Metabolic Engineering 30 (2015) 166–178

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/ymben

On the feasibility of growth-coupled product synthesis in microbial strains

Steffen Klamt [a,*], Radhakrishnan Mahadevan [b,c]

[a] Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstrasse 1, Magdeburg D-39106, Germany
[b] Department of Chemical Engineering and Applied Chemistry, University of Toronto, 200 College Street, Toronto, ON, Canada M5S 3E5
[c] Institute of Biomaterials and Biomedical Engineering, University of Toronto, 184 College Street, Toronto, ON, Canada, M5S3G9

ARTICLE INFO

Article history:
Received 9 December 2014
Received in revised form
11 March 2015
Accepted 8 May 2015
Available online 22 June 2015

Keywords:
Computational strain design
Escherichia coli
Growth-coupled product synthesis
Yield space
Elementary (flux) modes
Elementary (flux) vectors

ABSTRACT

Enforcing obligate coupling of growth with synthesis of a desired product has become a key principle for metabolic engineering of microbial production strains. Various methods from stoichiometric and constraint-based modeling have been developed to calculate intervention strategies by which a given microorganism can only grow if it synthesizes a desired compound as a mandatory by-product. However, growth-coupled synthesis is not necessarily feasible for every compound of a metabolic network and no rigorous criterion is currently known to test feasibility of coupled product and biomass formation (before searching for suitable intervention strategies). In this work, we show which properties a network must fulfill such that strain designs guaranteeing coupled biomass and product synthesis can exist at all. In networks without flux bounds, coupling is feasible if and only if an elementary mode exists that leads to formation of both biomass and product. Setting flux boundaries leads to more complicated inhomogeneous problems. Making use of the concept of elementary (flux) vectors, a generalization of elementary modes, a criterion for feasibility can also be derived for this situation. We applied our criteria to a metabolic model of *Escherichia coli* and determined for each metabolite, whether its net production can be coupled with biomass synthesis and calculated the maximal (guaranteed) coupling yield. The somewhat surprising result is that, under aerobic conditions, coupling is indeed possible for each carbon metabolite of the central metabolism. This also holds true for most metabolites under anaerobic conditions but consideration of ATP maintenance requirements implies infeasibility of coupling for certain compounds. On the other hand, ATP maintenance may also increase the maximal coupling yield for some metabolites. Overall, our work provides important insights and novel tools for a central problem of computational strain design.

1. Introduction

The rational redesign of the metabolism of microbial strains for the bio-based production of chemicals or biofuels is at the heart of metabolic engineering. Enforcing obligate coupling of growth with product synthesis has become a key principle for the targeted modification of microorganisms (Burgard et al., 2003; Campodonico et al., 2014; Feist et al., 2010; Fong et al., 2005; Hädicke and Klamt 2011; Kim and Reed 2010; Layton and Trinh 2014; Ranganathan et al., 2010, 2012; Ruckerbauer et al., 2014; Shen et al., 2011; Tepper and Shlomi 2010; Tervo and Reed 2012; Trinh et al., 2008, 2011; Xu et al., 2011; Yang et al., 2011; Yim et al., 2011; Zomorrodi et al., 2012). The goal is to delete some reactions such that the host organism can only grow if it synthesizes a desired compound as a mandatory by-product. One key advantage of growth-coupled product synthesis is that adaptive evolution can be used to further optimize the production rate (Conrad et al., 2011; Feist et al., 2010). Stoichiometric and constraint-based modeling techniques provide excellent tools to calculate intervention strategies enforcing growth-coupled product synthesis (Burgard et al., 2003; Hädicke and Klamt, 2010; Hädicke and Klamt, 2011; Kim and Reed 2010; Klamt et al., 2014; Ranganathan et al., 2010; Ruckerbauer et al., 2014; Tervo and Reed 2012; Trinh et al., 2008, 2009; von Kamp and Klamt, 2014; Yang et al., 2011; Zomorrodi et al., 2012). However, coupling biomass formation with synthesis of a metabolite may not be feasible for all compounds in a metabolic network. For example, Feist et al. (2010) applied the OptKnock and OptGene algorithms to identify growth-coupled designs for 11 compounds in the central metabolism of *Escherichia coli* and were able to identify designs in 8 out of the 11 cases whereas no suitable design

* Corresponding author at: Analysis and Redesign of Biological Networks, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstrasse 1, Magdeburg D-39106, Germany.
E-mail address: klamt@mpi-magdeburg.mpg.de (S. Klamt).

http://dx.doi.org/10.1016/j.ymben.2015.05.006
1096-7176/

could be found in the remaining three cases. Similarly, in a recent study, Campodonico et al. (2014) evaluated growth-coupled production for 20 commodity chemicals using RobustKnock and GDLS and showed that coupling was indeed possible in 15 cases but no suitable intervention strategies could be identified for 5 out of the 20 cases. If growth-coupled designs cannot be found, it is not clear whether the failure was due to computational limitations associated with the combinatorial complexity of the strain design problem or whether inherent properties in the network structure prevent growth-coupling. If the failure is due to computational limitations, one can focus on either tuning the parameters of the optimization problem (i.e., the number of allowed knockouts, minimal growth rate required etc.) or improving the computational algorithm or/and its parameters to identify solutions. If the failure is due to structural reasons, identifying a solution may not be possible even when the algorithm is allowed to knockout all genes in the network. Hence, it is important to understand the exact theoretical limitations of a given network structure for coupled product and biomass synthesis when using the various algorithms. To the best of our knowledge, such a study has not been conducted so far and general criteria to test feasibility of coupling are not available. Herein, based on techniques from computational geometry, we show which theoretical requirements must be fulfilled such that coupling of product and biomass formation can be enforced by knocking-out some suitable reactions/genes and how the maximal guaranteed coupling yields can be determined. These techniques are then applied to a metabolic model of *E. coli* where we determine for each metabolite in the model whether coupling is feasible and, if so, its maximal (guaranteed) coupling yield. This analysis reveals some interesting and non-intuitive results about feasibility of coupling strategies and underlines the value of our approach.

2. Theory

2.1. Definitions

A metabolic (stoichiometric) reaction network with m metabolites and n reactions is represented by its stoichiometric matrix $\mathbf{N}$ and a set *Irrev* containing the indices of irreversible reactions. We assume that the network is in steady state meaning that the net reaction rates $\mathbf{r}=(r_1, r_2, \ldots, r_n)^T$ fulfill

$$\mathbf{Nr}=\mathbf{0}, \quad r_i \geq 0 \forall i \in Irrev. \tag{1}$$

The vectors $\mathbf{r}$ satisfying (1) are called steady-state flux vectors and the complete set of steady-state flux vectors form a convex polyhedral cone (a special type of a polyhedron), herein referred to as flux cone $\boldsymbol{F}$ with $\boldsymbol{F}=\{\mathbf{r}\in\Re^n \mid \mathbf{Nr}=\mathbf{0},\ r_i\geq 0\ \ \forall i\in Irrev\}$. From computational geometry it is known that any polyhedral cone can be generated by a conical (non-negative linear) combination of the cone's (unique) extreme rays $\mathbf{x}^i$ superimposed by a linear combination of (non-unique) basis vectors $\mathbf{y}^j$ of the lineality space (Kelk et al., 2012):

$$\boldsymbol{F}=\{\mathbf{r}\in\Re^n \mid \sum_i \alpha_i \mathbf{x}^i + \sum_j \beta_j \mathbf{y}^j,\ \ \alpha_i \geq 0\}. \tag{2}$$

Elementary modes are particular steady-state flux vectors of $\boldsymbol{F}$ possessing a non-decomposability property (Gagneur and Klamt 2004; Schuster et al., 2000; Trinh et al., 2009). The support *supp* of a flux vector denotes the set of reaction indices being non-zero: $supp(\mathbf{r})=\{i \mid r_i\neq 0\}$. An elementary mode $\mathbf{e}$ is a support-minimal steady-state flux vector, that is, no steady-state vector $\mathbf{r}$ exists with $supp(\mathbf{r}) \subsetneq supp(\mathbf{e})$. EMs with the same support are considered equivalent; they only differ by a scalar factor. An EM whose support contains at least one irreversible reaction is called irreversible (can operate only in forward direction); otherwise it is a reversible EM (can run in both directions). We denote the sets of irreversible and reversible EMs by $\boldsymbol{E}^I$ and $\boldsymbol{E}^R$, respectively. Importantly, the set of EMs includes all extreme rays (always contained in $\boldsymbol{E}^I$) and covers the lineality space by the reversible EMs in $\boldsymbol{E}^R$. Therefore, the set of EMs also generates the flux cone meaning that any steady-state flux vector $\mathbf{r}$ can be obtained by a proper linear combination of reversible and irreversible EMs:

$$\boldsymbol{F}=\{\mathbf{r}\in\Re^n \mid \sum_{i:\ \mathbf{e}^i\in\boldsymbol{E}^I} \alpha_i \mathbf{e}^i + \sum_{j:\ \mathbf{e}^j\in\boldsymbol{E}^R} \beta_j \mathbf{e}^j,\ \ \alpha_i \geq 0\}. \tag{3}$$

(Note: for all reversible EMs $\mathbf{e}^j\in\boldsymbol{E}^R$ one might also add two irreversible EMs $\mathbf{e}^j$ and $-\mathbf{e}^j$ to $\boldsymbol{E}^I$ and then remove the sum term over $\boldsymbol{E}^R$ in (3)). However, the set of EMs is not necessarily a *minimal* generating set; it might contain more vectors than required to describe the flux cone (Llaneras and Picó 2010; Wagner and Urbanczik 2005). On the other hand, the set of EMs contains all flux vectors with minimal support (i.e., with a maximal set of zeros) which proved very useful for studying functional and combinatorial properties of metabolic networks (Klamt et al., 2014; Trinh et al., 2009).

Next we require a formal definition of *obligate coupling of biomass and product synthesis*. We assume that the metabolic network contains a substrate S (which can be taken up from the environment), a product P of interest (which can be excreted) and that biomass B can be synthesized with a certain (growth) rate. The biomass yield $Y_{\mathbf{r}}^{B/S}$ of a flux vector $\mathbf{r}$ is the ratio of the growth rate and substrate uptake rate in $\mathbf{r}$. Likewise, the product yield $Y_{\mathbf{r}}^{P/S}$ is defined as the ratio of product excretion and substrate uptake rate. We further assume that two thresholds have been specified, one for the minimum requested biomass yield ($Y_{\min}^{B/S}$) and one for the minimum requested product yield ($Y_{\min}^{P/S}$). We consider two different strengths of coupling, *weak* and *strong*:

Definition 1. (*Strong coupling of growth with product synthesis*): Given the thresholds $Y_{\min}^{B/S}$ and $Y_{\min}^{P/S}$, we say that biomass synthesis is strongly coupled with product synthesis if all steady-state flux vectors $\mathbf{r}$ have a product yield equal or larger than the product yield threshold ($Y_{\mathbf{r}}^{P/S} \geq Y_{\min}^{P/S}$) and if at least one $\mathbf{r}$ exceeds the biomass yield threshold ($Y_{\mathbf{r}}^{B/S} \geq Y_{\min}^{B/S}$). (See also Figure 1d.)

For some applications it might be useful to consider a weaker variant of coupling:

Definition 2. *(Weak coupling of growth with product synthesis):* Given the thresholds $Y_{\min}^{B/S}$ and $Y_{\min}^{P/S}$, we say that biomass synthesis is weakly coupled with product synthesis if all steady-state flux vectors $\mathbf{r}$ having a biomass yield above the requested minimum ($Y_{\mathbf{r}}^{B/S} \geq Y_{\min}^{B/S}$) exhibit a product yield equal or larger than the product yield threshold ($Y_{\mathbf{r}}^{P/S} \geq Y_{\min}^{P/S}$) and at least one such $\mathbf{r}$ exists. (See also Figure 1c.)

