Optimal rate pathways

*Metabolic states with maximal specific rate carry flux through an elementary flux mode*

In collaboration with:
Han Peters, Joost Hulshof, Frank J. Bruggeman, Bas Teusink

Adapted from: *FEBS Journal* (2014), 281(6):1547-1555
Chapter 4

Abstract

Specific product formation rates and cellular growth rates are important maximization targets in biotechnology and microbial evolution. Maximization of a specific rate, i.e. a rate expressed per unit biomass amount, requires expression of particular metabolic pathways at optimal enzyme concentrations. In contrast to the prediction of maximal product yields, any prediction of optimal specific rates at genome scale is currently computationally intractable, even if the kinetic properties of all enzymes were available. Here we characterize maximal-specific-rate states of metabolic networks of arbitrary size and complexity, including genome-scale kinetic models. We report that optimal states are elementary flux modes (EFMs), which are minimal metabolic networks operating at a thermodynamically-feasible steady state with one independent flux. Remarkably, EFMs rely only on reaction stoichiometry, yet they function as the optimal states of mathematical models incorporating enzyme kinetics. Our results pave the way for the optimization of genome-scale kinetic models as they offer huge simplifications to overcome the concomitant computational problems.

4.1 Introduction

Since enzyme concentrations co-determine the rate of metabolic processes and, ultimately, microbial fitness, they should be adjusted when changes occur in the environment of microorganisms. Thus, the concentrations of the hundreds of metabolic enzymes that are required for cellular growth (Burgard et al., 2001; Mushegian and Koonin, 1996) should be (precisely) tuned to achieve adaptation of metabolism. This is exemplified by the dependency of cellular growth rate and formation rate of metabolic products on enzyme concentrations (Stoebel et al., 2008; Dong et al., 1995; Dekel and Alon, 2005; Eames and Kortemme, 2012; Snoep et al., 1995). The interplay between the concentrations of the proteins expressed and a reaction rate of interest is captured by the specific rate (or specific flux) of that reaction. It has as unit mol product · hr⁻¹ · (gram total protein)⁻¹ and quantifies cellular productivity, as it expresses the reaction rate of interest per gram biomass, i.e. the catalytic machinery. Expression of enzymes that do not contribute to this reaction will therefore reduce the specific reaction rate, as is also shown experimentally (Stoebel et al., 2008; Dong et al., 1995; Dekel and Alon, 2005; Eames and Kortemme, 2012; Snoep et al., 1995).

The formation rate of a metabolic product expressed per gram biomass and the specific growth rate of a cell are both examples of specific (reaction) rates. The latter is often under natural selection in microbial evolution in the wild (Schuster et al., 2011), as an increased specific growth rate of a mutant cell causes it to outgrow the resident population of cells. Also in laboratory or biotechnological settings, such as serial dilutions, fed-batch or chemostats, faster growing mutants will increase their frequency and finally take over the population. However, not in all situations is the specific growth rate selected for. For instance, the gram biomass obtained per mol substrate, i.e. the biomass yield, can be under natural selection (Bachmann et al., 2013). This evolutionary scenario can be predicted by flux balance analysis (FBA; Orth et al. (2010)). Other than biomass yield or specific growth rate, network robustness and homeostasis can be alternative targets of natural selection (Kitano, 2004).
Our interest is in understanding the maximisation of specific reaction rate, both from a biotechnological and evolutionary perspective. This maximization requires the optimal partitioning of protein over metabolic processes, such that all proteins contribute to the target flux and at levels that maximize this flux (Heinrich and Klipp, 1996; Berkhout et al, 2013b; Dekel and Alon, 2005). What characterizes the topology of the metabolic subnetworks that maximize the specific flux of a target reaction is not known. This problem is solved here.

Maximization of a specific flux boils down to a complex nonlinear optimisation problem, which requires the kinetics of all metabolic enzymes involved. We emphasise that this optimization problem is quite different from those associated with popular stoichiometric modeling approaches, such as FBA. In FBA, a particular flux is optimised under the constraint of several other flux values, typically including the uptake rate (Schuster et al, 2008). Hence, in FBA no kinetics is involved and the system is linear with respect to the optimization variables—the fluxes—leading to computationally tractable problem for genome-scale metabolic networks. Thus, FBA predicts only maximal yield strategies. The metabolic network topologies that optimize yields have recently been characterised (Kelk et al, 2012). In this work, we simplify the nonlinear optimization of specific reaction rates in large reaction networks. We prove that the optimal network is an elementary flux mode (EFM), which is—surprisingly—a pathway defined by stoichiometry only.

