



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

Carlos Martinez, Eugenio Cinquemani, Hidde de Jong, Jean-Luc Gouzé

► To cite this version:

Carlos Martinez, Eugenio Cinquemani, Hidde de Jong, Jean-Luc Gouzé. Optimal protein production by a synthetic microbial consortium: Coexistence, distribution of labor, and syntrophy. *Journal of Mathematical Biology*, 2023, 87 (1), pp.1-37. 10.1007/s00285-023-01935-3 . hal-03525249

HAL Id: hal-03525249

<https://inria.hal.science/hal-03525249v1>

Submitted on 5 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



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

Carlos Martínez^{1,2} · Eugenio Cinquemani³ · Hidde de Jong³ · Jean-Luc Gouzé¹

Received: 14 December 2021 / Revised: 22 December 2022 / Accepted: 17 May 2023 /

Published online: 3 July 2023

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

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 mathematical model of such a synthetic community in a chemostat 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 (dis-

✉ Carlos Martínez
carlos.martinez@uv.cl

Eugenio Cinquemani
eugenio.cinquemani@inria.fr

Hidde de Jong
hidde.de-jong@inria.fr

Jean-Luc Gouzé
jean-luc.gouze@inria.fr

¹ Université Côte d'Azur, Inria, INRAE, CNRS, Sorbonne Université, Biocore Team, Sophia Antipolis, France

² Biology Centre of the Czech Academy of Sciences, Institute of Hydrobiology, Na Sádkách 7, 370 05 České Budějovice, Czech Republic

³ Univ. Grenoble Alpes, Inria, 38000 Grenoble, France

tribution 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

Mathematics Subject Classification 92B05 · 34C60 · 90C29

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). 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 (Luli and Strohl 1990; Eiteman and Altman 2006).

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 (Eiteman and Altman 2006; De Mey et al. 2007). These strategies have several inconveniences, 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 this new consortium, both *E. coli* strains, instead of only the producer strain, can produce recombinant proteins. The expression of the recombinant protein can be tuned for each strain (Izard et al. 2015; Miliadis-Argeitis et al. 2016), thus, adding a new dimension to the model studied by Mauri et al. (2020). We hypothesize that an appropriate tuning of protein expression may alleviate the productivity-yield trade-off in that some of the carbon diverted to the cleaners can be utilized for protein production.

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 (Smith and Waltman 1995). 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 (Heßeler et al. 2006; Sari et al. 2012; Harvey et al. 2014; Stump and Klausmeier 2016). The mixed culture studied in this work accounts for competition and syntrophy, so that establishing coexistence is not trivial. In the previous work by Mauri et al. (2020), coexistence was shown by numerical simulation. Although the model considered here is more complex, we provide analytical conditions for the occurrence of a coexistence equilibrium. This is achieved by slightly modifying the description of the uptake rates to make the model mathematically tractable.

Based on the coexistence 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 Sect. 2, we describe the chemostat model of the synthetic microbial community. In Sect. 3, we state Theorem 1 that gives necessary and sufficient conditions for the existence of a unique coexistence equilibrium. In Sect. 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

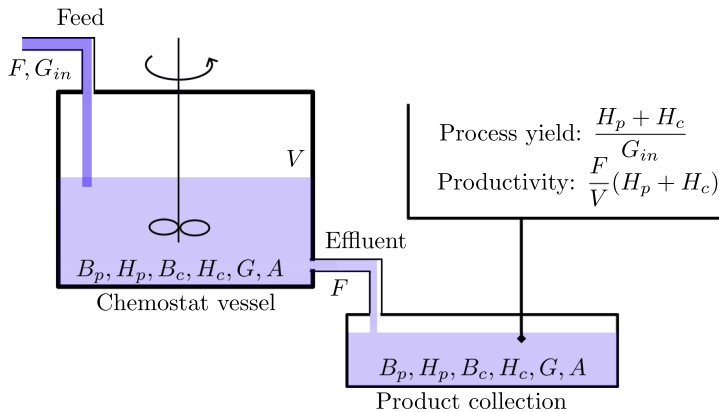


Fig. 1 Schematic diagram of a chemostat used for the production of recombinant proteins ($H_p + H_c$). The chemostat is fed at a rate F with a glucose concentration G_{in} . The reactor is emptied at the rate F keeping a constant volume V . The concentrations of producers (B_p) and cleaners (B_c), glucose (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

includes four Appendices with proofs and details of the algorithms to numerically determine the coexistence equilibrium and the solutions of the MOP.

2 Model description

The model of the *E. coli* community is a modified version of the model developed by Mauri et al. (2020). To facilitate the connection between the results presented in this work with those of Mauri et al. (2020), we will essentially use the same notation. Differences in the notation will be mentioned as they appear.

Consider a chemostat (see Fig. 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 by $r_{up,p}^g$ and $r_{up,c}^g$ for producers and cleaners, respectively. The glucose uptake rate by producers is given by

$$r_{up,p}^g(G, A) = k_g \frac{G}{G + K_g} \frac{\Theta_a}{A + \Theta_a}, \quad (1)$$

where k_g is the maximal uptake rate of glucose, K_g is a half saturation constant, A is the acetate concentration, and Θ 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

$$r_{up,c}^g(G, A) = k_{\Delta PTS} \frac{G}{G + K_g} \frac{\Theta_a}{A + \Theta_a}, \quad (2)$$

Table 1 Variables and rates in the model of producer-cleaner consortium given by the system (9)–(10). gDW stands for “gram Dry Weight”

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

where $k_{\Delta PTS} < k_g$. An equivalent way of describing $r_{up,c}^g$ is

$$r_{up,c}^g(G, A) = \beta r_{up,p}^g(G, A), \quad (3)$$

where we have introduced the parameter $\beta := \frac{k_{\Delta PTS}}{k_g}$ that lies on the interval $[0, 1]$. From a biological viewpoint, the parameter β reflects the presence of the gene *ptsG*. If $\beta = 1$, the gene *ptsG* is present and both strains take up glucose at the same rate. In the interest of mathematical analysis, we will refer to (3) in place of (2), since the relation between the two glucose uptake rates plays an important role in the analysis of the model. 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 , cells secrete acetate through overflow metabolism (Basan et al. 2015; Wolfe 2005). We note that the maximal glucose uptake rate of cleaners is given by βk_g , and according to parameters from Table 2, $\beta k_g \leq l$. Therefore, cleaners cannot secrete acetate through overflow metabolism. For producers, overflow metabolism is possible, and its rate is given by

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

where k_{over} is a proportionality constant.

In the model of Mauri et al. (2020), a term accounting for overflow metabolism in the cleaners was considered. However, the presence of this term has no impact on the dynamics because the authors use the same parameters as in Table 2. Assuming that cleaners cannot uptake glucose at a rate that supports overflow metabolism is consistent with the main motivation for using the producer-cleaner consortium: cleaners are primarily designed to remove acetate from the medium with low utilization of glucose. Therefore, we omit this term in our current analysis.

In the presence of glucose, *E. coli* cells cannot grow on acetate, a phenomenon known as Carbon Catabolite Repression (CCR) (Wolfe 2005; Kremling et al. 2015). 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

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

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 $r_{up,c}^a$. 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

$$r_{up,c}^a(G, A) = k_a \frac{A}{A + K_a} d(r_{up,c}^g(G, A)) + k_{AcS} \frac{A}{A + K_{AcS}}, \quad (5)$$

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 (2)) and $k_{AcS} = 0$. In this 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

$$\hat{d}(y) = \frac{\Theta_g}{\Theta_g + y}, \quad y = r_{up,p}^g, r_{up,c}^g \quad (6)$$

with Θ_g a constant. We modify the description of the down-regulation function (6) as follows:

$$d(y) = \max \left\{ 0, (1 + a) \frac{\Theta_g}{\Theta_g + y} - a \right\}, \quad y = r_{up,p}^g, r_{up,c}^g \quad (7)$$

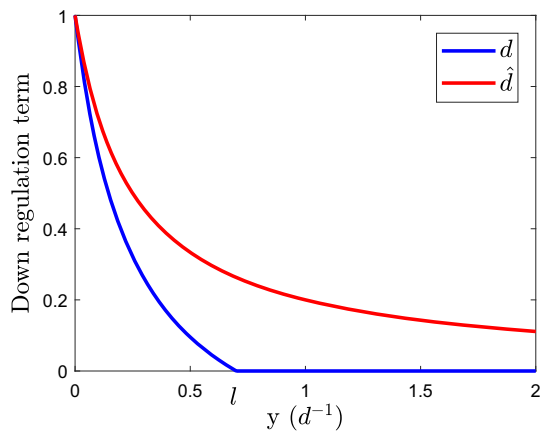
with $a = \Theta_g/l$. The main motivation to describe d as in (7) arises from the following mathematical property:

$$\text{if } y \geq l, \text{ then } d(y) = 0.^1 \quad (8)$$

¹ 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 (7) is that it facilitates the mathematical analysis of the model. Figure 2 compares the two alternatives of d . In Appendix 1, we show that our optimization results are insensitive to this model change.

