Essays in Biochemistry (2024) **68** 41–51
https://doi.org/10.1042/EBC20230045

Review Article

Understanding and computational design of genetic circuits of metabolic networks

Alicia Berkvens[1], Luis Salinas[1], Maaike Remeijer[1], Robert Planqué[2], Bas Teusink[1] and Frank J. Bruggeman[1]

[1]Systems Biology Lab, A-LIFE, AIMMS, VU University, Amsterdam, NL.; [2]Department of Mathematics, Amsterdam Center for Dynamics and Computation, VU University, Amsterdam, NL.

Correspondence: Frank J. Bruggeman (f.j.bruggeman@vu.nl)

The expression of metabolic proteins is controlled by genetic circuits, matching metabolic demands and changing environmental conditions. Ideally, this regulation brings about a competitive level of metabolic fitness. Understanding how cells can achieve a robust (close-to-optimal) functioning of metabolism by appropriate control of gene expression aids synthetic biology by providing design criteria of synthetic circuits for biotechnological purposes. It also extends our understanding of the designs of genetic circuitry found in nature such as metabolite control of transcription factor activity, promoter architectures and transcription factor dependencies, and operon composition (in bacteria). Here, we review, explain and illustrate an approach that allows for the inference and design of genetic circuitry that steers metabolic networks to achieve a maximal flux per unit invested protein across dynamic conditions. We discuss how this approach and its understanding can be used to rationalize *Escherichia coli*'s strategy to regulate the expression of its ribosomes and infer the design of circuitry controlling gene expression of amino-acid biosynthesis enzymes. The inferred regulation indeed resembles *E. coli*'s circuits, suggesting that these have evolved to maximize amino-acid production fluxes per unit invested protein. We end by an outlook of the use of this approach in metabolic engineering applications.

Introduction

Microbial cells adapt their protein expression to their environment to ensure that they achieve physiological states with a competitive fitness [1,2]. An emerging concept in the field is that the molecular circuits needed to cope with metabolic adaptation and current stresses are expressed such that growth is fast, stress tolerance is matching the environment, and/or adaptive changes to new conditions are rapid [3].

Such changes in physiology vary from minor adjustments – tuning of already-expressed protein concentrations – to large-scale qualitative changes such as shifts in metabolism (from respiration to fermentation) [4–6], from growth to dormancy [7,8], from division to spore formation [9], etc. Physiological adaptation is the result of regulatory circuitry that integrates external and internal cues to initiate an adaptive response [2]. This circuitry has been moulded by evolution, allowing fitter genotypes to outcompete inferior ones through mutation and selection [10].

Since cells have finite biosynthetic resources [11,12], suboptimal expression of needed proteins, such as heterologous pathways expressed for a biotechnological purpose [13], or expression of now-unneeded proteins [11], reduces growth rate and potentially leads to a suboptimal fitness in the long run. Random mutations may then lead to the adjustment of suboptimally expressed systems and fixation of better-adapted genotypes [14]. That cells have finite biosynthetic resources is illustrated by the observation that different cellular tasks trade off against each other [3,15]: fast growing cells are less stress tolerant and less adaptive than slow growing ones [16–18]. It, therefore, appears that regulatory circuitry is comprehensible in terms of its contribution to fitness. Such comprehension, however, is not straight-forward. First, we cannot easily evaluate the performance of regulatory circuits in the context of cellular fitness

Received: 30 October 2023
Revised: 28 March 2024
Accepted: 28 March 2024

Version of Record published:
30 April 2024

Essays in Biochemistry (2024) **68** 41–51
https://doi.org/10.1042/EBC20230045

because we lack sufficient information about the associated evolutionary history [3]. Moreover, fitness is not always directly related to the instantaneous growth rate of a cell, but, instead, is mathematically equivalent to the average growth rate over longer tracks of evolutionary history [3,10]. In addition, appreciating the fitness benefits of a regulatory circuit is complicated, because fitness is inherently dictated by whole-cell performance and not by any of the cell's subnetworks. As long as we cannot translate whole-cell fitness to a performance measure for subnetworks, evaluation of the fitness benefits of regulatory circuits is problematic. In this review, we bypass the first problem by focusing on constant conditions, for which fitness does equate to the instantaneous steady state growth rate [3], and we offer a solution for the second problem (in the third section of the Results section).