### 4.2 Results

Formally, the specific flux (or specific rate), $q_r$, of metabolic reaction $r$ is defined as (Berkhout et al, 2013b),

$$q_r = \frac{v_r}{e_T},$$  \hspace{1cm} (4.1)

where $e_T$ denotes the total protein content in the system in gram total protein, and $v_r$ is the flux value in mol · hr$^{-1}$. We note that scaling with total protein content is for some applications more useful, e.g. when studying a single metabolic pathway. However, for studying growth of cells it is more convenient to scale with respect to gram dry weight of biomass, which equals scaling with respect to total cellular protein when the protein density of cells is constant (which is a realistic scenario). Thus, in the case of the specific growth rate $e_T$ is the total cellular protein and $v_r$ the biomass production rate or, equivalently, the protein synthesis rate. In this work, we characterize the metabolic steady-state states that optimize the specific-flux of a target reaction, i.e. $q_r$, given boundary conditions such as enzyme kinetics, fixed nutrient concentrations and metabolic reaction stoichiometry. We emphasise that for prediction of such a specific flux, enzyme kinetics need to be considered (Molenaar et al, 2009).

As a first example, we discuss the maximization of the specific flux, $q_4$, of metabolite $W$ in a toy pathway depicted in Figure 4.1A. We assigned kinetic rate equations and corresponding parameters to each enzyme and computed the optimal states of the network when the specific flux is maximal for different sets of parameter values. Figure 4.1B shows the optimal states in flux space for different model parameters. To our initial surprise, all the optimal solutions gather along three lines, which correspond to the EFMs (Schuster et al, 2000) of this network (Figure 4.1C). This is remarkable, because EFMs rely only on reaction stoichiometry, irrespective of kinetic
properties. This is the main finding of this work: optimization of metabolism for a particular \( q_r \) is always achieved by an EFM that uses reaction \( r \), regardless of enzyme kinetics, e.g. reversibility, cooperativity and allosteric regulation. Furthermore, we see that only one of the three EFMs attains the highest yield (Figure 4.1D). Therefore, a yield optimization method such as FBA will always predict EFM 2, while, as we can see in Figure 1A, both EFM 1 and 3 can lead to optimal specific production rates for certain parameter sets. This means that the fact that EFMs 1 and 3 would be classified as suboptimal in a FBA does not mean that they cannot be a result of the optimization of the network for a specific rate.

In the main text we will discuss the understanding and implications of this result, while an outline of the mathematical proof can be found in Box 1. We start with a very tractable, branched network (Fig. 4.2A). Let us fix the input flux, \( J \), and describe reaction 1 and 2 with Michaelis-Menten kinetics; \( v_1 = e_1 f_1(X) \) and \( f_1(X) = \frac{k_{cat 1} X}{(K_M 1 + X)} \) and the same for enzyme 2. Next, the specific flux, \( \frac{J}{e_1 + e_2} \), is maximized by minimizing the total enzyme amount, \( e_1 + e_2 \). This system has two EFMs: One containing the reactions \( J \) and \( v_1 \) and the other containing the reactions \( J \) and \( v_2 \). Hence, in the optimal state, our proof makes us expect flux through one branch. We carry out the optimization graphically. We rewrite the objective as: \( e_1 + e_2 = \frac{v_1}{f_1} + \frac{v_2}{f_2} \) and the
Optimal rate pathways

steady-state flux relation $v_2 = J - v_1$ simplifies it to,

$$e_1 + e_2 = \alpha(X) \cdot v_1 + \beta(X),$$

with:

$$\alpha(X) = \frac{K_{M,1} + X}{k_{cat,1} \cdot X} - \frac{K_{M,2} + X}{k_{cat,2} \cdot X}$$

$$\beta(X) = \frac{J K_{M,2} + X}{k_{cat,2} \cdot X}$$

(4.2)

For any fixed value of $X$, the objective depends linearly on $v_1$. In Figure 4.2A, the invested enzyme amount is shown for three values of $X$. The maximal $\frac{J}{e_1 + e_2}$ is attained when $e_1 + e_2$ is minimal and this occurs either when $J$ equals $v_1$ or $v_2$. Since we can conclude this for any value of $X$, an optimal state with both branches active cannot occur: a single EFM is optimal. Different kinetic parameters do not influence this conclusion and neither does product inhibition, for which we show a detailed example in the Supplementary material.

