

# Optimal protein production by a synthetic microbial consortium: Coexistence, distribution of labor, and syntrophy

Carlos Martínez<sup>1,2</sup>, Eugenio Cinquemani<sup>3</sup>, Hidde de Jong<sup>3</sup>, Jean-Luc Gouze<sup>1</sup>

<sup>1</sup>Université Côte d'Azur, Inria, INRAE, CNRS, Sorbonne Université, Biocore Team, Sophia Antipolis, France (carlos.martinez, jean-luc.gouze@inria.fr)

<sup>2</sup>Institute of Hydrobiology, Biology Centre of the Czech Academy of Sciences, Na Sádkách 7, 370 05 České Budějovice, Czech Republic (carlos.martinez@hbu.cas.cz)

<sup>3</sup>Grenoble Alpes, Inria, 38000 Grenoble, France (eugenio.cinquemani, hidde.de-jong@inria.fr)

January 12, 2022

**Abstract** The bacterium *E. coli* is widely used to produce recombinant proteins such as growth hormone and insulin. One inconvenience with *E. coli* cultures is the secretion of acetate through overflow metabolism. Acetate inhibits cell growth and represents a carbon diversion, which results in several negative effects on protein production. One way to overcome this problem is the use of a synthetic consortium of two different *E. coli* strains, one producing recombinant proteins and one reducing the acetate concentration. In this paper, we study a chemostat model of such a synthetic community where both strains are allowed to produce recombinant proteins. We give necessary and sufficient conditions for the existence of a coexistence equilibrium and show that it is unique. Based on this equilibrium, we define a multi-objective optimization problem for the maximization of two important bioprocess performance metrics, process yield and productivity. Solving numerically this problem, we find the best available trade-offs between the metrics. Under optimal operation of the mixed community, both strains must produce the protein of interest, and not only one (distribution instead of division of labor). Moreover, in this regime acetate secretion by one strain is necessary for the survival of the other (syntrophy). The results thus illustrate how complex multi-level dynamics shape the optimal production of recombinant proteins by synthetic microbial consortia.

**Keywords:** Microbial consortia; Pareto optimality; Dynamical systems; Chemostat; Production of recombinant proteins

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Model description</b>	<b>4</b>
<b>3</b>	<b>Coexistence at steady states</b>	<b>9</b>
<b>4</b>	<b>Optimal performance at the coexistence equilibrium</b>	<b>12</b>
4.1	Impact of $G_{in}$ . . . . .	14
4.2	Comparison with the monoculture . . . . .	15
4.3	Cleaners must produce proteins for a better performance . . . . .	15
<b>5</b>	<b>Discussion and conclusions</b>	<b>16</b>
5.1	Characterization of coexistence . . . . .	16
5.2	Optimization: Coexistence, distribution of labor, and syntrophy . . . . .	17
<b>A</b>	<b>Proofs</b>	<b>18</b>
<b>B</b>	<b>Algorithm to find the coexistence equilibrium</b>	<b>27</b>
<b>C</b>	<b>Algorithm to solve the MOP</b>	<b>28</b>
<b>D</b>	<b>Robust choice of the down-regulation function</b>	<b>29</b>

## 1 Introduction

*Escherichia coli* (*E. coli*) is one of the most widely used bacteria for large-scale production of recombinant proteins such as insulin or the human growth hormone (Baeshen et al., 2015; Huang et al., 2012). The preferred carbon source for *E. coli*, as for many other bacteria, is glucose, supporting a faster growth rate compared to other sugars (Görke and Stülke, 2008; Wolfe, 2005). One problem with *E. coli* cultures grown on glucose is that fast growth leads to the secretion of acetate, a phenomenon known as overflow metabolism (Basan et al., 2015; Enjalbert et al., 2017). The resulting accumulation of acetate in the medium inhibits growth and represents a diversion of carbon, thus resulting in several negative effects on protein production (Eiteman and Altman, 2006; Luli and Strohl, 1990).

Different strategies have been proposed to overcome acetate formation. For example, some strategies prevent acetate overflow by deleting genes in acetate metabolism or by forcing cells to take up glucose at a rate below the overflow threshold (De Mey et al., 2007; Eiteman and Altman, 2006). These strategies have several inconveniencies, including suboptimal growth or secretion of other fermentation products. Other strategies aim to remove acetate from the culture medium. Acetate removal can be done, for example, with a dialysis reactor (Fuchs et al., 2002) or with macroporous ion-exchange resins (Huang et al., 2012). A promising alternative, inspired by naturally occurring syntrophic microbial consortia (Rosenzweig et al., 1994), consists in introducing an additional *E. coli* strain which has been metabolically engineered to consume acetate (Bernstein et al., 2012).

Along these lines, in recent theoretical work Mauri et al. (2020) proposed a coarse-grained mathematical model of a syntrophic consortium and investigated under which conditions it could improve the production of recombinant proteins as compared to mono-cultures. The

model accounts for two different *E. coli* strains growing together: one producing the protein of interest (producers), and one reducing the presence of acetate (cleaners). As highlighted by the authors, optimizing the production of recombinant proteins in such a mixed culture may lead to gains in productivity (gram product per hour), but at the expense of diverting substrate away from the producer to the cleaner, thus leading to a lower process yield (gram product per gram substrate). This trade-off between productivity and yield on the consortium level is a generalization of a well-known trade-off on the level of individual species, where high yield of protein production comes with an increased metabolic load that limits growth and therefore productivity (Kurland and Dong, 1996; Wu et al., 2016). Optimizing the performance of microbial consortia is a challenge for current industrial applications (Hays et al., 2015; Jagmann and Philipp, 2014; Pandhal and Noirel, 2014; Roell et al., 2019), because it requires the simultaneous optimization of individual microbial species and their interactions (Shong et al., 2012). Mathematical modeling is of great help in understanding and optimizing such complex multi-level dynamical systems.

In this work, we propose a theoretical study on the coexistence and optimization of a modified version of the above synthetic producer-cleaner consortium in which both *E. coli* strains produce recombinant proteins. We hypothesize that this may alleviate the productivity-yield trade-off in that some of the carbon diverted to the cleaners can be utilized for protein production. We first propose an extension of the mathematical model of Mauri et al. (2020) to account for the production of recombinant proteins by cleaners. Then, we study conditions for a coexistence equilibrium. In general, the long-term coexistence of two populations in a chemostat is not guaranteed. When two populations compete for a growth-limiting substrate, only one population survives (Hardin, 1960). Recent work on syntrophic relationships, in which one species produces an inhibitor (a metabolic by-product) that serves as a nutrient for the other species, have shown that coexistence is possible (Harvey et al., 2014; Heßeler et al., 2006; Sari et al., 2012; Stump and Klausmeier, 2016). The mixed culture studied in this work accounts for competition and syntrophy, so that establishing coexistence is not trivial. We determine necessary and sufficient mathematical conditions for a coexistence equilibrium.

Based on this equilibrium, we define a multi-objective optimization problem (MOP) for maximizing productivity and process yield. The solution of the MOP provides a graphical description, known as Pareto optimal front, of the combinations of productivity and process yield such that one metric cannot be improved without degrading the other (Miettinen, 2012). We show that, when cleaners are allowed to produce proteins as well, the Pareto optimal front is pushed towards higher levels of productivity and process yield. That is, the division of labor between a strain producing recombinant proteins and another strain cleaning up acetate is suboptimal as compared to the distribution of the production task over the two strains. Moreover, in order to attain the required growth rates along the Pareto optimal front, the cleaners need to take up not only glucose but also acetate secreted by the producers. In other words, the syntrophic relationship between the two strains is a necessary property for optimal production.

Our paper is organized as follows. In Section 2, we describe the chemostat model of the synthetic microbial community. In Section 3, we state Theorem 3.6 that gives necessary and sufficient conditions for the existence of a unique coexistence steady state. In Section 4, we define the process yield and productivity associated with each coexistence steady state. Then, we study the solutions of the multi-objective optimization problem (MOP) of simultaneously maximizing the process yield and productivity. The paper includes four Appendices with proofs and details of the algorithms to numerically determine the coexistence

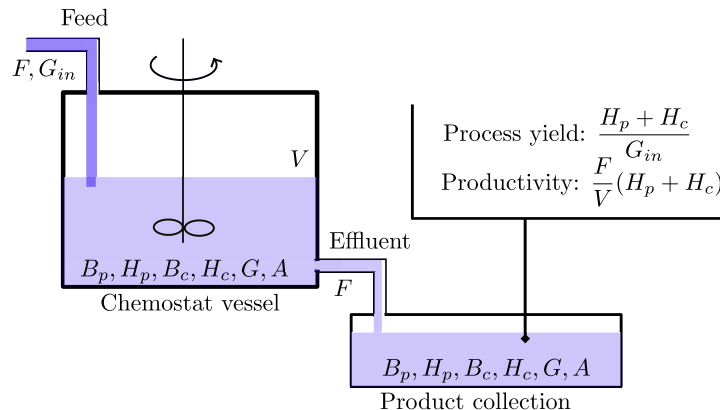


Figure 1: Schematic diagram of a chemostat used for the production of recombinant proteins ( $H_p + H_c$ ). The chemostat is fed at a rate  $F$  with a glucose concentration  $G_{in}$ . The reactor is emptied at the rate  $F$  keeping a constant volume  $V$ . The concentrations of producers ( $B_p$ ) and cleaners ( $B_c$ ), glucose ( $G$ ), and acetate ( $A$ ) are homogeneous in the medium. The dilution rate  $D$  is defined as  $F/V$ . See Table 1 for the units of the variables.

equilibrium and the solutions of the MOP.

## 2 Model description

The description of the model of the *E. coli* community follows that given by Mauri et al. (2020). Consider a chemostat (see Figure 1) in which two different strains of *E. coli* grow. We will refer to these strains as producers (with concentration  $B_p$ ) and cleaners (with concentration  $B_c$ ). Both *E. coli* strains can grow by taking up glucose (with concentration  $G$ ). The glucose uptake rates are denoted  $r_{up,p}^g$  and  $r_{up,c}^g$  for producers and cleaners, respectively. The glucose uptake rate by producers is given by

$$r_{up,p}^g(G, A) = k_g \frac{G}{G + K_g} \frac{\Theta_a}{A + \Theta_a},$$

where  $k_g$  is the maximal uptake rate of glucose,  $K_g$  is a half saturation constant,  $A$  is the acetate concentration, and  $\Theta$  is a constant representing the inhibitory effect of acetate. In cleaners, the gene *ptsG* is deleted. This gene encodes a major subunit of the glucose uptake system (Deutscher et al., 2006), so that its deletion reduces the glucose uptake rate.  $r_{up,c}^g$  is then given by

$$r_{up,c}^g(G, A) = \beta r_{up,p}^g(G, A), \quad (1)$$

where  $\beta \in [0, 1]$  is a dimensionless constant, reflecting the deletion of the gene *ptsG*. If  $\beta = 1$ , the gene *ptsG* is present and both strains take up glucose at the same rate. The units of the variables and rates are summarized in Table 1.

When the glucose uptake rate of *E. coli* is above a threshold rate  $l$ , then cells secrete acetate through overflow metabolism (Basan et al., 2015). The overflow rates are denoted by  $r_{over,p}^a$  and  $r_{over,c}^a$  for producers and cleaners, respectively. They are defined as

$$r_{over,p}^a = k_{over} \max\{0, r_{up,p}^g - l\} \quad \text{and} \quad r_{over,c}^a = k_{over} \max\{0, r_{up,c}^g - l\},$$

Variables	Unit	Description
$B_p$	gDW L <sup>-1</sup>	Biomass concentration of producers
$B_c$	gDW L <sup>-1</sup>	Biomass concentration of cleaners
$G$	g L <sup>-1</sup>	Concentration of glucose
$A$	g L <sup>-1</sup>	Concentration of acetate
$H_p$	gDW L <sup>-1</sup>	Concentration of recombinant proteins in producers
$H_c$	gDW L <sup>-1</sup>	Concentration of recombinant proteins in cleaners
Rates		
$r_{up,p}^g$	h <sup>-1</sup>	Specific glucose uptake rate of producers
$r_{up,c}^g$	h <sup>-1</sup>	Specific glucose uptake rate of cleaners
$r_{up,p}^a$	h <sup>-1</sup>	Specific acetate uptake rate of producers
$r_{up,c}^a$	h <sup>-1</sup>	Specific acetate uptake rate of cleaners
$r_{over,p}^a$	h <sup>-1</sup>	Specific acetate secretion rate of producers
$r_{over,c}^a$	h <sup>-1</sup>	Specific acetate secretion rate of cleaners

Table 1: Variables and rates in the model of producer-cleaner consortium given by the system (8)-(9).

where  $k_{over}$  is a proportionality constant. Note that the maximal glucose uptake rate of cleaners is given by  $\beta k_g$ . Thus, if

$$\beta k_g \leq l \quad (2)$$

then cleaners cannot secrete acetate. This property is satisfied by the parameters in Table 2, in accordance with the fact that cleaners are primarily designed to remove acetate from the medium.

In the presence of glucose, *E. coli* cells cannot grow on acetate, a phenomenon known as Carbon Catabolite Repression (CCR) (Kremling et al., 2015; Wolfe, 2005). The uptake rates of acetate are denoted by  $r_{up,p}^a$  and  $r_{up,c}^a$  for producers and cleaners, respectively. The acetate uptake rate by producers is given by

$$r_{up,p}^a(G, A) = k_a \frac{A}{A + K_a} d(r_{up,p}^g(G, A)), \quad (3)$$

where  $k_a$  is the maximal acetate uptake rate,  $K_a$  is a half-saturation constant, and  $d$  is a down-regulation function accounting for CCR. Before describing the function  $d$ , we define  $r_{up,p}^a$ . The assimilation of acetate by *E. coli* can be increased with a plasmid enabling the inducible overexpression of the native gene *acs*, coding for the enzyme acetyl-CoA synthetase (Lin et al., 2006). Hence, the uptake rate of acetate by cleaners is

$$r_{up,c}^a(G, A) = k_a \frac{A}{A + K_a} d(r_{up,c}^g(G, A)) + k_{Acs} \frac{A}{A + K_{Acs}}, \quad (4)$$

where  $k_{Acs}$  is the maximal acetate uptake rate due to overexpression of the gene *acs* and  $K_{Acs}$  is a half-saturation constant. Note that if cleaners are not genetically modified, then  $\beta = 1$  (see (1)) and  $k_{Acs} = 0$ . In that case, both acetate and glucose uptake rates, are equal for both strains.

The down-regulation function  $d$  represents the inhibitory effect of glucose uptake and decreases as the glucose uptake rate increases. In the work of Mauri et al. (2020), the down-regulation function is defined as

$$\hat{d}(y) = \frac{\Theta_g}{\Theta_g + y}, \quad y = r_{up,p}^g, r_{up,c}^g \quad (5)$$

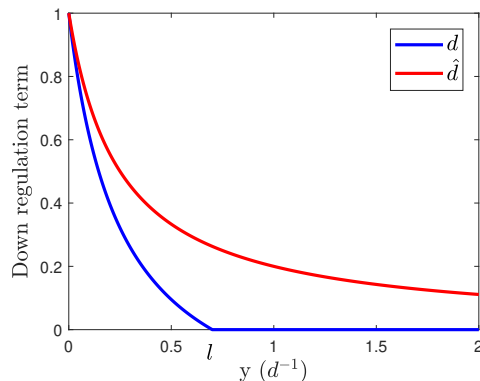


Figure 2: Down-regulation function  $d$  compared to the down-regulation function  $\Theta_g/(\Theta_g + y)$ . Parameters are taken from Table 2.

with  $\Theta_g$  a constant. We modify the description of the down-regulation function (5) as follows:

$$d(y) = \max \left\{ 0, (1 + a) \frac{\Theta_g}{\Theta_g + y} - a \right\}, \quad y = r_{up,p}^g, r_{up,c}^g \quad (6)$$

with  $a = \Theta_g/l$ . The main motivation to describe  $d$  as in (6) arises from the following property:

$$\text{if } y \geq l, \text{ then } d(y) = 0.^1 \quad (7)$$

This property states that if overflow metabolism occurs (*i.e.*, the glucose uptake rate is higher than  $l$ ) then there is no acetate uptake, which is qualitatively consistent with CCR (Wolfe, 2005). An important advantage of (6) is that it facilitates the mathematical analysis of the model (see below). Figure 2 compares the two alternatives of  $d$ . In Appendix D, we show that our optimization results are insensitive to this model change.

Both strains carry a plasmid for the expression of a recombinant protein. The concentration of proteins produced by producers and cleaners are denoted by  $H_p$  and  $H_c$ , respectively. The fraction of biomass production assigned to the synthesis of recombinant protein is determined by the (dimensionless) product yield constants  $Y_{h/p}$  and  $Y_{h/c}$ . In the model of Mauri et al. (2020), cleaners cannot produce heterologous proteins and they mainly serve to reduce the acetate concentration in the medium. In this new model, the carbon lost in the form of acetate can be recovered by cleaners and transformed into recombinant protein.

The specific rate of biomass production per unit of biomass, including  $B_p$  and  $H_p$ , is given by

$$f_p(G, A) = Y_g r_{up,p}^g(G, A) - Y_a r_{over,p}^a(G, A) + Y_a r_{up,p}^a(G, A),$$

with  $Y_a$  and  $Y_g$  yield coefficients. Note that  $Y_g r_{up,p}^g$  and  $Y_a r_{up,p}^a$  are gains of carbon due to glucose and acetate uptake, respectively, and  $Y_a r_{over,p}^a$  is a loss of carbon due to overflow metabolism. Similarly, the specific rate of cleaner biomass production per unit of biomass is given by

$$f_c(G, A) = Y_g r_{up,c}^g(G, A) - Y_a r_{over,c}^a(G, A) + Y_a r_{up,c}^a(G, A).$$

<sup>1</sup>This can be seen by noting that  $d$  is decreasing and that  $d(l) = 0$ .

The evolution of  $B_p, B_c, G$ , and  $A$  is now defined by

$$\begin{aligned}
 \frac{dB_p}{dt} &= (1 - Y_{h/p})f_p(\cdot)B_p - k_{deg}B_p - DB_p, \\
 \frac{dB_c}{dt} &= (1 - Y_{h/c})f_c(\cdot)B_c - k_{deg}B_c - DB_c, \\
 \frac{dG}{dt} &= D(G_{in} - G) - r_{up,p}^g(\cdot)B_p - r_{up,c}^g(\cdot)B_c, \\
 \frac{dA}{dt} &= -DA + [r_{over,p}^a(\cdot) - r_{up,p}^a(\cdot)]B_p + [r_{over,c}^a(\cdot) - r_{up,c}^a(\cdot)]B_c,
 \end{aligned} \tag{8}$$

and the evolution of the recombinant protein concentrations is given by

$$\begin{aligned}
 \frac{dH_p}{dt} &= Y_{h/p}f_p(\cdot)B_p - k_{deg}H_p - DH_p, \\
 \frac{dH_c}{dt} &= Y_{h/c}f_c(\cdot)B_c - k_{deg}H_c - DH_c.
 \end{aligned} \tag{9}$$

System (8)-(9) can be seen as an extension of the model of Mauri et al. (2020). Indeed, if  $Y_{h/c} = 0$  and  $d$  is taken as (5), then (8)-(9) is equivalent to the model of Mauri et al. (2020). Note that the model has been intentionally split into two subsystems, (8) and (9), such that (8) is uncoupled from (9). To establish the existence of coexistence steady states, we only need to study system (8), but for optimizing the production of recombinant proteins, we must study both.

Figure 3 shows the steady-state solution of (8) obtained by simulation from three different kinds of initial conditions: the absence of cleaners, the absence of producers, the presence of both. These simulations suggest that (8) admits at most three non-trivial equilibria:

- (I) An equilibrium with the presence of producers but absence of cleaners (see Figure 3A). In this equilibrium the acetate concentration in the bioreactor may be zero or positive.
- (II) An equilibrium with the presence of cleaners but absence of producers. Since cleaners do not produce acetate (because (2) holds), the acetate concentration is zero.
- (III) A coexistence equilibrium (see Figure 3C and 3D).

Note that, at equilibrium, the growth rate of any species present in the bioreactor (that is,  $(1 - Y_{h/p})f_p - k_{deg}$  for the producer, and  $(1 - Y_{h/c})f_c - k_{deg}$  for the cleaner) must be equal to  $D$ . We are especially interested in coexistence equilibria. We observe the existence of an interval of values of  $D$  (shaded area), which depends on the values of  $Y_{h/p}$  and  $Y_{h/c}$ , such that coexistence is possible. In the following section we determine necessary and sufficient conditions for the existence of a coexistence equilibrium.

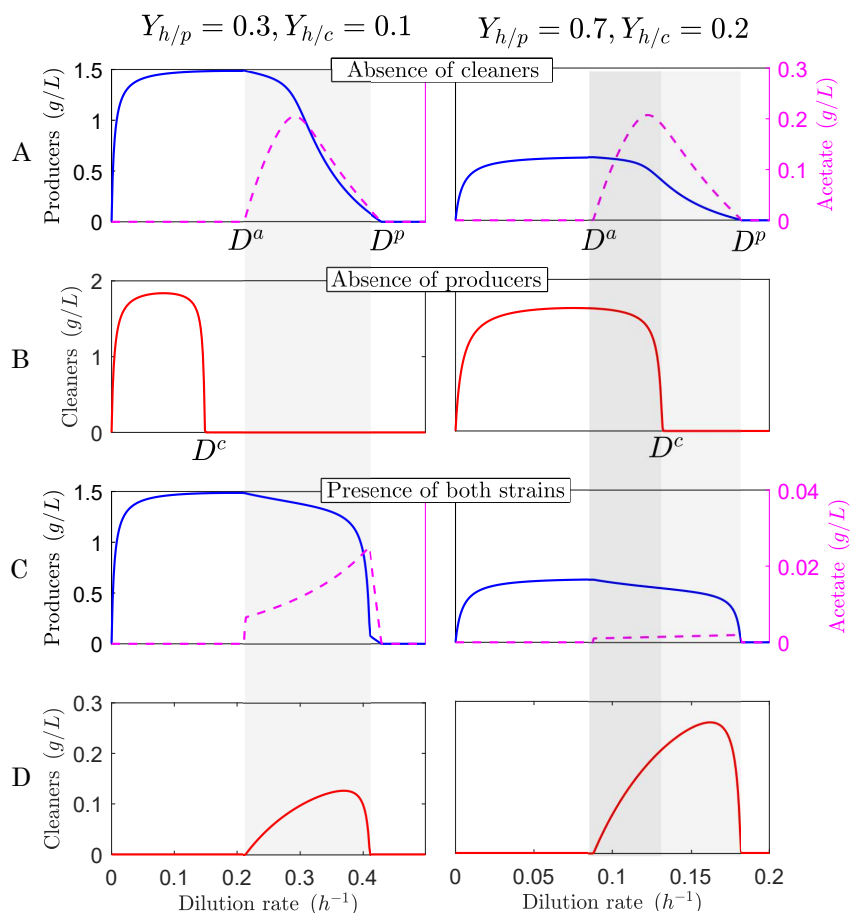


Figure 3: Steady-state solutions of (8) obtained by simulation for different values of the dilution rate and two different pairs of values for  $Y_{h/p}$  and  $Y_{h/c}$ .  $D^a$  is the dilution rate above which acetate overflow occurs when producers grow individually.  $D^p$  and  $D^c$  are the dilution rates at which producers and cleaners go extinct when growing individually. The dilution rates  $D^a$ ,  $D^p$ , and  $D^c$  are formally defined in Propositions 3.1 and 3.4. The shaded area represents the values of the dilution rate at which coexistence is possible. The darker shaded area corresponds to values of  $D$  lower than  $D^c$  and the lighter shaded area corresponds to values of  $D$  lower than  $D^p$  and higher than or equal to  $D^c$ . **A.**  $B_c(0) = 0, B_p(0) > 0$  (absence of cleaners). **B.**  $B_c(0) > 0, B_p(0) = 0$  (absence of producers). **C and D.**  $B_c(0), B_p(0) > 0$  (presence of producers and cleaners).



Parameter	Value	Unit	Description
$k_g$	1.53	$\text{h}^{-1}$	Maximal glucose uptake rate
$K_g$	0.09	$\text{g L}^{-1}$	Glucose half-saturation constant
$\Theta_a$	0.52	$\text{g/L}$	Acetate inhibition constant
$k_a$	0.97	$\text{h}^{-1}$	Maximal acetate uptake rate
$K_a$	0.5	$\text{g L}^{-1}$	Acetate half-saturation constant
$\Theta_g$	0.25	$\text{g L}^{-1}$	Glucose uptake inhibition constant
$k_{over}$	0.17	–	
$l$	0.7	$\text{h}^{-1}$	Overflow metabolism threshold rate
$Y_g$	0.44	$\text{gDW g}^{-1}$	Yield coefficient
$Y_a$	0.3	$\text{gDW g}^{-1}$	Yield coefficient
$\beta$	0.26	–	
$k_{Acs}$	1.46	$\text{h}^{-1}$	
$K_{Acs}$	0.012	$\text{g L}^{-1}$	
$Y_{h/p}$	0 - 1		
$Y_{h/c}$	0 - 1		
$k_{deg}$	0.0044	$\text{h}^{-1}$	Degradation rate of producers
$D$	-	$\text{h}^{-1}$	Dilution rate
$G_{in}$	-	$\text{g L}^{-1}$	Input glucose concentration

Table 2: Parameters and their values taken from the work of Mauri et al. (2020).

### 3 Coexistence at steady states

In this section, we characterize the existence of coexistence equilibria of (8). We first present some results on the existence of equilibria with the presence of only one bacterial population. Then, based on these equilibria, we state the main result on coexistence. To state our results, some assumptions are necessary. The first assumption is

$$Y_g - k_{over}Y_a > 0, \quad (10)$$

which is verified by the parameters in Table 2. The assumption implies that the function  $g(G, A) := Y_g r_{up,p}^g(G, A) - Y_a r_{over,p}^a(G, A)$  is strictly increasing with respect to  $G$  and strictly decreasing with respect to  $A$  (Martínez and Gouzé, 2021). This monotonicity of  $g$  is necessary to prove the uniqueness of the equilibrium with producers and no cleaners (Proposition 3.1). We also assume that

$$r_{up,p}^g(G_{in}, 0) > l, \quad (11)$$

in agreement with the parameter values listed in Table 2. During long-term operation of the bioreactor in presence of bacteria, the glucose concentration in the medium cannot be higher than  $G_{in}$ . Then, if (11) does not hold (*i.e.*,  $r_{up,p}^g(G_{in}, 0) \leq l$ ), overflow metabolism is impossible in the long-term, and the dynamics of (8) is reduced to that of a classical chemostat model with two competitors for one substrate (glucose) (Smith and Waltman, 1995). Note that assumptions (10) and (11) have been used by Martínez and Gouzé (2021) to study a simplified version of (8) without cleaners.

The following proposition characterizes the existence of steady states without cleaners. The proof is adapted from Proposition 1 by Martínez and Gouzé (2021), where a similar result is presented, but in the absence of acetate consumption ( $r_{up,p}^a = 0$ ).

**Proposition 3.1** (Producer steady state). *Assume that (10) and (11) hold. Let us define  $D^p := (1 - Y_{h/p})f_p(G_{in}, 0) - k_{deg}$ . We have:*

(a) *If  $D^p > D$ , then (8) admits a unique equilibrium with the presence of producers and absence of cleaners, denoted by  $E^p = (B_p^p, 0, G^p, A^p)$ . Moreover, for  $D^a = (1 - Y_{h/p})Y_g l - k_{deg}$  we have:*

- (I) *If  $D^a > D$ , then  $A^p = 0$  and  $r_{up,p}^g(G^p, 0) < l$ .*
- (II) *If  $D^a = D$ , then  $A^p = 0$  and  $r_{up,p}^g(G^p, 0) = l$ .*
- (III) *If  $D^a < D$ , then  $A^p > 0$  and  $r_{up,p}^g(G^p, A^p) > l$ .*

(b) *If  $D^p \leq D$ , then (8) has no equilibrium with the presence of producers and absence of cleaners.*

*Proof.* See Appendix A. □

*Remark 3.2* (Acetate secretion). Proposition 3.1 shows the existence of a threshold dilution rate characterizing the presence of acetate in the medium, denoted by  $D^a$  (see Figure 3A). As we will see later, a coexistence equilibrium for (8) is only possible if the dilution rate is higher than  $D^a$ . Proposition 3.1 also introduces  $D^p$ , the threshold dilution rate above which producers grow extinct. No coexistence equilibrium is possible for dilution rates higher than  $D^p$ .

*Remark 3.3* (Acetate consumption). Note that there is no acetate consumption by the bacteria at equilibrium (the acetate secreted, if any, is diluted out at the same rate). Indeed, if  $D \leq D^a$ , there is no acetate, so there is trivially no acetate consumption. Conversely, if  $D > D^a$ , there is overflow metabolism. Then, using (7), we have that there is no acetate consumption. The term  $r_{up,p}^a$  has therefore no influence on the steady state of producers. Thus, the non-trivial equilibrium determined by Martínez and Gouzé (2021) is exactly the same as the non-trivial equilibrium given by Proposition 3.1, even if the authors do not consider acetate consumption.

To study the existence of equilibria with cleaners and no producers, we assume that (2) holds. As discussed in Section 2, this implies that

$$r_{over,c}^a(G, A) = 0, \quad \text{for all } G, A \geq 0.$$

This means that cleaners cannot secrete acetate. This assumption is qualitatively consistent with the fact that cleaners are designed to mainly grow on acetate and is consistent with the parameters from Table 2.

**Proposition 3.4** (Cleaner steady state). *Assume that (2) holds and let us define  $D^c = (1 - Y_{h/c})f_c(G_{in}, 0) - k_{deg}$ . We have:*

(a) *If  $D^c > D$ , then (8) admits a unique equilibrium with the presence of cleaners and absence of producers, denoted by  $E^c = (0, B_c^c, G^c, 0)$ .*

(b) *If  $D^c \leq D$ , then (8) has no equilibrium with the presence of cleaners.*

*Proof.* See Appendix A. □

*Remark 3.5* (Cleaner washout). Proposition 3.4 introduces the threshold dilution rate  $D^c$  above which cleaners are washed out when grown without producers. As shown below, a coexistence equilibrium is possible at dilution rates higher than  $D^c$ .

The main result of this section characterizes the existence of coexistence steady states for (8). This characterization is based on the positive steady states of the cleaners and producers growing separately.

**Theorem 3.6** (Coexistence steady states). *Assume that (2), (10), and (11) hold. Let  $D^p$  and  $D^a$  be given by Proposition 3.1, and let  $D^c$  be given by Proposition 3.4.*

- (a) *If  $D \leq D^a$  or  $D \geq D^p$ , then there is no coexistence equilibrium.*
- (b) *If  $D^a < D < D^p$  and  $D < D^c$  then a (unique) coexistence equilibrium is possible, if and only if*

$$(1 - Y_{h/p})f_p(G^c, 0) - k_{deg} > D \text{ and } (1 - Y_{h/c})f_c(G^p, A^p) - k_{deg} > D,$$

*with  $G^p$  and  $A^p$  defined in Proposition 3.1, and  $G^c$  defined in Proposition 3.4.*

- (c) *If  $D^a < D < D^p$  and  $D \geq D^c$  then a (unique) coexistence equilibrium is possible, if and only if*

$$(1 - Y_{h/c})f_c(G^p, A^p) - k_{deg} > D,$$

*with  $G^p$  and  $A^p$  defined in Proposition 3.1.*

*Proof.* See Appendix A. □

**Remark 3.7** (Competition and syntrophy). Note that the difference between cases (b) and (c) in Theorem 3.6 is the value of the dilution rate with respect to  $D^c$ . If  $D < D^c$  (case (b)), according to Proposition 3.4, in the absence of producers, cleaners survive on glucose only and settle in an equilibrium. The presence of producers can be beneficial by providing acetate, which is known as protocoooperation (Roell et al., 2019), or harmful through glucose competition. This can be observed in Figures 3B and 3D (right-hand side). When  $D$  is higher than but close to  $D^a$ , cleaners reach a higher density when grown alone than when growing together with producers (competition). However, when  $D$  is lower than but close to  $D^c$ , cleaners reach a lower density when grown alone than when growing together with producers (protocoooperation). Now, if  $D \geq D^c$  (case (c)), cleaners go extinct when growing alone. Therefore, if coexistence is possible, this is due to the syntrophic relationship in which the producers provide acetate to cleaners. As we will see later, syntrophy seems to play an important role in the performance of the system.

The following corollary gives a characterization of the existence of coexistence equilibria based on dynamical properties of the non-coexistence equilibria.

**Corollary 3.8.** *Assume that (10), (11), and (2) hold. Let  $E^p$  and  $E^c$  be the equilibria given by Propositions 3.1 and 3.4, respectively, whenever they exist. Let  $D^a$  be given by Proposition 3.1 and assume that  $D > D^a$ .*

- (a) *If  $E^p$  does not exist, then there is no coexistence equilibrium.*
- (b) *If  $E^p$  exists and  $E^c$  does not exist, then there is a unique coexistence equilibrium if and only if  $E^p$  is hyperbolic and unstable.*
- (c) *If  $E^p$  and  $E^c$  exist, then there is a unique coexistence if and only if  $E^p$  and  $E^c$  are unstable and hyperbolic.*

*Proof.* The statement follows from applying Lemmas A.7 and A.8 in Appendix A. □

*Remark 3.9* (Dynamics on the boundary). According to Lemmas A.7 and A.8 in Appendix A, if  $E^p$  and  $E^c$  exist, the Jacobian matrices associated with them each have three negative eigenvalues. Therefore, the conditions in Corollary 3.8 state that the fourth eigenvalue must be negative. In such conditions, from a stable manifold argument,  $E^p$  and  $E^c$  can only be reached by solutions starting on  $\mathbb{R}_+^4/X$ , with  $X$  the set characterized by  $B_c > 0$  and  $B_p > 0$  (i.e.,  $X = (0, \infty)^2 \times \mathbb{R}_+^2$ ). Thus,  $E^p$  and  $E^c$ , whenever they exist, must repel solutions starting in  $X$  to ensure the existence of a coexistence equilibrium. This characterization can be compared with some classical results from the theory of persistence (Smith and Thieme, 2011), where the existence of a coexistence equilibrium can be determined from the dynamics on the boundary  $\mathbb{R}_+^4/X$  (i.e., the set where  $B_c = 0$  or  $B_p = 0$ ).

We end this section with a lemma stating how to determine the protein concentrations associated with the coexistence equilibrium given by Theorem 3.6.

**Lemma 3.10.** *Let  $(B_p^*, B_c^*, G^*, A^*)$  be a coexistence equilibrium of (8). Then, the protein concentrations associated with this equilibrium are given by*

$$H_p^* = \frac{Y_{h/p}}{1 - Y_{h/p}} B_p^* \quad \text{and} \quad H_c^* = \frac{Y_{h/c}}{1 - Y_{h/c}} B_c^*.$$

*Proof.* See Appendix A. □

We conclude the section by observing that, for a fixed set of parameter values (e.g., those in Table 2), the stability of the above equilibria can be verified by the same numerical/algebraic methods utilized in Mauri et al. (2020).

## 4 Optimal performance at the coexistence equilibrium

The necessary and sufficient conditions given by Theorem 3.6 for the existence of a coexistence equilibrium allow for a proper study of the optimization of (8)-(9) at steady state. While the statement of Theorem 3.6 ensures the existence of a non-empty region where coexistence is possible, its proof indicates how to construct an efficient algorithm for determining numerically the coexistence equilibria (see Appendix B). In this section, we define a multiobjective optimization problem (MOP) aiming to maximize the process yield (g product per g substrate) and protein volumetric productivity (g per L per h). The process yield is especially relevant in a biotechnological context when the selling price of the product is low, since the cost of glucose becomes a significant fraction of the value of the product. Volumetric productivity determines the rate at which the product can be formed, and thus dictates the overall volume needed for a given plant output (Van Dien, 2013).

The decision variables in the optimization problem are chosen to be the dilution rate ( $D$ ) and the product yield of each strain ( $Y_{h/p}$  and  $Y_{h/c}$ ). The dilution rate is well-known to be an important operational variable that is optimal at intermediate values (Doran, 1995). Similarly, the choice of product yields is not trivial. Low values of  $Y_{h/p}$  naturally lead to low protein production, while high values lead to more recombinant proteins but poor cell growth resulting again in low productivity. Another operational parameter is the input glucose concentration  $G_{in}$ . However, as we will show later, performance always improves with an increase of this parameter, making its choice trivial (see also the work of Mauri et al. (2020)).

In mathematical optimization, the feasible region corresponds to the set of all possible combinations of the decision variables. Since we are interested in coexistence equilibria,

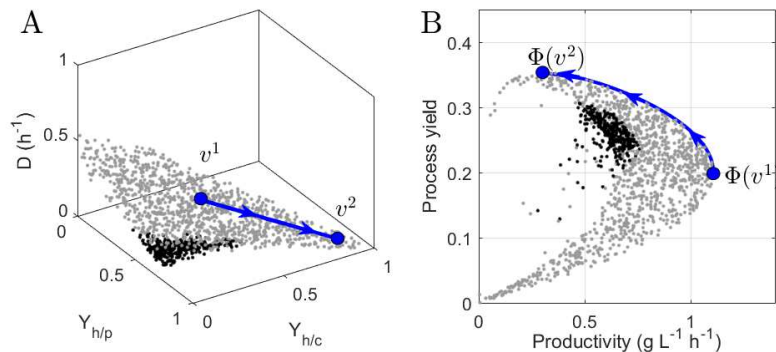


Figure 4: **A.** Region  $\Omega$  represented by 3000 points. Gray points correspond to  $D \geq D^c$  and black points correspond to  $D < D^c$  (see Remark 3.7). The point  $v^1 = (0.51, 0.34, 0.28)$  maximizes productivity ( $\Phi_1$ ) and  $v^2 = (0.91, 0.88, 0.04)$  maximizes the process yield ( $\Phi_2$ ). The blue curve is such that its image through  $\Phi$  is the POF. **B.** Function  $\Phi$  evaluated at the points represented in A. The blue curve is the POF.

we consider the feasible region  $\Omega \subset \mathbb{R}_+^3$ , such that for any  $v := (Y_{h/p}, Y_{h/c}, D) \in \Omega$ , (8) admits a coexistence equilibrium. The region  $\Omega$  can be determined by the conditions given by Theorem 3.6 and it looks like a cone with a vertex close to  $(1, 1, 0)$  (see Figure 4). For any  $v \in \Omega$ , we will denote by  $H_p^*(v)$  and  $H_c^*(v)$  the protein concentrations associated with producers and cleaners, respectively, at the coexistence steady state (see Lemma 3.10). We define the objective function  $\Phi : \Omega \rightarrow \mathbb{R}_+^2$  by

$$v = (Y_{h/p}, Y_{h/c}, D) \xrightarrow{\Phi} \left( DH^*(v), \frac{H^*(v)}{G_{in}} \right),$$

where  $H^*(v) = H_p^*(v) + H_c^*(v)$ . The quantities  $DH^*(v)$  and  $H^*(v)/G_{in}$  are the steady-state productivity and process yield, respectively. We want to solve the following MOP:

$$\begin{aligned} & \text{maximize} && \Phi(v) = \left( DH^*(v), \frac{H^*(v)}{G_{in}} \right) \\ & \text{subject to} && v \in \Omega. \end{aligned} \tag{12}$$

We look for Pareto optimal solutions, that is, solutions that cannot be improved in any of the objectives (process yield or productivity) without degrading the other objective. Generally, there is no single Pareto optimal solution optimizing both objectives, but a set of Pareto optimal solutions called the Pareto optimal front (POF). The structure of the sets  $\Omega$  and  $\Phi(\Omega)$  plays an important role in the success of the employed numerical method to solve (12). Pareto curves cannot be computed efficiently in many cases, especially in the non-convex case where methods such as  $\epsilon$ -constraint are necessary (Gunantara, 2018). Figure 4 reveals a favorable structure of  $\Omega$  and  $\Phi(\Omega)$  (scatter plots) for applying the weighting method (Miettinen, 2012). Accordingly, the problem (12) is transformed into the following so-called weighting problem:

$$\begin{aligned} & \text{maximize} && \lambda DH^*(v) + (1 - \lambda) \frac{H^*(v)}{G_{in}}, \\ & \text{subject to} && v \in \Omega, \end{aligned} \tag{13}$$

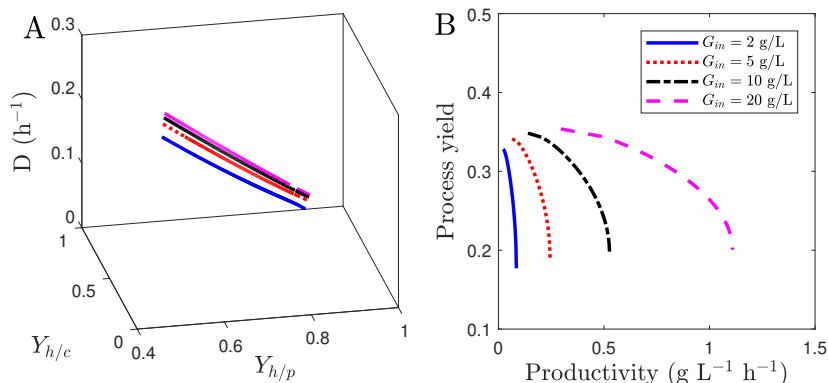


Figure 5: Influence of  $G_{in}$  on the POF. **A.** Decision variables which image through  $\Phi$  is a POF. **B.** POF for different values of  $G_{in}$ .

where  $\lambda \in [0, 1]$ . It is well known that if  $\Omega$  and  $\Phi(\Omega)$  are convex, then  $v^*$  is a Pareto optimal solution if there is  $\lambda \in [0, 1]$  such that  $v^*$  is a solution to the weighting problem (13) (see Corollary 3.1.8 by Miettinen (2012)).

For any  $\lambda \in [0, 1]$ , problem (13) is solved numerically with the interior point algorithm implemented in the toolbox *fmincon* of MATLAB (Byrd et al., 1999). Before using *fmincon*, some technical details must be addressed. For example, the feasible region must be a closed set and the objective function must be continuous on this set. For ease of reading, we discuss such technical details in Appendix C.

Figure 4 shows the POF obtained using the weighting method and the subset of  $\Omega$  whose image is the POF. From this figure we observe that:

- The weighting method accurately returns the POF.
- The set of all points  $v \in \Omega$  such that  $\Phi(v)$  belongs to the POF can be approximated by the line joining the points  $v^1$  and  $v^2$ , corresponding to maximum productivity and maximum yield, respectively.
- Along the POF, the process yield decreases as the productivity increases.

Most importantly, regarding the syntrophy-competition relationship between both strains, the image  $\Phi$  of all points  $(Y_{h/p}, Y_{h/c}, D) \in \Omega$  such that  $D < D^c$  is a dominated solution (black points). In other words, if cleaners can survive when growing individually, then the system is operated under suboptimal conditions (see Remark 3.7).

#### 4.1 Impact of $G_{in}$

Figure 5 shows the effects of varying the input glucose concentration  $G_{in}$ . According to Figure 5B, as  $G_{in}$  increases, the POF expands in such a way that the POF dominates the POFs associated with lower values of  $G_{in}$ . This shows why  $G_{in}$  is not a relevant decision variable for the MOP. The input glucose concentration must always be chosen as high as possible. We note that the maximal process yield slightly increases ( $\Phi_2$ ), while the maximal productivity increases almost linearly with  $G_{in}$  ( $\Phi_1$ ). Figure 5A shows that the values of the decision variables associated with each POF are not much affected by the value of  $G_{in}$ .

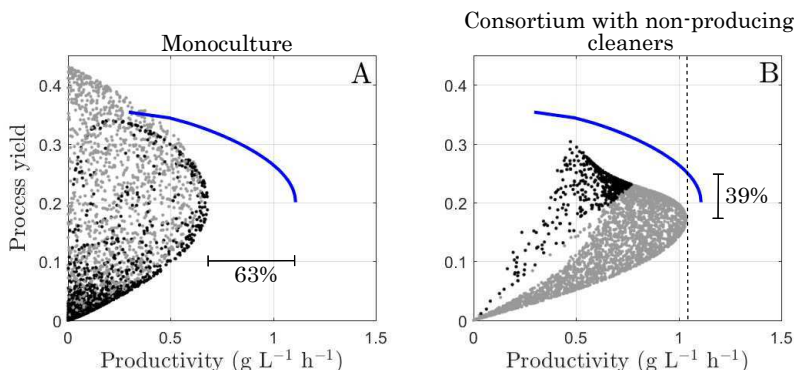


Figure 6: Scatter plots showing the performance of two alternative production methods, compared with Pareto-optimal performance of the consortium with producers and cleaners both synthesizing the target protein. **(A)** Monoculture of producers: scatter plot of  $(DH_p^p, H_p^p/G_{in})$ , with  $H_p^p$  the protein concentration associated with the non-trivial equilibrium given by Proposition 3.1. Black points correspond to  $D > D^a$  and gray points correspond to  $D \leq D^a$ . The continuous line is the POF given by Figure 4. **(B)** Consortium of producers and non-producing cleaners: scatter plot of  $\Phi$  with the protein yield  $Y_{h/c}$  fixed to 0. The scatter plot corresponds to  $\Phi(p)$  with  $v \in \Omega \cap \{v; Y_{h/c} = 0\}$  and, as in panel A, the continuous line is the POF from Figure 4. Black points correspond to  $D < D^c$  and gray points correspond to  $D \geq D^c$ .

## 4.2 Comparison with the monoculture

The potential of a consortium over a monoculture, in terms of productivity, has been demonstrated experimentally by Bernstein et al. (2012) and theoretically by Mauri et al. (2020), but limited to the case that only producers synthesize the target protein. Figure 6A confirms those findings in the more general case of both species contributing to the synthesis process, showing that productivity can increase 63% in case of a consortium of producers and cleaners. However, it also shows that the monoculture can reach a higher process yield, although of course at the expense of low productivity.

## 4.3 Cleaners must produce proteins for a better performance

In contrast with the work of Mauri et al. (2020), where cleaners cannot produce proteins, model (8)-(9) accounts for the production of recombinant proteins by cleaners. The question is whether there is a significant difference in terms of productivity when cleaners can produce proteins. To answer this question, we compare the POF obtained in Figure 4 with a scatter plot of the feasible objective region  $\Phi(\Omega)$  restricted to  $Y_{h/c} = 0$  (see Figure 6B). In other words, the scatter plot shows the different process yields and productivities that can be reached when cleaners are only dedicated to remove acetate. The first observation is that allowing cleaners to produce proteins increases the process yield by 39% for a productivity of 1.04 g L<sup>-1</sup> h<sup>-1</sup>, which is the maximal productivity for  $Y_{h/c} = 0$ . This increase is mainly explained by the fact that, when cleaners produce proteins, the problem of carbon diversion is addressed, that is, acetate secreted by producers is not wasted but transformed into proteins by cleaners.

## 5 Discussion and conclusions

### 5.1 Characterization of coexistence

We have established necessary and sufficient conditions for coexistence in a microbial consortium to synthesize recombinant protein. This consortium comprises two *E. coli* strains, the producer and the cleaner. Producers grow primarily on glucose while cleaners grow primarily on acetate that is secreted by producers as a fermentation by-product.

Conditions for coexistence are determined from analyzing the equilibria that each population can reach when growing separately. An important necessary condition for coexistence is that the producer secretes acetate at equilibrium, otherwise glucose is the only limiting resource and coexistence is impossible because of the competitive exclusion principle (Hardin, 1960). Assuming that this necessary condition holds, one simple description of our main result is obtained from the dynamical properties of the non-trivial equilibria on the boundary (*i.e.*, equilibria in which only one strain is present). If each non-trivial equilibrium on the boundary repels solutions such that both strains grow, then coexistence is possible (see Corollary 3.8).

Another convenient description of our main result follows from an invasion analysis (Chesson, 2000). In our case, an invasion analysis amounts to choosing the *E. coli* strains one at time as the invader. The invader density is set to zero and the other strain, the so-called resident, is allowed to reach equilibrium. If the invader is then allowed to grow and attains a positive growth rate, we say that the invader has succeeded. If both strains can succeed as an invader, then they are said to coexist. The conditions for coexistence in our main result, Theorem 3.6, ensure the success of each strain as invaders. For example, in case (b) of the theorem, each strain reaches a non-trivial equilibrium when playing the role of resident, corresponding to the satisfaction of the two inequalities defining the necessary and sufficient conditions for this case.

The predominant inter-species interaction during coexistence can be competition, proto-cooperation, or syntrophy (Roell et al., 2019). From Theorem 3.6, we distinguish two cases in which coexistence is possible. In case (b), cleaners can survive growing separately, that is, producers are not necessary for the survival of cleaners. In this case, when both strains are grown together, producers may have either a positive or negative effect on cleaners. If producers secrete acetate at a low rate, then both populations strongly compete for glucose, which results in low growth of cleaners. Note that coexistence is possible because of acetate secretion, but the culture is dominated by competition. When the culture is operated at conditions close to washout of the monoculture of cleaner, then producers enhance cleaner growth by supplying acetate. This is known as proto-cooperation (Roell et al., 2019). In case (c) of Theorem 3.6, cleaners cannot survive growing individually. Coexistence is possible due to a syntrophic relationship in which the survival of cleaners critically depends on the secretion of acetate by producers.

The range of conditions allowing coexistence is affected by the metabolic burden of recombinant protein production (Kurland and Dong, 1996; Wu et al., 2016), which is in our case modulated by the values of the production yield parameters  $Y_{h/p}$  and  $Y_{h/c}$ . A necessary condition for coexistence is that the monoculture of producers secretes acetate at equilibrium. Since the growth rate at equilibrium equals the dilution rate, and the secretion of acetate is directly related to the growth rate, this necessary condition can be put in terms of the dilution rate. Indeed, as shown in Proposition 3.1, the dilution rate must be higher than a threshold rate denoted by  $D^a$  and lower than a critical rate denoted  $D^p$ . It can be



proven from the definitions of  $D^p$  and  $D^a$  in Proposition 3.1 that

$$D^p - D^a = (1 - Y_{h/p})(Y_g - k_{over}Y_a)[r_{up,p}^g(G_{in}, 0) - l].$$

This shows how increasing  $Y_{h/p}$  decreases the range of dilution rates for which coexistence is possible.

## 5.2 Optimization: Coexistence, distribution of labor, and syntrophy

Using our results on the coexistence of the consortium, we numerically solved the multiobjective optimization problem (MOP) of maximizing both the process yield and productivity at the coexistence equilibrium. The performance of the consortium depends on different operational and strain design parameters. We chose as decision variables the dilution rate, a fundamental parameter of bioreactor operation, and product yields of the two strains, key parameters in the design and bioengineering of the synthetic community.

A first important question is whether the coexistence equilibrium is advantageous from a biotechnological point of view. As shown in Figure 6A, the *E. coli* consortium can reach a higher productivity than the monoculture of producers. However, the monoculture can reach a higher process yield. This is because of the inherent inefficiency associated with the loss of carbon through the intermediate resource (acetate). Indeed, the highest process yields for the monoculture are obtained in the absence of overflow metabolism (gray points in Figure 6A). One important observation is that all the elements of the Pareto optimal front are obtained when coexistence requires syntrophy, that is, when cleaners need the producers to survive. While this property holds in the case of the parameter values for the *E. coli* consortium (Mauri et al., 2020), we were not able to prove if it is a structural property of the system.

Division of labor refers to the execution of different tasks by different species in a consortium that are specialized for their respective tasks (Roell et al., 2019) and is often proposed as an effective design strategy in synthetic biology (Tsoi et al., 2018). Accordingly, Mauri et al. (2020) assumed that the production of proteins is the task of only one species, the producers. Another example is provided by Liu et al. (2018) who consider a syntrophic *E. coli* co-culture where only one strain synthesizes salidroside, the product of interest. Our results suggest that sharing the synthesis of the product of interest among the different species or strains enables higher yield and productivity. If only producers synthesize proteins, then the carbon lost through overflow metabolism is not eventually transformed into proteins. Alternatively, if only cleaners synthesize proteins, then carbon is inherently lost to producers through glucose competition. Distributing the task of producing proteins over the two strains reduces the loss of resources at the expense of increasing the metabolic burden of both populations. As a consequence, the trade-off between yield and productivity must be optimized simultaneously for both strains, leading to a global multi-level optimization problem. The example illustrates that, while division of labor may have advantages in terms of modularity and conceptual simplicity, it is not always optimal.

## Acknowledgements

This work was supported by the INRIA (IPL CoSy) and by the ANR projects Maximic (ANR-11-LABX-0028-01) and Ctrl-AB (ANR-20-CE45-0014). Additional support for

Carlos Martínez was provided by the European Union within ESIF in the framework of the Operational Programme "Research, Development and Education" (CZ.02.2.69/0.0/0.0/18.053/0016982).

## A Proofs

In this appendix we present the proofs of all the statements in Section 3 and we also present a series of technical results needed for their proofs. Throughout the appendix, we assume that (10), (11), and (2) hold.

*Proof.* (of Proposition 3.1) First, note that if (11) holds, then  $r_{up,p}^a(G_{in}, 0) = 0$  and  $r_{over,p}^a(G_{in}, 0) = k_{over}(r_{up,p}^g(G_{in}, 0) - l)$ . Consequently,

$$D^p = (1 - Y_{h/p})[\alpha r_{up,p}^g(G_{in}, 0) + k_{over}Y_a l] - k_{deg}, \quad (14)$$

with  $\alpha = Y_g - k_{over}Y_a > 0$  (see (10)). We also note that (see (11)):

$$D^p - D^a = (1 - Y_{h/p})\alpha[r_{up,p}^g(G_{in}, 0) - l] > 0,$$

hence  $D^p > D^a$ . Now, the steady states of (8) in absence of cleaners are given by the solution of the following system

$$\begin{aligned} 0 &= (1 - Y_{h/p})f_p(\cdot) - k_{deg} - D, \\ 0 &= D(G_{in} - G) - r_{up,p}^g(\cdot)B_p, \\ 0 &= -DA + [r_{over,p}^a(\cdot) - r_{up,p}^g(\cdot)]B_p. \end{aligned} \quad (15)$$

Following the same arguments of Proposition 1 by Martínez and Gouzé (2021), it is possible to show that

- (1) If  $D \geq D^a$ , then any solution of (15) satisfies  $r_{up,p}^g(G, A) \geq l$ .
- (2) If  $D < D^a$ , then any solution of (15) satisfies  $A = 0$  and  $r_{up,p}^g(G, A) < l$ .

If (1) holds, according to the property (7), we have that  $r_{up,p}^a(G, A) = 0$ . Consequently, (15) is equivalent to

$$\begin{aligned} 0 &= (1 - Y_{h/p})[\alpha r_{up,p}^g(G, A) + kY_a l] - k_{deg} - D, \\ 0 &= D(G_{in} - G) - r_{up,p}^g(G, A)B_p, \\ 0 &= -DA + r_{up,p}^a(G, A)B_p. \end{aligned} \quad (16)$$

Again, following the proof of Proposition 1 by Martínez and Gouzé (2021), we have that

$$G = G_{in} - \beta_g B_p, \quad \beta_g = \frac{D + k_{deg} - k_{over}(1 - Y_{h/p})Y_a l}{D(1 - Y_{h/p})\alpha}, \quad (17)$$

$$A = \beta_a B_p, \quad \beta_a = \frac{(1 - Y_{h/p})Y_g \beta_g - 1 - \frac{k_{deg}}{D}}{(1 - Y_{h/p})Y_a}. \quad (18)$$

Replacing (17) and (18) in the first equation of (16), we obtain the following equation for  $B_p$ :

$$\underbrace{(1 - Y_{h/p})[\alpha r_{up,p}^g(G_{in} - \beta_g B_p, \beta_a B_p) + k_{over} Y_a l]}_{g(B_p)} - k_{deg} - D = 0 \quad (19)$$

Since  $g$  is strictly decreasing,  $g(0) = D^p - D$ , and

$$\begin{aligned} g(G_{in}/\beta_g) &= (1 - Y_{h/p})k_{over} Y_a l - k_{deg} - D \\ &< (1 - Y_{h/p})Y_g l - k_{deg} - D \\ &= D^a - D \\ &< 0, \end{aligned}$$

we conclude that (19) admits a unique solution  $B_p^p > 0$  if  $D < D^p$ , and has no positive solution if  $D \geq D^p$ . In particular, this proves (b).

If (2) holds, then  $A = 0$  and  $r_{over,p}^a(G, A) = 0$ . Replacing  $A$  and  $r_{over,p}^a(G, A)$  by 0 in (15), we obtain

$$0 = \underbrace{(1 - Y_{h/p})Y_g r_{up,p}^g(G, 0) - k_{deg} - D}_{h(G)} \quad (20)$$

$$0 = D(G_{in} - G) - r_{up,p}^g(G, 0)B_p.$$

As in a classical chemostat model (note that  $G \mapsto h(G)$  is strictly increasing), (20) admits a unique positive steady state if, and only if,  $h(G_{in}) > 0$ . Since  $D < D_a$  and (11) holds, we have

$$\begin{aligned} h(G_{in}) &> (1 - Y_{h/p})Y_g r_{up,p}^g(G_{in}, 0) - k_{deg} - D^a \\ &= D^p - D^a \\ &> 0. \end{aligned}$$

This completes the proof.  $\square$

*Proof.* (of Proposition 3.4) Any equilibrium in absence of producers is a solution of

$$\begin{aligned} 0 &= (1 - Y_{h/p})f_c(\cdot) - k_{deg} - D, \\ 0 &= D(G_{in} - G) - r_{up,c}^g(\cdot)B_c, \\ 0 &= -DA - r_{up,c}^g(\cdot)B_c. \end{aligned} \quad (21)$$

It is clear that  $A = 0$ . Consequently, we have

$$0 = \underbrace{(1 - Y_{h/p})Y_g r_{up,c}^g(G, 0) - k_{deg} - D}_{h(G)} \quad (22)$$

$$0 = D(G_{in} - G) - r_{up,c}^g(G, 0)B_c.$$

As in a classical chemostat model (Smith and Waltman, 1995), we have that (22) admits a solution, which is unique, if and only if  $h(G_{in}) > 0$ . The rest of the proof follows from noting that  $h(G_{in}) = D^c - D$ .  $\square$

**Lemma A.1.** *Let  $D^a$  and  $D^p$  be defined by Proposition 3.1. If  $D \leq D^a$  and  $(1 - Y_{h/p}) \neq \beta(1 - Y_{h/c})$ , or  $D \geq D^p$ , then (8) has no coexistence equilibrium.*

*Proof.* If  $D \leq D^a$ , we claim that there is no coexistence equilibrium with acetate. Indeed, by contradiction, if  $(B_p^*, B_c^*, G^*, A^*)$  is a coexistence equilibrium of (8) with  $A^* > 0$ , from the fourth equation in (8) we obtain that

$$-DA^* + [r_{over,p}^a(G^*, A^*) - r_{up,p}^a(G^*, A^*)]B_p^* - r_{over,c}^a(G^*, A^*)B_c^* = 0,$$

and therefore

$$r_{over,p}^a(G^*, A^*) \geq DA^*/B_p^* > 0. \quad (23)$$

This implies that overflow metabolism occurs, hence, using (7), we have that there cannot be acetate uptake,  $r_{up,p}^a(G^*, A^*) = 0$ . From the first equation in (8) and the definition of a coexistence equilibrium, we have that

$$(1 - Y_{h/p})(Y_g - k_{over}Y_a)r_{over,p}^a(G^*, A^*) + D^a - D = 0. \quad (24)$$

However,  $D^a \geq D$  and  $(1 - Y_{h/p})(Y_g - k_{over}Y_a)r_{over,p}^a(G^*, A^*) > 0$ , which contradicts (24). This proves that any coexistence equilibrium has no acetate. Therefore, any coexistence equilibrium is of the form  $(B_p^*, B_c^*, G^*, 0)$  satisfying

$$\begin{aligned} 0 &= ((1 - Y_{h/p})Y_g r_{up,p}^g(G^*, 0) - k_{deg} - D)B_p^*, \\ 0 &= ((1 - Y_{h/c})\beta Y_g r_{up,p}^g(G^*, 0) - k_{deg} - D)B_c^*, \\ 0 &= D(G_{in} - G^*) - r_{up,p}^g(G^*, 0)B_p - \beta r_{up,p}^g(G^*, 0)B_c. \end{aligned} \quad (25)$$

However, this implies that  $B_p = 0$  or  $B_c = 0$ . Thus, there cannot be a coexistence equilibrium when  $D \leq D^a$ . This completes the proof.  $\square$

**Lemma A.2.** *If (8) admits a coexistence equilibrium, say  $(B_p^*, B_c^*, G^*, A^*)$ , then*

$$0 < A^* \text{ and } r_{up,p}^g(G^*, A^*) > l.$$

*Proof.* Let us assume that  $D > D^a$  and that (8) admits a coexistence equilibrium  $(B_p^*, B_c^*, G^*, A^*)$ . If  $A^* = 0$ , from the differential equation for  $A$  in (8), we obtain  $r_{over,p}^a(G^*, 0) = 0$  and consequently  $r_{up,p}^g(G^*, 0) \leq l$ . This implies that (see ODE for  $B_p$ ):

$$(1 - Y_{h/p})Y_g r_{up,p}^g(G^*, 0) - k_{deg} = D > D^a = (1 - Y_{h/p})Y_g l - k_{deg}.$$

Hence,  $r_{up,p}^g(G^*, 0) > l$ , which is a contradiction. Therefore  $A^*$  must be positive, which implies that  $r_{over,p}^a(G^*, A^*) > 0$  (same argument used to obtain (23)).  $\square$

To continue studying the existence of a coexistence equilibrium, it is convenient to define two functions,  $\gamma$  and  $\phi$ .

From Lemma A.1, coexistence is possible only when  $D \in (D^a, D^p)$ , with  $D^a$  and  $D^p$  defined in 3.1. Under such conditions, overflow metabolism occurs at any equilibrium with producers (see Proposition 3.1 and Lemma A.2). Therefore, the glucose uptake rate ( $r_{up,p}^g$ ) at any equilibrium with the presence of producers ( $B_p > 0$ ) has the same value and is obtained from the following equation

$$(1 - Y_{h/p})[Y_g r_{up,p}^p - k_{over}(r_{up,p}^p - l)] - D - k_{deg} = 0. \quad (26)$$

Let us define the function  $\gamma : (D^a, D^p) \rightarrow \mathbb{R}_+$  by

$$\gamma(D) = r_{up,p}^g, \quad (27)$$

with  $r_{up,p}^g$  given by (27). It can be shown that

$$\gamma(D) = \frac{D - D^a}{(Y_g - k_{over}Y_a)(1 - Y_{h/p})}. \quad (28)$$

We also define the function  $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}$  by

$$\phi(A) := (1 - Y_{h/c}) \left( Y_g \beta \gamma(D) + k_a \frac{A}{K_a + A} d(\beta \gamma(D)) + k_{Acs} \frac{A}{K_{Acs} + A} \right) - k_{deg} - D. \quad (29)$$

Note that if  $A^*$  is the acetate concentration at a coexistence equilibrium, then  $\phi(A^*) = 0$ .