Since evaluation of the fitness benefits of regulatory circuitry is challenging, we generally lack a satisfactory understanding of regulatory circuity of metabolism, even in the best understood microbes such as *Escherichia coli* and for the best understood biochemical systems such as glycolysis. In particular, we generally do not know what the exact objective is of regulatory networks. Human-made devices, such as a thermostat that keeps the average temperature in a room constant, have clear objectives – in this case, minimizing the deviation of the actual room temperature from the set-point. Accordingly, any satisfactory understanding of a regulatory network involves understanding its (regulatory) objective (or function or the fitness purpose). Without such a clear objective for regulatory circuits, we can only infer how they work, but not why they work as they do.

For instance, why are nearly all amino-acid biosynthesis pathways inhibited by their products [19]? Even if the reason is homeostasis of those product concentrations, as is often hypothesized and can be shown to be in agreement with enzyme kinetics [20], we lack an understanding of how this homeostasis is beneficial for whole-cell fitness. Why is pyruvate kinase [21], and often also lactate dehydrogenase [22], activated by fructose l,6-bisphosphate in so many species? Why is *E. coli*'s isocitrate dehydrogenase (in the TCA cycle) regulated by (de)phosphorylation by an associated phosphatase and kinase, respectively activated and inhibited by 3-phosphoglycerate (a glycolytic intermediate) [23]? Why does 3-phosphoglycerate not directly regulate isocitrate dehydrogenase? Why is pyruvate dehydrogenase regulated by (de)phosphorylation in *Saccharomyces cerevisiae* [24] and humans, while it is regulated by the NADH/NAD ratio in many bacteria (e.g., *E. coli* [25]), giving rise to effectively the same metabolic regulation? Why is phosphofructokinase inhibited by citrate in so many species [26]?

What are the objectives of these regulations and how do they contribute to the fitness of a cell? To address this, do we need to consider the entire metabolism of the cell, or can we obtain significant understanding by focusing only on the kinetics of subnetworks, for example, glycolysis and the citric acid cycle?

Regulation is an inherently dynamic concept [27,28], as it involves steering a circuit in a direction associated with reaching an objective. Studying regulation therefore requires dynamic data and dynamic models of metabolism. This implies that current approaches to metabolism modeling, such as flux balance analysis [29] and steady-state kinetic models [30], are not suited for understanding regulation. We also need corresponding dynamic experimental data of fluxes and concentrations of metabolites and enzymes, but this is experimentally still difficult (pioneering exceptions exist, e.g., [31–33]). As long as we lack such dynamic data and understanding of regulation, biotechnology and quantitative physiology is stuck with steady-state thinking and concepts.

In this review, we present a theoretical, model-based approach to rationalize and design the regulation of metabolic-gene expression by assuming as objective the maximization of metabolic flux per unit invested enzyme (specific flux maximization) [34] or, equivalently, the minimization of the amount of enzyme needed to achieve a target value of a metabolic flux (enzyme-cost minimization) [35]. We focus on three examples, based on [36–39]: a toy model to illustrate the main concepts, regulation of ribosomal expression and protein synthesis rate, and regulation of amino-acid biosynthesis flux.

We have chosen to discuss concrete small examples, to make the approach accessible and easy to follow. The theory, however, not only works for small networks. It can be formulated in general terms, for large metabolic networks that incorporate kinetics of all participating enzymes [38,39]. This 'qORAC' theory, for q (denoting a specific flux with unit $mol(gr\ drw)^{-1}hr^{-1}$) Optimization by Robust Adaptive Control, gives a general way to infer gene regulatory circuits that can steer a metabolic network to optimal states of maximal flux per unit invested protein, while remaining robust to changes in conditions. The inputs to the regulatory circuit are concentrations of metabolic intermediates which bind transcription factors. The qORAC theory indicates that the number of such metabolites must be at least equal to the number of environmental parameters the network needs to be robust to. For example, if a nutrient concentration and temperature are the two environmental parameters then two metabolic intermediates suffice as inputs to a gene regulatory circuit controlling metabolic enzyme expression.

Essays in Biochemistry (2024) **68** 41–51
https://doi.org/10.1042/EBC20230045

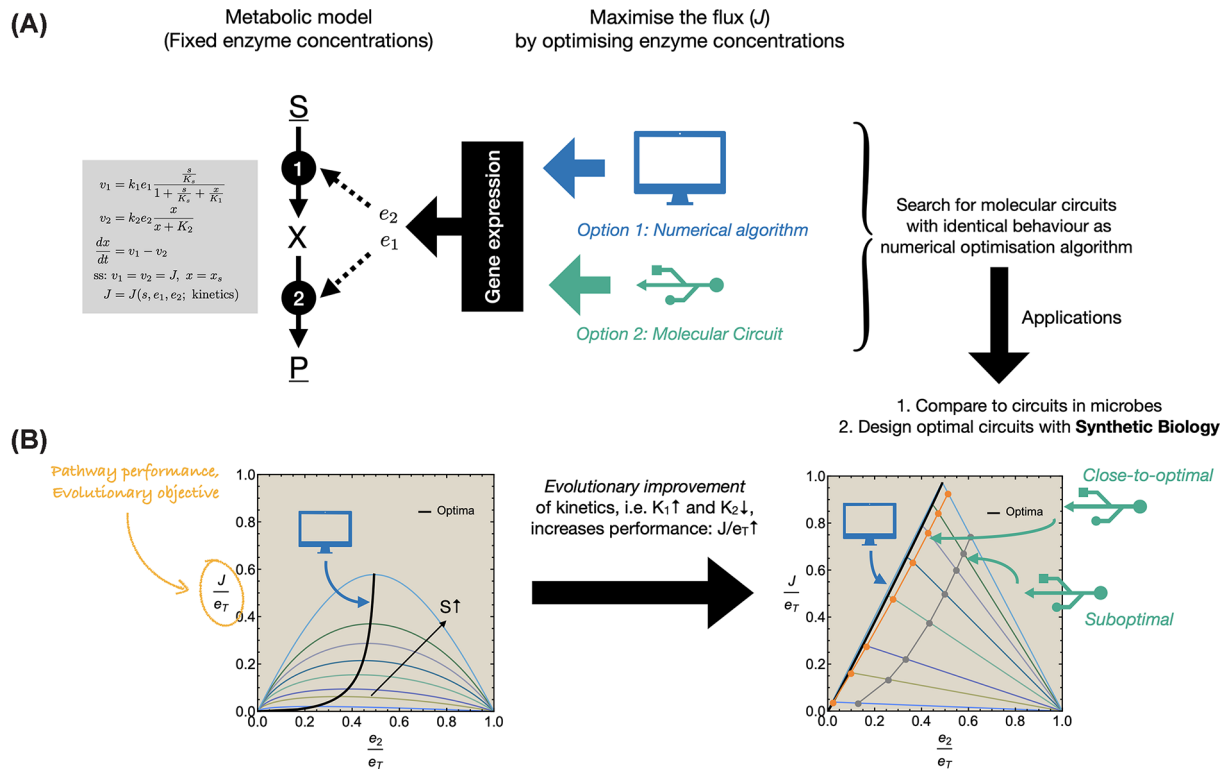

Figure 1. Illustration of the maximization of steady-state pathway flux per unit invested protein by optimal gene expression, either with a numerical algorithm or a molecular genetic circuit

(**A**) Overview of the rationale. We consider a mathematical model of a metabolic network in which an intermediate metabolite *X* inhibits its own production and stimulates its conversion into end product *P*. We aim to maximize its steady-state performance (J/e_T) as function of an environmental condition (e.g., a nutrient concentration) by optimizing the expression of its enzymes. The optimization can be carried out by an optimization algorithm (symbolized by the computer) or, remarkably, also by a molecular genetic circuit (symbolized by a circuit) that controls the expression of the pathway's enzymes as a function of one intermediate of the metabolic pathway. (**B**) Illustration of simulation results. We consider in the left and right figure the relationship between J/e_T and e_2/e_T (with e_2 as the concentration of enzyme 2) for different values of the environmental parameter *S*. The black lines connect the optimal protein fractions, as a function of *S*. These plots were obtained by calculating the steady state of the network as function of e_2 (with $e_1 = e_T - e_2$) for different values of S. The black line connects the maximal values of J/e_T when e_2/e_T has attained its optimal level. The figure on the right indicates higher maximal values of J/e_T; the underlying metabolic network therefore outperforms the network considered in the left figure. The best network differs from the suboptimal network only in two parameter values: the better network has a first enzyme that is inhibited less by *X* and a second enzyme that has a higher affinity for *X*. The orange and gray lines in the right figure illustrate the performance of a genetic circuit, which can approximate the optimal behavior (black line) either very well (orange line) or less well (gray line), depending on parameter values in the genetic circuit. See Figure 2 for details on the genetic circuit, and the Appendix for parameter values and more detail.

