam.