Antonio Espuña, Moisès Graells and Luis Puigjaner (Editors), Proceedings of the 27th European Symposium on Computer Aided Process Engineering – ESCAPE 27
October 1st - 5th, 2017,Barcelona, Spain
http://dx.doi.org/10.1016/B978-0-444-63965-3.50478-5

Metabolic Network design of *Synechocystis* sp. PCC 6803 to obtain bioethanol under autotrophic conditions

Romina Lasry Testa[a], Claudio Delpino[a], Vanina Estrada[a], M. Soledad Diaz[a]

[a]*Planta Piloto de Ingeniería Química (PLAPIQUI), Universidad Nacional del Sur (UNS)-CONICET, Camino La Carrindanga Km 7,Bahía Blanca 8000, Argentina*
sdiaz@plapiqui.edu.ar

Abstract

In this work, we propose a genomic scale metabolic network model of the genetically engineered *Synechocystis* sp. PCC 6803 within a bilevel programming framework to study ethanol photosynthetic production. The model is studied under carbon limiting conditions with restricted photon flux. Maximum biomass and ethanol theoretical productions are obtained using flux balance analysis for the decoupled case. Furthermore, we formulate a bilevel programming problem, reformulated into a mixed integer linear problem (MILP), to study the possibility of coupling cell growth with ethanol production. Models are formulated within an equation-oriented framework in GAMS.

Numerical results provide useful insights on ethanol production by this strain within the context of genomic-scale cyanobacterial metabolism.

Keywords: *Synechocystis*, Ethanol, Bilevel Programming, Knock-outs.

1. Introduction

Cyanobacteria constitute a group of photoautotrophic prokaryote microorganisms capable of performing oxygenic photosynthesis, therefore using water as electron source, solar energy as photon source and carbon dioxide as carbon source. These characteristics have made them widely studied as candidates to becoming microbial cell factories for the production of renewable energy and high value chemicals, as they can grow on inexpensive raw materials. Within cyanobacteria, *Synechocystis* is considered a model photosynthetic microorganism due to its capacity of growing photoautotrophically during light periods and heterotrophically at the expense of reduced carbon sources during dark periods (Nogales et al., 2012). A few authors have studied *Synechocystis* for bioethanol production, modifying its genome by including *Zymomonas mobilis* (an obligate ethanol producer) genes that codify for pyruvate decarboxylase and alcohol dehydrogenase.

Despite experimental efforts made to improve photoautotrophic ethanol production by cyanobacteria (from the initial 0.0082 to the actual 0.552 g $L^{-1}day^{-1}$,Dexter et al. (2015)), the productivity rate is still far from heterotrophic production (Hjersted & Henson, 2008). One strategy for enhancing microbial production of several compounds is to genetically modify cells to couple growth with production of the target compound. In this way, maximization of the growth rate implies the consequent increase in the synthesis of the target compound. In recent years, much effort has been made to find in silico strategies that couple growth with product formation in microorganisms. The mathematical

procedures focus on finding reactions that, when eliminated from the network, couple the production of a target compound with cell growth. The formulation of a bilevel optimization problem, as proposed by Burgard et al. (2003) seeks to find a minimal set of reaction removals to achieve the desired coupling, while the strategy proposed by Ranganathan et al. (2010) identifies reaction removals and up and down regulations of fluxes. The two procedures are based on a bilevel programming approach and represent discrete decisions associated to gene knock-outs by binary variables, resulting in mixed integer linear problems (MILP). These approaches were used to couple succinate production with growth in *Escherichia coli* with positive results (Chowdhury et al., 2015). In the case of photosynthetic cyanobacteria, there have been studies that present coupled production of Limonene, 1-Butanol and 1-Octanol with biomass formation (Shabestary and Hudson, 2016), but there are no studies on coupling ethanol production to cell growth in photosynthetic microorganisms.

In this work, we present an extended genomic scale model and a strategy to couple ethanol production to biomass formation in *Synechocystis* within a bilevel programming structure based on the available procedures. Numerical results provide useful insights on ethanol production by this strain, while also presenting the use of optimization tools over genome scale reconstructions of cyanobacterial metabolism.