Whereas strong coupling demands high product yield (above $Y_{\min}^{P/S}$) for all non-zero flux vectors, weak coupling demands high product yield only if the cell grows with high biomass yield ($\geq Y_{\min}^{B/S}$). Obviously, strong coupling implies weak coupling. Note also that both strong and weak coupling state that product synthesis is mandatory when the cell grows (in case of weak coupling if it grows with high biomass yields) but nothing is said for the other direction: solutions may exist, where the cell synthesizes the product without growing.

For illustrating weak and strong coupling and for studying the relationships between biomass and product yields it is very helpful to investigate the *yield space* of a metabolic network which is a projection of all non-zero steady state flux vectors on their specific biomass (x-axis) and product (y-axis) yields. Yield spaces, in the following denoted by $\boldsymbol{Y}$, normally form a bounded two-dimensional area. A typical example is shown in Fig. 1a. For

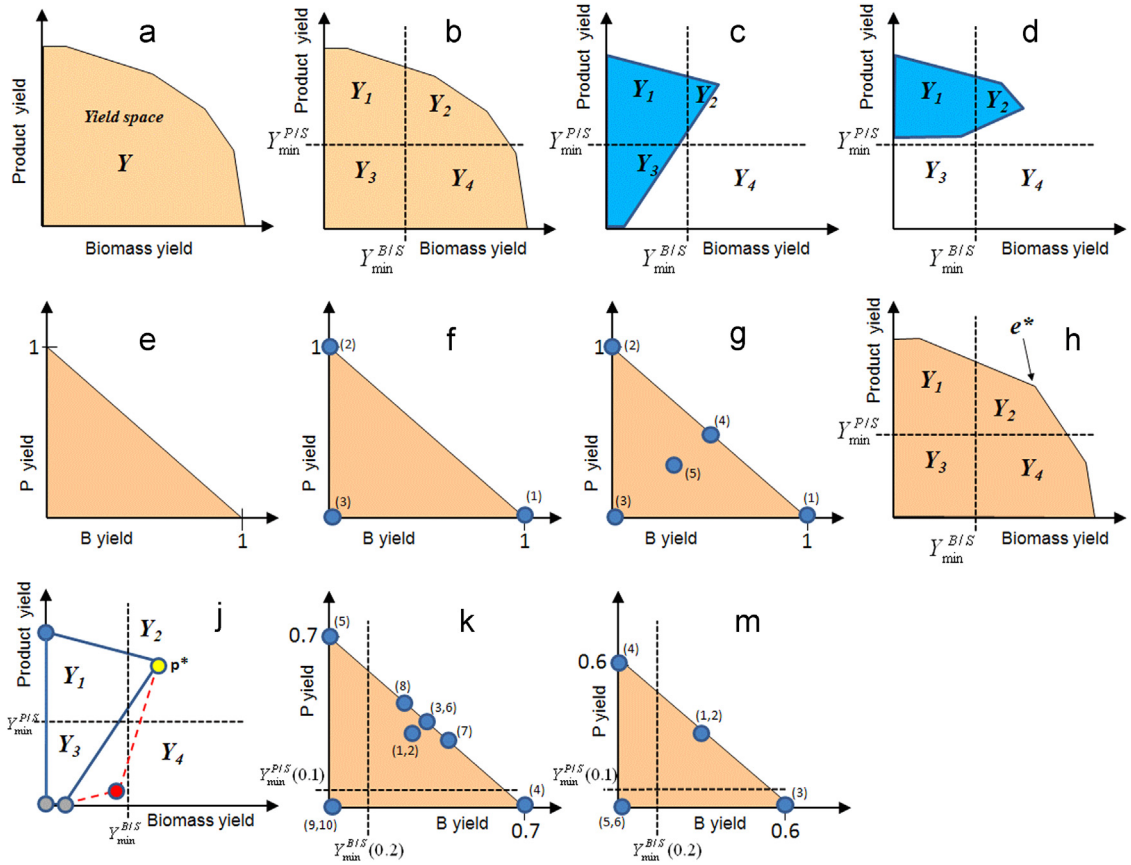

Fig. 1. Examples of yield spaces. (a) Typical shape of a yield space. (b) Introducing minimum thresholds for product and biomass yield dissects the yield space into four regions. (c) Example of a yield space with weak coupling of biomass and product synthesis. (d) Example of a yield space with strong coupling of biomass and product synthesis where high product synthesis is attained also under low biomass yields. (e) The (identical) yield space of network N1 and N2. (f) Yield space of network N1 together with the projection of the EMs of N1. The numbers refer to the EMs in Table 1. (g) Yield space of network N2 together with the projection of the EMs of N2. The numbers refer to the EMs in Table 1. (h) Example for a yield space where a 'corner' in region $\boldsymbol{Y_2}$ indicates existence of $\mathbf{e}^*$ and thus feasibility of coupling. (j) Example of a yield space and EVs where the yellow EV $\mathbf{p}^*$ fulfills Eq. (9) for the two grey EVs from $\boldsymbol{Y_3}$ but not for the red EV (the line from the red EV to the yellow p* crosses $\boldsymbol{Y_4}$). The red EV must therefore be deleted to obtain weak coupling, i.e., to ensure that no solution remains in $\boldsymbol{Y_4}$. (k) Yield space and yield space projections of the EVs in network N2 in the inhomogeneous scenario 2 in Table 1 (the numbers given at the projected EVs correspond to their numbers in Table 1). **(m)** Yield space and yield space projections of the EVs in network N2 in the inhomogeneous scenario 3 in Table 1. See text for further explanations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

well-posed networks we may assume that any feasible non-zero flux vector in the network involves uptake of the substrate (such that biomass and product yields are well-defined for all flux vectors). $\boldsymbol{Y}$ forms a convex set (i.e., the line between any two points from $\boldsymbol{Y}$ lies also within $\boldsymbol{Y}$) and each point in $\boldsymbol{Y}$ corresponds to one or several steady-state flux distributions. The yield space can be determined by different methods, for example, by flux variability analysis (see below) or by calculating elementary modes: $\boldsymbol{Y}$ can then be generated by the convex hull of the yield space projection of all the EMs. Note that each EM has a unique biomass and product yield, irrespective of the scaling of the EM.

The two thresholds $Y^{B/S}_{\min}$ and $Y^{P/S}_{\min}$ used in Definitions 1 and 2 divide the yield space into four regions $\boldsymbol{Y_1}$ … $\boldsymbol{Y_4}$ (Fig. 1b). $\boldsymbol{Y_1}$ contains solutions where the product yield is high but the biomass yield is below the threshold $Y^{B/S}_{\min}$. $\boldsymbol{Y_2}$ contains the "desired" phenotypes because biomass and products are produced with high yields. $\boldsymbol{Y_3}$ and $\boldsymbol{Y_4}$ exhibit low product yield, in case of $\boldsymbol{Y_3}$ also low biomass yield. Fig. 1c and d depict examples of yield spaces where growth and product synthesis are weakly/strongly coupled in the sense of Definitions 2 and 1. In both cases, no solution is allowed to exist in $\boldsymbol{Y_4}$ and at least one must lie in $\boldsymbol{Y_2}$. Weak coupling (Fig. 1c) still allows solutions to exist in region $\boldsymbol{Y_3}$ which may exhibit low or even zero product synthesis under low biomass yield. Consideration of weak coupling is motivated by the fact that microorganisms often evolve to a metabolic state with high biomass yield which, under weak coupling, is then contained in the desired region $\boldsymbol{Y_2}$ ensuring high product yield. In contrast, the yield space shown in Fig. 1d indicates strong coupling: solutions are not allowed to exist in $\boldsymbol{Y_3}$ or in $\boldsymbol{Y_4}$. Hence all feasible flux distributions in the network synthesize the product with a high yield above $Y^{P/S}_{\min}$. Knockout strategies leading to strong coupling and thus to yield spaces as in Fig. 1d can therefore guarantee high product yield also for suboptimal growth yields.

The yield space is related with the *production envelope* where absolute rates instead of yields are considered (Feist et al., 2010). In fact, Definitions 1 and 2 could also be formulated with minimal desired production and growth rates (instead of minimal product and biomass yields). Technical problems may arise if these rates are not limited in the model (as in the homogeneous case; see below), which is not a problem in the yield space. However, in all other cases, the methods presented herein can be analogously applied to absolute rates and production envelopes as well.

We also mention here that even weaker variants of coupling could be formulated. For example, some bilevel optimization approaches (as (Tepper and Shlomi, 2010)) demand for coupling only that a (high) product yield is obtained when the cell grows with maximal biomass production. Hence, low or even zero product yields are allowed for all flux vectors as long as the

biomass yield is below the maximal achievable value (which demands less than weak coupling in Definition 2). Although we will concentrate on the two definitions given above, other variants could be treated analogously with the same techniques as used herein.

Based on Definitions 1 and 2, we can now formulate the existence problem of intervention strategies enforcing coupled biomass and product synthesis:

Definition 3. (*Feasibility of strong/weak coupling of growth with product synthesis*): Given the thresholds $Y^{B/S}_{\min}$ and $Y^{P/S}_{\min}$, we say that growth can be strongly/weakly coupled with product synthesis if there exists a set of reaction knockouts such that growth and product synthesis in the reduced network is strongly/weakly coupled according to Definitions 1 and 2.

The goal of this work is to investigate requirements for feasibility of coupling (according to Definition 3), that is, to derive conclusive criteria which unambiguously prove/disprove the existence of intervention (knockout) strategies enforcing weak or/and strong coupling of biomass and product synthesis. Existence of appropriate intervention strategies will certainly depend on the two thresholds $Y^{B/S}_{\min}$ and $Y^{P/S}_{\min}$. If one just wants to test for the existence of *any* suitable strategy enforcing coupling with *any* non-zero biomass and product yields one can set these thresholds to arbitrarily small non-zero values.

2.2. Feasibility of coupling in networks without flux bounds

We initially consider only the (homogeneous) constraint (1) for steady state flux vectors; the more complicated inhomogeneous case with additional flux capacity constraints will be taken into account at a later step.

The yield space (Fig. 1a) contains in most cases solutions where only biomass or only the product is synthesized; these solutions lie on the *x*- and *y*-axis, respectively. Due to the convexity of $\boldsymbol{Y}$, all lines between (i.e. convex combinations of) these extreme solutions are contained in $\boldsymbol{Y}$ as well. Hence, whenever biomass *B* and *P* can be produced independently, $\boldsymbol{Y}$ will display "desired" solutions from $\boldsymbol{Y_2}$ where growth and product synthesis occur concurrently. However, existence of flux distributions $\mathbf{r}$ with $Y^{B/S}_{\mathbf{r}} \geq Y^{B/S}_{\min}$ and $Y^{P/S}_{\mathbf{r}} \geq Y^{P/S}_{\min}$ alone can neither *guarantee* coupled biomass and product synthesis (because solutions in $\boldsymbol{Y_4}$ usually exist which will preferably be taken by the cell) nor can they *guarantee* that intervention strategies (combinations of reaction deletions) exist that would induce strong or weak coupling. We can also conclude that weak coupling is only feasible if an appropriate knockout strategy exist such that (i) all solutions contained in $\boldsymbol{Y_4}$ will be deleted and (ii) at least one from $\boldsymbol{Y_2}$ remains feasible. For strong coupling, suitable intervention strategies must delete all solutions in both $\boldsymbol{Y_3}$ *and* $\boldsymbol{Y_4}$. Below we will see that a *single criterion* can be derived which is sufficient and necessary for feasibility of *both* weak and strong coupling in homogeneous systems.