In Fig. 4.2B we visualise the optimal solution space for an example with allosteric regulation. Every straight line in this figure represents a steady-state solution, not necessarily optimal, for one set of steady-state metabolite concentrations. Therefore, the combination of all lines represents the entire solution space. The lower boundary (thick lines in Figure 4.2B) expresses the maximal specific-flux solution as function of the fractional contribution of each EFM. This figure again illustrates the optimality of the usage of one EFM as only with one EFM the minimal enzyme amount is attained in Fig. 4.2B. In some situations, such as the top panel in Fig. 2B, another EFM can have a nearly identical optimum but this depends on the exact kinetic parameterisation. What is surprising is that mixtures of EFMs are in some cases even less optimal than a single, suboptimal EFM.
Figure 4.2. Optimal specific pathway flux in branched networks is achieved when only one branch is active.

A A simple network with a fixed flux $J$ that produces the metabolite $X$, which is consumed by $v_1$ and $v_2$. The right panel shows the required $e_T = e_1 + e_2$ to reach the fixed flux $J$. The dots indicate where $e_T$ is minimal at a given, fixed $X$. The rates $v_1$ and $v_2$ are modeled with irreversible Michaelis Menten (MM) kinetics ($v = v_{\text{cat}} \cdot \frac{X}{K_M + X}$) with $J = 10$, $K_{M,1} = 5$, $K_{M,2} = 0.5$, $k_{\text{cat},1} = 2$ and $k_{\text{cat},2} = 1$).

B A branched network with allosteric inhibition of $X_1$ on $v_3$ where the pathway substrate ($X$) is considered fixed. The right panel shows the required $e_T$ to achieve a fixed objective flux ($v_1$), where the different lines correspond to different, fixed metabolite concentrations of $X_1$ and $X_2$. Depending on the kinetic parameters (upper panel: $K_{i,3} = 4$ and $k_{\text{cat},4} = 2$ and lower panel: $K_{i,3} = 10$ and $k_{\text{cat},4} = 7$) the minimal $e_T$ as a function of the ratio of the branches is either peaked (concave-down) or monotonically decreasing (still concave). The dashed lines are calculated with the optimal metabolite concentrations for $v_3$, the slope of which indicates whether mixed strategies are least optimal. The rates are modeled with irreversible MM kinetics, with product inhibition for $v_1$ and $v_2$ (a term $\frac{1}{1+I/K_i}$ is added to the numerator) and allosteric inhibition for $v_3$ (the rate equation is multiplied by $\frac{1}{1+I/K_i}$), with $K_{M,1} = 5$, $K_{i,1} = 20$, $k_{\text{cat},1} = 3$, $K_{M,2} = 0.5$, $K_{i,2} = 5$, $k_{\text{cat},2} = 2$, $K_{M,3} = 8$, $k_{\text{cat},3} = 8$ and $K_{M,4} = 0.5$. The left bottom panel shows that at low substrate concentrations, the network can reach a higher specific flux, $q_1$, with branch $v_3$, while at a high substrate concentration branch $v_4$ can lead to a higher $q_1$. $q_i$ indicates $q_i$ with the use of branch $i$). A slightly simplified explanation is the following: because the affinity of $e_3$ for its substrate is much lower than that of $e_4$, an increased pathway substrate concentration benefits $v_4$ more. Parameter settings: $K_{i,3} = 3$, $K_{M,3} = 0.1$, $k_{\text{cat},3} = 6$, $K_{M,4} = 40$, $k_{\text{cat},4} = 8$. 
Optimal rate pathways

BOX 1: EMFs are optimal metabolic states for specific flux optimizations

To optimize a specific flux, \( q_v \), we can fix the rate \( v_i \) and minimize the \( e_f \) needed to attain this level of \( q_v \) (4.1) (Heinrich and Schuster, 1996). We consider the enzyme kinetics, reaction stoichiometry, steady-state requirement and reaction thermodynamics. The entire optimization problem is formulated as,

\[
\text{min}_{x \in \mathbb{R}^n} \left\{ \sum_{i=1}^{r} e_i \right\}, \quad \text{subject to:} \begin{align*}
N \cdot v &= 0, \quad \forall v : v_i = e_i f_i(x), \quad \forall v : e_i \geq 0, \quad v_r = 1
\end{align*}
\]

where \( v_i \) means for all \( i \). The thermodynamic constraints (i.e. \( v_i \geq 0 \) or \( v_i \leq 0 \) for certain \( i \)) are included in the enzyme kinetics as the corresponding \( f_i \)-functions take only positive or negative values.

This minimization problem states that the total enzyme concentration is minimized by finding optimal values for the variables, the metabolite (x) and enzyme concentrations (e), to reach an objective flux \( v_r \) equal to 1 given the constraints. Biochemistry dictates that the rate of each enzyme depends linearly on the enzyme concentration; \( v_i = e_i f_i(x) \), with as exception the occurrence of metabolite channeling (Huang et al., 2001). The optimization problem can also be stated in terms of the variables \( x \) and \( v \) (using \( e_i = \frac{v_i}{f_i(x)} \)),

\[
\min_{x \in \mathbb{R}^n} \left\{ \sum_{i=1}^{r} e_i \right\}, \quad \text{subject to:} \begin{align*}
N \cdot v &= 0, \quad \forall v : \frac{v_i}{f_i(x)} \geq 0, \quad v_r = 1
\end{align*}
\]

To characterize the optimal state, it is instructive to study the optimization problem when the metabolite concentrations are at their (globally) optimal concentrations, denoted by the vector \( x_o \). Then, the inverse kinetic functions \( c_i = 1/f_i(x) \) become fixed and the reaction rates, \( v \), remain as optimization variables,

\[
\min_{x \in \mathbb{R}^n} \left\{ \sum_{i=1}^{r} c_i v_i \right\}, \quad \text{subject to:} \begin{align*}
N \cdot v &= 0, \quad \forall v : c_i v_i \geq 0, \quad v_r = 1
\end{align*}
\]

This is a linear program (LP). We can simplify this LP by splitting the reversible reactions into a forward and backward rate and defining all rates as positive. This introduces a new stoichiometry matrix \( \tilde{N} \) and rate vector \( \tilde{v} \),

\[
\min_{x \in \mathbb{R}^n} \left\{ \sum_{i=1}^{r} c_i v_i \right\}, \quad \text{subject to:} \begin{align*}
\tilde{N} \cdot \tilde{v} &= 0, \quad \forall \tilde{v} : c_i \tilde{v} \geq 0, \quad v_r = 1
\end{align*}
\]

The forward and backward rate of each reversible reaction, \( j \), obtains the \( c \) value of the original reversible rate, \( c_j \) and \( f \) equals \( r \) plus the number of reversible reactions. In the optimal state, the forward and backward reaction will never be used simultaneously because this increases, rather than reduces, the objective. The feasible flux space is the cone \( C \) defined in equation B4 (Figure 4.3AB). \( C \) is characterized by its extreme rays. Gagneur and Klamt (2004) proved that these rays are the EMFs of the original metabolic network. The intersection \( C \cap \{ v_r = 1 \} \) defines the solution space, (Figure 4.3CD),

\[
P = C \cap \{ v_r = 1 \}
\]

The extremal points of the polyhedron \( P \) are the extremal rays of \( C \): the EMFs (Figure 4.3D). Next, the linear function \( \sum_{i=1}^{r} c_i v_i \), is minimized over \( P \). The minimum of this function occurs at an extremal point of \( P \). Thus, the optimal state is an EMF of the original metabolic network that contains \( v_r > 0 \). This is the key result of this work and it holds regardless of the complexity and the kinetics of the metabolic network. The polyhedron \( P \) can in principle be unbounded if cycles occur in the network; generally, the target flux is an efflux and cycles will not be relevant. Also, when multiple EMFs have the same objective value, a new polyhedron describes the optimal solution space. However, this is very unlikely as the objective value depend on the kinetics of all the active enzymes within the EMF. Hence, we limit ourselves to one optimal EMF. (Our method does identify alternative solutions, if they occur.) The optimal state of an EMF can be calculated. A defining property of any EMF is that one flux value is required to determine all its flux values (Gagneur and Klamt, 2004). Since we set \( v_r = 1 \), all the rate values, \( c_i \), can be determined:

\[
\forall v_i \in \text{EFM} : e_i = \frac{v_i}{f_i(x)}
\]

Next, we can determine the optimal metabolite vector \( x_o \) that minimizes the objective \( \sum_{i=1}^{r} e_i / f_i(x) \), by numerical optimization and the optimal enzyme levels follow from: \( \forall v : e_i = \frac{v_i}{f_i(x)} \).

The optimal EMF has maximal \( q_v \). In the Supplementary material, we provide more detailed mathematical definitions and proofs.
Figure 4.3. Schematic representation of the optimization of a specific pathway flux $q_r$. A We have decomposed all reactions in a forward and backward reaction to make all reaction rates ($\bar{v}_1 \ldots \bar{v}_r$) positive. To sketch the solution space we draw the planes for $\bar{v}_i = 0$. B The solution space is a pointed cone $C$ and its extreme rays are the intersections of the planes $\bar{v}_i = 0$. The extreme rays coincide with the EFMs. C To optimize the specific pathway flux, we fix the objective flux $\bar{v}_r = 1$ and intersect the cone with $\bar{v}_r = 1$ to obtain the solution space $P$, a polyhedron. D Next, we minimize the total enzyme concentration necessary to obtain $\bar{v}_r = 1$. $eT$ is a linear function of $e$ and therefore also in $\bar{v}$ (at $x_0$) and its minimum is obtained at an extreme point of the polyhedron $P$ (red dots), which is at an extremal ray of the cone, which is an EFM with $\bar{v}_r = 1$.

With hindsight it is intuitive that optimal specific-flux-states are EFMs. EFMs are minimal routes in the sense that no reaction is redundant; no reaction can be removed without violating the steady-state requirement. This partially explains why EFMs are the optimal states for specific flux maximization, because networks with redundant enzymes can attain a higher specific flux by redistributing protein over the minimal set of enzymes to sustain the target process at steady state.

We have shown that optimal specific-flux-states are EFMs and we will next reason which EFM out of all is optimal. The EFM with the minimal number of enzymes does not need to be the optimal EFM. If an EFM contains enzymes with low catalytic capacities ($k_{cat}$'s), a longer EFM with high $k_{cat}$'s can attain a higher flux for the same amount of invested enzyme. Nor does the EFM with the highest yield of the target product has to be the optimal network for the specific flux (as we discussed above). A shorter EFM with a lower yield could have a higher flux for the same amount of invested enzyme, which is illustrated in Figure 1D where the yields are given for each EFM. Finally, EFMs are not necessarily linear pathways. For instance, the EFMs that produce
Optimal rate pathways

biomass from extracellular nutrients will be highly branched. We conclude that all EFMs can in principle be optimal specific-flux-states and kinetic—not stoichiometric—properties define which EFM attains a maximal specific flux.

A consequence of the previous reasoning is that when nutrient concentrations change, switches between EFMs may occur to preserve optimality. This is evident from Figure 1B, where the optimal states at different nutrient levels (green dots) spread over the three EFMs, and from Figure 3B, which shows the relative specific flux as function of the nutrient concentration. Metabolic switches are ubiquitous in microbiology (Goel et al., 2012) and could, therefore, be related to specific growth-rate maximization (Molenaar et al., 2009). In addition, catabolite repression, which leads to the use of only one carbon source in stead of mixtures of carbon sources, could be the result of specific growth-rate maximization. However, mixed EFM usage could contribute to fitness in changing environments or result from non-genetic phenotypic stochasticity (Eldar and Elowitz, 2010).