2. Metabolic Network of *Synechocystis* sp. PCC 6803

Several genome-scale reconstructions of *Synechocystis* sp.PCC6803 (*Synechocystis*) have been published in the last ten years. In this work, we extend the model proposed by Knoop et al. (2013), available in SBML format(Hucka et al., 2003), as it is the latest and most complete. As we study ethanol production, we have included the heterologous gene *pdc* from *Zymomonas mobilis*. This gene encodes the enzyme pyruvate decarboxylase (PDC), which catalyses the non-oxidative decarboxylation of pyruvate (PYR) to produce acetaldehyde (ACAL) and CO_2. We added exchange reactions for metabolites that can be released to the culture medium, like acetate, pyruvate, 2-oxoglutarate, fumarate, succinate, citrate and (S)-malate (Nogales et al., 2012), and we turned the reaction corresponding to the acetyl-CoA synthetase enzyme to its reversible form (Ihlenfeldt and Gibson, 1977). After these modifications, the model comprises 671 reactions and 523 metabolites. The reconstruction includes reversible and irreversible reactions. All changes in the model were made using the functions provided in the COBRA Toolbox (Schellenberger et al., 2011) running in MATLAB.

In all cases, we consider autotrophic growth and carbon limited states. Carbon is taken up as bicarbonate and limited with an upper bound of 3.7 mmolgDW^{-1} h^{-1}. The photon flux is limited with an upper bound of 60 mmol gDW^{-1} h^{-1} to control the light input to the network. Glucose and other possible carbon sources uptakes are set to zero to ensure autotrophic conditions. The ATP flux for maintenance is fixed to 0.13 mmol gDW^{-1} h^{-1}. Following Knoop et al. (2013), basal respiratory activity in the light through the terminal oxidase, as well as the conversion of NADPH and O_2 to H_2O and NADP (Mehler-like reaction) are assumed to take up 10% of O_2 evolution of photosystem II. To account for oxidative stress, 0.5% of the respective electron flow through both photosystems (PSI and PSII) is assumed for the production of superoxide by the plant-type Mehler reactions.

3. Model Formulation and conditions

In previous work (Lasry Testa *et al*., 2016), we studied the behaviour of the network under different limiting conditions showing that for the strain considered, growth and ethanol production are decoupled.

To obtain the maximal theoretical growth rate and ethanol production we use the Flux Balance Analysis approach (FBA) as suggested by Savinell and Palsson(1992). This approach is also used to study flux distributions presented in this work.

Bilevel Problem

$$\min_{y_k} \sum_j y_k$$

$$s.t.$$

$$\left\{ \begin{array}{lll} \min v_{ethanol} & & \\ s.t. & & \\ v_{ethanol} \geq v_{ethanol}^{\min} & & (1) \\ v_{biomass} \geq v_{biomass}^{\min} & & (2) \\ \sum_j S_{i,j} v_j = 0 & \forall i \in M & (3) \\ v_j = v_j^{fixed} & \forall j \in R^{fixed} & (4) \\ v_j \geq LB_j (1 - y_k) & \forall j \in R & (5) \\ v_j \leq UB_j (1 - y_k) & \forall j \in R & (6) \\ LB_j \leq v_j \leq UB_j & \forall j \in R & (7) \end{array} \right\}$$

$$v_j \in \mathbf{R}, \quad y_k \in \{0,1\}$$

Single level MILP Problem

$$\min_{\substack{v_j, w_j, w_i^{MB}, x_j^{LB}, \\ x_j^{UB}, \alpha_j, \alpha_j^{LB}, \alpha_j^{UB}, y_k}} \sum_j y_k$$

$$s.t.$$

$$-v_{ethanol} = \sum_{j \in fixed} w_j v_j^{fixed} + \ldots$$

$$\ldots + \sum_{j \in candidates} \left[LB_j \left(\alpha_j^{LB} - x_j^{LB} \right) + UB_j \left(x_j^{UB} - \alpha_j^{UB} \right) \right] \quad (8)$$