We consider the two example networks N1 and N2 in Fig. 2. Here, B represents biomass (or a biomass component), S the substrate, P a product of interest, and A another central metabolite such as ATP. Indicated by dashed arcs, B, P, and A can be excreted, in the case of biomass (B) this corresponds to "growth", and in the case of A it mimics intracellular consumption by other processes. It is easy to see that N1 is a sub-network of N2 and that only N2 contains pathways (e.g. R8→R9→R10→R11) along which biomass and product synthesis can be coupled. Interestingly, both networks have exactly the same yield space with respect to B and P (Fig. 1e). Hence, from the shape of the yield space alone we cannot conclude whether biomass and product coupling can be achieved or not. However, projecting also the elementary modes (EMs) onto

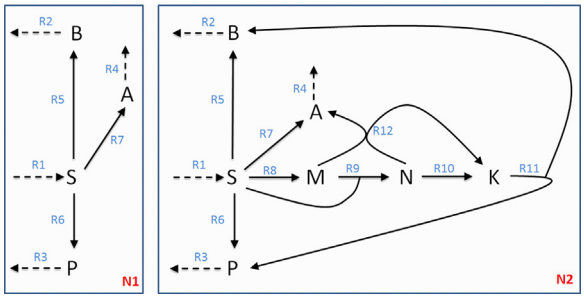

Fig. 2. Example Networks N1 and N2. Note that N1 is a subnetwork of N2.

the yield space brings about an important difference (Fig. 1f and g). The EMs of network N1 and N2 with their coefficients are shown in Table 1 and one can verify that Fig. 1f and g correctly display the yield-space projections of these EMs (and, as mentioned earlier, that the convex hull of the yield space projections of EMs generates $\boldsymbol{Y}$). Both networks give rise to EM1–EM3 whereas only network N2 contains two additional EMs (EM4 and EM5) which, in contrast to EM1–EM3, do not lie on one of the two axes. Nevertheless, the exhibited biomass and product yields of EM4 and EM5 can also be obtained in network N1. For example, the biomass and product yield of 0.5 of EM4 in N2 can be obtained by combining EM1 and EM2 in N1 each being weighted with a factor of 0.5. However, it is the existence of an EM such as EM4 (or EM5) that guarantees existence of an intervention strategy inducing obligate coupling of production of B with P. A corresponding proof can be constructed as follows: if an EM $\mathbf{e}^*$ exists with

$$Y^{B/S}_{\mathbf{p}^*} \geq Y^{B/S}_{\min} \text{ and } Y^{P/S}_{\mathbf{p}^*} \geq Y^{P/S}_{\min} \tag{4}$$

(i.e., if projection of at least one EM onto the yield space is contained in $\boldsymbol{Y_2}$) then we can easily construct an intervention strategy fulfilling the requirements of weak *and* strong coupling (Definitions 2 and 1), namely by deleting all reactions not involved in the EM $\mathbf{e}^*$. Due to the non-decomposability property of EMs, this ensures that no other flux distribution than the EM itself will be maintained in the network. Hence, the yield space would be reduced to a single point contained in $\boldsymbol{Y_2}$ and all flux vectors in $\boldsymbol{Y_4}$ (and even in $\boldsymbol{Y_1}$ and $\boldsymbol{Y_3}$) would be blocked. We can furthermore show that existence of an EM $\mathbf{e}^*$ contained in $\boldsymbol{Y_2}$ is not only sufficient but even necessary to render weak/strong coupling of growth and product synthesis feasible: if such an EM $\mathbf{e}^*$ does not exist then all steady state flux vectors, whose projections onto the yield space lie in the desired region $\boldsymbol{Y_2}$, can only be generated by combinations of EMs contained in $\boldsymbol{Y_1}$, $\boldsymbol{Y_3}$, and $\boldsymbol{Y_4}$ (contributions from EMs in $\boldsymbol{Y_1}$ and $\boldsymbol{Y_4}$ will be essential). Since we need to block all flux vectors contained in $\boldsymbol{Y_4}$ to get weak or strong coupling, all flux vectors with projections contained in $\boldsymbol{Y_2}$ would be blocked as well and coupling becomes infeasible. We have thus proven:

Criterion 1. *(Feasibility of weak/strong coupling in homogeneous systems)*: Weak and strong coupling of biomass and product synthesis in the homogeneous system (Eq. (1)) is feasible if and only if an EM $\mathbf{e}^*$ exists fulfilling (4).

One interesting and important consequence of Criterion 1 is that feasibility of weak coupling immediately implies feasibility of strong coupling (again, the other way around is trivial by definition). However, a larger number of knockouts might be required to achieve strong coupling. The maximal (guaranteed) product yield under weak or strong coupling is given by the EM with the highest product yield among all EMs with the demanded minimum biomass yield. Hence, we can conclude that the maximal achievable guaranteed product yield under coupling in network N2 is 0.5 given by EM4 (Table 1, Fig. 1g). Furthermore, if we set the

Table 1
Elementary modes and elementary vectors in the networks N1 and N2 (see also Fig. 2). The crosses in C1, C2, and C3 indicate which of the Criteria 1, 2, and 3 hold for the respective EMs/EVs and can thus be used to prove strong or/and weak coupling (assumed thresholds: $Y^{B/S}_{min} = 0.2$ and $Y^{P/S}_{min} = 0.1$).

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	$Y^{B/S}$	$Y^{P/S}$	$Y^{A/S}$	C1	C2	C3
Elementary modes of network N1																		
EM1	1	1			1								1	0	0			
EM2	1		1			1							0	1	0			
EM3	1			1			1						0	0	1			
Elementary modes of network N2																		
EM1	1	1			1								1	0	0			
EM2	1		1			1							0	1	0			
EM3	1			1			1						0	0	1			
EM4	2	1	1					1	1	1	1		0.5	0.5	0	X		
EM5	3	1	1	1				1	1		1	1	0.33	0.33	0.33	X		
Scenario 1: Elementary vectors in network N2 with inhomogeneous constraints $r_{R1} \leq 10$																		
EV1	10	10			10								1	0	0			
EV2	10		10			10							0	1	0			
EV3	10			10			10						0	0	1			
EV4	10	5	5					5	5	5	5		0.5	0.5	0	X		
EV5	10	3.33	3.33	3.33				3.33	3.33		3.33	3.33	0.33	0.33	0.33	X		
Scenario 2: Elementary vectors in network N2 with inhomogeneous constraints $r_{R1} \leq 10$ and $r_{R4} \geq 3$																		
EV1	9	3	3	3				6	3		3	3	0.33	0.33	0.33		X	X
EV2	10	3.33	3.33	3.33				6.66	3.33		3.33	3.33	0.33	0.33	0.33		X	X
EV3	10	3.5	3.5	3			3	3.5	3.5	3.5	3.5		0.35	0.35	0.3			X
EV4	10	7		3	7		3						0.7	0	0.3			
EV5	10		7	3		7	3						0	0.7	0.3			
EV6	10	3.5	3.5	3				6.5	3.5	0.5	3.5	3	0.35	0.35	0.3		X	X
EV7	10	4	3	3	1			6	3		3	3	0.4	0.3	0.3		X	X
EV8	100	3	4	3		1		6	3		3	3	0.3	0.4	0.3		X	X
EV9	10			10			10						0	0	1			
EV10	3			3			3						0	0	1			
Scenario 3: Elementary vectors in network N2 with inhomogeneous constraints $r_{R1} \leq 10$ and $r_{R4} \geq 4$																		
EV1	10	3	3	4			4	3	3	3	3		0.3	0.3	0.4			X
EV2	10	3	3	4			1	6	3		3	3	0.3	0.3	0.4			X
EV3	10	6		4	6		4						0.6	0	0.4			
EV4	10		6	4		6	4						0	0.6	0.4			
EV5	10			10			10						0	0	1			
EV6	4			4			4						0	0	1			

minimum demanded product yield $Y^{P/S}_{min}$ above 0.5, no EM would exist in $\boldsymbol{Y_2}$ and coupling becomes infeasible (similar arguments hold for the minimum demanded biomass yield $Y^{B/S}_{min}$). We also note that it is usually not necessary to cut all reactions not involved in the EM $\mathbf{e}^*$ to achieve coupling. Assume that we choose $Y^{P/S}_{min} = 0.4$ and $Y^{B/S}_{min} = 0.2$ for network N2. EM4 in N2, which exceeds these thresholds and thus represents the $\mathbf{e}^*$, does not involve reactions R4, R5, R6, R7, and R12 and we know that removal of these reactions thus induces (strong) coupling. However, deletion of R5, R6, R7, and R12 is already sufficient to ensure that the only remaining flux vector would be EM4 lying in $\boldsymbol{Y_2}$. Knockout of R6 can also be omitted for weak and strong coupling if we accept that EM2 (product yield is 1, biomass yield is 0) needs not to be blocked. In fact, only with R6 product yields above the guaranteed 0.5 would be feasible, namely if the cell would grow with suboptimal biomass yield by using combinations of EM2 and EM4. Finally, would we reduce our demanded minimum product yield $Y^{P/S}_{min}$ from 0.4 to 0.3 then EM5 would become an acceptable phenotype and removal of R12 is also not needed anymore (leaving only R5 and R7 as required reaction deletions).

Enumeration of all (including the smallest/most efficient) intervention strategies that enforce coupling with the given thresholds can conveniently be done by computing constrained Minimal Cut Sets (Hädicke and Klamt, 2011). One first specifies a set $\boldsymbol{T}$ of (target) EMs with undesired phenotypes. For example, for strong coupling, $\boldsymbol{T}$ comprises all EMs from $\boldsymbol{Y_4}$ and $\boldsymbol{Y_3}$ (optionally, also from $\boldsymbol{Y_1}$). Specifying additionally a second set $\boldsymbol{D}$ of desired behaviors (normally all EMs from $\boldsymbol{Y_2}$) constrained minimal cut sets (cMCSs) can be computed which are support-minimal sets of reaction knockouts such that (i) all EMs from $\boldsymbol{T}$ (and thus all steady state flux vectors from yield space regions $\boldsymbol{Y_3}$ and $\boldsymbol{Y_4}$ (and, optionally, $\boldsymbol{Y_1}$)) are blocked and, (ii) that at least one EM (and thus at least one steady state flux vector) from $\boldsymbol{Y_2}$ remains feasible. Existence of a cMCS with these properties proves feasibility of coupled growth and product synthesis (in fact, such an cMCS is the intervention strategy demanded by Definition 3). A similar cut-set approach can be used to enumerate all intervention strategies inducing weak coupling.

If the EMs of a given network can be determined, it is thus easy to check whether a network can be redesigned to obtain coupled biomass and product synthesis. In large-scale networks, enumeration of EMs is often not possible. However, even in those cases tools are available that can be used to check for the existence of an EM $\mathbf{e}^*$ with the desired property (4). A direct approach is based on the work of Pey and Planes (2014) who showed that single EMs fulfilling certain properties (including yield thresholds) can be readily calculated in genome-scale networks. A second approach uses an indirect proof: we have recently shown how smallest constrained minimal cut sets (cMCSs) can be calculated in genome-scale networks, in particular also those that lead to coupled growth and product synthesis (von Kamp and Klamt, 2014). Since, as described above, existence of one such cMCS is equivalent with feasibility of coupling this implies that an EM $\mathbf{e}^*$ with property (4) must exist. Hence, calculation of one single (e.g., the shortest) cMCS for growth-coupled product synthesis with the requested yield thresholds proves existence of $\mathbf{e}^*$. A third method uses another geometric criterion. Even if the EMs cannot be determined, the yield space can normally be quickly constructed by flux variability analysis (FVA (Mahadevan and Schilling, 2003)) in large networks. If the yield space contains a

'corner' at the upper borderline (the Pareto front) in region $\boldsymbol{Y_2}$ this must correspond to an EM $\mathbf{e}^*$ which thus proves feasibility of growth-coupled product synthesis (Fig. 1h). However, even if such a corner does not exist (i.e., if the yield space has a triangle shape as in Fig. 1e, f, and g), coupling might be feasible. An example for this situation is given by network N2: here, the EM $\mathbf{e}^*$ lies directly on the line connecting the two extreme points of $\boldsymbol{Y_1}$ and $\boldsymbol{Y_4}$. In fact, EMs indicating feasibility of coupling may also lie in the inner region of the yield space.