**Lemma A.3.** *Let  $D^a$  and  $D^p$  be defined by Proposition 3.1. If  $D \in (D^a, D^p)$ , then (8) admits at most one coexistence equilibrium. Moreover, if a coexistence equilibrium exists, say  $(B_p^*, B_c^*, G^*, A^*)$ , then*

$$0 < A^* < A^p, \quad 0 < G^* < G^p, \quad \phi(A^p) > 0 \text{ and } \phi(0) < 0,$$

where  $G^p$  and  $A^p$  are defined in Proposition 3.1.

*Proof.* Let us assume that (8) admits a coexistence equilibrium  $(B_p^*, B_c^*, G^*, A^*)$ . From the equation for  $B_p$  in (8), we have  $r_{up,p}^g(G^*, A^*) = \gamma(D)$ , with  $\gamma$  defined by (27). From the definition of  $G^p$  and  $A^p$ , we also have  $r_{up,p}^g(G^p, A^p) = \gamma(D)$ . Similarly, we obtain that

$$r_{over,p}^a(G^p, A^p) = r_{over,p}^a(G^*, A^*) = \gamma_{over}(D) := k_{over}(\gamma(D) - l).$$

Now, we prove that  $A^* < A^p$ . By contradiction, assume that  $A^p \leq A^*$ . Using the monotonicity of  $r_{up,p}^g$  and the fact that  $r_{up,p}^g(G^p, A^p) = r_{up,p}^g(G^*, A^*)$ , we obtain that  $G^p \leq G^*$ . From the third equation in (8) we obtain that:

$$D(G_{in} - G^p) - \gamma(D)B_p^p = 0 < r_{up,c}^g(G^*, A^*)B_c^* = D(G_{in} - G^*) - \gamma(D)B_p^*,$$

from where  $0 \leq D(G^* - G^p) < \gamma(D)(B_p^p - B_p^*)$  which implies

$$B_p^p > B_p^*. \quad (30)$$

From the fourth equation in (8) we obtain that:

$$-DA^p + \gamma_{over}(D)B_p^p < r_{up,c}^a(G^*, A^*)B_c^* = -DA^* + \gamma_{over}(D)B_p^*,$$

from where  $0 \leq D(A^* - A^p) < \gamma_{over}(D)(B_p^* - B_p^p)$  which implies  $B_p^* > B_p^p$ . This contradicts (30). Then,  $A^p > A^*$  and consequently  $G^p > G^*$ . Now, from the equation for  $B_c$  in (8), we obtain  $\phi(A^*) = 0$ , with  $\phi$  defined by (29). Since  $\phi$  is strictly increasing, we have the uniqueness of  $A^c$ . The rest of the proof is straightforward.  $\square$

**Lemma A.4.** *Let  $D^p$  and  $D^a$  be defined in Proposition 3.1 and assume that  $D^a < D < D^p$ . Let  $D^c$  and  $G^c$  be defined by Proposition 3.1 and assume that  $D < D^c$ . If  $(1 - Y_{h/p})f_p(G^c, 0) - k_{deg} \leq D$ , then (8) has no coexistence equilibrium.*

*Proof.* Let us assume that (8) admits a coexistence equilibrium  $(B_p^*, B_c^*, G^*, A^*)$ . By definition we have

$$\begin{aligned} (1 - Y_{h/c})f_c(G^c, 0) - k_{deg} &= D \\ (1 - Y_{h/c})f_c(G^*, A^*) - k_{deg} &= D. \end{aligned} \quad (31)$$

From Lemma A.3, we have that  $A^* > 0$ , hence  $r_{up,c}^a(G^*, A^*) > 0$ . Recalling the definition of  $f_c$ , from (31) we have that

$$\begin{aligned} Y_g r_{up,c}^g(G^c, 0) &= Y_g r_{up,c}^g(G^*, A^*) + Y_a r_{up,c}^a(G^*, A^*) \\ &> Y_g r_{up,c}^g(G^*, A^*). \end{aligned}$$

Consequently  $r_{up,c}^g(G^c, 0) > r_{up,c}^g(G^*, A^*)$ . Since  $r_{up,c}^g$  is proportional to  $r_{up,p}^g$ , we have that  $r_{up,p}^g(G^c, 0) > r_{up,p}^g(G^*, A^*)$ . Now it is not difficult to show that

$$f_p(G^c, 0) = \alpha r_{up,p}^g(G^c, 0) + k_{over} Y_a l > \alpha r_{up,p}^g(G^*, A^*) + k_{over} Y_a l = f_p(G^*, A^*).$$

Now, if  $(1 - Y_{h/p})f_p(G^c, 0) - k_{deg} \leq D$ , then  $(1 - Y_{h/p})f_p(G^*, A^*) - k_{deg} < D$  which contradicts the definition of the coexistence equilibrium.  $\square$

**Lemma A.5.** *Let  $D^p$ ,  $D^a$ , and  $A^p$  be given by Proposition 3.1. If  $D^a < D < D^p$ , then (8) admits a coexistence equilibrium if and only if*

$$\phi(A^p) > D \text{ and } \phi(0) < 0. \quad (32)$$

*Proof.* From Lemma A.3 we have that (32) is a necessary condition for the existence of a coexistence equilibrium. We prove here that it is also a sufficient condition. If (32) holds, since  $\phi$  is strictly increasing, there exists a unique  $A^* \in (0, A^p)$  such that  $\phi(A^*) = 0$ . From the first equation in (8), we have  $\varphi(G) = 0$  with  $\varphi(G) = (1 - Y_{h/p})f_p(G, A^*) - k_{deg} - D$ . It is clear that  $\varphi$  is strictly increasing and that  $\varphi(0) = -D - k_{deg} < 0$  and  $\varphi(G^p) > (1 - Y_{h/p})f_p(G^p, A^p) - k_{deg} - D = 0$ . Consequently, there is a unique  $G^* \in (0, G^p)$  such that  $\varphi(G^*) = 0$ . It remains to prove the positiveness of the unique solution of the following linear system for  $(B_p, B_c)$ :

$$\begin{aligned} 0 &= D(G_{in} - G^*) - r_{up,p}^g(G^*, A^*)B_p - r_{up,c}^g(G^*, A^*)B_c, \\ 0 &= -DA^* + r_{over,p}^a(G^*, A^*)B_p - r_{up,c}^a(G^*, A^*)B_c. \end{aligned} \quad (33)$$

This system can be rewritten as

$$\begin{aligned} B_p &= \frac{r_{up,c}^a(G^*, A^*)B_c + DA^*}{r_{over,p}^a(G^*, A^*)} \text{ and} \\ \underbrace{G_{in} - G^* - A^* \frac{r_{up,p}^g(G^*, A^*)}{r_{over,p}^a(G^*, A^*)}}_{\kappa} &= \frac{1}{D} \left( r_{up,c}^g(G^*, A^*) + \frac{r_{up,c}^a(G^*, A^*)r_{up,p}^g(G^*, A^*)}{r_{over,p}^a(G^*, A^*)} \right) B_c. \end{aligned}$$

Since  $A^p > A^*$  and  $G^p > G^*$ , we have that

$$\kappa > G_{in} - G^p - A^p \frac{r_{up,p}^g(G^*, A^*)}{r_{over,p}^a(G^*, A^*)}.$$

Finally, since  $r_{up,p}^g(G^*, A^*) = r_{up,p}^g(G^p, A^p)$ , from (15) we conclude that  $\kappa = 0$ . Then (33) has a positive solution.  $\square$

*Proof.* (of Theorem 3.6) Part (a) follows directly from Lemma A.2. Let us assume that  $D^a < D < D^p$ . Using Lemma A.5, the proof of parts (b) and (c) follows from proving that

- (I) If  $D < D^c$  and  $(1 - Y_{h/p})f_p(G^c, 0) - D - k_{deg} > 0$ , then  $\phi(0) < 0$ ,
- (II) If  $D \geq D^c$ , then  $\phi(0) < 0$ ,
- (III) If  $(1 - Y_{h/c})f_c(G^p, A^p) - k_{deg} - D > 0$ , then  $\phi(A^p) > 0$ .

For (I), from the hypotheses, we have that  $(1 - Y_{h/p})f_p(G^c, 0) - k_{deg} > D^a$ . Using the definition of  $D^a$  we obtain  $Y_g l < f_p(G^c, 0)$ , and hence

$$l < r_{up,p}^g(G^c, 0). \quad (34)$$

Now, from the definition of  $G^p$  and  $A^p$  (see ODE for  $B_p$ ), and from the hypotheses, we have that

$$(1 - Y_{h/p})f_p(G^p, A^p) = D + k_{deg} < (1 - Y_{h/p})f_p(G^c, 0). \quad (35)$$

Since  $\gamma(D) = r_{up,p}^g(G^p, A^p)$  and  $r_{up,p}^g(G^c, 0) > l$ , from (35) we conclude that

$$\gamma(D) < r_{up,p}^g(G^c, 0), \quad (36)$$

From the definition of  $G^c$  and (36) we obtain that

$$(1 - Y_{h/c})Y_g \beta \gamma(D) < (1 - Y_{h/c})Y_g \beta r_{up,p}^g(G^c, 0) = k_{deg} + D,$$

which implies  $\phi(0) < 0$ .

For (II), we note that

$$f_c(G_{in}, 0) = Y_g \beta r_{up,p}^g(G_{in}, 0) > Y_g \beta r_{up,p}^g(G^p, A^p) = Y_g \beta \gamma(D),$$

which implies

$$\phi(0) < (1 - Y_{h/c})f_c(G_{in}, 0) - k_{deg} - D = D^c - D < 0.$$

Finally, for (III), it is straightforward to verify that

$$\phi(A^p) = (1 - Y_{h/c})f_c(G^p, A^p) - k_{deg} - D > 0.$$

This completes the proof.  $\square$

The following lemma, shows that the equilibrium with producers given by Proposition 3.1 is locally stable when seen as an equilibrium of (8) without cleaners.

**Lemma A.6.** *Assume that  $D > D^a$ . Let  $E^p = (B_p^p, 0, G^p, A^p)$  be the equilibrium given by Proposition 3.1. Then  $(B_p^p, G^p, A^p)$  is locally stable with respect to the following system:*

$$\begin{aligned} \frac{dB_p}{dt} &= (1 - Y_{h/p})f_p(\cdot)B_p - k_{deg}B_p - DB_p, \\ \frac{dG}{dt} &= D(G_{in} - G) - r_{up,p}^g(\cdot)B_p, \\ \frac{dA}{dt} &= -DA + [r_{over,p}^a(\cdot) - r_{up,p}^a(\cdot)]B_p. \end{aligned} \quad (37)$$

*Proof.* The proof follows the same idea of the proof of Proposition 2 by Martínez and Gouzé (2021). Along the proof, we will write  $r$  instead of  $r_{up,p}^g$ . According to Proposition 3.1,  $(B_p^p, G^p, A^p)$  is in the region  $\{(B_p, G, A); r(G, A) \geq l\}$ . Therefore, using the definition of  $r_{over,p}^g$  and the property (7), we can study the local stability of  $(B_p^p, G^p, A^p)$  in the following system

$$\begin{aligned}\frac{dB_p}{dt} &= \alpha' r(\cdot) B_p + \eta B_p - D B_p, \\ \frac{dG}{dt} &= D(G_{in} - G) - r(\cdot) B_p, \\ \frac{dA}{dt} &= -D A + k_{over}(r(\cdot) - l) B_p,\end{aligned}\tag{38}$$

with  $\alpha' = (1 - Y_{h/p})(Y_g - k_{over} Y_a)$  and  $\eta = (1 - Y_{h/p})k_{over} Y_a l - k_{deg}$ . The change of variables  $U = B_p + \alpha' G$  and  $W = B_p + (1 - Y_{h/p}) Y_g G + (1 - Y_{h/p}) Y_a A$  leads (38) to

$$\begin{aligned}\frac{dB_p}{dt} &= \alpha' \hat{r}(\cdot) B_p + \eta B_p - D B_p, \\ \frac{dU}{dt} &= \eta B_p + D(\alpha' G_{in} - U), \\ \frac{dW}{dt} &= D(Y'_g G_{in} - W) - k_{deg} B_p,\end{aligned}\tag{39}$$

with

$$\hat{r}(B_p, U, W) = r \left( \frac{1}{\alpha'} (U - B_p), \frac{1}{Y'_a} \left( W - \frac{Y'_g}{\alpha'} U \right) + \frac{k_{over}}{\alpha'} B_p \right),\tag{40}$$