Results

A simple example of metabolic-flux maximization by optimal gene expression for a two-enzyme metabolic pathway

To illustrate the key ideas of the following sections, we consider a simple, analytically solvable example of a metabolic pathway (Figure 1A), inspired by published models [37,40–42] and motivated by experimental data on the 'growth law' [11,43,44], showing that the ribosomal protein fraction of *E. coli* increases linearly with its (balanced) growth rate.

The model considers two enzymes in a sequence, converting an external nutrient *S* into product *P*, via the metabolic intermediate *X*. The catalysis rate of each enzyme is described with basic enzyme kinetics (Figure 1A). A feature of enzyme kinetics is that the rate of an enzyme is proportional to its concentration, i.e., its expression level [45]. The concentrations of the two enzymes are denoted by e_1 and e_2. We consider the concentrations of *S* and *P* fixed to allow

Essays in Biochemistry (2024) **68** 41–51
https://doi.org/10.1042/EBC20230045

for the occurrence of a steady state. At steady state, the enzymatic rates v_1 and v_2 balance and equal the steady-state pathway flux: $v_1 = v_2 = J$.

Our aim is to design a genetic network that induces steady-state concentrations of enzyme l and enzyme 2 such that at all steady states of the pathway, as a function of S, the pathway flux attains its corresponding maximal value, while the total concentration of enzyme $e_1 + e_2 = e_T$ expressed in this pathway remains fixed. Since reaction rates are proportional to enzyme concentrations, we thus need to consider the maximization of J/e_T [34,35]. One could think of this as an optimization of the return (J) of the (biosynthetic resource) investment (e_T).

Clearly, given a concentration of S and a total enzyme concentration e_T, the pathway flux J depends on the concentrations of enzyme l and 2, e_1 and e_2: if one of these concentrations is zero then the flux is zero; if both e_1 and e_2 are positive, the flux is positive. Hence, a maximum value for the flux exists as function of the enzyme concentrations (the blue lines in Figure 1B). The maximal flux value and corresponding optimal protein fractions (e_1/e_T and e_2/e_T) depend on the concentration of S. Thus, to achieve a maximum of J/e_T as function of S, the enzyme concentrations need to be adjusted optimally. This is shown in two plots of Figure 1B by the black lines, they indicate the relation between the maximal flux and the corresponding optimal protein fraction of the second enzyme.

The optimum is also dependent on the kinetics of the enzymes. For instance, the two plots in Figure 1B differ in the value of two kinetic parameters. As discussed in detail in the caption of that figure, the performance of the pathway (its optimal J/e_T value) is increased when enzyme l is inhibited less by X and enzyme 2 has a higher affinity for X.

We can find the optimal enzyme concentration by a numerical optimization using a computer algorithm (black lines in Figure 1B), but as we show in the next section, we can also design a genetic circuit that gives rise to a (close-to) optimal performance that is (nearly) equally good. This illustrates that the biochemistry of gene regulation can optimize metabolic performance as captured by the measure J/e_T. That this is possible is illustrated in the right plot of Figure 1B. The orange and gray lines show the performance of a genetic circuit (modeled with differential equations, see Figure 2C and Appendix). The orange line indicates a parameterization of it that gives close-to-optimal behavior. How one can design such an optimally performing molecular genetic circuit is discussed in the next sections.

E. coli's control of ribosomal gene expression by ppGpp suggests that it aims to maximize its growth rate

Our current understanding of the expression of ribosomal versus biosynthesis genes can be described in terms of the coarse-grained model introduced in Figure 1A [37,40,42]. Protein synthesis is of course a complex system, comprising much more than a two-step pathway. However, a reduction to a two-enzyme system is believed to capture the essential properties of this system: in this view, the central tradeoff in metabolism is between proteins involved in the routes from nutrients to amino acids, whilst ribosomal protein is invested into translation into new protein. Despite its simplified view, it has shed significant light on the organization of balanced growth in bacteria [11,37,40,42].