Generally, inspection of the yield space with its EMs can be used to find out what the maximum *guaranteed* product yield under coupling is and how the two thresholds must possibly be moved such that coupling becomes feasible.

2.3. Networks with inhomogeneous constraints

So far we have considered the network without any capacity constraints, that is, without any limitations or restrictions on specific flux values (except that irreversible reactions cannot be negative). However, the boundaries of some exchange fluxes, in particular of the substrate or oxygen uptake rate, are often known and therefore included as constraints in stoichiometric models. Moreover, some fluxes may have minimum values above zero. A prominent example for the latter is the non-growth associated ATP maintenance demand of the cell specifying the minimum amount of ATP to maintain basic processes not covered in the metabolic model such as turnover of macromolecules, maintenance of electrical membrane potential, unspecific energy-consuming transport processes and others. The ATP maintenance demand is usually modeled as a pseudo-reaction in the network drawing constantly some minimum (non-zero) amount of ATP per unit biomass and time.

In the most general case, incorporation of such constraints extends system (1) to

$$\begin{aligned}&\mathbf{Nr}=\mathbf{0}\\&\mathbf{Ar}=\mathbf{a}\\&\mathbf{Br}\le\mathbf{b}\\&r_i\ge 0\quad \forall i\in Irrev\end{aligned}\tag{5}$$

$\mathbf{A}$ is a $t\times n$ matrix and $\mathbf{a}$ is a $t\times 1$ vector encoding together t additional equality constraints, and $\mathbf{B}$ is a $h\times n$ matrix and $\mathbf{b}$ an $h\times 1$ vector encoding h additional inequality constraints. For example, if we know for the ith reaction that $r_i\ge 5$ then a row with zeros except a -1 at the ith position will be added to $\mathbf{B}$ as well as a -5 in the same line in the vector $\mathbf{b}$. Note that, in principle, also more complicated constraints than simple flux bounds can be expressed in this way (e.g., bounds for *linear combinations* of fluxes).

The inclusion of such inhomogeneous constraints (inhomogeneous because $\mathbf{a}$ and $\mathbf{b}$ usually contain non-zero values) changes the shape of the solution space of flux vectors $\mathbf{r}$ satisfying (5). It is not a (flux) cone anymore, but still a (flux) polyhedron – in the following denoted by $\boldsymbol{P}$ – and we will now investigate how feasibility of coupled biomass and product synthesis can be studied in this space. Whereas cones can be generated by extreme rays and lineality space basis vectors (Eq. (2)), a polyhedron can be generated by a *convex* combination of its extreme points (vertices) $\{\mathbf{p}^k\}_{k\in K}$ superimposed by, first, a conical (non-negative) combination of its extreme rays $\{\mathbf{x}^i\}_{i\in I}$ and, second, by a linear combination of lineality space basis vectors $\{\mathbf{y}^j\}_{j\in J}$ (Bertsimas and Tsitsiklis, 1997; Kelk et al., 2012):

$$\begin{aligned}\boldsymbol{P}=\{\mathbf{r}\in\Re^n\,|\,\mathbf{r}=&\sum_{k\in K}\gamma_k\mathbf{p}^k+\sum_{i\in I}\alpha_i\mathbf{x}^i+\sum_{j\in J}\beta_j\mathbf{y}^j,\\&\gamma_k\ge 0\,,\ \sum_{k\in K}\gamma_k=1,\ \alpha_i\ge 0\}\end{aligned}\tag{6}$$

Note that the convex combination of extreme points requires $\sum\gamma_k=1$. Since convex polyhedral cones represent a particular class of polyhedra, the generation of cones in (2) represents indeed a special case of (6): polyhedral cones have either no extreme point (namely if the lineality space is not empty) or the zero vector is the only vertex. In both these cases the term with the linear combinations of extreme points $\mathbf{p}^k$ can be removed from (6) yielding (2).

Urbanczik and Wagner (2005), Urbanczik (2007) showed that, similar as for the flux cone where EMs provide a (not necessarily minimal but) very useful set of generating vectors, it is reasonable to not restrict oneself to a *minimal* set of generators of the flux polyhedron. They proposed to consider instead *elementary (flux) vectors* (EVs). With the definition of Urbanczik and Wagner, EVs can be seen as a generalization of EMs for the inhomogeneous case: EVs cover not only all extreme points and extreme rays as well as the lineality space but in addition also all support-minimal vectors. We call a vector $\mathbf{v}$ support-minimal w.r.t. $\boldsymbol{P}$ if no other non-zero vector $\mathbf{r}$ contained in $\boldsymbol{P}$ (fulfilling (5)) exists with $supp(\mathbf{r})\subsetneq supp(\mathbf{v})$. As we will see below, similar as for EMs, not all support-minimal vectors of the flux polyhedron $\boldsymbol{P}$ are essentially required to generate $\boldsymbol{P}$. However, as in the case of EMs in the homogeneous system, the set of EVs greatly facilitates deletion or knockout studies as considered herein because after deletion of one reaction the new set of EVs in the remaining network can be obtained by just taking all EVs from the full network and discarding all those that involved the deleted reaction with non-zero rate.

Formally, Urbanczik and Wagner defined the set of elementary vectors of a polyhedron $\boldsymbol{P}$ as the set of all vectors that lie on an extreme ray or extreme point of the polyhedron obtained by intersecting $\boldsymbol{P}$ with one of the 2^n orthants of $\Re^n$. In the special case where $\boldsymbol{P}$ is a flux cone, the definition of EVs coincides with that of EMs. We noticed that the definition of EVs by Urbanczik and Wagner is not fully consistent with the EVs defined by Rockafellar (1970), who only considered support-minimal vectors of linear subspaces instead of polyhedra. We nevertheless stick herein to the term of EV to not introduce yet another term.

An algorithm for calculating elementary vectors (EVs) in inhomogeneous systems was given in Urbanczik (2007). In the supplementary material we introduce another approach by transforming the problem of calculating EVs into an equivalent problem of determining EMs allowing one to use existing high-performance algorithms for EM calculations (Gagneur and Klamt 2004; Hunt et al., 2014; Terzer and Stelling 2008) to compute EVs. The algorithm outputs three groups of EVs: (i) the bounded EVs $\{\tilde{\mathbf{p}}^k\}_{k\in\tilde{K}}$ (which include all extreme points $\mathbf{p}^k$ of the polyhedron $\boldsymbol{P}$ used in (6)), (ii) the extreme directions $\{\tilde{\mathbf{x}}^i\}_{i\in\tilde{I}}$ (which includes all extreme rays of $\boldsymbol{P}$), and (iii) all support-minimal vectors $\{\tilde{\mathbf{y}}^j\}_{j\in\tilde{J}}$ of the lineality space (which also span the lineality space of $\boldsymbol{P}$). Analogous to (6) the flux polyhedron $\boldsymbol{P}$ defined by (5) can then be generated from these three classes of EVs:

$$\begin{aligned}\boldsymbol{P}=\{\mathbf{r}\in\Re^n\,|\,\mathbf{r}=&\sum_{k\in\tilde{K}}\gamma_k\tilde{\mathbf{p}}^k+\sum_{i\in\tilde{I}}\alpha_i\tilde{\mathbf{x}}^i+\sum_{j\in\tilde{J}}\beta_j\tilde{\mathbf{y}}^j,\\&\gamma_k\ge 0\,,\ \sum_{k\in\tilde{K}}\gamma_k=1,\ \alpha_i\ge 0\}\end{aligned}\tag{7}$$

Again, the key difference to (6) is that the three sets of EVs encompass not only the minimal generating set of $\boldsymbol{P}$ (composed of extreme points, extreme rays, and a lineality space basis) but also all support-minimal vectors of $\boldsymbol{P}$ not contained yet in the minimal generating set.

In many relevant cases, the zero point represents an extreme point of the flux cone or flux polyhedron, although often not explicitly listed in metabolic network studies. In fact, the zero point is always part of the flux cone but omitted when plotting the

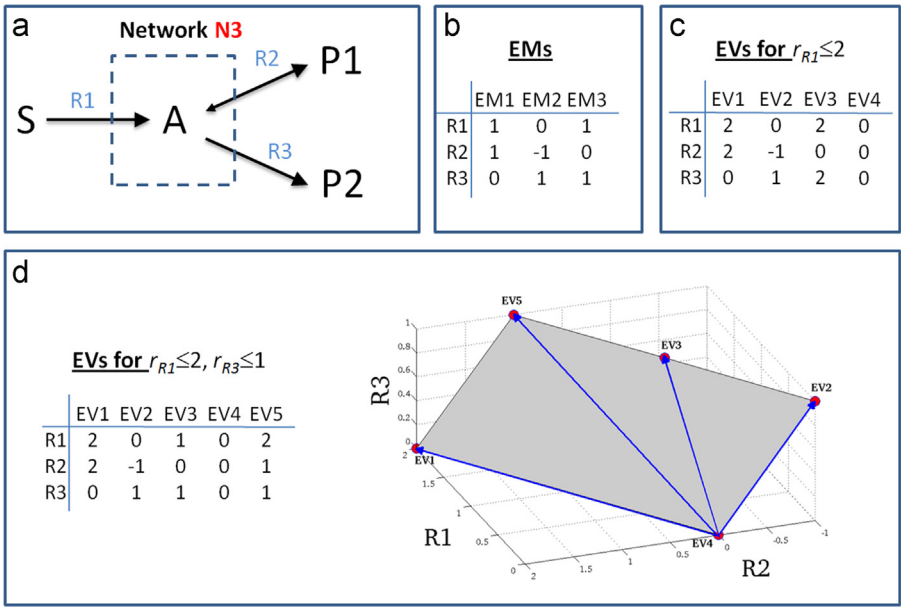

Fig. 3. (a) Example Network N3. (b) Elementary modes of the network. (c) Elementary vectors for the inhomogeneous scenario with $r_{R1} \leq 2$. (d) Elementary vectors for the inhomogeneous scenario with $r_{R1} \leq 2$ and $r_{R3} \leq 1$. The resulting solution space (flux polyhedron) spanned by the five elementary vectors is also shown.

yield space (as in Fig. 1), since biomass and product yields are not defined at this trivial solution. Whereas the zero vector is relevant for the formally correct generation of flux polyhedra in Eqs. (6) or (7), it needs not be taken into account when studying the yield space as we are usually only interested in non-trivial flux vectors carrying some flux (we could also enforce this behaviour by setting a minimal substrate uptake rate through which the zero vector disappears as solution from the flux polyhedron).