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

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



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 yields $Y_{h/p}$ and $Y_{h/c}$. These parameters take values between 0 and 1, and dictate the distribution of biomass toward either population growth or production of proteins. Thus, $Y_{h/p} = 1$ means zero bacterial growth and $Y_{h/p} = 0$ means maximal growth supported by the medium. In the model of Mauri et al. (2020), cleaners cannot produce recombinant proteins (i.e., $Y_{h/c} = 0$) and they mainly serve to reduce the acetate concentration in the medium. In this new model, the carbon not utilized by the producer and secreted in the form of acetate can be recovered by cleaners and transformed into recombinant protein.

The product yield of a suitably engineered bacterial strain can be tuned by changing the expression of the recombinant protein. This may involve a modification of the promoter of the gene encoding the heterologous protein, or a change in its activity through a change in the concentration of a chemical inducer (Izard et al. 2015) or through optogenetic control (Miliás-Argeitis et al. 2016). As discussed by Mauri et al. (2020), the coexistence of both species depend on the choice of $Y_{h/p}$.

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. The yield coefficients Y_g and Y_a characterize the conversion of glucose and acetate, respectively, into biomass. Thus, $Y_g r_{up,p}^g$ and $Y_a r_{up,p}^a$ correspond to synthesis of new biomass from the uptake of glucose and acetate, respectively, and $Y_a r_{over,p}^a$ represents the quote of biomass that is not synthesized in response to acetate uptake due to acetate overflow. Similarly, the specific rate of cleaner biomass production per unit of biomass is given by

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

The biomass of both strains decays with a specific rate k_{deg} . This same rate applies to the recombinant proteins, given that normally expressed proteins form the major component of bacterial biomass (Bremer and Dennis 2013). Biomass degradation is introduced to account for non-growth-related maintenance costs (Mauri et al. 2020; Pirt 1965). That is, substrate uptake must exceed the maintenance costs for net growth to occur.

Considering that the chemostat is operated at a dilution rate D with glucose concentration G_{in} in the inflow, the evolution of B_p , B_c , and G and A is given by

$$\begin{aligned}\frac{dB_p}{dt} &= (1 - Y_{h/p})f_p(G, A)B_p - k_{deg}B_p - DB_p, \\ \frac{dB_c}{dt} &= (1 - Y_{h/c})f_c(G, A)B_c - k_{deg}B_c - DB_c, \\ \frac{dG}{dt} &= D(G_{in} - G) - r_{up,p}^g(G, A)B_p - r_{up,c}^g(G, A)B_c, \\ \frac{dA}{dt} &= -DA + [r_{over,p}^a(G, A) - r_{up,p}^a(G, A)]B_p - r_{up,c}^a(G, A)B_c,\end{aligned}\tag{9}$$

and the evolution of the recombinant protein concentrations is given by

$$\begin{aligned}\frac{dH_p}{dt} &= Y_{h/p}f_p(G, A)B_p - k_{deg}H_p - DH_p, \\ \frac{dH_c}{dt} &= Y_{h/c}f_c(G, A)B_c - k_{deg}H_c - DH_c.\end{aligned}\tag{10}$$

System (9)–(10) 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 (6), then (9)–(10) is equivalent to the model of Mauri et al. (2020). Note that the model has been intentionally split into two subsystems, (9) and (10), such that (9) is uncoupled from (10). To establish the existence of coexistence equilibria, we only need to study system (9), but for optimizing the production of recombinant proteins, we must study both.

Figure 3 shows the steady-state solution of (9) 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 (9) admits at most three non-trivial equilibria:

- (I) An equilibrium with the presence of producers but absence of cleaners (see Fig. 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, the acetate concentration is zero.
- (III) A coexistence equilibrium (see Fig. 3C and D).

Note that, at a coexistence equilibrium, we have that $(1 - Y_{h/p})f_p(G, A) - k_{deg} = D$ and $(1 - Y_{h/c})f_c(G, A) - k_{deg} = D$. We observe the existence of an interval of

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

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

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.

3 Coexistence at steady states

In this section, we characterize the existence of coexistence equilibria of (9). 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

$$Y_g - k_{over}Y_a > 0, \quad (11)$$

which is verified by the parameters in Table 2. This 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

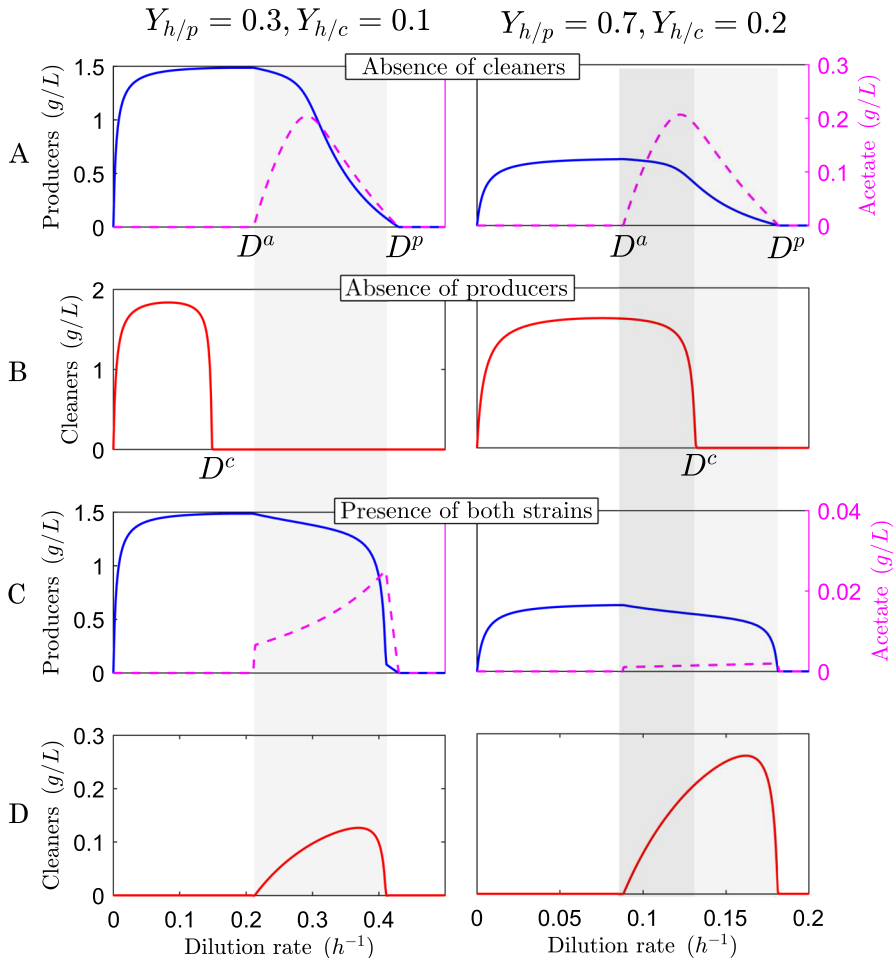


Fig. 3 Steady-state solutions of (9) obtained by simulation for different values of the dilution rate and two different pairs of values for $Y_{h/p}$ and $Y_{h/c}$. D^a is the dilution rate above which acetate overflow occurs when producers grow individually, respectively. D^p and D^c are the dilution 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 1 and 2. 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 higher than or equal to D^c . Note that the bacterial concentrations are shown on the left-hand side and acetate concentrations (dashed lines) are shown on the right-hand side. **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)

and no cleaners (Proposition 1). We also assume that

$$r_{up,p}^g(G_{in}, 0) > l, \quad (12)$$

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 (12) 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 (9) is reduced to that of a classical chemostat model with two competitors for one substrate (glucose) (Smith and Waltman 1995). Note that assumptions (11) and (12) have been used by Martínez and Gouzé (2021) to study a simplified version of (9) without cleaners.

The following proposition characterizes the existence of steady states without cleaners.