$Y'_g = (1 - Y_{h/p}) Y_g$ , and  $Y'_a = (1 - Y_{h/p}) Y_a$ . The characteristic polynomial associated with the Jacobian matrix of (39) is given by

$$p(\lambda) = \det \begin{bmatrix} \alpha' B_p^p \frac{\partial \hat{r}}{\partial B_p} - \lambda & \alpha' B_p^p \frac{\partial \hat{r}}{\partial U} & \alpha' B_p^p \frac{\partial \hat{r}}{\partial W} \\ \eta & -D - \lambda & 0 \\ -k_{deg} & 0 & -D - \lambda \end{bmatrix}.$$

It can be shown that  $p(\lambda)$  can be written as

$$p(\lambda) = -(D + \lambda) \det \begin{bmatrix} \alpha' B_p^p \frac{\partial \hat{r}}{\partial B_p} - \lambda & B_p^p \left( \eta \frac{\partial \hat{r}}{\partial U} - k_{deg} \frac{\partial \hat{r}}{\partial W} \right) \\ \alpha' & -D - \lambda \end{bmatrix}.$$

Thus, one root of  $p$  is  $-D$  and the other two roots are the eigenvalues of the matrix

$$J = \begin{bmatrix} \alpha' B_p^p \frac{\partial \hat{r}}{\partial B_p} & B_p^p \left( \eta \frac{\partial \hat{r}}{\partial U} - k_{deg} \frac{\partial \hat{r}}{\partial W} \right) \\ \alpha' & -D \end{bmatrix}.$$



We note that:

$$\begin{aligned}\frac{\partial \hat{r}}{\partial B_p} &= -\frac{1}{\alpha} \frac{\partial r}{\partial G} + \frac{k_{over}}{\alpha'} \frac{\partial r}{\partial A} < 0, \\ \frac{\partial \hat{r}}{\partial U} &= \frac{1}{\alpha'} \frac{\partial r}{\partial G} - \frac{Y'_g}{\alpha' Y'_a} \frac{\partial r}{\partial A} > 0, \\ \frac{\partial \hat{r}}{\partial W} &= \frac{1}{Y'_a} \frac{\partial r}{\partial A} < 0.\end{aligned}$$

Using the fact that  $D > Y'_g l$ , we have

$$\begin{aligned}\frac{\det(J)}{B_p^p} &= D \underbrace{\left( \frac{\partial r}{\partial G} - k_{over} \frac{\partial r}{\partial A} \right)}_{>0} + \left( \eta \frac{Y'_g}{Y'_a} + \frac{k_{deg}}{Y'_a} \right) \frac{\partial r}{\partial A} - \eta \frac{\partial r}{\partial G} \\ &> Y'_g l \left( \frac{\partial r}{\partial G} - k_{over} \frac{\partial r}{\partial A} \right) + \left( \eta \frac{Y'_g}{Y'_a} + \frac{k_{deg}}{Y'_a} \right) \frac{\partial r}{\partial A} - \eta \frac{\partial r}{\partial G} \\ &= (Y'_g l - \eta) \frac{\partial r}{\partial G} + (-Y'_g Y'_a l k_{over} + \eta Y'_g + \alpha' k_{deg}) \frac{1}{Y'_a} \frac{\partial r}{\partial A} \\ &= (\alpha' + k_{deg}) \frac{\partial r}{\partial G} - k_{over} \frac{\partial r}{\partial A} \\ &> 0.\end{aligned}$$

Therefore  $\det(J) > 0$ . Now, it is clear that  $Tr(J) < 0$ , hence the eigenvalues of  $J$  have negative real part, and all the roots of  $p$  have negative real part. This completes the proof.  $\square$

**Lemma A.7.** *Assume that  $D^a < D < D^p$  and let  $E^p$  be given by Proposition 3.1. Then the Jacobian matrix associated with (8) evaluated at  $E^p$  has three eigenvalues with negative real part, and one eigenvalue is given by:*

$$(1 - Y_{h/c})f_c(G^p, A^p) - k_{deg} - D.$$

*Proof.* The evolution of  $B_p, B_c, G$ , and  $A$  is given by

$$\begin{aligned}\frac{dB_p}{dt} &= \alpha' r_{up,p}^g(\cdot) B_p + \eta B_p - D B_p, \\ \frac{dB_c}{dt} &= (1 - Y_{h/c}) f_c(\cdot) B_c - D B_c, \\ \frac{dG}{dt} &= D(G_{in} - G) - r_{up,p}^g(\cdot) B_p - r_{up,c}^g(\cdot) B_c, \\ \frac{dA}{dt} &= -DA + k_{over} [r_{up,p}^g(\cdot) - l] B_p - r_{up,c}^a(\cdot) B_c,\end{aligned}\tag{41}$$

with  $\alpha' = (1 - Y_{h/p})(Y_g - k_{over} Y_a)$  and  $\eta = (1 - Y_{h/p})k_{over} Y_a l + k_{deg}$ . The Jacobian matrix,  $J$ , associated with (41) evaluated at  $E^p$  takes the following form

$$\begin{bmatrix} 0 & 0 & \alpha' B_p \frac{\partial r_{up,p}^g}{\partial G} & \alpha' B_p \frac{\partial r_{up,p}^g}{\partial A} \\ 0 & (1 - Y_{h/c})f_c - k_{deg} - D & 0 & 0 \\ -r_{up,p}^g & -r_{up,c}^g & -D - B_p \frac{\partial r_{up,p}^g}{\partial G} & -B_p \frac{\partial r_{up,p}^g}{\partial A} \\ k_{over}(r_{up,p}^g - l) & -r_{up,c}^a & k_{over} B_p \frac{\partial r_{up,p}^g}{\partial G} & -D + k_{over} B_p \frac{\partial r_{up,p}^g}{\partial A} \end{bmatrix}.$$

Using the properties of the determinant, we have that one eigenvalue is  $(1 - Y_{h/c})f_c(G^p, A^p) - k_{deg} - D$ , while the other three eigenvalues are exactly the same as those of the Jacobian matrix associated with (37) and evaluated at  $(B_p^p, G^p, A^p)$ . From Lemma A.6, we conclude that these three eigenvalues are negative. Thus,  $E^p$  is unstable and hyperbolic, if and only if,  $(1 - Y_{h/c})f_c(G^p, A^p) - k_{deg} - D < 0$ .  $\square$

**Lemma A.8.** *Assume that  $D < D_c$  and let  $E^c$  be the equilibrium given by Proposition 3.4. Then the Jacobian matrix associated with (8) evaluated at  $E^c$  has three eigenvalues with negative real part, and one eigenvalue given by*

$$(1 - Y_{h/p})f_p(G^c, 0) - k_{deg} - D.$$

*Proof.* The Jacobian matrix of (8) evaluated at  $E^c$  is given by

$$J = \begin{bmatrix} (1 - Y_{h/p})f_p - k_{deg} - D & 0 & 0 & 0 \\ 0 & 0 & (1 - Y_{h/p})B_c \frac{\partial f_c}{\partial G} & (1 - Y_{h/p})B_c \frac{\partial f_c}{\partial A} \\ -r_{up,p}^g & -r_{up,c}^g & -D - B_c \frac{\partial r_{up,c}^g}{\partial G} & -B_c \frac{\partial r_{up,c}^g}{\partial A} \\ k_{over}(r_{up,p}^g - l) & 0 & 0 & -D - B_c \frac{\partial r_{up,c}^a}{\partial A} \end{bmatrix}.$$

The characteristic polynomial associated with  $J$  takes the following form

$$p(\lambda) = (\lambda_1 - \lambda) \det \begin{bmatrix} -\lambda & (1 - Y_{h/p})B_c \frac{\partial f_c}{\partial G} & (1 - Y_{h/p})B_c \frac{\partial f_c}{\partial A} \\ -r_{up,c}^g & -D - B_c \frac{\partial r_{up,c}^g}{\partial G} - \lambda & -B_c \frac{\partial r_{up,c}^g}{\partial A} \\ 0 & 0 & -D - B_c \frac{\partial r_{up,c}^a}{\partial A} - \lambda \end{bmatrix},$$

with  $\lambda_1 = (1 - Y_{h/p})f_p(G^c, 0) - k_{deg} - D$ . Then we can write

$$p(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \det(J' - \lambda I),$$

with  $\lambda_2 = -D - B_c \frac{\partial r_{up,c}^a}{\partial A} < 0$  and

$$J' = \begin{bmatrix} 0 & (1 - Y_{h/p})B_c \frac{\partial f_c}{\partial G} \\ -r_{up,c}^g & -D - B_c \frac{\partial r_{up,c}^g}{\partial G} \end{bmatrix}.$$

Since  $Tr(J') < 0$  and  $\det(J') > 0$ ,  $J'$  has two negative eigenvalues. Thus,  $E^c$  is hyperbolic and unstable if and only if  $\lambda_1 > 0$ .  $\square$

*Proof.* (of Lemma 3.10) Note that the variables  $U_p := \frac{Y_{h/p}}{1-Y_{h/p}}B_p - H_p$  and  $U_c := \frac{Y_{h/c}}{1-Y_{h/c}}B_c - H_c$  satisfy the following differential equations:

$$\frac{dU_p}{dt} = -(k_{deg} + D)U_p \quad \text{and} \quad \frac{dU_c}{dt} = -(k_{deg} + D)U_c.$$

We conclude the proof noting that  $U_p$  and  $U_c$  asymptotically approach zero. □

## B Algorithm to find the coexistence equilibrium

Let  $D^a$  and  $D^p$  be given by Proposition 3.1. From now on, we assume that  $D \in (D^a, D^p)$ , otherwise there is no coexistence equilibrium (see Theorem 3.6). The first step is to determine the equilibrium  $E^p = (B_p^p, 0, G^p, A^p)$  given by Proposition 3.1. The instructions on how to do so are dictated by the proof of Proposition 3.1. Indeed,  $B_p^p$  is obtained as the solution of

$$g(B_p) = 0, \tag{42}$$

with  $g$  defined by (19). This equation has a unique solution on the interval  $[0, G_{in}/\beta_g]$ , with  $\beta_g$  defined in (17). Moreover,  $g(0) > 0$  and  $g(G_{in}/\beta_g) < 0$ , which provides an interval to look for the solution. Thus, equation (42) can be easily solved, for example, with the solver *fzero* in MATLAB. The values of  $A^p$  and  $G^p$  are obtained from

$$G^p = G_{in} - \beta_g B_p^p \quad \text{and} \quad A^p = \beta_a B_p^p,$$

where  $\beta_a$  is defined in (18).

We also need the value of  $G^c$ , the glucose concentration associated with the equilibrium  $E^c$  given by Proposition 3.4. Let  $D^c$  be given by Proposition 3.4. If  $D < D^c$ , then  $G^c$  is the unique solution of  $(1 - Y_{h/c})f_c(G, 0) - k_{deg} - D = 0$ . This equation is easily solved explicitly. If  $D \geq D^c$ , we will take  $G^c$  as  $G_{in}$ . This is useful to distinguish the cases (b) and (c) in Theorem 3.6.

Now, to determine the coexistence equilibrium we use the following algorithm:

- 1) Determine  $c_1 = (1 - Y_{h/p})f_p(G^p, A^p) - k_{deg} - D$  and  $c_2 = (1 - Y_{h/c})f_c(G^c, 0) - k_{deg} - D$ .
- 2) If  $c_1 \leq 0$  or  $c_2 \leq 0$ , then there is no coexistence equilibrium. The algorithm ends. However, if  $c_1$  and  $c_2$  are positive, go to the next step.
- 3) Find  $A^* \in [0, A^p]$  as the unique solution of  $\phi(A) = 0$ , with  $\phi$  defined by (29). Note that  $\phi(0) < 0$  and  $\phi(A^p) > 0$ .
- 4) Find  $G^* \in [0, G_{in}]$  as the unique solution of  $\varphi(G) = 0$  with  $\varphi(G) := (1 - Y_{h/p})f_p(G, A^*) - k_{deg} - D$ . Note that  $\varphi(0) < 0$  and  $\varphi(G_{in}) > 0$ .
- 5) Find  $B_p^*$  and  $B_c^*$  as the unique solution of the following linear system:

$$\begin{bmatrix} r_{up,p}^g(G^*, A^*) & r_{up,c}^g(G^*, A^*) \\ r_{over,p}^a(G^*, A^*) & r_{up,c}^a(G^*, A^*) \end{bmatrix} \begin{bmatrix} B_p \\ B_c \end{bmatrix} = \begin{bmatrix} D(G_{in} - G^*) \\ DA^* \end{bmatrix}.$$

## C Algorithm to solve the MOP

Problem (13) is solved numerically with the interior point algorithm implemented in the toolbox *fmincon* of MATLAB (Byrd et al., 1999). To use *fmincon*, the objective function must be continuous on the feasible region which must be a closed set. In the following, we rewrite (13) in the standard form to use *fmincon*.