We illustrate the concept of EVs with the example network N3 in Fig. 3 consisting of one internal metabolite and three reactions (thereof R2 being reversible). The network has three EMs (Fig. 3b). If we constrain $r_{R1} \leq 2$ (Fig. 3c) then we obtain four EVs (including the zero flux vector). Three of them (EV1, EV3, EV4) are bounded with EV1 and EV4 representing extreme points (EV3 can be obtained by combining EV1 and EV2), whereas EV2 is an extreme ray of the polyhedron (a lineality space does not exist). Adding also the constraint $r_{R3} \leq 1$ results in five EVs (Fig. 3d), all of which are bounded. These five EVs include the four extreme points of the polyhedron (EV1, EV2, EV4, EV5) and in addition EV3 which is not an extreme point as it can be generated by the convex combination 0.5•EV2+0.5•EV5. This can also be seen by the resulting solution space where EV3 lies on the line connecting EV2 and EV5 (Fig. 3d). As mentioned, for our considerations it will be essential to keep also the elementary vector EV3 in the generating set. For example, if we deleted reaction R2, the new set of EVs would be given by deleting all EVs that involved this reaction, thus, in fact, only EV3 and the zero vector EV4 would remain as the new set of EVs. Having only the extreme points instead of the full set of EVs would require recalculation of the new generating set. Note that EV5 *is* an extreme point although it can be obtained by 1•EV1+1•EV2, because this is not a *convex* linear combination of EVs (where all factors must sum up to 1).

Using the framework of EVs, we can now try to tackle the question of coupling growth and product synthesis in the inhomogeneous case. To simplify the discussion, we assume (i) that the substrate uptake rate is limited by a finite upper boundary and, as in the homogeneous case, (ii) that any feasible (non-zero) flux vector in $\boldsymbol{P}$ (fulfilling Eq. (5)) involves uptake of the substrate. Both assumptions are reasonable and fulfilled in most networks since substrate uptake cannot be arbitrarily large and any (non-zero) flux distribution not consuming substrate will represent thermodynamically infeasible cycles (Kelk et al., 2012). Even if such EVs exist in the network, we may neglect them in the following because they will have a weight of 0 in Eq. (7) due to infeasibility.

With these two assumptions we always obtain a bounded polyhedron (also called polytope): all non-trivial flux vectors have a non-zero substrate uptake rate. Since the latter is bounded, no flux vector can be scaled arbitrarily and neither extreme rays nor a lineality space can exist. Thus, the flux polyhedron (5) is then generated by a convex combination of its extreme points; here we use again the extended set of bounded elementary vectors $\{\tilde{\mathbf{p}}^k\}_{k\in\tilde{K}}$:

$$\boldsymbol{P} = \{\mathbf{r}\in\Re^n \,|\, \mathbf{r} = \sum_{k\in\tilde{K}} \gamma_k \tilde{\mathbf{p}}^k \quad \gamma_k \geq 0\,, \quad \sum_{k\in\tilde{K}} \gamma_k = 1\} \tag{8}$$

With the same argumentation as in the homogeneous case, we might scan the set of EVs (excluding the zero vector) to find a special EV $\mathbf{p}^* \in \{\tilde{\mathbf{p}}^k\}_{k\in\tilde{K}}$ that fulfills the condition of Criterion 1, that is $Y^{B/S}_{\mathbf{p}^*} \geq Y^{B/S}_{\min}$ and $Y^{P/S}_{\mathbf{p}^*} \geq Y^{P/S}_{\min}$, and to show then that the reactions with a zero rate in this EV form a proper knockout set enforcing coupling. However, whereas Criterion 1 (applied to EMs) was generally sufficient in the homogeneous case, it is not in the case of flux polyhedra. As explained above, the set of EVs includes all support-minimal vectors of the flux polyhedron $\boldsymbol{P}$. However, it may also contain vectors that are not support-minimal, that is for two non-zero EVs $\mathbf{u}$ and $\mathbf{v}$ it may happen that $supp(\mathbf{u}) \subset supp(\mathbf{v})$. Why this can happen is explained in the supplementary material and it is exemplified by Fig. 3d: EV5 uses a superset of the reactions involved in EV1–EV3. And in fact, although EV5 exhibits simultaneous synthesis of P2 and P3 in this example, no knockout set exists that would obligatorily couple synthesis of P2 and P3.

As the set of bounded EVs may contain vectors that are not support-minimal, simply deleting all reactions having a zero rate in an EV $\mathbf{p}$ will not necessarily lead to deletion of all other EVs, and undesired solutions without coupling might still exist. For the inhomogeneous case we therefore need stronger conditions which are now distinct for strong and weak coupling:

Criterion 2. *(Feasibility of strong coupling in inhomogeneous systems):* strong coupling of biomass and product synthesis in an

inhomogeneous system (Eq. (5)) is feasible if and only if an EV $\mathbf{p}^*$ exists with

(a) $Y^{B/S}_{\mathbf{p}^*} \geq Y^{B/S}_{\min}$ and $Y^{P/S}_{\mathbf{p}^*} \geq Y^{P/S}_{\min}$, and
(b) For all non-zero EVs $\mathbf{u}$ from $\boldsymbol{Y_3}$ and $\boldsymbol{Y_4}$ (i.e., for all $\mathbf{u}$ with $Y^{P/S}_{\mathbf{u}} \leq Y^{P/S}_{\min}$) it must hold that $supp(\mathbf{u}) \not\subseteq supp(\mathbf{p}^*)$.

A proof can be given as follows: criteria 2(a) and 2(b) ensure that the zero reactions in $\mathbf{p}^*$ mark a proper knockout set blocking all flux vectors in the yield space regions $\boldsymbol{Y_3}$ and $\boldsymbol{Y_4}$. Recall that, in the homogeneous case, (b) is not required because this condition is automatically fulfilled due to the support-minimal nature of EMs. Furthermore, with the same argument as for EMs in the homogeneous case, existence of an appropriate $\mathbf{p}^*$ fulfilling 2 (a) and 2(b) is not only sufficient but also necessary: if no such $\mathbf{p}^*$ exists, then all (desired) flux vectors in $\boldsymbol{Y_2}$ are merely arising as convex combinations of EVs from $\boldsymbol{Y_1}$, $\boldsymbol{Y_3}$, and $\boldsymbol{Y_4}$ (with essential contributions from $\boldsymbol{Y_1}$ and $\boldsymbol{Y_4}$) and no suitable knockout strategy enforcing coupling can exist.

For weak coupling, the situation becomes a bit more complicated. Clearly, Criterion 2a must hold as well and Criterion 2b must be fulfilled for the EVs $\mathbf{u}$ lying in $\boldsymbol{Y_4}$. Special care must be taken for EVs in $\boldsymbol{Y_3}$: only those EVs $\mathbf{u}$ in $\boldsymbol{Y_3}$ must be disabled where the line connecting $\mathbf{u}$ with $\mathbf{p}^*$ intersects with $\boldsymbol{Y_4}$. This is illustrated in Fig. 1j: if we consider the yellow EV in $\boldsymbol{Y_2}$ as $\mathbf{p}^*$, then the two grey EVs are not critical whereas the red EV is since the line connecting the red EV with $\mathbf{p}^*$ crosses $\boldsymbol{Y_4}$. It is easy to calculate that a line from a point $\mathbf{u}$ in $\boldsymbol{Y_3}$ to a $\mathbf{p}^*$ in $\boldsymbol{Y_2}$ intersects with $\boldsymbol{Y_4}$ if

$$Y^{P/S}_{\mathbf{u}} + \frac{(Y^{B/S}_{\min} - Y^{B/S}_{\mathbf{u}})(Y^{P/S}_{\mathbf{p}^*} - Y^{P/S}_{\mathbf{u}})}{(Y^{B/S}_{\mathbf{p}^*} - Y^{B/S}_{\mathbf{u}})} \leq Y^{P/S}_{\min} \quad (9)$$

The term on the left-hand side of (9) is the product yield at the point where the line from $\mathbf{u}$ to $\mathbf{p}^*$ crosses the biomass threshold $Y^{B/S}_{\min}$. We accordingly obtain:

Criterion 3. *(Feasibility of weak coupling in inhomogeneous systems)*: Weak coupling of biomass and product synthesis in an inhomogeneous system (Eq. (5)) is feasible if and only if an EV $\mathbf{p}^*$ exists with

(a) $Y^{B/S}_{\mathbf{p}^*} \geq Y^{B/S}_{\min}$ and $Y^{P/S}_{\mathbf{p}^*} \geq Y^{P/S}_{\min}$, and
(b) For all non-zero EVs $\mathbf{u}$ in $\boldsymbol{Y_3}$ ($Y^{B/S}_{\mathbf{u}} \leq Y^{B/S}_{\min}$ and $Y^{B/S}_{\mathbf{u}} \leq Y^{P/S}_{\min}$) that satisfy (9) and for all $\mathbf{u}$ in $\boldsymbol{Y_4}$ ($Y^{B/S}_{\mathbf{u}} \geq Y^{B/S}_{\min}$and $Y^{B/S}_{\mathbf{u}} \leq Y^{P/S}_{\min}$) it must hold that $supp(\mathbf{u}) \not\subseteq supp(\mathbf{p}^*)$.

Although Criterion 3 for feasibility of weak coupling appears a bit more complicated, it actually demands less than Criterion 2 for strong coupling because only some EVs of $\boldsymbol{Y_3}$ need be deleted. It still holds that strong coupling implies weak coupling. Furthermore, once the EVs are known, we can determine the maximal (guaranteed) coupling yield $Y^{P/S}_{\min}$ for which strong (or weak) coupling is feasible: we just search for the $\mathbf{p}^*$ with the highest product yield that fulfills the properties in Criterion 2 (or 3).

Criteria 2 and 3 will be illustrated in Section 3.1 (toy example) and then be applied to examine feasibility of growth-coupled synthesis of metabolites in *E. coli* (Section 3.2).

We have not yet discussed how Criteria 2 and 3 can be tested efficiently. In medium-scale networks where the EVs can be enumerated (such a functionality is, for example, provided by *CellNetAnalyzer* (Klamt et al., 2007)), this can be done in a straightforward manner by searching for EVs $\mathbf{p}^*$ that satisfy Criteria 2 and 3. As a side product of the considerations made herein, we can also generalize the method for calculating cMCSs in homogeneous systems (Hädicke and Klamt, 2011) to apply it also to the computation of interventions strategies enforcing coupling in the inhomogeneous case. For example, for strong coupling we proceed as follows: after calculating the EVs, the set of undesired EVs (in $\boldsymbol{Y_3}$ and $\boldsymbol{Y_4}$) and the set of desired EVs (in $\boldsymbol{Y_2}$) are specified. Afterwards, the adapted *Berge* algorithm (Hädicke and Klamt, 2011) is employed to enumerate all cMCSs that solve the intervention problem. Satisfaction of Criterion 2(b) is implicitly ensured by the Berge algorithm for constrained MCSs. Thus, intervention strategies for inhomogeneous problems can be calculated by applying the Berge algorithm to the given undesired and desired set of EVs. In a similar manner, cMCSs enforcing weak coupling can be calculated as well.

The situation in genome-scale networks is more complicated. The algorithm of Pey and Planes (2014) cannot be used here, because it was designed for EMs, not EVs. One may indirectly show that coupling is feasible by determining a single cMCS inducing the desired behavior, for example, by the duality-based algorithm where inhomogeneous constraints can straightforwardly be integrated (von Kamp and Klamt, 2014). Future work will show whether an algorithm can be elaborated for large-scale networks that can directly determine an EV $\mathbf{p}^*$ satisfying Criterions 2 and 3 in a faster manner than the indirect way via proving the existence of any suitable cMCS.

3. Results

3.1. Examples for testing feasibility of coupling in inhomogeneous systems

We illustrate treatment of inhomogeneous systems by example scenarios in network N2 (cf. Fig. 2 and Table 1).