Proposition 1 (Producer steady state) *Assume that (11) and (12) 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 (9) admits a unique equilibrium with the presence of producers and absence of cleaners, denoted by $E^p = (B^p, 0, G^p, A^p)$. Moreover, for $D^a = (1 - Y_{h/p})Y_{gl} - k_{deg}$ we have:*

(I) *If $D > D^a$, then $A^p > 0$ and $r_{up,p}^g(G^p, A^p) \geq l$.*

(II) *If $D \leq D^a$, then $A^p = 0$ and $r_{up,p}^g(G^p, 0) \leq l$.*

(b) *If $D^p \leq D$, then (9) has no equilibrium with the presence of producers and absence of cleaners.*

Proof The proof follows directly from Proposition 4 in Appendix 1. Note that in Appendix 1, we write r_g instead of $r_{up,p}^g$. $\square$

Remark 1 (Acetate secretion) Proposition 1 shows the existence of a threshold dilution rate characterizing the presence of acetate in the medium, denoted by D^a (see Fig. 3A). As we will see later, a coexistence equilibrium for (9) is only possible if the dilution rate is higher than D^a . Proposition 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 2 (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 $D > D^a$, acetate secretion occurs at equilibrium, which means that overflow metabolism is active. In this regime, cells are growing fast on glucose. Because of carbon catabolite repression, in this regime, no acetate is consumed (see (8)).

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

(a) *If $D^c > D$, then (9) admits a unique equilibrium with the presence of cleaners and absence of producers, denoted by $E^c = (0, B^c, G^c, 0)$.*

(b) *If $D^c \leq D$, then (9) has no equilibrium with the presence of cleaners.*

Proof The proof follows directly from Proposition 5 in Appendix 1. Note that in Appendix 1, we write r_g instead of $r_{up,p}^g$. $\square$

Remark 3 (Cleaner washout) Proposition 2 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 following proposition states some basic properties of any coexistence equilibrium.

Proposition 3 Let D^a , D^p , G^c , and A^c be given by Proposition 1. If (9) admits a coexistence equilibrium (B_p^*, B_c^*, G^*, A^*) , then:

- (a) $D \leq D^a$ implies $A^* = 0$ and $r_{up,p}^g(G^*, 0) \leq l$.
- (b) $D^a < D < D^p$ implies $0 < A^* < A^p$ and $r_{up,p}^g(G^*, A^*) > l$.

Proof The proof follows directly from Proposition 6 in Appendix 1. $\square$

Remark 4 (Acetate reduction under coexistence) Proposition 3 shows that the acetate concentration is lower at a coexistence equilibrium than at a non-coexistence equilibrium with producers. Let us fix the dilution rate at a value between D^a and D^p . From Proposition 1, we know that producers will secrete acetate at equilibrium when grown without cleaners. Part (b) of Proposition 3 states that if cleaners coexist with producers, then the acetate concentration at equilibrium will be lower. This situation can be observed in Fig. 3. When both species grow together, acetate concentration at equilibrium is lower than in a monoculture of producers.

The main result of this section characterizes the existence of coexistence steady states for (9). This characterization is based on the positive steady states of the cleaners and producers growing separately.

Theorem 1 (Coexistence steady states) Assume that (11) and (12) hold. Let D^a and D^p be given by Proposition 1, and let D^c be given by Proposition 2.

- (a) If $D \leq D^a$ then (9) admits a coexistence equilibrium, if and only if

$$(1 - Y_{h/p}) = \beta(1 - Y_{h/c}).$$

- (b) If $D^a < D < D^p$ and $D < D^c$ then (9) admits a (unique) coexistence equilibrium, 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 1, and G^c defined in Proposition 2.

- (c) If $D^a < D < D^p$ and $D \geq D^c$ then (9) admits a (unique) coexistence equilibrium, 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 1.

- (d) If $D \geq D^p$, then (9) has no coexistence equilibrium.

Proof The proof follows from the results presented in Appendix 1. Parts (a) and (d) follows directly from Lemma 4 (see notation in (17)) and Lemma 5, respectively. For parts (b) and (c), note that according to the notation of Appendix 1, we have

$$\begin{aligned} \mu_p(0, r_g(G^c, 0)) &= (1 - Y_{h/p})f_p(G^c, 0) - k_{deg}, \\ \mu_c(0, \gamma) &= (1 - Y_{h/c})f_c(G^p, A^p) - k_{deg}. \end{aligned}$$

Thus, (b) and (c) follow directly from Proposition 7. $\square$

Remark 5 (*Competition and syntrophy*) Note that the difference between cases (b) and (c) in Theorem 1 is the value of the dilution rate with respect to D^c . If $D < D^c$ (case (b)), according to Proposition 2, 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 proto-cooperation (Roell et al. 2019), or harmful through glucose competition. This can be observed in Fig. 3B and D (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 (proto-cooperation). 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.

Remark 6 (*Coexistence in absence of acetate*) Part (a) in Theorem 1 is an unlikely situation in which parameters are exactly balanced to allow coexistence. From Proposition 3, if $D \leq D^a$, producers do not secrete acetate at coexistence equilibrium. Therefore, both species compete for a single substrate (glucose), and coexistence is possible if both species admit the same break-even concentration (Smith and Waltman 1995). Although unlikely, such coexistence with perfectly balanced parameters has been observed in chemostat experiments (Hansen and Hubbell 1980). However, in our study, such parameter values are not of interest because cleaners lose their property of “cleaning” (removing acetate) and only consume glucose.

Remark 7 (*Theorem notation*) The concentration G^c is defined when $D < D^c$ (see Proposition 2). We can extend the definition of G^c as follows:

$$G^c = \begin{cases} \text{Given by Proposition 2, if } D < D^c \\ G_{in}, & \text{if } D \geq D^c. \end{cases} \quad (13)$$

Thus, we can replace parts (b) and (c) in Theorem 1 by the following statement:

If $D^a < D < D^p$ then (9) admits a (unique) coexistence equilibrium, 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 1, and G^c defined in Proposition 2.

This notation simplifies the application of Theorem 1 and is the notation that we will use to determine numerically the coexistence equilibria in the next section.

The following lemma states how to determine the protein concentrations associated with the coexistence equilibrium given by Theorem 1.

Lemma 1 *Let (B_p^*, B_c^*, G^*, A^*) be a coexistence equilibrium of (9). 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 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. $\square$

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 1 for the existence of a coexistence equilibrium allow for a proper study of the optimization of (9)–(10) at steady state. While the statement of Theorem 1 gives necessary and sufficient conditions for coexistence, its proof indicates how to construct an efficient algorithm for determining numerically the coexistence equilibria (see Appendix 1). 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 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, we consider the feasible region $\Omega \subset \mathbb{R}_+^3$, such that for any $v := (Y_{h/p}, Y_{h/c}, D) \in \Omega$, (9) admits a coexistence equilibrium. Since $Y_{h/p}, Y_{h/c} \in [0, 1]$

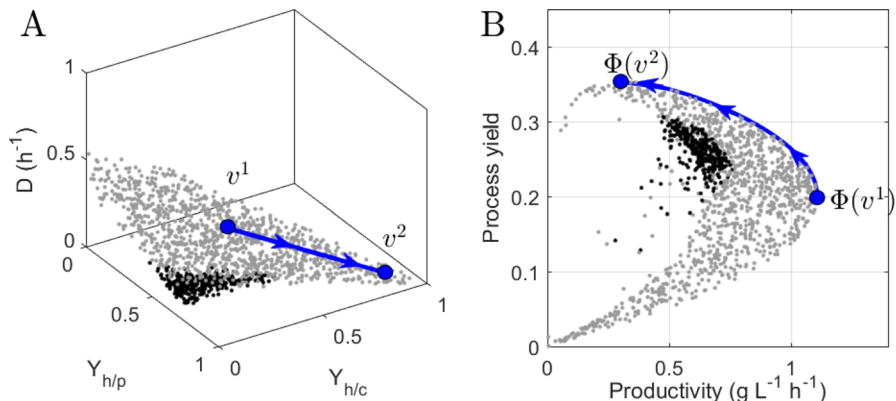


Fig. 4 **A** Region Ω represented by 3000 points. Gray points correspond to $D \geq D^c$ and black points correspond to $D < D^c$ (see Remark 5). The point $v^1 = (0.51, 0.34, 0.28)$ maximizes productivity (first component of Φ) and $v^2 = (0.91, 0.88, 0.04)$ maximizes the process yield (second component of Φ). The blue curve is such that its image through Φ is the POF. **B** Function Φ evaluated at the points represented in A. The blue curve is the POF

and coexistence is only possible for $D < D^p$ (see Theorem 1), we have that Ω is bounded. The region Ω can be determined by the conditions given by Theorem 1. Indeed, using the parameters from Table 2, Ω looks like a cone with a vertex close to $(1, 1, 0)$ (see Fig. 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 1). 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:

$$\begin{aligned} &\text{maximize } \Phi(v) = \left(DH^*(v), \frac{H^*(v)}{G_{in}} \right) \\ &\text{subject to } v \in \Omega. \end{aligned} \quad (14)$$

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 Ω and $\Phi(\Omega)$ plays an important role in the success of the employed numerical method to solve (14). Pareto curves cannot be computed efficiently in many cases, especially in the non-convex case where methods such as ϵ -constraint are necessary (Gunantara 2018). Figure 4 reveals a favorable structure of Ω and $\Phi(\Omega)$ (scatter plots) for applying the weighting method (Miettinen 2012). Accordingly, the problem

(14) is transformed into the following so-called weighting problem:

$$\begin{aligned} & \text{maximize } \lambda D H^*(v) + (1 - \lambda) \frac{H^*(v)}{G_{in}}, \\ & \text{subject to } v \in \Omega, \end{aligned} \quad (15)$$

where $\lambda \in [0, 1]$. It is well known that if Ω 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 (15) (see Corollary 3.1.8 by Miettinen (2012)).

