

## Optimization of *Clostridium acetobutylicum* metabolism for biobutanol production using *in silico* tools.

C. Portela <sup>1</sup>, S. Freitas <sup>1</sup>, I. Rocha <sup>2</sup>

<sup>1</sup> Brazilian Bioethanol Science and Technology Laboratory – CTBE, Campinas, São Paulo

<sup>2</sup> Departamento de Engenharia Biológica, Universidade do Minho, Braga, Portugal

E-mail contact: carla.portela@bioetanol.org.br, portela.carla@gmail.com

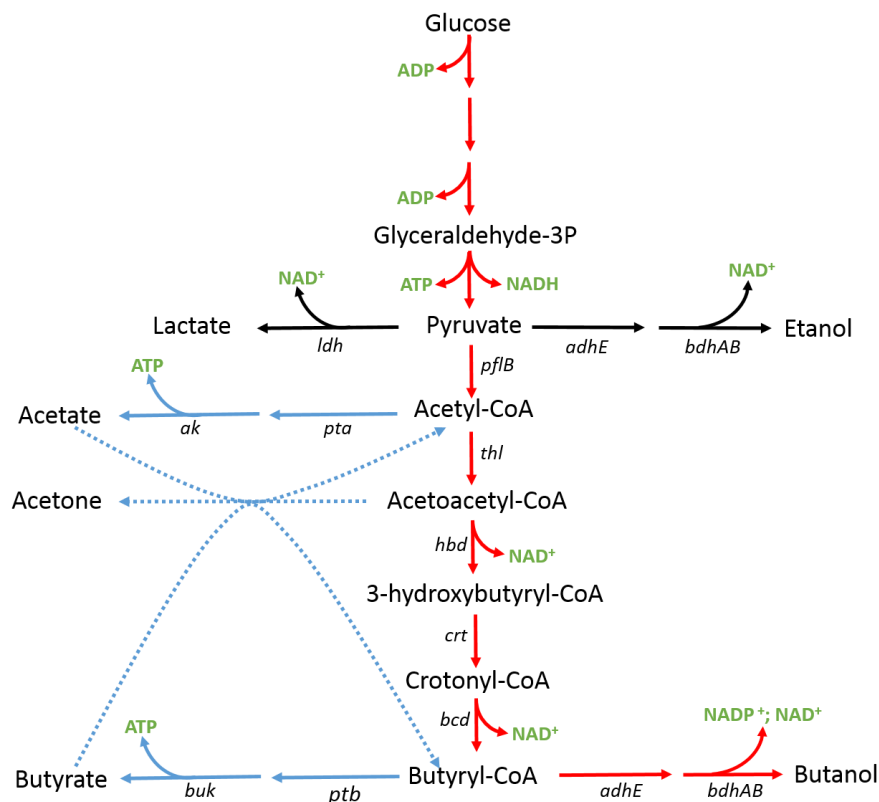
**ABSTRACT** – Alternative biofuels to petroleum-derived compounds are in high demand nowadays. Ethanol has played an important role in the picture as the major biofuel currently produced [1], [2]. However, advances in technology and the current search for alternatives biofuels, brought butanol into the scenario. The characteristics of this primary alcohol make it an interesting alternative to ethanol as it can directly replace gasoline without the need of engine adaptation. Moreover, it is an excellent diluent and solvent for manufacturing intermediates of chemical industry augmenting the potential of this compound. With the most recent Genome-scale Model (GEM) of *C. acetobutylicum* available, new *in silico* strategies can be further explored to improve butanol production. By modeling clostridial metabolism, new insights on the acidogenic-solventogenic metabolism are debated in this work.

### 1. INTRODUCTION

Research on biofuels has been growing over the years due to the pressure caused by the cost and availability of fossil-derived fuels along with an environmental conscience related to an increasing biomass waste produced from agricultural activities [3]. Sustainable technologies for alternative fuel production are on demand in the globalized world we live in.