3.1.1. Scenario 1: $r_{R1} \leq 10$

As demanded by our assumption for inhomogeneous systems, we first introduce a maximal rate for the substrate uptake reaction R1 ($r_{R1} \leq 10$). A general finding one can make at this point is that merely adding an upper boundary or even a fixed value for the substrate uptake flux without setting any further inhomogeneous constraint will not change any of the results from the section on homogeneous systems. In those cases, we just truncate the (unbounded) cone and the resulting EVs are (up to the zero vector) in one-to-one relationship with the original EMs of the flux cone. With $r_{R1} \leq 10$ we therefore obtain 5 EVs (plus the zero vector not shown in Table 1) corresponding to the 5 EMs in the homogeneous case except that they were all scaled such that the boundary of the substrate uptake rate is exactly met (see Table 1). This change, however, does not affect the yields, hence, as long as the inhomogeneous constraints concern only the substrate uptake, we may still rely on the simpler analysis of EMs as described in the previous section. Accordingly, EM4 and EM5 prove feasibility of weak and strong coupling and the maximal guaranteed coupling yield is $Y^{P/S}_{\min} = 0.5$ (if $Y^{B/S}_{\min} \leq 0.5$).

3.1.2. Scenario 2: $r_{R1} \leq 10$, $r_{R4} \geq 3$

In addition to setting a maximum substrate uptake rate, we now assume that metabolite A mimics the role of ATP and that some constant (maintenance) demand of A is required that must be provided by the network. We therefore introduce a lower boundary of 3 for reaction R4. The resulting 10 EVs are shown in Table 1 (note that the zero vector is not part of the flux polyhedron anymore since a non-zero flux for R4 is required) and their projection onto the yield space is depicted in Fig. 1k. We want to examine whether any coupled synthesis of B with P is feasible and use therefore rather small thresholds $Y^{B/S}_{\min} = 0.2$ and $Y^{P/S}_{\min} = 0.1$ From the resulting yield space partitioning (Fig. 1k) we see that EV4, EV9, and EV10 must be deleted for strong coupling; for weak

coupling deletion of EV4 is also essential, whereas the necessity to remove EV9 and EV10 is not clear yet. There are several EVs producing both B and P with yields above the minimum thresholds thus satisfying (a) of Criterion 2 (and 3). Among those is EV3 which clearly fulfills Criteria 2a and 3a but also 3b: for EV4 from $\boldsymbol{Y_4}$ it holds *supp*(EV4) $\not\subseteq$ *supp*(EV3) and the lines connecting EV3 with the two EVs from $\boldsymbol{Y_3}$ (EV9 and EV10) do not cross $\boldsymbol{Y_4}$. Therefore, EV3 represents the $\mathbf{p}^*$ demanded by Criterion 3 proving feasibility of weak coupling (and knocking-out the reactions not involved in EV3 will induce weak coupling). However, EV3 violates Criterion 2 (b) for feasibility of strong coupling since the support of EV10, which lies in $\boldsymbol{Y_3}$ and produces neither B nor P, is a subset of the support of EV3 (hence, knockout of the unused reactions in EV3 is not sufficient to block EV10). In contrast, EV1, EV2, EV6, EV7, and EV8 really ensure feasibility also of strongly (and again also weakly) coupled synthesis of B and P as they all fulfill both (a) and (b) of Condition 2. Note that the supports of EV6, EV7, and EV8 are supersets of the support of EV1 and EV2 and that EV1 and EV2 have even identical support, however, these EVs still fulfill Condition 2 and 3. The common property of all these 5 coupling-enabling EVs is that they do not involve reaction R7 which can therefore be deleted ensuring that none of the EVs with low product yield will remain. In fact, as long as R7 is active, the substrate could be completely converted to A implying a zero yield for B and P thus preventing strong coupling. An interesting effect of the inhomogeneous constraint on A is that reaction R5 need not be deleted anymore, in contrast to the homogeneous case. The reason is that after inactivating R7 the only remaining pathway for the required synthesis of A implies production of P and B. In a sense, the production of P and B are coupled to the generation of A which is here assumed to play a similar role as ATP for non-growth associated maintenance demand. Similar effects can be expected in realistic networks when product synthesis generates ATP in excess which can then be used for maintenance processes. One example is production of ethanol or lactate under anaerobic conditions. Hence, ATP maintenance demand may help to reduce the number of required knockouts, as was recently also discussed for cyanobacteria producing biofuels (Erdrich et al., 2014).

3.1.3. Scenario 3: $r_{R1} \leq 10$, $r_{R4} \geq 4$

In this last scenario we increase the constant cellular demand of A from 3 to 4 and use again the thresholds $Y^{B/S}_{\min} = 0.2$; $Y^{P/S}_{\min} = 0.1$. The number of EVs reduces from 10 to 6 (Table 1); the resulting yield space is shown in Fig. 1m. EV1 and EV2 exhibit coupled synthesis of B and P and fulfill Criterion 2(a) and 3(a), however, both violate Criterion 2(b) for strong coupling as the support of EV5 and EV6 producing neither B nor P are subsets of those from EV1 and EV2. This reflects that the pathway via R12 cannot produce sufficient amount of A due to low stoichiometric yield. Therefore, R7 becomes essential which in turn implies that, in principle, all substrate taken up could be converted to A by which strong coupling of B and P synthesis becomes infeasible. A different result is obtained for weak coupling: although the support of EV5 and EV6 are subsets of EV1 and EV2, they are not critical as they do not satisfy Eq. (9). For example, with $\mathbf{u}$=EV5 (or $\mathbf{u}$=EV6) and $\mathbf{p}^*$=EV1 the left-hand side of Eq. (9) is 0.2 which is larger than $Y^{P/S}_{\min} = 0.1$. Hence, the line connecting EV5 (or EV6) with EV1 does not cross $\boldsymbol{Y_4}$ in Fig. 1m. A suitable intervention strategy enforcing weak coupling would be to delete just R5. We can then assume that, under growth with higher biomass yields, the cell will not waste all substrate for producing exclusively A and instead use the pathway via K to produce the biomass component B coupled with synthesis of P. Clearly, had we chosen $Y^{P/S}_{\min} = 0.25$ then even weak coupling would not be possible with this threshold.

We mention here that fixing both the substrate uptake as well as the maintenance demand of A to the (exact) values of 10 and 4, respectively, (instead of using upper/lower boundaries) would re-enable also feasibility of strong coupling. This emphasizes the importance of accurate adjustment and formulation of those constraints in metabolic models.

3.2. Systematic evaluation of feasibility of growth-coupled synthesis of metabolites in E. coli

We applied our theoretical framework to a metabolic model of the central metabolism of *E. coli* to test for which metabolites intervention strategies exist that could strongly or weakly couple synthesis of these metabolites with growth. The model was taken from Hädicke and Klamt (2010) and uses glucose as substrate. We set the maximum glucose uptake rate to 10 mmol/(gDW h) and the minimum requested biomass yield $Y^{B/S}_{\min}$ to 0.01 gDW/mmol glucose. We considered aerobic as well as anaerobic growth, first without inhomogeneous constraints up to an upper boundary for the substrate uptake rate. As described in the Theory section, this situation can be treated as a homogeneous system and Criterion 1 can be applied for testing feasibility of both weak and strong coupling. For each metabolite M in the model, we added temporarily an export reaction for M and determined the resulting EMs in this network. From Criterion 1 we know that any EM having a non-zero yield of M and a biomass yield above $Y^{B/S}_{\min}$ guarantees existence of an intervention strategy for coupled growth and synthesis of M. We determined for each metabolite M the EM that has the highest product yield (and a biomass yield above $Y^{B/S}_{\min}$) thus exhibiting the maximal product yield that can be guaranteed under coupled growth by using a suitable (and definitely existing) knockout strategy. A somewhat surprising result was that production of *all* carbon metabolites can be coupled with biomass synthesis both under aerobic as well as anaerobic conditions (Table 2). Especially for a metabolite like glucose-6-phosphate (G6P), which, in the network, is close to the substrate glucose, feasibility of coupling was unexpected. Coupling for G6P is feasible because uptake of glucose involves consumption of ATP or PEP, and this dependency can be exploited to find interventions in the model enforcing coupled biomass and G6P synthesis. We also notice that the theoretical maximal coupling yield under aerobic growth is higher than under anaerobic conditions for most metabolites. For example, for glyoxylate we obtain 1.815 and 0.462 mmol/mmol glucose under aerobic and anaerobic growth, respectively. Some exceptions can be found for fermentation products such as ethanol or lactate where the anaerobic maximum coupling yield is close or even equal to the aerobic value. Note that the maximum aerobic coupling yield cannot be smaller than for the anaerobic case since the EMs of the aerobic scenario build a superset of the anaerobic EMs. Hence, any suitable anaerobic knockout strategy can also be obtained for the aerobic case by just deleting in addition the oxygen uptake reaction.

Taking into account ATP maintenance demand (we used the commonly applied value of 8.39 mmol/(gDW h)) gives rise to an inhomogeneous system. For each metabolite M as a virtual product, we determined the resulting elementary vectors for both aerobic and anaerobic conditions. As described in the Theory section, in contrast to EMs in the homogeneous system, not all EVs with non-zero biomass and product yield imply feasibility of coupling because EVs may exist that are not support-minimal. We therefore have to apply the more complicated test described by Criterions 2 and 3 to test feasibility of strong/weak coupling and to calculate the maximal guaranteed coupling yields for both. For the aerobic case we found again that synthesis of all metabolites can still be coupled with growth, however, ATP maintenance decreases the maximal coupling yields for most metabolites, in a few cases

Table 2
Maximal (guaranteed) coupling yields [mmol/mmol glucose] of carbon metabolites that can be obtained by appropriate intervention strategies enforcing coupled synthesis of the metabolite and biomass during growth of *E. coli* on glucose. Results were calculated from a model of the central metabolism of *Escherichia coli* (Hädicke and Klamt, 2010). Four different scenarios were considered: aerobic and anaerobic growth, each considered with and without ATP maintenance demand. Minimum desired biomass yield was set to 0.01 gDW/mmol glucose, maximum glucose uptake rate to 10 mmol/gDW/h and ATP maintenance demand (if included) to 8.39 mmol/gDW/h. For the two inhomogeneous cases (where ATP maintenance is included in the model), we calculated the maximum product yields for strong as well as for weak coupling. For the homogeneous case (where ATP maintenance is not included in the model), the maximum yields for strong and weak coupling coincide.