For any $\lambda \in [0, 1]$, problem (15) 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 should be defined through non-strict inequalities and the objective function must be continuous on this set. For ease of reading, we discuss such technical details in Appendix 1.

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

- (a) The weighting method accurately returns the POF.
- (b) 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.
- (c) Along the POF, the process yield decreases as the productivity increases.

Most importantly, regarding the syntrophy-competition relationship between both strains, the image Φ 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 5).

4.1 Impact of G_{in}

Figure 5 shows the effects of varying the input glucose concentration G_{in} . According to Fig. 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 (Φ_2), while the maximal productivity increases almost linearly with G_{in} (Φ_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} .

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

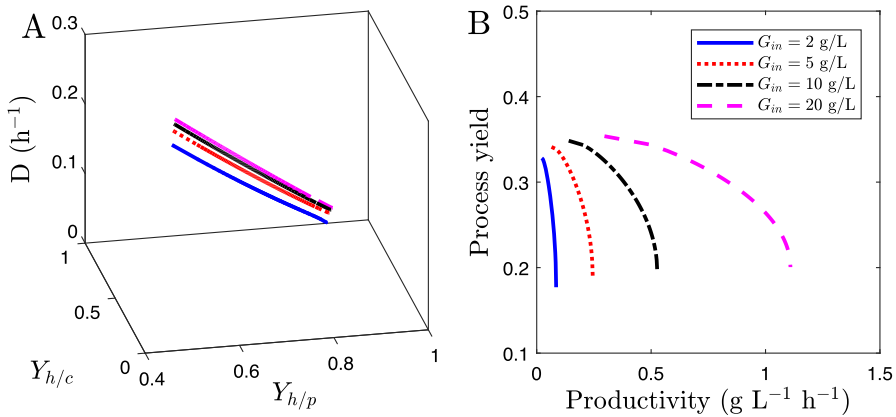


Fig. 5 Influence of G_{in} on the POF. **A** Decision variables which image through Φ is a POF. **B** POF for different values of G_{in}

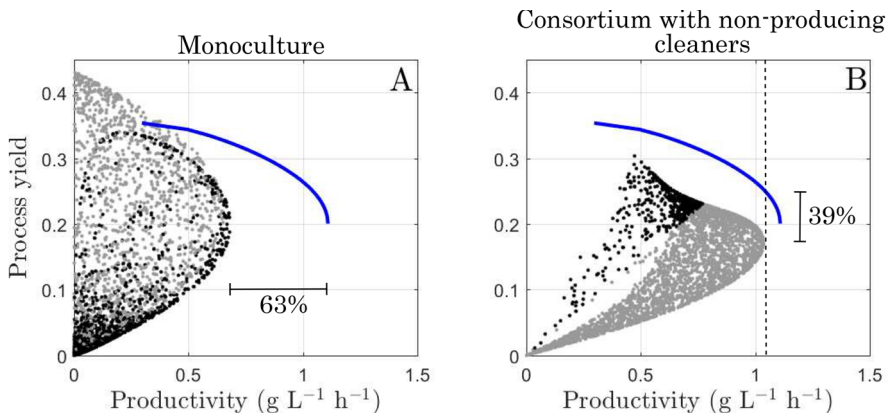


Fig. 6 Scatter plots showing the performance of two alternative production methods, compared with Pareto-optimal performance of the consortium with producers and cleaners both synthesizing the target protein. **A** Monoculture of producers: scatter plot of $(DH_p^p, H_p^p/G_{in})$, with H_p^p the protein concentration associated with the non-trivial equilibrium given by Proposition 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 Fig. 4. **B** Consortium of producers and non-producing cleaners: scatter plot of Φ 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 Fig. 4. Black points correspond to $D < D^c$ and gray points correspond to $D \geq D^c$ (color figure online)

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 (9)–(10) 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 Fig. 4 with a scatter plot of the feasible objective region $\Phi(\Omega)$ restricted to $Y_{h/c} = 0$ (see Fig. 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 $1.04 \text{ g L}^{-1} \text{ h}^{-1}$, 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 almost impossible because of the competitive exclusion principle (see Remark 6).

A 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 a 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 1, ensure the success of each strain as an invader. For example, in case (b) of Theorem 1, each strain reaches a non-trivial equilibrium when playing the role of resident. If the two inequalities defining the necessary and sufficient conditions are satisfied, then both strains succeed as an invader.

The predominant inter-species interaction during coexistence can be competition, proto-cooperation, or syntrophy (Roell et al. 2019). From Theorem 1, 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 1, 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 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 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 Fig. 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 Fig. 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).

Data availability statement Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

A Proofs

A.1 Change of notation

In this appendix, we present the proofs of all the statements from Sect. 3, and we present a series of technical results needed for their proofs. We begin by simplifying the notations presented in Sect. 2 and recalling the hypotheses that are necessary for the proofs.

To simplify the notation of uptake rates, we denote $r_{up,p}^g$ by r_g and we define

$$r_a(A, r_g) := \begin{cases} -k_{over}(r_g - l), & \text{if } r_g > l, \\ k_a \frac{A}{A+k_a} d(r_g), & \text{if } r_g \leq l. \end{cases} \quad (16)$$

Using property (8) and the definition of $r_{over,p}^a$, we have that

$$r_a(A, r_g(G, A)) = r_{up,p}^a(G, A) + r_{over,p}^a(G, A).$$

Thus, r_a represents an acetate flux between producers and the medium, which can be negative or positive depending on the sign of $r_g - l$. We also define

$$\rho_a(A, r_g) := k_a \frac{A}{A + K_a} d(\beta r_g) + k_{Acs} \frac{A}{A + K_{Acs}}.$$

It is important to note that $\rho_a(0, r_g) = 0$ and $\rho_a(A, r_g) \geq 0$ for all $A, r_g \geq 0$, and that $A \mapsto \rho_a(A, r_g)$ is strictly increasing. The following parameters simplify the

notations of product yields:

$$\begin{aligned} Y_{g/p} &:= (1 - Y_{h/p})Y_g, & Y_{a/p} &:= (1 - Y_{h/p})Y_a, \\ Y_{g/c} &:= (1 - Y_{h/c})Y_g, & Y_{a/c} &:= (1 - Y_{h/c})Y_a. \end{aligned} \quad (17)$$

Finally, to simplify the notation of (9), we define

$$\begin{aligned} \mu_p(A, r_g) &:= Y_{g/p}r_g + Y_{a/p}r_a(A, r_g) - k_{deg}, \\ \mu_c(A, r_g) &:= Y_{g/c}\beta r_g + Y_{a/c}\rho_a(A, r_g) - k_{deg}. \end{aligned} \quad (18)$$

The system (9) can be rewritten as follows:

$$\begin{aligned} \frac{dB_p}{dt} &= [\mu_p(A, r_g(G, A)) - D]B_p, \\ \frac{dB_c}{dt} &= [\mu_c(A, r_g(G, A)) - D]B_c, \\ \frac{dG}{dt} &= D(G_{in} - G) - r_g(G, A)B_p - \beta r_g(G, A)B_c, \\ \frac{dA}{dt} &= -DA - r_a(A, r_g(G, A))B_p - \rho_a(A, r_g(G, A))B_c, \end{aligned} \quad (19)$$

This new notation is a compromise between a simple notation to address the proofs and a notation that is closed to that presented in Sect. 2. Note that we have defined μ_p , μ_c , r_a , and ρ_a in terms of A and r_g . The dilution rates defined in Propositions 1 and 2 can be rewritten as:

$$\begin{aligned} D^p &= \mu_p(0, r_g(G_{in}, 0)), \\ D^c &= \mu_c(0, r_g(G_{in}, 0)), \\ D^a &= Y_{g/p}l - k_{deg}. \end{aligned} \quad (20)$$

We also define the following constant:

$$\alpha := Y_{g/p} - k_{over}Y_{a/p}. \quad (21)$$

We recall the hypotheses of the model in terms of the new notation. Hypothesis (11) is equivalent to

$$\alpha > 0, \quad (22)$$

and Hypothesis (12) is equivalent to

$$r_g(G_{in}, 0) > l. \quad (23)$$

From now on, we assume that (22) and (23) are true.