Butanol has been regarded as an interesting alternative because it can be directly used in the existing engines without need for adaptation, it has a higher energy density and better performance than ethanol [4]. Additionally, butanol and its derivatives can also be used for surface coating, plasticizers, diluent and solvent for antibiotics and hormones, confirming the versatility and market interest of this compound [5].

*Clostridium* genera are of particular interest in this field since some species are natural producers of butanol, such as *Clostridium acetobutylicum* and *Clostridium beijerinckii* with butanol titers as high as 18g/L [3].



**Figure 1. Central metabolism of *C. acetobutylicum* for acids and solvent production. The numbering from (1) to (3) are flux ratios defined later in the results section.**

*C. acetobutylicum* is a spore forming, gram positive, anaerobic bacterium. Solventogenic organisms such as *C. acetobutylicum* can utilize a wide variety of substrates for growth with concomitant production of solvents. These organisms normally carry a biphasic fermentation, the first phase being known as acidogenesis and the second as solventogenesis (Figure 1). Acidogenic phase is characterized by the degradation of sugars via glycolysis coupled to acid production. The major end products are acetate and butyrate along with production of ethanol, hydrogen and CO<sub>2</sub>. Due to acid production, the pH drops, which triggers the *Clostridium* cells to switch its metabolism into a solventogenic phase which is characterized by the re-consumption of the acids and further conversion into solvents, namely butanol, acetone and ethanol.

With the complete sequenced genome of *C. acetobutylicum*, *in silico* genome-scale models (GEM) have been reconstructed [6]–[8] in order to better understand its metabolism and accurately predict the *in vivo* phenotypes using mathematical tools.

This model was analyzed and compared with the published data using Optflux software which is a computational tool developed to perform *in silico* simulations, manipulate and analyze GEM in an user friendly interface [9]. This software gives the researcher the flexibility required to perform different tasks such as to perform simulations and/or metabolic engineering optimizations.

The purpose of this work was to gather the *in silico* knowledge so far obtained from different

studies related to *C. acetobutylicum* and use mathematical approaches to explore the metabolism of this organism in search for new insights and strategies to enhance butanol production.

## 2. Metodology

The GEM of *C. acetobutylicum* was loaded in the Optflux software using the available repositiorium. Next, some errors related to repeated reactions ID's which lead to incongruences in the simulations (each reaction must have an unique ID) were corrected.

Performed simulations analyzed and compared the published results since the model is intended be further used to predict new capabilities in terms of butanol production enhancement. One of the approaches used in this work included the addition of flux ratios constraints (FBrAtios) as defined in the published work [8]. FBrAtios are flux ratios imposed to the model to force flux distribution according to what is described *in vivo* about the metabolism. That limits the flux values and force the model to find alternatives to maximize the objective function.

The other approach consisted in minimizing the set of constraints imposed to the model relaxing the model (in terms of degrees of freedom) while still predicting with accuracy the qualitative behavior of the organism. The simulations used parsimonius FBA (pFBA) which is a variant of flux balance analysis (FBA). FBA consists in maximizing an objective function (generally biomass production) assuming a steady state approximation [10]. In pFBA, the same principles apply, but the objective function is the minimization of the sum of the fluxes. The applied constraint exerts a pressure over the maximum theoretical biomass production limiting this value. Therefore, it is expected to observe a distribution of the fluxes over the other expected end products.

Results were analyzed in terms of the predicted capabilities of the model and the potential to design new metabolic engineering strategies.

## 3. Results

### 3.1. FBrAtios with metabolic flux analysis (MFA)

The first approach uses FBrAtios and assumes the inclusion of flux ratios to constraint the space of possible solutions in order to predict qualitatively the expected phenotype. Additionally, glucose uptake was qualitatively defined to be  $10 \text{ mmol.gDW}^{-1}.\text{h}^{-1}$ . The simulations were performed using metabolic flux analysis (MFA) strategy, i.e., the analysis of flux distributions using pFBA approach.