Metabolite full name	Species ID	Aerobic without ATPmaint	Aerobic with ATPmaint	Aerobic with ATPmaint	Anaerobic without ATPmaint	Anaerobic with ATPmaint	Anaerobic with ATP maint
		max Yield strong+weak coupling	max Yield strong coupling	max Yield weak coupling	max Yield strong+weak coupling	max Yield strong coupling	max Yield weak coupling
Glucose-6-phosphate	G6P	0.8437	0.6504	0.6504	0.4872	0	0
Fructose-6-phosphate	F6P	0.8437	0.6584	0.6584	0.4872	0.1707	0.1707
Fructose-1,6-bisphosphat	F16P	0.8024	0.7626	0.7626	0.4560	0	0.0969
Dihydroxyaceton-phosphate	DHAP	1.5957	1.5251	1.5251	0.9120	0	0.1545
Glycerol-3-phosphat	Glyc3P	1.3209	0.9900	0.9900	1.0520	0.4377	0.4377
Glyceraldehyde-3-phosphate	G3P	1.5957	1.5251	1.5251	0.9120	0	0.0188
Diphosphoglycerate	DPG	1.8348	1.4613	1.4613	0.4553	0.2213	0.2213
3-Phosphoglycerate	3PG	1.8243	1.8450	1.8450	0.9228	0.4407	0.4407
2-Phosphoglycerate	2PG	1.8243	1.8450	1.8450	0.9228	0.4407	0.4407
Phosphoenolpyruvate	PEP	1.8243	1.8450	1.8450	0.9228	0.4407	0.4407
Pyruvate	Pyr	1.8446	1.8466	1.8466	0.9132	0.9204	0.9204
Acetyl-CoA	AcCoA	1.8455	1.8466	1.8466	0.9132	0.9204	0.9204
Citrate	Cit	0.9196	0.9196	0.9196	0.4622	0.4635	0.4635
Isocitrate	ICit	0.9196	0.9196	0.9196	0.4622	0.4635	0.4635
alpha-ketoglutarate	alKG	0.9188	0.8953	0.8953	0.3692	0.3708	0.3708
Succinyl-CoA	SuccCoA	1.4945	1.3869	1.3869	1.2935	0.9923	0.9923
Succinate	Succ	1.5850	1.5862	1.5862	1.5850	1.5862	1.5862
Fumarate	Fum	1.6859	1.3949	1.3949	1.2517	0.9004	0.9004
Malate	Mal	1.6859	1.3949	1.3949	1.2517	0.9004	0.9004
Oxalacetate	OxA	1.8243	1.8450	1.8450	0.9228	0.5982	0.5982
Glyoxylate	Glyox	1.8153	1.8466	1.8466	0.4622	0.4635	0.4635
Ribose-5-phosphate	R5P	1.0006	0.9217	0.9217	0.5826	0.3002	0.3002
Ribulose-5-phosphate	Rl5P	1.0006	0.9236	0.9236	0.5826	0.3142	0.3142
Erythrose-4-phosphate	E4P	1.2287	1.1660	1.1660	0.7048	0.2979	0.3240
Xylulose-5-phosphate	X5P	1.0006	0.9236	0.9236	0.5826	0.3142	0.3142
Seduheptulose-7-phosphate	S7P	0.7215	0.6586	0.6586	0.4328	0.1772	0.2228
6-Phospho-gluconolactone	PGlac	0.9174	0.5621	0.5621	0.4937	0	0.2438
6-Phospho-gluconate	PGluc	0.9174	0.5621	0.5621	0.4937	0	0.2438
2-Keto-3-deoxy-6-gluconate	2KD6PG	0.9174	0.5621	0.5621	0.4937	0	0.2438
CO2	CO2	5.5508	5.5736	5.5736	1.8495	1.8263	1.8263
Acetyl-phosphate	AcP	1.8455	1.8466	1.8466	0.9132	0.9204	0.9204
Acetate	Ac	1.8359	1.8446	1.8446	0.9238	0.9046	0.9046
Formate	Form	3.3958	3.6131	3.6131	2.2528	2.2496	2.2496
Lactate	Lac	1.8060	1.8416	1.8416	1.7407	1.8222	1.8222
Acetaldehyde	Adh	1.8446	1.8421	1.8421	1.7407	1.8222	1.8222
Ethanol	Eth	1.8431	1.8416	1.8416	1.8431	1.8189	1.8189
Chorismate	Chor	0.5161	0.4540	0.4540	0.2842	0.0982	0.1361
Phosphoribosyl-1-Pyrophosphate	PRPP	0.8525	0.7983	0.7983	0.3280	0.1250	0.1658
Aspartate-4-Semialdehyde	AspSAld	1.3587	1.2228	1.2228	0.6328	0.4208	0.4208
Homoserine	HSer	1.2025	0.9225	0.9225	0.8007	0.3347	0.3347
Alanine	Ala	1.8388	1.8445	1.8445	1.8355	1.7840	1.7840
Cysteine	Cys	0.8562	0.6413	0.6413	0.2640	0.1930	0.1930
Aspartate	Asp	1.6444	1.5498	1.5498	1.0659	0.5534	0.5534
Glutamate	Glu	0.9196	0.9197	0.9197	0.4622	0.4635	0.4635
Phenylalanine	Phe	0.4886	0.4525	0.4525	0.2831	0.1707	0.1707
Glycine	Gly	3.2530	3.0997	3.0997	1.8982	1.1068	1.1068
Histidine	His	0.8562	0.6227	0.6227	0.2444	0	0.1610
Isoleucine	Ile	0.6656	0.6391	0.6391	0.3690	0.2169	0.2169
Lysine	Lys	0.6289	0.6694	0.6694	0.4657	0.2297	0.2297
Leucine	Leu	0.6076	0.6138	0.6138	0.3707	0.3708	0.3708
Methionine	Met	0.4957	0.4829	0.4829	0.1585	0.1194	0.1194
Asparagine	Asn	1.3936	1.3714	1.3714	0.5353	0.2006	0.2006
Proline	Pro	0.9147	0.8372	0.8372	0.5211	0.2610	0.2610
Glutamine	Gln	0.9192	0.9138	0.9138	0.4614	0.2186	0.2186
Arginine	Arg	0.7517	0.7144	0.7144	0.2506	0.1157	0.1157
Serine	Ser	1.8262	1.8438	1.8438	0.9038	0.5442	0.5442

Table 2 (*continued*)

Metabolite full name	Species ID	Aerobic without ATPmaint	Aerobic with ATPmaint	Aerobic with ATPmaint	Anaerobic without ATPmaint	Anaerobic with ATPmaint	Anaerobic with ATP maint
		max Yield strong+weak coupling	max Yield strong coupling	max Yield weak coupling	max Yield strong+weak coupling	max Yield strong coupling	max Yield weak coupling
Threonine	Thr	0.9257	0.9265	0.9265	0.6229	0.3216	0.3216
Valine	Val	0.9194	0.9222	0.9222	0.9178	0.8920	0.8920
Tryptophan	Trp	0.3806	0.3645	0.3645	0.1613	0.0923	0.0923
Tyrosine	Tyr	0.4908	0.4555	0.4555	0.2710	0.1384	0.1384
ATP_for_RNA_Synthesis	rATP	0.4477	0.3931	0.3931	0.1253	0.0241	0.0825
GTP_for_RNA_Synthesis	rGTP	0.4370	0.3891	0.3891	0.1141	0.0022	0.0626
CTP_for_RNA_Synthesis	rCTP	0.4928	0.4643	0.4643	0.1585	0	0.0940
UTP_for_RNA_Synthesis	rUTP	0.5218	0.5182	0.5182	0.1832	0	0.1109
ATP_for_DNA_Synthesis	dATP	0.4249	0.3885	0.3885	0.1166	0.0238	0.0825
GTP_for_DNA_Synthesis	dGTP	0.4216	0.3886	0.3886	0.1118	0.0022	0.0683
CTP_for_DNA_Synthesis	dCTP	0.4607	0.4627	0.4627	0.1213	0	0.0414
TTP_for_DNA_Synthesis	dTTP	0.4044	0.3795	0.3795	0.1629	0.0022	0.0463
OH_myristic_acid	OH_myr_ac	0.2632	0.2142	0.2142	0.1799	0.0785	0.0785
C_14:0_fatty_acid	C14_0_FS	0.2557	0.2144	0.2144	0.1652	0.0921	0.0921
Diaminopimelate	di_am_pim	0.6536	0.6694	0.6694	0.4657	0.2297	0.2297

Note: It is assumed that phosphate, nitrogen, oxygen, water etc. are provided in sufficient amounts. Furthermore, for acetyl-CoA and succinyl-CoA, the stoichiometric demand for synthesis of coenzyme A (CoA) was not taken into account (since the CoA synthesis pathway was not part of the model). Possible energy requirements for exporting metabolites to the extracellular space were not considered.

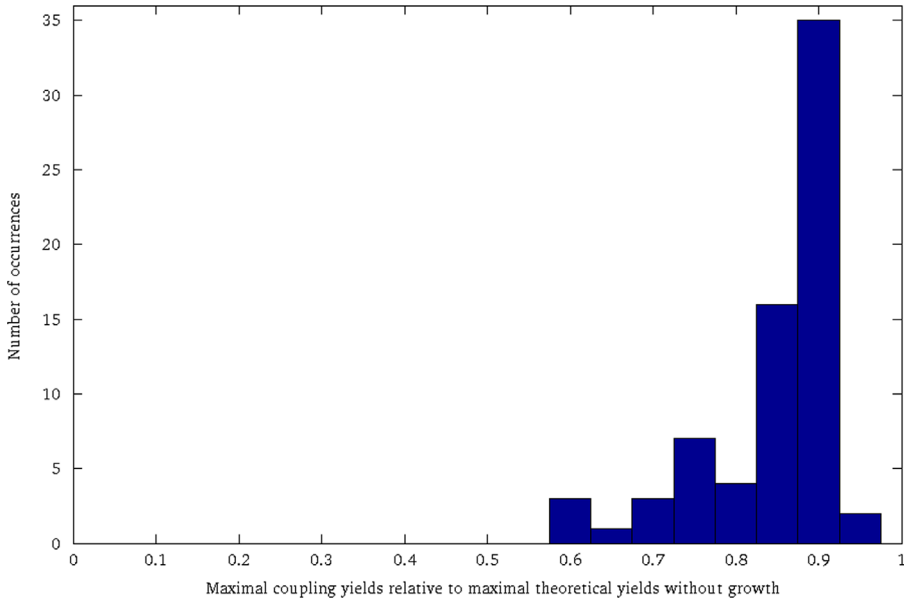

Fig. 4. Distribution of maximal guaranteed coupling yields of metabolites in the *Escherichia coli* core network relative to the maximal product yield achievable (without growth) under aerobic conditions with consideration of ATP maintenance (the values hold for both weak and strong coupling).

drastically. Interestingly, for some metabolites a slight increase can be observed. No differences in the maximal coupling yields can be seen for weak and strong coupling under aerobic conditions and for most metabolites the maximal coupling yields are close to the theoretical maximum yield obtainable for each metabolite without growth (Fig. 4). This result indicates that intervention strategies exist that enforce coupling with only a minor effect on the maximal product yields (unless $Y^{B/S}_{\min}$ is set to larger values). For anaerobic conditions, consideration of ATP maintenance reduces the yield of most metabolites further and growth-coupled synthesis of some compounds becomes now infeasible (especially if we demand strong coupling; feasibility and maximal coupling yields now differ for strong and weak coupling for some metabolites under these conditions). But as for the aerobic case, ATP maintenance demand may also have a positive effect on the attainable coupling yield, for instance, the maximal guaranteed coupling yield of lactate increases from 1.74 to 1.82 mmol/(mmol glucose) suggesting the value of expressing futile cycles for certain products.

We repeated the same calculation with another model of *E. coli*'s central metabolism (Orth et al., 2010) and obtained basically the same results. The maximal coupling yields for the compounds do slightly differ due to differences in the stoichiometries for the biomass synthesis reaction. One should also keep in mind that the calculated coupling yields depend on the chosen threshold for $Y^{B/S}_{\min}$.