Our main task is to prove Theorem 1, that is, to find necessary and sufficient conditions for the existence of a coexistence equilibrium of (19). Therefore, we must study the existence and uniqueness of solutions to

$$\begin{aligned} 0 &= \mu_p(A, r_g(G, A)) - D, \\ 0 &= \mu_c(A, r_g(G, A)) - D, \\ 0 &= D(G_{in} - G) - r_g(G, A)B_p - \beta r_g(G, A)B_c, \\ 0 &= -DA - r_a(A, r_g(G, A))B_p - \rho_a(A, r_g(G, A))B_c, \end{aligned} \quad (24)$$

with $B_p, B_c > 0$ and $G, A \geq 0$.

A.2 Results

Our first lemma establishes some properties of any non-negative solution to (24) with $B_p > 0$. These properties will be repeatedly used along the appendix.

Lemma 2 *Let $D > 0$, and let B_p, B_c, A , and r_g be such that*

$$\begin{aligned} \mu_p(A, r_g) - D &= 0, \\ -DA - r_a(A, r_g)B_p - \rho_a(A, r_g)B_c &= 0, \end{aligned} \quad (25)$$

with $B_p > 0$ and $B_c, A, G, r_g \geq 0$. Let D^a and α be given by (20) and (21), respectively, and let us define

$$\gamma := l + \frac{1}{\alpha}(D - D^a). \quad (26)$$

The following statements are equivalent:

- (a) $A > 0$,
- (b) $r_a(A, r_g) < 0$,
- (c) $r_g > l$,
- (d) $\mu_p(A, r_g) = \alpha(r_g - l) + D^a$ and $r_g > l$,
- (e) $r_g = \gamma$ and $\gamma > l$,
- (f) $D > D^a$.

Proof If (a) holds, then $DA > 0$ and $\rho_a(A, r_g)B_c \geq 0$. From the second equation in (25), we obtain that $r_a(A, r_g)B_p < 0$. Since $B_p > 0$, we conclude that $r_a(A, r_g) < 0$, and therefore (b) holds. If (b) holds, then r_g cannot be equal or lower than l , otherwise otherwise $r_a(A, r_g)$ would be non-negative. Thus, (c) holds. Now, if (c) holds, by definition, $r_a(A, r_g) = -k_{over}(r_g - l)$. Using the definition of μ_p , we obtain

$$\mu_p(A, r_g) = Y_{g/p}r_g - Y_{a/p}k_{over}[r_g - l] - k_{deg},$$

which is equivalent to

$$\mu_p(A, r_g) = \alpha[r_g - l] + D^a,$$

and (d) holds. If (d) holds, from the first equation in (25), we obtain that

$$\alpha[r_g - l] + D^a = D,$$

from where $r_g = \gamma$, and (e) holds. Now, if (e) holds, we have that $\gamma > l$, or equivalently

$$l + \frac{1}{\alpha}(D - D^a) > l,$$

from where $D > D^a$, and (f) holds. We prove the last part indirectly, that is, we prove that if (a) is false, then (f) is also false. Indeed, if $A = 0$, then $\rho_a(A, r_g) = 0$, and from the second equation in (25), we conclude that $r_a(A, r_g) = 0$. Thus, necessarily $r_g \leq l$, otherwise $r_a(A, r_g)$ is negative. Consequently, by definition of μ_p ,

$$\mu_p(A, r_g) = Y_{g/p}r_g - k_{deg} \leq Y_{g/p}l - k_{deg} = D^a.g$$

Combining the previous equation with the first equation in (25), we conclude that $D \leq D^a$. This completes the proof. $\square$

The following lemma states some properties related to D^p .

Lemma 3 *Let γ be given by (26) and let D^a and D^p be given by (20). We have:*

- (a) $D^p = \alpha[r_g(G_{in}, 0) - l] + D^a$.
- (b) $D^p > D^a$.
- (c) *If $D < D^p$, then $\gamma < r_g(G_{in}, 0)$.*

Proof From (23), we have that $r_g(G_{in}, 0) > l$. Then, we have that

$$r_a(0, r_g(G_{in}, 0)) = -k_{over}(r_g(G_{in}, 0) - l).$$

From the previous equation and the definitions of D^p (see (20)) and μ_p , we obtain that

$$D^p = \alpha[r_g(G_{in}, 0) - l] + D^a, \quad (27)$$

which proves (a) and implies

$$D^p - D^a = \alpha[r_g(G_{in}, 0) - l] > 0,$$

and (b) is proved. For part (c), if $D < D^p$, using (27) and the definition of γ , we have that

$$\gamma = l + \frac{1}{\alpha}(D - D^a) < l + \frac{1}{\alpha}(D^p - D^a) = l + \frac{\alpha[r_g(G_{in}, 0) - l]}{\alpha} = r_g(G_{in}, 0).$$

This completes the proof. $\square$

The proof of Proposition 1 in Sect. 3 follows directly from the following proposition.

Proposition 4 *Let D^a and D^p be given by (20). We have:*

- (a) *If $D^p > D$, then (19) admits a unique equilibrium of the form $E^p = (B_p^p, 0, G^p, A^p)$, with $B_p^p > 0$. Moreover:*
- (I) *If $D > D^a$, then $A^p > 0$ and $r_g(G^p, A^p) > l$.*
- (II) *If $D \leq D^a$, then $A^p = 0$ and $r_g(G^p, 0) \leq l$.*
- (b) *If $D^p \leq D$, then (19) has no equilibrium with $B_p > 0$ and $B_c = 0$.*

Proof The equilibria of (19) with $B_c = 0$ and $B_p > 0$ are given by the solutions of the following system

$$\begin{aligned} 0 &= \mu_p(A, r_g(G, A)) - D, \\ 0 &= D(G_{in} - G) - r_g(G, A)B_p, \\ 0 &= -DA - r_a(A, r_g(G, A))B_p. \end{aligned} \quad (28)$$

From the equivalence among (a), (c) and (f) in Lemma 2, we have that

- (1) If $D > D^a$, then any solution of (28) satisfies $A > 0$ and $r_g(G, A) > l$.
- (2) If $D \leq D^a$, then any solution of (28) satisfies $A = 0$ and $r_g(G, A) \leq l$.

If statement (1) holds, using Lemma 2 part (d), we obtain that (28) is equivalent to

$$\begin{aligned} 0 &= \alpha(r_g(G, A) - l) + D^a - D, \\ 0 &= D(G_{in} - G) - r_g(G, A)B_p, \\ 0 &= -DA + k_{over}[r_g(G, A) - l]B_p. \end{aligned} \quad (29)$$

From the first equation in (29) (or Lemma 2 part (e)), we have that $r_g(G, A) = \gamma$, with γ defined by (26). Thus, from the second and third equations in (29), we have

$$G = G_{in} - \beta_g B_p, \quad \beta_g = \frac{\gamma}{D}, \quad (30)$$

$$A = \beta_a B_p, \quad \beta_a = \frac{k_{over}(\gamma - l)}{D}. \quad (31)$$

Replacing (30) and (31) in the first equation of (29), we obtain the following equation for B_p :

$$\underbrace{\alpha[r_g(G_{in} - \beta_g B_p, \beta_a B_p) - l] + D^a - D}_{g(B_p)} = 0 \quad (32)$$

From the monotonicity of r_g , we have that g is strictly decreasing. Since $g(0) = D^p - D$ (see Lemma 3 part (a)), and $g(G_{in}/\beta_g) = D^a - D < 0$, we conclude that (32) 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 statement (2) holds, after replacing A by 0, we obtain that (28) is equivalent to

$$\begin{aligned} 0 &= Y_{g/p}r_g(G, 0) + Y_{a/p}r_a(0, r_g(G, 0)) - k_{deg} - D, \\ 0 &= D(G_{in} - G) - r_g(G, 0)B_p, \\ 0 &= -r_a(0, r_g(G, 0))B_p. \end{aligned} \quad (33)$$

From the third equation in (33), we obtain that $r_a(G, A) = 0$. Thus, (33) is reduced to:

$$\begin{aligned} 0 &= \underbrace{Y_{g/p}r_g(G, 0) - k_{deg} - D}_{h(G)}, \\ 0 &= D(G_{in} - G) - r_g(G, 0)B_p. \end{aligned} \quad (34)$$

From the second equation in (34), we have that $G \in [0, G_{in})$, otherwise $B_p \leq 0$. Now, since $G \mapsto h(G)$ is strictly increasing and $h(0) = -k_{deg} - D < 0$, (34) admits a unique positive equilibrium if and only if $h(G_{in}) > 0$. We have

$$\begin{aligned} h(G_{in}) &= Y_{g/p}r_g(G_{in}, 0) - k_{deg} - D, \quad (\text{use (23)}) \\ &> Y_{g/p}l - k_{deg} - D, \quad (\text{use definition of } D^a) \\ &= D^a - D, \quad (\text{apply statement (2)}) \\ &\geq 0. \end{aligned}$$

Therefore, $h(G_{in}) > 0$ and the proof is completed. $\square$

The following proposition is equivalent to Proposition 2 in Sect. 3.