$$v_{ethanol} \geq v_{ethanol}^{\min} \quad (1)$$

$$v_{biomass} \geq v_{biomass}^{\min} \quad (2)$$

$$\sum_j S_{i,j} v_j = 0 \qquad \forall i \in M \quad (3)$$

$$v_j = v_j^{fixed} \qquad \forall j \in R^{fixed} \quad (4)$$

$$v_j \geq LB_j (1 - y_k) \qquad \forall j \in R \quad (5)$$

$$v_j \leq UB_j (1 - y_k) \qquad \forall j \in R \quad (6)$$

$$LB_j \leq v_j \leq UB_j \qquad \forall j \in R \quad (7)$$

$$\sum_i S_{i,j} w_i^{MB} \geq -1 \qquad \vee j = ethanol \quad (9)$$

$$\sum_i S_{i,j} w_i^{MB} + w_j = 0 \qquad \forall j \in R^{fixed} \quad (10)$$

$$\sum_i S_{i,j} w_i^{MB} - x_j^{LB} + x_j^{UB} = 0 \qquad \forall j \in R^{candidates} \quad (11)$$

$$\left. \begin{array}{l} M_j y_k \geq \alpha_j^{LB} \\ \alpha_j^{LB} \geq x_j^{LB} - M_j (1 - y_k) \\ x_j^{LB} \geq \alpha_j^{LB} \end{array} \right\} \forall j \in R \quad (12.1),\ (12.2),\ (12.3)$$

$$\left. \begin{array}{l} M_j y_k \geq \alpha_j^{UB} \\ \alpha_j^{UB} \geq x_j^{UB} - M_j (1 - y_k) \\ x_j^{UB} \geq \alpha_j^{UB} \end{array} \right\} \forall j \in R \quad (13.1),\ (13.2),\ (13.3)$$

$$v_j \in \mathbf{R},\ w_j \in \mathbf{R},\ w_i^{MB} \in \mathbf{R},\ x_j^{LB} \in \mathbf{R}^+,\ x_j^{UB} \in \mathbf{R}^+,\ y_k \in \{0;1\}$$

To study the possibility of coupling ethanol production to biomass formation in *Synechocystis,* we formulate a bi-level problem, as suggested by Ranganathan et al. (2010). The inner level problem minimizes ethanol production subject to the metabolic network mass balances, and the outer level problem minimizes the number of knock-outs required to achieve coupling of ethanol production to cell growth. The inner problem constraints come from the FBA formulation. Binary variables (y_j) are introduced to represent knock-outs when equal to 1. We set a lower bound on ethanol flux to ensure a reasonable minimal production (Eq. (1)). We consider a set of essential reactions that

cannot be knocked-out as they are related to essential genes (~38% of the total genes, Nogales et al. 2012) which, if eliminated, cause zero growth for the *in silico* case and death of the microorganism for the *in vivo* case. We fix binary variables values to 0 for genes belonging to this set. It is necessary to set a lower bound on biomass production to avoid a possible solution with zero growth rate (Eq.(2)). We also consider a set of fixed reactions (Eq.(6)) that correspond to reactions set to zero for the FBA and to the maintenance reaction fixed at 0.13 mmol gDW^{-1} h^{-1}. Fluxes can take negative values in the case of reversible reactions. The lower bounds (LB) and upper bounds (UB) for all the reactions except biomass and the ones already set by the FBA constraints (Carbon uptake, photon flux) are obtained by solving a Flux Variability Analysis (FVA) for each of the reactions.

The bi-level problem is reformulated into a single level MILP problem using duality theory (Ignizio and Cavalier, 1994). We augment the constraints of the inner and outer problem with the dual constraints of the inner problem and impose the strong duality condition (Eq.(8)). By doing this, bilinear continuous-discrete terms appear in the dual objective function, which are reformulated via the exact linearizations proposed by Glover (1975), giving rise to the constraints in Eqs.(12) and Eqs.(13). All problems are solved with CPLEX in GAMS 24.2.3 with a 2.6 GHz Intel i7 processor and 16 GB RAM.

4. Numerical Results and Discussion

The maximal growth rate obtained by FBA for the case of the carbon limited conditions considered in this work was of 0.073 h^{-1}. At this state of maximal growth rate, there is no ethanol production. To obtain maximal ethanol production for this microorganism we maximized its production also by FBA, rendering 1.85 mmol$gDW^{-1}h^{-1}$ with no biomass production.