4. Discussion and conclusions

Many approaches for computational strain design seek to identify genetic intervention strategies that enforce obligate

coupling of growth with product synthesis. What has been missing so far is a systematic investigation of which topological requirements must be fulfilled such that any suitable knockout strategy enforcing coupling can exist. Here we introduced the notions of strong and weak coupling and criteria which rigorously prove or disprove feasibility of these couplings without explicit calculation of the interventions. Whereas the existence of an EM with desired minimum biomass and product yield is necessary and sufficient for both weak and strong coupling in homogeneous systems (no flux boundaries defined except for substrate uptake), (dis)proving feasibility of coupling in inhomogeneous systems (with some non-zero flux bounds) is more complicated and requires the notion of elementary flux vectors. The latter represent a generating set of the resulting flux polyhedron which includes not only all extreme points (vertices) and extreme rays, and a basis for the lineality space but also all support-minimal flux vectors. We showed that feasibility of coupling can be tested by means of EVs using Criteria 2/3. The concept of EVs, originally introduced by Urbanczik and Wagner (2005), Urbanczik (2007), turns out to be a very useful approach to deal with inhomogeneous constraints. Similar as for EMs in flux cones, the EVs generate the polyhedral solution space and include all support-minimal vectors of the solution space. Although a subset of the EVs might be sufficient to generate the solution space the full set of EVs brings many advantages, especially when analyzing the effects of interventions (knockouts) as done herein. We believe that other applications analyzing flux polyhedra could very much benefit from the approach of EVs, for instance, when analyzing the optimal solution space of genome-scale networks which has so far been done only at the level of minimal generating vector sets (Kelk et al., 2012) thus neglecting potentially important support-minimal vectors of this space. As a side product of this work, we presented an alternative algorithm for calculating EVs by determining EMs in a transformed system. This method should increase the applicability of the EV approach since many tools available for calculating EMs can then also be used for computing EVs. Another side result of this work is the generalization of the EM-based calculation of constrained minimal cut sets (cMCSs) originally introduced for homogeneous systems: the modified Berge algorithm described in Hädicke and Klamt (2011) can straightforwardly be applied to calculate also cMCSs from EVs in inhomogeneous systems by selecting appropriate sets of desired and undesired EVs. All mentioned algorithms are available in the MATLAB toolbox *CellNetAnalyzer*.

Our theoretical findings provide valuable novel tools and important results for computational strain design and metabolic engineering. They give insights which properties of metabolic networks render coupling of product and biomass synthesis feasible. With our results one may exclude feasibility of coupling or determine maximal feasible growth-coupled product yields before starting complicated algorithms searching for suitable intervention strategies. The criteria for testing feasibility of coupling can directly be applied in medium-scale networks and can also be used in genome-scale networks with homogeneous constraints. In inhomogeneous large-scale systems, indirect proof via existence of a suitable cMCS could be used and future work will be dedicated to elaborate algorithms that can directly determine EV $\mathbf{p}^*$ satisfying Criteria 2/3. Our method can straightforwardly be generalized for production envelops (instead of yield spaces) and for other definitions of coupling.

Application of our criteria to *E. coli* unexpectedly revealed that, with glucose as substrate, synthesis of all carbon metabolites can, in principle, be coupled to growth, in the majority of cases with high yields close to the theoretical maximum. Coupling becomes only infeasible for some metabolites under anaerobic growth with increased ATP maintenance demand. Importantly, if we consider the employed core model as a sub-network of a genome-scale network (with simplified biomass synthesis reaction) then the EMs/EVs of the core model will remain valid at genome-scale implying that the calculated maximum coupling yields from the core model represent lower boundaries at genome-scale.

In summary, we believe that the results outlined in this paper form the basis for understanding the theoretical limitations of product yields associated with growth-coupled strain design and will be a valuable tool in guiding the development of improved strain design methods. In addition, these results may help to identify potential structural changes in the network through the addition of certain (heterologous) reactions that can enable growth-coupled production for cases where the existing network does not allow it. Finally, the maximal guaranteed coupling yields that can be calculated with our method facilitate the objective evaluation of the technoeconomic potential of bio-based production of chemicals.

Acknowledgments

SK acknowledges funding from the German Federal Ministry of Education and Research (projects CYANOSYS II (FKZ 0316183D) and CASCOO (FKZ: 031A180B)) and the Federal State of Saxony-Anhalt (Research Center “Dynamic Systems: Biosystems Engineering”). RM would like to acknowledge funding from Alexander von Humboldt foundation, NSERC and BioFuelNet. We also thank Oliver Hädicke and Axel von Kamp (MPI Magdeburg) and Arne Reimers (FU Berlin) for fruitful discussions.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ymben.2015.05.006.

References

Bertsimas, D., Tsitsiklis, J.N., 1997. Linear optimization. Athena Scientific, Belmont, Massachusetts.
Burgard, A.P., Pharkya, P., Maranas, C.D., 2003. OptKnock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnol. Bioeng. 84, 647–657.
Campodonico, M.A., Andrews, B.A., Asenjo, J.A., Palsson, B.Ø., Feist, A.M., 2014. Generation of an atlas for commodity chemical production in *Escherichia coli* and a novel pathway prediction algorithm, GEM-Path. Metab. Eng. 25, 140–158.
Conrad, T.M., Lewis, N.E., Palsson, B.Ø., 2011. Microbial laboratory evolution in the era of genome-scale science. Mol. Syst. Biol. 7, 509.
Erdrich, P., Knoop, H., Steuer, R., Klamt, S., 2014. Cyanobacterial biofuels: new insights and strain design strategies revealed by computational modeling. Microb. Cell Factories 13, 128.
Feist, A.M., Zielinski, D.C., Orth, J.D., Schellenberger, J., Herrgard, M.J., Palsson, B.Ø., 2010. Model-driven evaluation of the production potential for growth-coupled products of *Escherichia coli*. Metab. Eng. 12, 173–186.
Fong, S.S., Burgard, A.P., Herring, C.D., Knight, E.M., Blattner, F.R., Maranas, C.D., Palsson, B.Ø., 2005. In silico design and adaptive evolution of *Escherichia coli* for production of lactic acid. Biotechnol. Bioeng. 91, 643–648.
Gagneur, J., Klamt, S., 2004. Computation of elementary modes: a unifying framework and the new binary approach. BMC Bioinf. 5, 175.
Hädicke, O., Klamt, S., 2010. CASOP: a computational approach for strain optimization aiming at high productivity. J. Biotechnol. 147, 88–101.
Hädicke, O., Klamt, S., 2011. Computing complex metabolic intervention strategies using constrained minimal cut sets. Metab. Eng. 13, 204–213.
Hunt, K.A., Folsom, J.P., Taffs, R.L., Carlson, R.P., 2014. Complete enumeration of elementary flux modes through scalable demand-based subnetwork definition. Bioinformatics 30, 1569–1578.
von Kamp, A., Klamt, S., 2014. Enumeration of smallest intervention strategies in genome-scale metabolic networks. PLoS Comput. Biol. 10, e1003378.
Kelk, S.M., Olivier, B.G., Stougie, L., Bruggemann, F.J., 2012. Optimal flux spaces of genome-scale stoichiometric models are determined by a few subnetworks. Sci. Reports 2, 580.
Kim, J., Reed, J.L., 2010. OptORF: optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. BMC Syst. Biol. 4, 53.
Klamt, S., Saez-Rodriguez, J., Gilles, E.D., 2007. Structural and functional analysis of cellular networks with CellNetAnalyzer. BMC Syst. Biol. 1, 2.

Klamt, S., Hädicke, O., von Kamp, A., 2014. Stoichiometric and constraint-based analysis of biochemical reaction networks. In: Benner, P, Findeisen, R, Flockerzi, D, Reichl, U, Sundmacher, K (Eds.), Large-Scale Networks in Engineering and Life Sciences. Springer, pp. 263–316.

Layton, D.S., Trinh, C.T., 2014. Engineering modular ester fermentative pathways in *Escherichia coli*. Metab. Eng. 26, 77–88.

Llaneras, F., Picó, J., 2010. Which metabolic pathways generate and characterize the flux space? A comparison among elementary modes, extreme pathways and minimal generators. J. Biomed. Biotechnol. 2010, 753904.

Mahadevan, R., Schilling, C.H., 2003. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metab. Eng. 5, 264–276.

Orth, J.D., Fleming, R.M., Palsson, B.Ø, 2010. 1—Reconstruction and use of microbial metabolic networks: the core Escherichia coli metabolic model as an educational guide. In: Karp, PD (Ed.), EcoSal—Escherichia coli and Salmonella Cellular and Molecular Biology. ASM Press, Washington DC 10.2.1.

Pey, J., Planes, F.J., 2014. Direct calculation of elementary flux modes satisfying several biological constraints in genome-scale metabolic networks. Bioinformatics 30, 2197–2203.

Ranganathan, S., Suthers, P.F., Maranas, C.D., 2010. OptForce: an optimization procedure for identifying all genetic manipulations leading to targeted overproductions. PLoS Comput. Biol. 6, e1000744.

Ranganathan, S., Tee, T.W., Chowdhury, A., Zomorrodi, A.R., Yoon, J.M., Fu, Y., Shanks, J.V., Maranas, C.D., 2012. An integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. Metab. Eng. 14, 687–704.

Rockafellar, R.T., 1970. Convex Analysis. Princeton University Press, Princeton.

Ruckerbauer, D.E., Jungreuthmayer, C., Zanghellini, J., 2014. Design of optimally constructed metabolic networks of minimal functionality. PLoS One 9, e92583.

Schuster, S., Fell, D.A., Dandekar, T., 2000. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. Nat. Biotechnol. 18, 326–332.

Shen, C.R., Lan, E.I., Dekishima, Y., Baez, A., Cho, K.M., Liao, J.C., 2011. Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. Appl. Environ. Microbiol. 77, 2905–2915.

Tepper, N., Shlomi, T., 2010. Predicting metabolic engineering knockout strategies for chemical production: accounting for competing pathways. Bioinformatics 26, 536–543.

Tervo, C.J., Reed, J.L., 2012. FOCAL: an experimental design tool for systematizing metabolic discoveries and model development. Genome Biol. 13, R116.

Terzer, M., Stelling, J., 2008. Large-scale computation of elementary flux modes with bit pattern trees. Bioinformatics 24, 2229–2235.

Trinh, C.T., Unrean, P., Srienc, F., 2008. Minimal *Escherichia coli* cell for the most efficient production of ethanol from hexoses and pentoses. Appl. Environ. Microbiol. 74, 3634–3643.

Trinh, C.T., Wlaschin, A., Srienc, F., 2009. Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism. Appl. Environ. Microbiol. 81, 813–826.

Trinh, C.T., Li, J., Blanch, H.W., Clark, D.S., 2011. Redesigning *Escherichia coli* metabolism for anaerobic production of isobutanol. Appl. Environ. Microbiol. 77, 4894–4904.

Urbanczik, R., 2007. Enumerating constrained elementary flux vectors of metabolic networks. IET Syst. Biol. 1, 274–279.

Urbanczik, R., Wagner, C., 2005. Functional stoichiometric analysis of metabolic networks. Bioinformatics 21, 4176–4180.

Wagner, C., Urbanczik, R., 2005. The geometry of the flux cone of a metabolic network. Biophys. J. 89, 3837–3845.

Xu, P., Ranganathan, S., Fowler, Z.L., Maranas, C.D., Koffas, M.A., 2011. Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA. Metab. Eng. 13, 578–587.

Yang, L., Cluett, W.R., Mahadevan, R., 2011. EMILiO: a fast algorithm for genome-scale strain design. Metab. Eng. 13, 272.

Yim, H., Haselbeck, R., Niu, W., Pujol-Baxley, C., Burgard, A., Boldt, J., Khandurina, J., Trawick, J.D., Osterhout, R.E., Stephen, R., Estadilla, J., Teisan, S., Schreyer, H.B., Andrae, S., Yang, T.H., Lee, S.Y., Burk, M.J., Van Dien, S., 2011. Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. Nat. Chem. Biol. 7, 445–452.

Zomorrodi, A.R., Suthers, P.F., Ranganathan, S., Maranas, C.D., 2012. Mathematical optimization applications in metabolic networks. Metab. Eng. 14, 672–686.