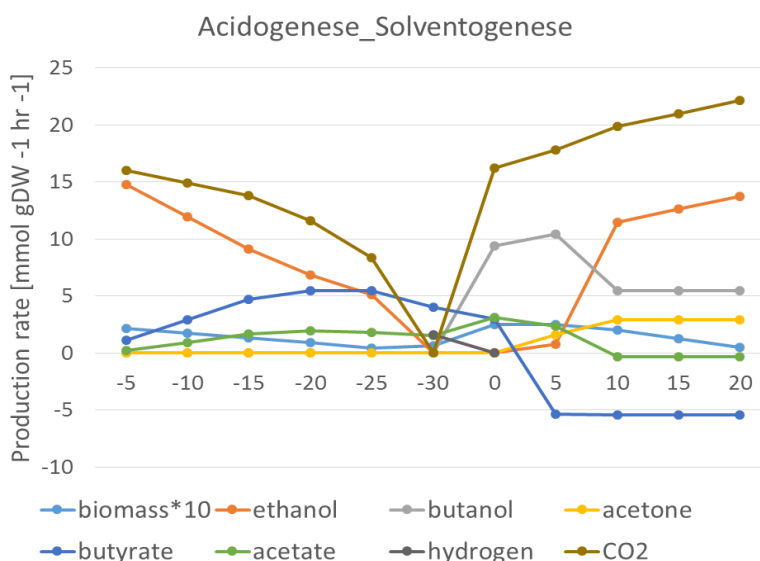
However, when including all the flux ratios and constraints defined in the published paper, it was observed that the model, was not able to simulate. Therefore, the flux ratios were included one-by-one and the simulations performed for varying SPF values.

The flux ratios are illustrated in Figure 1 and were defined as following:

- (1)  $f(r_{THL})/f(r_{PTA})=2$ , which forces twice the flux towards the acetoacetyl-CoA formation rather than acetate formation.
- (2)  $f(r_{CoAT,butyrate})/f(r_{CoAT,acetate})=0.63$ , which defines a higher flux over acetate consumption rather than butyrate. The value was defined from previously published data [11].
- (3)  $f(r_{PFO})/f(r_{LDH})=10$ , which defines a higher flux from pyruvate towards the formation of the acids rather than lactate production.

Interestingly, the inclusion of these three flux ratios were enough to predict qualitatively the metabolism of *C. acetobutylicum*. Nevertheless, the results differ from those published by McAnulty *et al.*, 2012, which can be explained by the different algorithms used to solve the stoichiometric matrix.

**Erro! Fonte de referência não encontrada.** displays the simulation results obtained. It can be seen the results qualitatively fit the wild type metabolism. In the acidogenic phase there is a continuous production of butyrate and acetate which are later re-consumed in the solventogenic phase. Butyrate levels are higher than acetate levels which is also observed *in vivo*.



**Figure 2. Simulations of the shift between acidogenesis and solventogenesis. Simulations were performed considering the given environmental conditions: (i) glucose uptake  $10 \text{ mmol.gDW}^{-1}.\text{h}^{-1}$ , (ii) varied SPF, (iii) FBrAtios:  $f(r_{THL})/f(r_{PTA})=2$ ,  $f(r_{CoAT,butyrate})/f(r_{CoAT,acetate})=0.63$ ,  $f(r_{PFO})/f(r_{LDH})=10$ .**

During solventogenic phase, it is observed an increase in butanol production with a maximum level at  $\text{SPF}=5 \text{ mmol.gDW}^{-1}.\text{h}^{-1}$ . The butanol production decreases thereafter, but remains higher than acetone production.

Ethanol profile however, does not match the *in vivo* profile expected. This metabolite is

produced during acidogenesis as a sink to balance redox cofactor  $\text{NAD}^+/\text{NADH}$  and it is not expected to be re-consumed like the acids for solvent production as it seems to be occurring *in silico*.