Proposition 5 (Cleaner steady state) *Let D^c be given by (20). We have:*

- (a) *If $D^c > D$, then (19) admits a unique equilibrium with $B_c > 0$ and $B_p = 0$, which is of the form $E^c = (0, B_c^c, G^c, 0)$.*
- (b) *If $D^c \leq D$, then (19) has no equilibrium with $B_c > 0$ and $B_p = 0$.*

Proof Any equilibrium with $B_c > 0$ and $B_p = 0$ is a solution of (replace B_p by 0 in (24)):

$$\begin{aligned} 0 &= \mu_c(A, r_g(G, A)) - D, \\ 0 &= D(G_{in} - G) - \beta r_g(G, A)B_c, \\ 0 &= -DA - \rho_a(A, r_g(G, A))B_c. \end{aligned} \quad (35)$$

From the third equation in (35), we have that $A = 0$, otherwise we have an equality between a positive and a negative number. Using the definition of μ_c and the fact that $\rho_a(0, r_g) = 0$, (35) can be rewritten as

$$0 = \underbrace{Y_{g/c}\beta r_g(G, 0) - k_{deg} - D}_{h(G)}, \quad (36)$$

$$0 = D(G_{in} - G) - \beta r_g(G, 0)B_c.$$

From the second equation in (36), we have that $G \in [0, G_{in})$, otherwise $B_c \leq 0$. Now, since h is strictly increasing and $h(0) = -k_{deg} - D < 0$, we have that (36) admits a solution, which is unique, if and only if $h(G_{in}) > 0$. The rest of the proof follows from noting that $h(G_{in}) = D^c - D$. $\square$

The following lemma shows that if $D \leq D^a$, then there is a coexistence equilibrium if and only if some parameters are perfectly balanced. Note that this lemma does not ensure the uniqueness of a coexistence equilibrium.

Lemma 4 *Let D^a and D^p be given by (20). If $D \leq D^a$, then (19) admits a coexistence equilibrium if, and only if*

$$Y_{g/p} = \beta Y_{g/c}.$$

Proof Since $D \leq D^a$, from the equivalence among the statements (a), (c), and (f) in Lemma 2, we have that for any solution to (24) with $B_p > 0$, $A = 0$ and $r_g(G, 0) \leq l$. Therefore, $r_a(0, r_g(G, 0)) = \rho_a(0, r_g(G, 0)) = 0$ and (24) is equivalent to

$$\begin{aligned} 0 &= Y_{g/p}r_g(G, 0) - k_{deg} - D, \\ 0 &= Y_{g/c}\beta r_g(G, 0) - k_{deg} - D, \\ 0 &= D(G_{in} - G) - r_g(G, 0)B_p - \beta r_g(G, 0)B_c. \end{aligned} \quad (37)$$

If $Y_{g/p} \neq \beta Y_{g/c}$, then the first and second equation cannot be satisfied at the same time. Thus, there cannot be a coexistence equilibrium. On the other hand, if $Y_{g/p} = \beta Y_{g/c}$, then G is a solution of $h(G) = 0$ with $h(G) := Y_{g/p}r_g(G, 0) - k_{deg} - D$. Note that h is strictly increasing, $h(0) = -k_{deg} - D < 0$, and

$$\begin{aligned} h(G_{in}) &= Y_{g/p}r_g(G_{in}, 0) - k_{deg} - D, \quad (\text{use (23)}) \\ &> Y_{g/p}l - k_{deg} - D, \quad (\text{use definition of } D^a) \\ &= D^a - D, \quad (\text{apply assumption on } D) \\ &\geq 0. \end{aligned}$$

Thus $h(G) = 0$ admits a unique solution $G^* \in (0, G_{in})$. Replacing G by G^* in the third equation in (37), we obtain the existence of infinity coexistence equilibria. $\square$

The following lemma shows that if the dilution rate is too high then coexistence is impossible.

Lemma 5 *Let D^p be given by (20). If $D \geq D^p$, then (19) has no coexistence equilibrium.*

Proof We prove this by contradiction. Let us assume that $D \geq D^p$ and that (19) admits a coexistence equilibrium (B_p^*, B_c^*, G^*, A^*) . Since $D^p > D^a$ (see Lemma 3 part (b)), we can use the equivalence between (d) and (f) from Lemma 2 to conclude that

$$\mu_p(A^*, r_g(G^*, A^*)) = \alpha[r_g(G^*, A^*) - l] + D^a. \quad (38)$$

From Lemma 2 part (a), we obtain that

$$D^p = \alpha[r_g(G_{in}, 0) - l] + D^a. \quad (39)$$

From the third equation in (24), we have that $G^* < G_{in}$. Using the monotonicity of r_g , we can combine (38) and (39) to obtain $\mu_p(A^*, r_g(G^*, A^*)) < D^p$. Finally, from the first equation in (24), we conclude that $D < D^p$, which contradicts our initial hypothesis ($D \geq D^p$). This completes the proof. $\square$

The following establishes some necessary conditions for coexistence.

Lemma 6 *Let D^a , D^p , and D^c be given by (20) and let G^c be given by Proposition 5. Let us assume that $D^a < D < D^p$ and $D < D^c$. If (19) admits a coexistence equilibrium, then*

$$\mu_p(0, r_g(G^c, 0)) > D. \quad (40)$$

Proof Let us assume that (19) admits a coexistence equilibrium (B_p^*, B_c^*, G^*, A^*) . Since $D > D^a$, from Lemma 2 part (e), we know that $r_g(G^*, A^*) = \gamma$ with γ defined by (26). Now, from the second equation in (24) and the definition of G^c we have

$$\begin{aligned} \mu_c(0, \gamma) &= D, \\ \mu_c(0, r_g(G^c, 0)) &= D, \end{aligned}$$

which is equivalent to

$$\begin{aligned} Y_{g/c}\beta\gamma + Y_{a/c}\rho_a(A^*, \gamma) &= k_{deg} + D \\ Y_{g/c}\beta r_g(G^c, 0) &= k_{deg} + D. \end{aligned} \quad (41)$$

Since $D > D^a$, from Lemma 2, we have that $A^* > 0$, hence $\rho_a(A^*, \gamma) > 0$. Consequently, from (41), we conclude that

$$r_g(G^c, 0) > \gamma = l + \frac{1}{\alpha}(D - D^a), \quad (42)$$

which implies

$$\alpha(r_g(G^c, 0) - l) + D^a > D. \quad (43)$$

Since $r_g(G^c, 0) > l$ (see (42)), from the definitions of r_a and μ_p , we conclude that

$$\mu_p(0, r_g(G^c, 0)) = \alpha(r_g(G^c, 0) - l) + D^a.$$

Combining the previous equation with (43), we conclude that (40) holds and the proof is completed. $\square$

The following lemma gives upper bounds for any coexistence equilibrium.

Lemma 7 *Let D^a and D^p be given by (20). Let us assume that $D^a < D < D^p$. Then, any solution of (24) with $B_p, B_c > 0$ satisfies*

$$0 < A < A^p \text{ and } 0 < G < G^p, \quad (44)$$

with A^p and G^p given by Proposition 4.

Proof Let us assume that (19) admits a solution (B_p^*, B_c^*, G^*, A^*) . Since $D > D^a$, we have that $\gamma = r_g(G^*, A^*) = r_g(G^p, A^p)$ (see Lemma 2). Thus, we have the following equations:

$$\begin{aligned} 0 &= D(G_{in} - G^p) - \gamma B_p^p, \\ 0 &= -DA^p - r_a(A^p, \gamma) B_p^p, \end{aligned} \quad (45)$$

and

$$\begin{aligned} 0 &= D(G_{in} - G^*) - \gamma B_p^* - \beta \gamma B_c^*, \\ 0 &= -DA^* - r_a(A^*, \gamma) B_p^* - \rho_a(A^*, \gamma) B_c^*. \end{aligned} \quad (46)$$

We prove that $A^* < A^p$ by contradiction. Let us assume that $A^* \geq A^p$. Using the monotonicity of r_g and the fact that $r_g(G^p, A^p) = r_g(G^*, A^*)$, we obtain that $G^p \leq G^*$. From the first equation in (46) and the first equation in (45), we obtain that:

$$D(G_{in} - G^p) - \gamma B_p^p = 0 < \beta \gamma B_c^* = D(G_{in} - G^*) - \gamma B_p^*,$$

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

$$B_p^p > B_p^*. \quad (47)$$

From the second equation in (46) and the second equation in (45), we obtain that:

$$-DA^p - r_a(A^p, \gamma) B_p^p < \rho_a(A^*, \gamma) B_c^* = -DA^* - r_a(G^*, A^*) B_p^*. \quad (48)$$

Now, since $\gamma > l$, we have that $r_a(A^*, \gamma) = r_a(A^p, \gamma) = -k_{over}(\gamma - l) < 0$. Thus, from (48), we obtain that

$$0 \leq D(A^* - A^p) < k_{over}(\gamma - l)(B_p^* - B_p^p),$$

which implies $B_p^* > B_p^p$. This contradicts (47). Then, $A^p > A^*$ and consequently $G^p > G^*$. $\square$

The following proposition is equivalent to Proposition 3 in Sect. 3.