We recall the set  $\Omega$  defined in Section 4. Consider the compact set  $\bar{\Omega} \subset \mathbb{R}_3^+$  which can be described such that each element  $p = (Y_{h/p}, Y_{h/c}, D)$  on  $\bar{\Omega}_s$  satisfies

$$\begin{array}{ll} D & \leq D^p, \\ D & \geq D^c, \\ D & \geq D^a, \\ D & \leq (1 - Y_{h/p})f_p(G^c, 0) - k_{deg}, \\ Y_{h/p}, Y_{h/c} & \in [0, 1]. \end{array} \quad \text{or} \quad \begin{array}{ll} D & \leq D^p, \\ D & < D^c, \\ D & \geq D^a, \\ D & \leq (1 - Y_{h/c})f_c(G^p, A^p) - k_{deg}, \\ D & \leq (1 - Y_{h/p})f_p(G^c, 0) - k_{deg}, \\ Y_{h/p}, Y_{h/c} & \in [0, 1]. \end{array}$$

where  $D^a$ ,  $D^p$ ,  $G^p$ , and  $A^p$  are given by Proposition 3.1, and  $D^c$  and  $A^c$  are given by Proposition 3.4. To extend the definition of  $\Phi$  (defined in Section 4) on the boundary of  $\bar{\Omega}$  note that:

- If  $D \geq D^c$ , then cleaners can never survive (Proposition 3.4). Hence, we expect that solutions of (8) converges either to  $(B_p^p, 0, G^p, A^p)$  or to  $(0, 0, G_{in}, 0)$ .
- There is no element of the boundary satisfying  $D = D^a$ . Indeed, as in the proof of Theorem 3.6, it can be proven that if  $D < D^c$ , (34) holds. Now, in  $\bar{\Omega}$ , we have that  $D \leq (1 - Y_{h/p})f_p(G^c, 0) - k_{deg}$ . Combining this equation with (34), it is straightforward to show that  $D > D^a$ .
- If  $D < D^c$  and  $D = D^p$ , then there is no equilibrium with producers.
- If  $D < D^c$ ,  $D < D^p$  and  $D = (1 - Y_{h/p})f_p(G^c, 0) - k_{deg}$ , then the equilibrium with producers is unstable.

Then we define

$$\hat{\Phi} = \begin{cases} (DH_p^p, H_p^p/G_{in}) & \text{if } D < D^p \text{ and } D \geq D^c, \\ (0, 0) & \text{if } D = D^p \text{ and } D \geq D^c, \\ (DH_c^c, H_c^c/G_{in}) & \text{if } D = (1 - Y_{h/p})f_p(G^c, 0) - k_{deg} \text{ and } D < D^c, \\ (DH_p^p, H_p^p/G_{in}) & \text{if } D = (1 - Y_{h/c})f_c(G^p, A^p) - k_{deg} \text{ and } D < D^c, \end{cases}$$

where  $H_p^p$  and  $H_c^c$  are the protein concentration associated with producers and cleaners, respectively. Then, we solve numerically the following problem instead of (13):

$$\begin{aligned} \max \quad & \lambda \hat{\Phi}(v) + (1 - \lambda) \hat{\Phi}(v) \\ & Av \leq b \\ & c(v) \leq 0 \\ & ub \leq v \leq up, \end{aligned} \tag{43}$$

where

$$A = \begin{bmatrix} f_p(G_{in}, 0) & 0 & 1 \\ -Y_{gl} & 0 & -1 \end{bmatrix}, b = \begin{bmatrix} f_p(G_{in}, 0) - k_{deg} \\ k_{deg} - Y_{Gl} \end{bmatrix}, lb = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

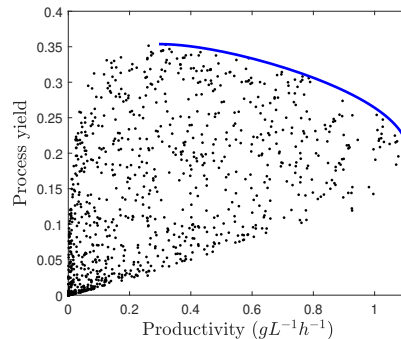


Figure 7: Scatter plot of  $\Phi$  using the steady states of model (8)-(9) with  $d$  replaced by (5). The continuous line is the POF obtained in Figure 4

$$ub = \begin{bmatrix} 1 \\ 1 \\ f_p(G_{in}, 0) - k_{deg} \end{bmatrix}, \text{ and } c(v) = \begin{bmatrix} v_3 + k_{deg} - (1 - v_1)f_p(G^c(v), 0) \\ v_3 + k_{deg} - (1 - v_2)f_c(G^p(v), A^p(v)) \end{bmatrix}.$$

$G^c(v)$  is defined as  $G_{in}$  when  $v_3 \geq D^c$ .

## D Robust choice of the down-regulation function

Let us assume that in (8) the down-regulation function  $d$  is replaced by  $\hat{d}$ , defined in (5). We then run long-term simulations of (8)-(9) until a steady state is detected (which is always the case).<sup>2</sup> Then, we evaluate the process yield and the productivity of the system at the time at which steady state has been reached. Figure 7 shows the result of this experiment for 1000 different values of  $(Y_{h/p}, Y_{h/c}, D)$ . As we can see, the POF obtained when  $d$  is given by (6) represents a good approximation of the POF when  $d$  is given by (5). This shows that the choice of  $d$  in this paper is adequate to study the model proposed by Mauri et al. (2020).

## References

- Baeshen MN, Al-Hejin AM, Bora RS, Ahmed M, Ramadan H, Saini KS, Baeshen NA, Redwan EM (2015) Production of biopharmaceuticals in *E. coli*: current scenario and future perspectives. *J Microbiol Biotechnol* 25(7):953–962
- Basan M, Hui S, Okano H, Zhang Z, Shen Y, Williamson JR, Hwa T (2015) Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature* 528(7580):99
- Bernstein HC, Paulson SD, Carlson RP (2012) Synthetic *Escherichia coli* consortia engineered for syntrophy demonstrate enhanced biomass productivity. *J Biotechnol* 157(1):159–166

<sup>2</sup>Let  $\xi$  be a state variable of the model. A steady-state is detected when  $|\xi(t + \Delta t) - \xi(t)| < \delta$ . In our simulations we choose  $\delta = 10^{-6}$  g L<sup>-1</sup> and  $\Delta t = 10$  days.

- Byrd RH, Hribar ME, Nocedal J (1999) An interior point algorithm for large-scale nonlinear programming. *SIAM J Optim* 9(4):877–900
- Chesson P (2000) Mechanisms of maintenance of species diversity. *Annu Rev Ecol Syst* 31(1):343–366
- De Mey M, De Maeseneire S, Soetaert W, Vandamme E (2007) Minimizing acetate formation in *E. coli* fermentations. *J Ind Microbiol Biotechnol* 34(11):689–700
- Deutscher J, Francke C, Postma PW (2006) How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol Mol Biol Rev* 70(4):939–1031
- Doran PM (1995) *Bioprocess Engineering Principles*. Elsevier
- Eiteman MA, Altman E (2006) Overcoming acetate in *Escherichia coli* recombinant protein fermentations. *Trends Biotechnol* 24(11):530–536
- Enjalbert B, Millard P, Dinclaux M, Portais JC, Létisse F (2017) Acetate fluxes in *Escherichia coli* are determined by the thermodynamic control of the Pta-AckA pathway. *Scientific reports* 7(1):1–11
- Fuchs C, Köster D, Wiebusch S, Mahr K, Eisbrenner G, Märkl H (2002) Scale-up of dialysis fermentation for high cell density cultivation of *Escherichia coli*. *J Biotechnol* 93(3):243–251
- Görke B, Stülke J (2008) Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat Rev Microbiol* 6(8):613–24
- Gunantara N (2018) A review of multi-objective optimization: Methods and its applications. *Cogent Eng* 5(1):1502242
- Hardin G (1960) The competitive exclusion principle. *Science* 131(3409):1292–1297
- Harvey E, Heys J, Gedeon T (2014) Quantifying the effects of the division of labor in metabolic pathways. *J Theor Biol* 360:222–242
- Hays SG, Patrick WG, Ziesack M, Oxman N, Silver PA (2015) Better together: engineering and application of microbial symbioses. *Curr Opin Biotechnol* 36:40–49
- Heßeler J, Schmidt JK, Reichl U, Flockerzi D (2006) Coexistence in the chemostat as a result of metabolic by-products. *J Math Biol* 53(4):556–584
- Huang CJ, Lin H, Yang X (2012) Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J Ind Microbiol Biotechnol* 39(3):383–399
- Jagmann N, Philipp B (2014) Design of synthetic microbial communities for biotechnological production processes. *J Biotechnol* 184:209–218
- Kremling A, Geiselmann J, Ropers D, de Jong H (2015) Understanding carbon catabolite repression in *Escherichia coli* using quantitative models. *Trends Microbiol* 23(2):99–109
- Kurland C, Dong H (1996) Bacterial growth inhibition by overproduction of protein. *Mol Microbiol* 21(1):1–4

- Lin H, Castro NM, Bennett GN, San KY (2006) Acetyl-CoA synthetase overexpression in *Escherichia coli* demonstrates more efficient acetate assimilation and lower acetate accumulation: a potential tool in metabolic engineering. *Appl Microbiol Biotechnol* 71(6):870–874
- Liu X, Li XB, Jiang J, Liu ZN, Qiao B, Li FF, Cheng JS, Sun X, Yuan YJ, Qiao J, et al. (2018) Convergent engineering of syntrophic *Escherichia coli* coculture for efficient production of glycosides. *Metab Eng* 47:243–253
- Luli GW, Strohl WR (1990) Comparison of growth, acetate production, and acetate inhibition of *Escherichia coli* strains in batch and fed-batch fermentations. *Appl Environ Microbiol* 56(4):1004–1011
- Martínez C, Gouzé JL (2021) Global dynamics of the chemostat with overflow metabolism. *J Math Biol* 82(3):1–21
- Mauri M, Gouzé JL, De Jong H, Cinquemani E (2020) Enhanced production of heterologous proteins by a synthetic microbial community: Conditions and trade-offs. *PLoS Comput Biol* 16(4):e1007795
- Miettinen K (2012) *Nonlinear Multiobjective Optimization*, vol 12. Springer Science & Business Media
- Pandhal J, Noirel J (2014) Synthetic microbial ecosystems for biotechnology. *Biotechnol Lett* 36(6):1141–1151
- Roell GW, Zha J, Carr RR, Koffas MA, Fong SS, Tang YJ (2019) Engineering microbial consortia by division of labor. *Microb Cell Factories* 18(1):1–11
- Rosenzweig RF, Sharp R, Treves DS, Adams J (1994) Microbial evolution in a simple unstructured environment: genetic differentiation in *Escherichia coli*. *Genetics* 137(4):903–917
- Sari T, El Hajji M, Harmand J (2012) The mathematical analysis of a syntrophic relationship between two microbial species in a chemostat. *Math Biosci Eng* 9(3):627–645
- Shong J, Diaz MRJ, Collins CH (2012) Towards synthetic microbial consortia for bioprocessing. *Curr Opin Biotechnol* 23(5):798–802
- Smith HL, Thieme HR (2011) *Dynamical Systems and Population Persistence*, vol 118. American Mathematical Soc.
- Smith HL, Waltman P (1995) *The Theory of the Chemostat: Dynamics of Microbial Competition*, vol 13. Cambridge University Press
- Stump SM, Klausmeier CA (2016) Competition and coexistence between a syntrophic consortium and a metabolic generalist, and its effect on productivity. *J Theor Biol* 404:348–360
- Tsoi R, Wu F, Zhang C, Bewick S, Karig D, You L (2018) Metabolic division of labor in microbial systems. *Proc Natl Acad Sci USA* 115(10):2526–2531
- Van Dien S (2013) From the first drop to the first truckload: commercialization of microbial processes for renewable chemicals. *Curr Opin Biotechnol* 24(6):1061–1068

Wolfe AJ (2005) The acetate switch. *Microbiol Mol Biol Rev* 69(1):12–50

Wu G, Yan Q, Jones JA, Tang YJ, Fong SS, Koffas MA (2016) Metabolic burden: cornerstones in synthetic biology and metabolic engineering applications. *Trends Biotechnol* 34(8):652–664