Although during acidogenic phase,  $\text{CO}_2$  concentration appears to be high with small SPF values, in the solventogenesis metabolism the production of  $\text{CO}_2$  concomitant with acetone production is expected [8]. It seems the model is able to produce  $\text{CO}_2$  and acids (during acidogenesis) which indicates the model is producing  $\text{CO}_2$  in other reactions of the model not associated with acetone formation.

### 3.2. Wild type simulations using pFBA

The second approach aimed to simulate the model minimizing the number of constraints imposed and still being able to predict a qualitative phenotype of the wild type.

In this approach, the model was optimized for acid production with a maximum glucose intake of  $10 \text{ mmol.gDW}^{-1}.\text{h}^{-1}$ . The constraints imposed to the model were defined to block the flux to acetone and butanol formation since these are not produced in acidogenesis, therefore:

- (1)  $-10 < f(\text{glucose}) < 1000 \text{ mmol.gDW}^{-1}.\text{h}^{-1}$ ;
- (2)  $0 < f(\text{acetone}) < 0 \text{ mmol.gDW}^{-1}.\text{h}^{-1}$ ;
- (3)  $0 < f(\text{butanol}) < 0 \text{ mmol.gDW}^{-1}.\text{h}^{-1}$ .

Flux distributions indicate the production of either acetate ( $4.725 \text{ mmol.gDW}^{-1}.\text{h}^{-1}$ ) or butyrate ( $4.720 \text{ mmol.gDW}^{-1}.\text{h}^{-1}$ ) (Table 1). These end products are not produced together in the same simulation, most likely due to a mathematical “limitation”. The solver algorithm used maximizes the objective function independently of the end products formed. If the production of both acetate and butyrate is not mandatory for biomass optimization but rather a regulatory mechanism, the end product formation will be independent of the expected *in vivo* phenotype. Nevertheless, this result does not affect the following simulations, as the environmental constraints defined next could cope with that limitation.

**Table 1. Simulations using pFBA. (A) Acetate production, (B) Butyrate production.**

| <b>(A)</b>    |  |                      |  | <b>(B)</b>    |  |                      |  |
|---------------|--|----------------------|--|---------------|--|----------------------|--|
| Consumed      | $f/\text{mmol.gDW}^{-1}.\text{h}^{-1}$ | Products             | $f/\text{mmol.gDW}^{-1}.\text{h}^{-1}$ | Consumed      | $f/\text{mmol.gDW}^{-1}.\text{h}^{-1}$ | Products             | $f/\text{mmol.gDW}^{-1}.\text{h}^{-1}$ |
| D-Glucose     | 10.00                                  | Succinate            | 0.491                                  | Ortho-P       | 0.648                                  | Succinate            | 0.491                                  |
| $\text{NH}_3$ | 3.485                                  | $\text{H}_2\text{O}$ | 2.934                                  | $\text{NH}_3$ | 3.483                                  | <b>Butanoic acid</b> | <b>4.720</b>                           |
| Sulfate       | 0.042                                  | Ethanol              | 9.372                                  | D-Glucose     | 10.00                                  | Ethanol              | 4.653                                  |
| Ortho-P       | 0.648                                  | $\text{H}^+$         | 1.631                                  | Sulfate       | 0.042                                  | $\text{H}_2\text{O}$ | 7.647                                  |
|               |  | $\text{CO}_2$        | 13.980                                 |               |  | $\text{CO}_2$        | 13.983                                 |
|               |  | Hydrogen             | 9.060                                  |               |  | $\text{H}^+$         | 1.635                                  |
|               |  | <b>Acetate</b>       | <b>4.725</b>                           |               |  | Hydrogen             | 9.066                                  |
|               |  | <b>Biomass</b>       | <b>0.405</b>                           |               |  | Acetate              | 0.008                                  |
|               |  |                      |  |               |  | <b>Biomass</b>       | <b>0.405</b>                           |

Therefore, the next simulation aimed to maximize solventogenesis. Since the model was able to produce either acetate or butyrate when maximizing acidogenesis, we defined a new set of environmental conditions that accounted for that capability while also induced the consumption of the previously produced acetate and butyrate:

- (1)  $-10 < f(\text{glucose}) < 1000 \text{ mmol.gDW}^{-1}.\text{h}^{-1}$ ;
- (2)  $-4.72 < f(\text{acetate}) < -4.72 \text{ mmol.gDW}^{-1}.\text{h}^{-1}$ ;
- (3)  $-4.72 < f(\text{butyrate}) < -4.72 \text{ mmol.gDW}^{-1}.\text{h}^{-1}$ ;
- (4)  $0 < f(\text{biomass}) < 0.4 \text{ mmol.gDW}^{-1}.\text{h}^{-1}$ .

These constraints will induce the consumption of the acids while simultaneously limit the biomass formation. This artifact will ensure that the glucose will be responsible for biomass formation while the acids will have to be channeled to other end products such as, acetone and butanol as expected *in vivo*.

As observed in Table 2, with a minimal set of constraints it is also possible to simulate qualitatively the expected *in vivo* phenotype. Additionally, it is also clear that the acids can be converted into the corresponding solvents at the expense of a decrease in biomass levels. Since solventogenesis causes the cells to enter in a stationary phase, a decrease in biomass would be also expected.

This approach seems to represent better the expected phenotype without the need to manipulate the SPF values.

**Table 2. Simulations using pFBA with solventogenic environmental conditions defined to simulate *in vivo* phenotype.**

| Consumed | $f/\text{mmol.gDW}^{-1}.\text{h}^{-1}$ | Products         | $f/\text{mmol.gDW}^{-1}.\text{h}^{-1}$ |
|----------|--|------------------|--|
| 0,438    | Ortho-P                                | Succinate        | 0,381                                  |
| 2,357    | NH <sub>3</sub>                        | <b>Biomass</b>   | <b>0,274</b>                           |
| 10,00    | D-Glucose                              | (S)-Malate       | 0,062                                  |
| 4,710    | <b>Butanoic acid</b>                   | <b>Ethanol</b>   | <b>0,150</b>                           |
| 0,028    | Sulfate                                | H <sub>2</sub> O | 9,788                                  |
| 4,710    | <b>Acetate</b>                         | CO <sub>2</sub>  | 9,541                                  |
|          |  | (S)-Lactate      | 10,978                                 |
|          |  | <b>Acetone</b>   | <b>4,715</b>                           |
|          |  | <b>1-Butanol</b> | <b>4,710</b>                           |

## 4. Discussion and Conclusion

The overall results indicate that the *C. acetobutylicum* model can be used to predict some *in vivo* capabilities of the microorganism.

It was observed that the methodology described in the work of McNulty *et al.*, 2012 is an interesting approach to reduce the space of possibilities when trying to simulate an expected

phenotype [8]. However, we were not successful when using the same constraints to simulate the phenotype described in the published paper, which may be due to some missing information or simply because the approach utilized to solve the stoichiometric matrix differed. Nevertheless, it is our understanding that multiple constraints imposed in the model with the FBrAtio approach may be a limiting methodology when aiming at performing optimizations, since the cone of possible solutions is reduced with the increasing amount of constraints. With such approach, although one can mimic more accurately a specific phenotype, it will also reduce the space of solutions where a possible knock-out option will no longer be available due to the tight constraining methodology applied which could ultimately influence an overproduction of our compound of interest.

For that reason, we have also tried a different approach where an opposite approach based on the minimal number of constraints were applied.

We observed that when imposing some constraints to the model such as forcing the consumption of the maximum theoretical values obtained for butyrate and acetate production during acidogenesis, we were able to produce the corresponding solvents.

The limitation in this approach relies on the fact that, the values are most likely over estimated since we know from the literature that butyrate and acetate are never completely depleted [12].

Nevertheless, by simple creating an artifact limiting the maximum theoretical biomass production, we redirected the system to convert the carbon sources into other possible products, namely the production of acetone and butanol.