Our theory is not limited to the optimization of one specific-flux. If two fluxes, say $j$ and $k$, should be maximized at a fixed ratio, e.g. $v_j/v_k = \beta$, the theory applies to the original metabolic network with those two reactions replaced by one aggregate reaction equation (with reaction equation we mean the chemical conversion balance: e.g. $3x + 2y \rightleftharpoons 2z$). This new reaction equation has a stoichiometry that derives from summing the reaction equations of $j$ and $k$ as: $j + \beta k$. The specific flux of this new reaction then becomes the optimization target. The optimal EFM is an EFM of the modified metabolic network. In this manner, the specific growth rate of an organism can be maximized under the condition that a particular product is made at a fixed yield. This is a relevant extension for biotechnological applications that aim to "uncouple" growth and product formation.

An example of the application of this method to a more realistic case is shown in a simplified kinetic network of succinate production (Fig. 4.4). We show a more extended example in the Supplementary materials that does not reduce metabolic pathways to single supra-reactions. We will explain the realistic scenario of how to find the optimal protein allocation to maximize the specific production flux of succinate $q_8$. First, we determine the EFMs of this network (Fig. 4.4B). Next, we optimize the protein allocation for every EFM separately by a numerical minimization of $e_T$ with $v_8 = 1$ and only the metabolite concentrations as variables, which is a very fast computation. Next, we compare the values of the resulting $q_8$, obtained by dividing $v_8 = 1$ by the result of the minimized value of $e_T$, for every EFM. Finally, we select the EFM with the highest $q_8$. Without the use of our theory the optimization of this network would already be troublesome, because the entire network would have to be optimized at once instead of every EFM individually. Again, we can observe a shift in the optimal EFM upon change in an extracellular substrate concentration. We can distinguish different cases, such as: (i) an elementary mode that includes the increased substrate can be preferred over one that does not (Fig. 4.4C top panel) or (ii) a different elementary mode that uses the same substrate can become the preferred EFM (Fig. 4.4C bottom panel).
**Figure 4.4. Lumped example of succinate production illustrates condition-dependent sugar preference.** 

A Network topology of the network for the production of Succinate from either glucose (Glu), galactose (Gal) or lactate (Lac) with objective succinate production flux $v_8$. Fixed external metabolites are underlined. We have assumed transhydrogenase activity such that the pentose phosphate pathway (PPP) can be assumed to produce NADH. We have assumed that one C3 molecule from glycolysis is used for growth. B Elementary modes. Thick arrows indicate a double flux. C Maximal specific glucose flux of each EFM as fraction of the optimal EFM for different sugar concentrations. An EFMs using galactose becomes optimal upon an increase in galactose (top panel). Upon increase in glucose, the optimal EFM switches from one mode using glucose to another mode using glucose (bottom panel). (For the simulations a low substrate concentration is 1% of the $K_M$ and a high substrate concentrations is 10 times the $K_M$. All simulations use a different set of parameters.

**4.3 Discussion**

In this work we have achieved two main results: i. we have characterized optimal specific-flux states of metabolic networks as single EFMs and ii. we have shown how we can use this result for optimization of kinetic models of metabolic networks. Next, we will discuss the influence of activation of a reaction in one EFM by a metabolite used in another EFM (cross-EFM activation), additional physicochemical constraints, and identical EFMs.

Firstly, it might seem that cross-EFM activation (e.g. Fig. S4 EFM3), is not covered by our theory. This would suggest that flux through two EFMs can be optimal, where one EFM is just there to produce the activating metabolite. This can be achieved by investing a negligible amount of enzyme in the EFM that produces the activating metabolite. The required enzyme amount is determined by the dilution of the metabolite concentration by cell growth. As fluxes of metabolites in metabolism are generally much faster than the dilution of growth, the additional enzyme investment is negligible.

Secondly, cells might be constrained by additional physicochemical constraints, such as intracellular or membrane space. These constraints can lead to mixed metabolic strategies when the biomass yield is optimised in FBA (Shlomi et al, 2011; Vazquez et al, 2010). Schuster et al (2011) showed in a toy model with linear kinetics and constraints on fluxes that an EFM gives rise to the optimal yield state, but that a mixed strategy can optimize specific growth rate. Generally, we can consider those constraints into our formalism, e.g. including a kinetic description of lipid biosynthesis pathways in the case of a membrane constraint (Molenaar et al, 2009). This leads to a new optimisation problem, which we can write in the form of equation B1, and...
therefore the optimal solution is again a single EFM.