By solving the proposed MILP problem with different lower bounds on biomass production rate, we determined five mutants with coupled biomass and ethanol productions. The approximate solution time for the five solutions was of 8 hours. The first MILP, with 0.01 h^{-1}as lower bound for biomass growth rate, suggested a set of six reactions that had to be eliminated. This set contained the enzyme catalase-peroxidase, which catalyses the decomposition of hydrogen peroxide into water and oxygen. Due to physiological reasons, this reaction cannot be eliminated in vivo, as accumulation of hydrogen peroxide is toxic to the cell, so the first mutant is not considered a viable option. A second MILP, with 0.07h^{-1}as lower bound for biomass growth rate, suggested a set of eight reactions to be eliminated. One of the reactions for this mutant is the periplasmatic ATP synthase which is codified by the same genes that the tilacoidal ATP synthase. Thus, it is not experimentally possible to eliminate this reaction without the associated elimination of the ability to produce ATP by photosynthesis. Similar results were obtained by Shabestary and Hudson (2016) while trying to couple growth to the 1-octanol production in *Synechocystis*. Finally, an MILP considering 0.05 h^{-1}as lower bound for biomass growth rate,(Fig.(1))determined a set of five reactions to be eliminated. In this case all of them are viable to be eliminated in vivo. In this *in silico* viable mutant the production of ethanol is achieved by acetaldehyde synthesis from acetate. Acetate is produced through the arginine, glutamate, glutamine and proline biosynthesis pathway. Two knock-outs eliminate the reversible production of acetate to acetyl-CoA (AcCoa) (ACS) and the extracellular elimination of acetate, in both cases in order to redirect this metabolite to the ethanol synthesis. The latter is also allowed by the elimination of the serine hydroxymethyltransferase (SHMT) of the glycine, serine and threonine

metabolism and the ornithine aminotransferase (OAT) of the arginine, glutamate, glutamine metabolism. Elimination of the reaction catalysed by glutamate-5-semialdehyde dehydrogenase cuts off carbon flux to the succinate synthesis. The elimination of this set of reactions causes a decrease on the maximal growth rate, but successfully couples ethanol production. The new maximal biomass production is 0.065h^{-1}, showing a drop of ~14%. Ethanol production obtained for the new maximal growth is 0.046 mmol gDW^{-1} h^{-1}, which compared to the theoretical maximum obtained is considerably low.

The *in silico* results show equal ethanol production through the acetate natural pathway contained in *Synechocystis* and through the heterologous pyruvate decarboxylase pathway, i.e., if the model is forced to use the PDC pathway, the knock-outs and productions obtained are the same (Results not shown). There is no experimental information about the flow capacity of each pathway to incorporate to the model, although there is experimental evidence of the production of acetate in *Synechocystis* (Xiong et al. 2015).

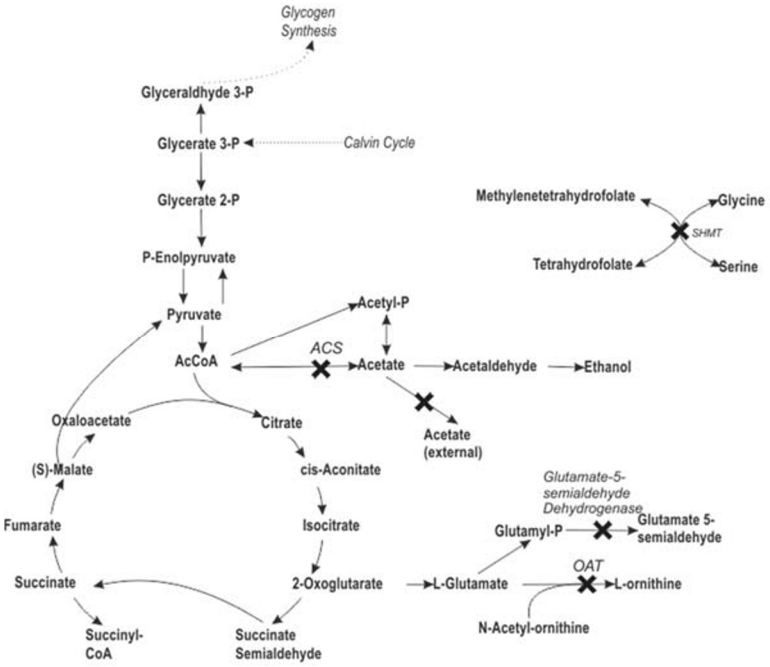

Figure (1): Reaction network showing the knockouts needed for the coupling. Crosses indicate the 5 knock outs proposed by the model.