Finally, this strategy seems to better fit the global purpose of our work, which is the optimization of *C. acetobutylicum* GEM in order to find new metabolic engineering targets to increase butanol production. It seems reliable to assume that the lesser the number of constraints imposed to the model the higher the probability of not cutting off possible solutions that can ultimately prove to be interesting to increase butanol production.

## 5. REFERÊNCIAS

- [1] S. Prasad, A. Singh, and H. C. Joshi, "Ethanol as an alternative fuel from agricultural, industrial and urban residues," *Resour. Conserv. Recycl.*, vol. 50, no. 1, pp. 1–39, Mar. 2007.
- [2] M. O. S. Dias, A. V. Ensinas, S. a. Nebra, R. Maciel Filho, C. E. V. Rossell, and M. R. W. Maciel, "Production of bioethanol and other bio-based materials from sugarcane bagasse: Integration to conventional bioethanol production process," *Chem. Eng. Res. Des.*, vol. 87, no. 9, pp. 1206–1216, Sep. 2009.
- [3] Y.-S. Jang, A. Malaviya, C. Cho, J. Lee, and S. Y. Lee, "Butanol production from renewable biomass by clostridia," *Bioresour. Technol.*, vol. 123, pp. 653–63, Nov. 2012.



- [4] E. M. Green, "Fermentative production of butanol--the industrial perspective.," *Curr. Opin. Biotechnol.*, vol. 22, no. 3, pp. 337–43, Jun. 2011.
- [5] S. Y. Lee, J. H. Park, S. H. Jang, L. K. Nielsen, J. Kim, and K. S. Jung, "Fermentative butanol production by *Clostridia*.,” *Biotechnol. Bioeng.*, vol. 101, no. 2, pp. 209–28, Oct. 2008.
- [6] R. S. Senger and E. T. Papoutsakis, "Genome-scale model for *Clostridium acetobutylicum*: Part I. Metabolic network resolution and analysis.,” *Biotechnol. Bioeng.*, vol. 101, no. 5, pp. 1036–52, Dec. 2008.
- [7] J. Lee, H. Yun, A. M. Feist, B. Ø. Palsson, and S. Y. Lee, "Genome-scale reconstruction and in silico analysis of the *Clostridium acetobutylicum* ATCC 824 metabolic network.,” *Appl. Microbiol. Biotechnol.*, vol. 80, no. 5, pp. 849–62, Oct. 2008.
- [8] M. J. McNulty, J. Y. Yen, B. G. Freedman, and R. S. Senger, "Genome-scale modeling using flux ratio constraints to enable metabolic engineering of clostridial metabolism in silico.,” *BMC Syst. Biol.*, vol. 6, no. 1, p. 42, Jan. 2012.
- [9] I. Rocha, P. Maia, P. Evangelista, P. Vilaça, S. Soares, J. P. Pinto, J. Nielsen, K. R. Patil, E. C. Ferreira, and M. Rocha, "OptFlux: an open-source software platform for in silico metabolic engineering.,” *BMC Syst. Biol.*, vol. 4, no. 1, p. 45, Jan. 2010.
- [10] M. Rocha, P. Maia, R. Mendes, J. P. Pinto, E. C. Ferreira, J. Nielsen, K. R. Patil, and I. Rocha, "Natural computation meta-heuristics for the in silico optimization of microbial strains.,” *BMC Bioinformatics*, vol. 9, p. 499, Jan. 2008.
- [11] R. P. Desai, L. K. Nielsen, and E. T. Papoutsakis, "Stoichiometric modeling of *Clostridium acetobutylicum* fermentations with non-linear constraints,” *J. Biotechnol.*, vol. 71, pp. 191–205, 1999.
- [12] W. Andersch, H. Bahl, and G. Gottschalk, "Level of Enzymes Involved in Acetate , Butyrate , Acetone and Butanol Formation by *Clostridium acetobutylicum*,” *Appl. Microbiol. Biotechnol.*, vol. 3, no. 18, pp. 327–332, 1983.