Lastly, in theory we cannot exclude that two EFMs have exactly identical optimality properties. Since we include enzyme kinetics, we consider this situation as very unlikely. We do not expect it to occur in any realistic case. However, it is interesting to note that when the differences in maximal specific flux between EFMs are small, there might not be a strong selection pressure on expressing the optimal EFM. Therefore we might expect to find suboptimal EFMs in this case, and perhaps even combinations of the two EFMs. Also in FBA the focus is solely on optimal yield solutions, while close to optimal solutions can easily arise (Schuetz et al., 2012).

Applications of our method to genome-scale models require EFM enumeration followed by optimization of the kinetic models of the EFMs. EFM enumerations are a computationally hard task due to the enormous number of EFMs (Terzer and Stelling, 2008). The number of EFMs that we require, however, is much smaller than their total number; we require only EFMs defined at one growth condition and then only those that use the target reaction, while classical EFM computation considers all growth conditions and reactions. How to compute all EFMs with a specific target flux is still an open question.

Besides yields of products or biomass, specific production rates are of biotechnological interest. Also in studies of metabolic adaptation strategies, such as in cancer or laboratory evolution experiments, specific rates, and specific growth rate in particular, are the targets of selection. Our findings indicate that selection for growth rate forces cells to use the "cheapest" and "fastest" EFM. We again emphasize that stoichiometric approaches cannot find optimal specific flux states, as these models contain no kinetic information and rely on input flux constraints. Our findings enable the optimization of kinetic models, even at genome scale. Without this result, the entire model needs to be optimized, which is a huge nonlinear constraint optimization problem that is likely impossible to solve with current numerical methods. Our result simplifies this task significantly to the optimization of only EFMs that use the target reaction. This reduces the optimization by several orders of magnitude in the number of algebraic equations. Moreover, because the optimisation problem is concave (as shown in Figure 3B), there is a realistic danger of reaching a local optimum with a naive optimization strategy. Once the optimal state is identified, the optimal metabolite concentrations belonging to an EFM can be independently determined from the optimal enzyme levels (see Box).

Independent from us, Müller et al. (2013) have also concluded that EFMs are the optimal metabolic networks for specific-flux maximisation (Müller et al., 2014). Their approach requires understanding of more complex mathematics, i.e. oriented matroid theory. In contrast, our approach follows in a few steps from the mathematical optimisation problem and definition of elementary flux modes (see Box). We think that both papers give valuable insight into a basic feature of metabolic networks optimised for a specific-flux.

In conclusion, our theoretical results allow the identification of optimal specific-flux states for large, realistic systems, which is now only limited by the availability of kinetic data. Further technological developments should provide such kinetic data at larger scale. Meanwhile, Monte Carlo parameter sampling can be used to deal with parameter uncertainties, which again, is only feasible with the results in this paper. Hence, kinetic optimisation of relevant parts of metabolism,
such as central metabolism and potentially engineered branches to interesting products, is now within reach.

Acknowledgements

We would like to thank Evert Bosdriesz, Michael Ferris, Pınar Kahraman, Timo Maarleveld, Douwe Molenaar, Brett Olivier and Robert Planqué for critical discussions. This project was carried out within the research programme of the Kluyver Centre for Genomics of Industrial Fermentation, which is part of the Netherlands Genomics Initiative / Netherlands Organization for Scientific Research. FJB acknowledges funding by NWO-VIDI number 864.11.011. HP was supported by a SP3-People Marie Curie Actions grant in the project Complex Dynamics (FP7-PEOPLE-2009-RG, 248443).

Supporting information

Additional supporting information may be found in the online version of this article at the publisher’s website:

- Doc. S1. Theoretical background.
- Doc. S2. Additional models and model descriptions.
- Doc. S3. Additional and detailed proofs.
- Fig. S1. Example toy network with product inhibition.
- Fig. S2. Example of a more extensive metabolic network with multiple EFMs and allosteric interactions.
- Fig. S3. Illustrative graphical examples of maximization of the specific flux $q_r$.
- Table S1. Parameters for the main text model of Fig. 4.1.