Proposition 6 Let G^c and A^c be given by Proposition 4 and let D^a and D^p be given by (20). If (19) admits a coexistence equilibrium (B_p^*, B_c^*, G^*, A^*) , then:

- (a) $D \leq D^a$ implies $A^* = 0$ and $r_g(G^*, 0) \leq l$.
- (b) $D^a < D < D^p$ implies $0 < A^* < A^p$ and $r_g(G^*, A^*) > l$.

Proof From the equivalence among (a), (c), and (f) in Lemma 2, we obtain that immediately

- (I) $D \leq D^a$ implies $A^* = 0$ and $r_g(G^*, A^*) \leq l$.
- (II) $D^a < D < D^p$ implies $0 < A^*$ and $r_g(G^*, A^*) > l$.

And from Lemma 7, we obtain that $D^a < D < D^p$ implies $A^* < A^p$. This completes the proof. $\square$

The following result states some necessary and sufficient conditions for the coexistence of a unique coexistence equilibrium.

Lemma 8 Let D^p , D^a , and A^p be given by Proposition 4 and let γ be defined by (26). If $D^a < D < D^p$, then (9) admits a coexistence equilibrium (unique) if, and only if,

$$\mu_c(A^p, \gamma) > D \text{ and } \mu_c(0, \gamma) < D. \quad (49)$$

Proof Since $D > D^a$, from Lemma 2, we know that any solution to (24) satisfies

$$r_g(G, A) = \gamma. \quad (50)$$

Thus, from the second equation in (24), we obtain that $\phi(A) := \mu_c(A, \gamma) - D = 0$. We note that (49) is equivalent to $\phi(A^p) > 0$ and $\phi(0) < 0$. If (19) has a coexistence equilibrium, say (B_p^*, B_c^*, G^*, A^*) , then $\phi(A^*) = 0$. Since ϕ is strictly increasing and $0 < A^* < A^p$ (see Lemma 7), we conclude that $\phi(A^p) > 0$ and $\phi(0) < 0$. Thus, (49) is a necessary condition for the existence of a coexistence equilibrium.

We prove now that (49) is also a sufficient condition, that is, if (49) holds, then (24) admits a unique solution. Since (49) holds and ϕ is strictly increasing, there exists a unique $A^* \in (0, A^p)$ such that $\phi(A^*) = 0$.

Now, replacing A by A^* in (50), we obtain the following equation for G :

$$\varphi(G) := r_g(G, A^*) - \gamma = 0.$$

It is clear that φ is strictly increasing and that $\varphi(0) = -\gamma < 0$. Let G^p be given by Proposition 4. Since $A^p > A^*$ (see Lemma 7), we have that $\varphi(G^p) > r_g(G^p, A^p) - \gamma$. From the definition of G^p and A^p (e.g. see (45)), we have that $r_g(G^p, A^p) - \gamma = 0$, and hence $\varphi(G^p) > 0$. Consequently, there is a unique $G^* \in (0, G^p)$ such that $\varphi(G^*) = 0$. It remains to prove the existence and uniqueness of a positive solution of the following linear system for (B_p, B_c) obtained from the third and fourth equations in (24):

$$\begin{aligned} 0 &= D(G_{in} - G^*) - \gamma B_p - \beta \gamma B_c, \\ 0 &= -DA^* - k_{over}(\gamma - l)B_p - \rho_a(A^*, \gamma)B_c. \end{aligned} \quad (51)$$

From the first equation in (51) we have that

$$B_p = \frac{\rho_a(A^*, \gamma)B_c + DA^*}{k_{over}(\gamma - l)}.$$

Combining the previous equation with the second equation in (51), we obtain:

$$\underbrace{G_{in} - G^* - A^* \frac{\gamma}{k_{over}(\gamma - l)}}_{\kappa} = \frac{\gamma}{D} \left(\beta + \frac{\rho_a(A^*, \gamma)}{k_{over}(\gamma - l)} \right) B_c.$$

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

$$\kappa > G_{in} - G^p - A^p \frac{\gamma}{k_{over}(\gamma - l)}.$$

Finally, if we isolate B_p^p in the second equation in (45), and we replace it in the first equation, we obtain that

$$G_{in} - G^p - A^p \frac{\gamma}{k_{over}(\gamma - l)} = 0.$$

This shows that $\kappa > 0$. Then (51) has a positive solution. $\square$

The following result states the conditions presented in statements (b) and (c) in Theorem 1.

Proposition 7 *Let D^p , D^a , and D^c be given by (20) and let γ be defined by (26). We have that:*

- (a) *If $D^a < D < D^p$ and $D < D^c$ then (9) admits a (unique) coexistence equilibrium, if and only if*

$$\mu_p(0, r_g(G^c, 0)) > D \text{ and } \mu_c(A^p, \gamma) > D,$$

with G^p and A^p defined in Proposition 4, and G^c defined in Proposition 5.

(b) If $D^a < D < D^p$ and $D \geq D^c$ then (9) admits a (unique) coexistence equilibrium, if and only if

$$\mu_c(A^p, \gamma) > D,$$

with G^p and A^p defined in Proposition 4.

Proof Let us assume that $D^a < D < D^p$. From Lemmas 6 and 8, we know that (a) and (b) provide necessary conditions for the existence of a coexistence equilibrium. It remains to prove that they also provide sufficient conditions. Using Lemma 8, we know that $\mu_c(A^p, \gamma) > D$ and $\mu_c(0, \gamma) < D$ are sufficient conditions. In statements (a) and (b), it is direct to see that $\mu_c(A^p, \gamma) > D$ holds, therefore we must prove that each statement, (a) and (b), also implies $\mu_c(0, \gamma) < D$. Thus, the proof of parts (a) and (b) follows from proving that

(I) if $D < D^c$, $D^a < D < D^p$, and $\mu_p(0, r_g(G^c, 0)) > D$, then $\mu_c(0, \gamma) < D$, and
 (II) if $D \geq D^c$ and $D^a < D < D^p$, then $\mu_c(0, \gamma) < D$,

respectively.

For (I), from the hypotheses, we have that $\mu_p(0, r_g(G^c, 0)) > D^a$. Using the definition of μ_p and D^a , we have

$$Y_{g/p}r_g(G^c, 0) + Y_{a/p}r_a(0, r_g(G^c, 0)) - k_{deg} > Y_{g/p}l - k_{deg},$$

which can be rearranged as

$$Y_{g/p}[r_g(G^c, 0) - l] > -Y_{a/p}r_a(0, r_g(G^c, 0)). \quad (52)$$

From the definition of r_a (see (16)), we have that $r_a(0, r_g(G^c, 0))$ cannot be positive (evaluate $r_a(A, r_g)$ at $A = 0$). Therefore, (52) implies that $r_g(G^c, 0) > l$. Using again the definition of r_a , we have that $r_a(0, r_g(G^c, 0)) = -k_{over}[r_g(G^c, 0) - l]$. Hence, from the definition of μ_p (see (18)), we have

$$\mu_p(0, r_g(G^c, 0)) = \alpha[r_g(G^c, 0) - l] + D^a, \quad (53)$$

with α defined by (21). Now, since $D > D^a$, according to Proposition 4, $r_g(G^p, A^p) > l$. Thus, using Lemma 2 and the definition of γ , we have that

$$\mu_p(A^p, \gamma) = \alpha[\gamma - l] + D^a. \quad (54)$$

By definition of G^p and A^p , we have that $\mu_p(G^p, A^p) = D$. Thus, from the hypotheses we have that

$$\mu_p(G^c, 0) > \mu_p(A^p, \gamma). \quad (55)$$

Combining (53), (54), and (55), we conclude that

$$r_g(G^c, 0) > \gamma. \quad (56)$$

Now, using the definition of G^c (dilution rate equal to growth rate) and (56), we have that

$$D = \mu_c(0, r_g(G^c, 0)) = Y_{g/c}\beta r_g(G^c, 0) - k_{deg} > Y_{g/c}\beta\gamma - k_{deg} = \mu_c(0, \gamma),$$

and (I) is proved.