5. Conclusions

In this work, we propose a strategy to couple cell growth to ethanol production in *silico*, by solving a bilevel programming problem within the context of genomic scale metabolic networks. Among others, the model predicts knock-outs over the acetate production pathway and the results are the same taking into account the natural or the heterologous ethanol production pathway. Experimental evidence of the activities of these two possible pathway is needed. Numerical results show that, although the objective of coupling is achievable, ethanol productivity is very low as compared to the case of other

microorganisms. However; the exploration of new strategies to optimize ethanol production through acetate pathway will lead to improve productivities.

References

Burgard, A.P., Pharkya, P., Maranas, C.D., 2003. OptKnock: A Bilevel Programming Framework for Identifying Gene Knockout Strategies for Microbial Strain Optimization. Biotechnol. Bioeng. 84, 647–657. doi:10.1002/bit.10803

Chowdhury, A., Zomorrodi, A.R., Maranas, C.D., 2015. Bilevel optimization techniques in computational strain design. Comput. Chem. Eng. 72, 363–372. doi:10.1016/j.compchemeng.2014.06.007

Dexter, J., Armshaw, P., Sheahan, C., Pembroke, J.T., 2015. The state of autotrophic ethanol production in Cyanobacteria. J. Appl. Microbiol. 119, 11–24. doi:10.1111/jam.12821

Glover, F., 1975. Improved linear integer programming formulations of nonlinear integer problems. Manag. Sci. 22, 455–460. doi:10.1287/mnsc.22.4.455

Hjersted, J.L., Henson, M.A., 2008. Optimization of Fed-Batch Saccharomyces cerevisiae

Fermentation Using Dynamic Flux Balance Models. Biotechnol. Prog. 22, 1239–1248. doi:10.1021/bp060059v

Hucka, M., Finney, A., Sauro, H.M., Bolouri, H., Doyle, J.C., 2003. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models 19, 524–531. doi:10.1093/bioinformatics/btg015

Ignizio and Cavalier, 1994. Linear Programming. Pearson.

Ihlenfeldt, M.J.A., Gibson, J., 1977. Acetate Uptake by the Unicellular Cyanobacteria Synechococcus and Aphanocapsa. Arch. Microbiol. 113, 231–241.

Knoop, H., Gründel, M., Zilliges, Y., Lehmann, R., Hoffmann, S., Lockau, W., Steuer, R., 2013. Flux Balance Analysis of Cyanobacterial Metabolism: The Metabolic Network of Synechocystis sp. PCC 6803. PLoS Comput. Biol. 9. doi:10.1371/journal.pcbi.1003081

Lasry Testa, R., Delpino, C., Estrada, V., Diaz, M.S., 2016, "Metabolic Network Design for Ethanol Production by Synechocystis sp. PCC 6803", AIChE Annual meeting, 13-18 Nov., San Francisco, United States

Nogales, J., Gudmundsson, S., Knight, E.M., Palsson, B.O., Thiele, I., 2012. Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis. Proc. Natl. Acad. Sci. 109, 2678–2683. doi:10.1073/pnas.1117907109

Ranganathan, S., Suthers, P.F., Maranas, C.D., 2010. OptForce : An Optimization Procedure for Identifying All Genetic Manipulations Leading to Targeted Overproductions. PLoS Comput. Biol. 6, e1000744. doi:10.1371/journal.pcbi.1000744

Savinell, J.M., Palsson, B.O., 1992. Network analysis of intermediary metabolism using linear optimization. I. Development of mathematical formalism. J. Theor. Biol. 154, 421–454. doi:10.1016/S0022-5193(05)80161-4

Schellenberger, J., Que, R., Fleming, R.M.T., Thiele, I., Orth, J.D., Feist, A.M., Zielinski, D.C., Bordbar, A., Lewis, N.E., Rahmanian, S., Kang, J., Hyduke, D.R., Palsson, B.Ø., 2011. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nat. Protoc. 6, 1290–1307. doi:10.1038/nprot.2011.308

Shabestary, K., Hudson, E.P., 2016. Computational metabolic engineering strategies for growth-coupled biofuel production by Synechocystis. Metab. Eng. Commun. 3, 216–226. doi:10.1016/j.meteno.2016.07.003