In Figure 1B, we concluded that reduced inhibition of enzyme l and enhanced saturation of enzyme 2 leads to a linear relation between the flux per unit invested protein and the protein fraction of enzyme 2 (as found experimentally for the ribosomal protein fraction [11,43]. Since

$$\frac{J}{e_T} = \frac{v_2}{e_T} = \frac{k_2 e_2 f_2(x)}{e_T},$$

the specific flux is proportional to the enzyme-2 fraction (i.e. $J/e_T \propto e_2/e_T$) when $f_2(x)$ is approximately constant. So, in order to observe this proportional relationship, the saturation degree of the second enzyme ($f_2(x)$) needs to be kept constant. In the experimental case [11,43], this implies that the ribosome's saturation with substrate is remarkably constant across conditions [37]. This is apparently what the optimising gene circuit achieves: it adjusts the protein concentration such that this happens, independent of the value of S. Note that the maximal flux is achieved in the limit where the affinity constant K_1 approaches infinity (so that $f_1(s) \to 1$) and K_2 approaches 0 (leading to $f_2(x) \to 1$). This underlies the improved performance in Figure 1B.

Interestingly, when one plots the ribosome fraction of *E. coli* as function of the growth rate then also a linear relation is found (similar to the right figure of Figure 1B) [11,43]. This relation is known as the 'growth law' [11] (we note that the growth rate equals the protein synthesis rate divided by the total protein content [40,50], so it is also a J/e_T measure). The growth law suggests that the saturation degree of the ribosome is held fixed. Note two things here: (i) the experimental relation is linear and not proportional: an offset exists [11], and (ii) the relation is at slow growth no longer linear but for intermediate to high growth rates it is [52]. We omit these details here.

It is been well known that ribosomal saturation in *E. coli* is regulated by ppGpp [46–49]. To keep the saturation degree of the ribosome constant, we thus added a genetic circuit that incorporates ppGpp as the main regulator (Figure