For (II), using the definition of D^c and μ_c , and Lemma 3 part (c), we have

$$D^c = Y_{g/c}\beta r_g(G_{in}, 0) - k_{deg} > Y_{g/c}\beta\gamma - k_{deg} = \mu_c(0, \gamma).$$

Now, using the fact that $D \geq D^c$, from the previous equation, we conclude that $D > \mu_c(0, \gamma)$. This completes the proof. $\square$

B Algorithm to find the coexistence equilibrium

Let D^a and D^p be given by Proposition 1. From now on, we assume that $D \in (D^a, D^p)$, otherwise there is no coexistence equilibrium (see Theorem 1). The first step is to determine the equilibrium $E^p = (B_p^p, 0, G^p, A^p)$ given by Proposition 1. The instructions on how to do so are dictated by the proof of Proposition 1. Indeed, B_p^p is obtained as the solution of

$$g(B_p) = 0, \quad (57)$$

with g defined by

$$g(B_p) = \alpha[r_{up,p}^g(G_{in} - \beta_g B_p, \beta_a B_p) - l] + D - D^a,$$

with $\alpha = (1 - Y_{h/p})(Y_g - k_{over}Y_a)$, $\beta_a = \gamma/D$, $\beta_g = k_{over}(\gamma - l)/D$, and

$$\gamma = l + \frac{D - D^a}{(1 - Y_{h/p})(Y_g - k_{over}Y_a)}. \quad (58)$$

Equation (57) has a unique solution on the interval $[0, G_{in}/\beta_g]$. Moreover, $g(0) > 0$ and $g(G_{in}/\beta_g) < 0$, which provides an interval to look for the solution. Thus, equation (57) can be easily solved, for example, with the solver `fzero` in MATLAB. The values of A^p and G^p are obtained from

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

We also need the value of G^c , the glucose concentration associated with the equilibrium E^c given by Proposition 2. Let D^c be given by Proposition 2. 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 1 (see Remark 7).

Now, we have to determine the coexistence equilibrium. The instructions on how to do so are dictated by the proof of Lemma 8:

- 1) Determine $c_1 = (1 - Y_{h/p})f_p(G^c, 0) - k_{deg} - D$ and $c_2 = (1 - Y_{h/c})f_c(G^p, A^p) - k_{deg} - D$.
- 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 ϕ defined by

$$\phi(A) = (1 - Y_{h/c}) \left[Y_g \beta \gamma + Y_a \left(k_a \frac{A}{A + K_a} d(\beta \gamma) + k_{Acs} \frac{A}{A + K_{Acs}} \right) \right] - k_{deg} - D,$$

with γ given by (58). We have 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 $\varphi(G) := (1 - Y_{h/p})f_p(G, A^*) - k_{deg} - D$. We have 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 (15) 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 defined through equalities and non-strict inequalities. In the following, we adapt (15) to use *fmincon*.

Following Theorem 1, the set Ω defined in Sect. 4 can be described such that each element $v = (Y_{h/p}, Y_{h/c}, D)$ on Ω satisfies

$$\begin{aligned} D &< D^p, \\ D &> D^a, \\ D &< (1 - Y_{h/c})f_c(G^p, A^p) - k_{deg}, \\ D &< (1 - Y_{h/p})f_p(G^c, 0) - k_{deg}, \\ Y_{h/p}, Y_{h/c} &\in [0, 1], \end{aligned} \quad (59)$$

where D^a , D^p , G^p , and A^p are given by Proposition 1, D^c and A^c are given by Proposition 2, and G^c is given by (13). We define the set $\tilde{\Omega}$ such that each element $v = (Y_{h/p}, Y_{h/c}, D)$ on $\tilde{\Omega}$ satisfies

$$\begin{aligned} D &\leq D^p, \\ D &\geq D^a, \\ D &\leq (1 - Y_{h/c})f_c(G^p, A^p) - k_{deg}, \\ D &\leq (1 - Y_{h/p})f_p(G^c, 0) - k_{deg}, \\ Y_{h/p}, Y_{h/c} &\in [0, 1]. \end{aligned} \quad (60)$$

The function Φ defined in Sect. 4 is defined on Ω and we must extend its definition on $\bar{\Omega}$. If $v \in \Omega$, we can determine the coexistence equilibrium using the algorithm presented in Appendix B, and therefore evaluate the function Φ defined in Sect. 4. For $v \in \bar{\Omega} - \Omega$, we solve the system (9)–(10) with an initial condition satisfying $B_p(0), B_c(0) > 0$. We run then the model in the long-term until an equilibrium is reached. We will denote the total protein concentration associated to this equilibrium by $\hat{H}^*(v)$. Thus, the extension of Φ on $\bar{\Omega}$, is given by:

$$\hat{\Phi} = \begin{cases} (DH^*(v), H^*(v)/G_{in}) & \text{if } v \in \Omega, \\ (D\hat{H}^*(v), \hat{H}^*(v)/G_{in}) & \text{if } v \in \bar{\Omega} - \Omega. \end{cases}$$

The continuity of $\hat{\Phi}$ can be observed in Figs. 3C and D. In the region of coexistence (shaded region), we have that $v \in \Omega$, while on the boundary of this interval, we have that $v \in \bar{\Omega} - \Omega$. In this situation, we observe a continuous transition between a coexistence equilibrium and a non-coexistence equilibrium as D approaches the boundary of this region.

We solve numerically the following problem instead of (15):

$$\begin{aligned} \max \quad & \lambda \hat{\Phi}(v) + (1 - \lambda) \hat{\Phi}(v) \\ & Av \leq b \\ & c(v) \leq 0 \\ & ub \leq v \leq up, \end{aligned} \quad (61)$$

where

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

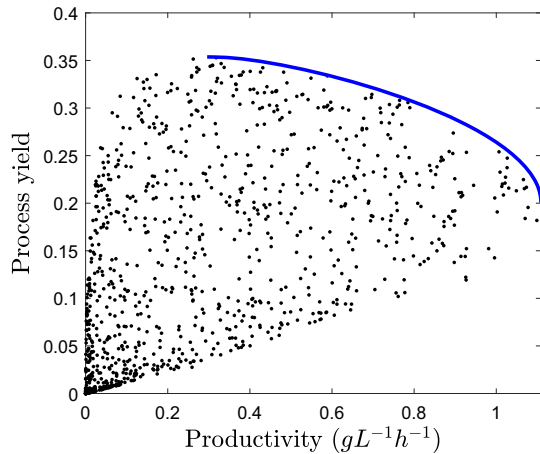
$$ub = \begin{bmatrix} 1 \\ 1 \\ f_p(G_{in}, 0) - k_{deg} \end{bmatrix}, \text{ and } c(v) = \begin{bmatrix} D + k_{deg} - (1 - Y_{h/p}) f_p(G^c(v), 0) \\ D + k_{deg} - (1 - Y_{h/c}) f_c(G^p(v), A^p(v)) \end{bmatrix}.$$

D Robust choice of the down-regulation function

Let us assume that in (9) the down-regulation function d is replaced by $\hat{d}$, defined in (6). We then run long-term simulations of (9)–(10) until an equilibrium is reached (which is always the case).² Then, we evaluate the 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 (7) represents a good approximation of the POF when d is given by (6). This shows that the choice of d in this paper is adequate to study the model proposed by Mauri et al. (2020).

² Let ξ 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 $\delta = 10^{-6} \text{ g L}^{-1}$ and $\Delta t = 10$ days.

Fig. 7 Scatter plot of Φ using the equilibria of model (9)–(10) with d replaced by (6). The continuous line is the POF obtained in Fig. 4



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 *E. 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
- Bremer H, Dennis P (2013) Modulation of chemical composition and other parameters of the cell at different exponential growth rates. In: Slauch J (ed) *Ecosal plus: cellular and molecular biology of E. coli, Salmonella, and the Enterobacteriaceae*, ASM Press, Washington, DC
- 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, Amsterdam
- 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. *Sci Rep* 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
- Hansen SR, Hubbell SP (1980) Single-nutrient microbial competition: qualitative agreement between experimental and theoretically forecast outcomes. *Science* 207(4438):1491–1493
- 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
- Izard J, Gomez Balderas CD, Ropers D, Lacour S, Song X, Yang Y, Lindner AB, Geiselmann J, de Jong H (2015) A synthetic growth switch based on controlled expression of rna polymerase. *Mol Syst Biol* 11(11):840
- 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
- Miliás-Argeitis A, Rullan M, Aoki S, Buchmann P, Khammash M (2016) Automated optogenetic feedback control for precise and robust regulation of gene expression and cell growth. *Nat Commun* 7:12546
- Pandhal J, Noirel J (2014) Synthetic microbial ecosystems for biotechnology. *Biotechnol Lett* 36(6):1141–1151
- Pirt S (1965) The maintenance energy of bacteria in growing cultures. *Proc R Soc Lond B Biol Sci* 163(991):224–31
- 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, Waltman P (1995) The theory of the chemostat: dynamics of microbial competition, vol 13. Cambridge University Press, Cambridge
- 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

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval, sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com