Essays in Biochemistry (2024) **68** 41–51
https://doi.org/10.1042/EBC20230045

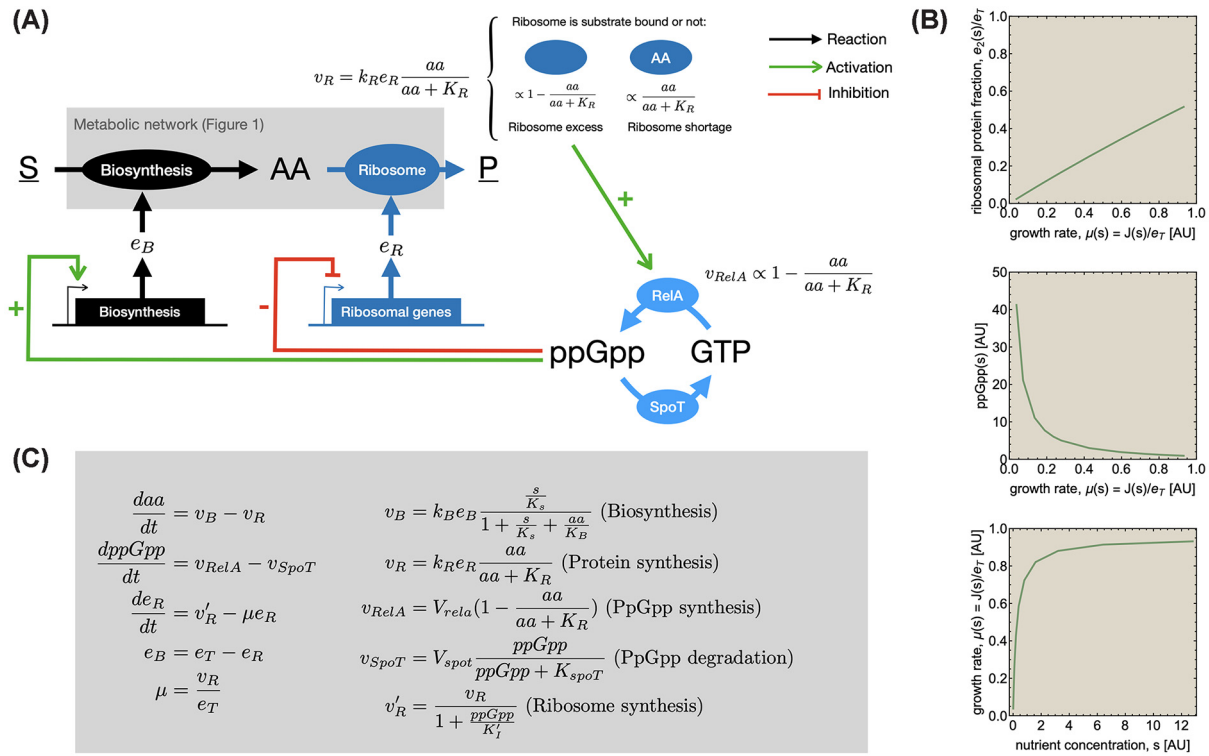

Figure 2. Ribosomal expression regulation in *E. coli* by ppGpp approximates growth-rate maximizing behavior

(**A**) A coarse-grained model of ribosomal gene expression regulation for *E. coli*. In *E. coli* ribosomal genes are regulated by the concentration of the molecule ppGpp [46,47]. ppGpp is made by the protein RelA when it is bound to a ribosome with an unloaded tRNA in its A site [48]. This happens under ribosome excess and biosynthetic enzyme-shortage and accordingly ppGpp can bind to RNA polymerase, thereby reducing its affinity for ribosomal promoters and enhancing it for biosynthesis promoters [46,47,49]. This leads to a regulation of the ribosomal activity that aims to keep its saturation with its substrate, loaded tRNAs, fixed. (**B**) Top figure: the known experimental linear relation between the ribosomal fraction and the growth rate – known as the 'growth law' [40,43,50,51] – is reproduced by the model. Middle figure: the model relation between the steady-state ppGpp concentration and the growth rate, which is in qualitative agreement with experimental data. Bottom figure: the dependency of the growth rate and the nutrient concentration agrees with the expected Monod relationship. (**C**) Model equations of the system shown in (A) that give the data shown in (B). Note that the model is surprisingly simple. A key ingredient is that the rate of ppGpp synthesis by RelA is proportional to 1 - *aa*/(*aa* + K_R) with *aa*/(*aa* + K_R) as the saturation level of the ribosome with substrate. As shown in Figure 1B, orange line, changing the V_{rela} parameter in $v_{Rel\,A}$ dramatically improves the ability of the gene network to approximate optimal performance. Motivation and model parameters are given in the Appendix.

2) [37,40]. The results of this model are in qualitative agreement with experimental data (Figure 2B) and the model equations are remarkably simple (Figure 2C) [37,40].

The ppGpp-mediated control of ribosomal gene expression is a powerful control system: it tracks and adjusts the metabolic state of the cell by sensing only the substrate saturation of ribosomes [37,48,52] and this appears to be sufficient for maximization of the steady state growth rate. The extent to which this gene network is able to keep ribosomal saturation constant and thus track the maximal growth state, however, depends on its parameterization. In particular, changing the maximal rate of RelA (i.e., V_{RelA}), responsible for ppGpp production, transforms the control system from a suboptimal circuit (gray line in Figure 1B) to a nearly optimal one (orange line). This indicates that tuning of kinetic parameters of the control system is an important (evolutionary) aspect of optimal regulation of gene expression.

Since the behavior of the model is so close to experimental behavior, one could argue that the ribosomal-expression control system of *E. coli* indeed aims to maximize a flux per unit total protein J/e_T: in this case it is the growth rate as this equals (protein synthesis rate/cellular protein content = v_R/e_T). The condition that we described for linearity above implies in the current context that the ribosome saturation is kept constant with tRNAs loaded with amino

Essays in Biochemistry (2024) **68** 41–51
https://doi.org/10.1042/EBC20230045

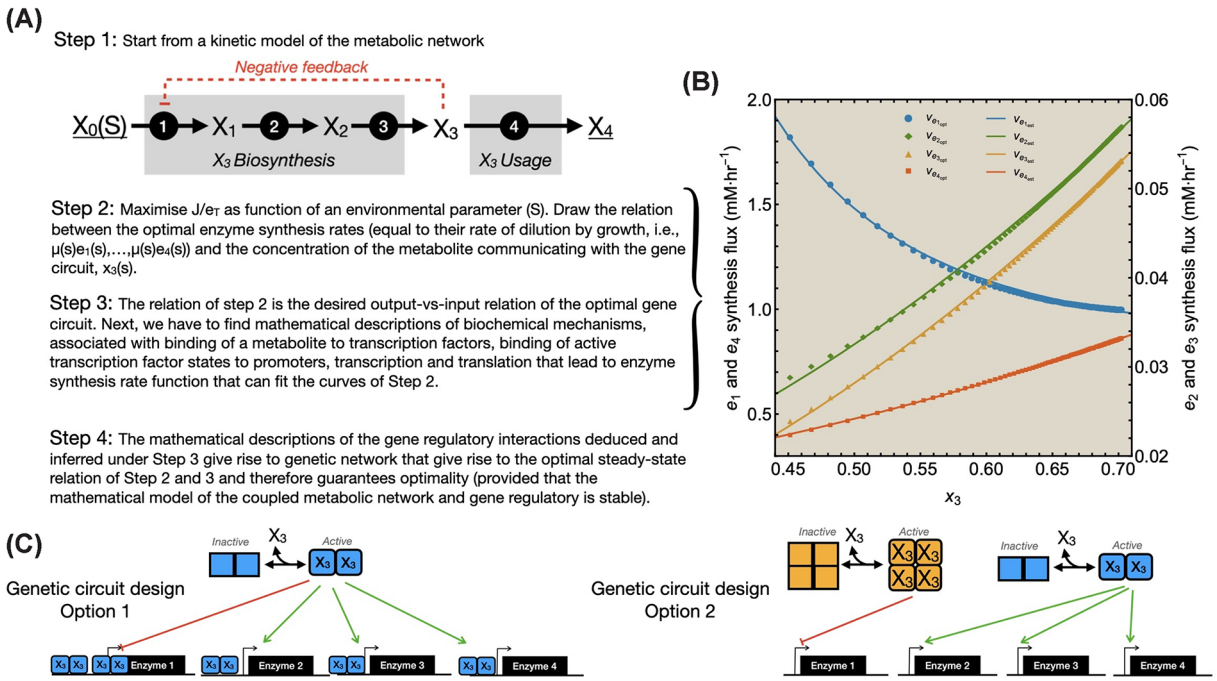

Figure 3. Approach and illustration of inferring of an genetic network given a kinetic model of the associated metabolic network

The steps (**A**) and the outcome (**B,C**) of the approach described in the main text are explained in this figure. The plot in (**B**) indicates the optimization results as dots and the fitted regulatory circuitry as lines. The two genetic circuits shown in (**C**). can both fit the optimization results, indicating that alternative regulatory network designs are possible. Model details and parameters may be found in the Appendix.

acids. A priori, this situation can occur at low saturation levels, with many ribosomes that are not bound to substrate, but that would lead to low growth rates. Experimental data, however, indicate that the saturation level is around 0.8 [52,53], remarkably high given the maximum saturation level of l (only attainable at infinite substrate concentrations of loaded tRNAs). The constantly high saturation level thus suggests that *E. coli* has a strategy to maximize growth rate by optimal expression of biosynthetic versus ribosomal genes. We refer the reader to [37,40] for a more detailed analysis.

Designing an optimal gene circuit for a metabolic pathway with a negative feedback

In the previous sections, we considered maximization of pathway performance (J/e_T) both from the perspective of a numerical algorithm and a known genetic circuit which approximates optimal behavior. Next, we will consider the design of a genetic circuit that can optimize the performance of a metabolic pathway as function of external conditions by sensing a metabolic intermediate and, given this information, optimally expresses metabolic enzymes (for another example, see [36]). For a design of such a genetic circuit we only require a mathematical model of the pathway and its numerical optimization.

As an example, we consider a metabolic pathway with a negative feedback loop in the context of a cell that aims to maximize its growth rate by optimization of its metabolic enzyme concentrations. The pathway resembles an amino-acid biosynthesis pathway as we find them in bacteria (step l in Figure 3A).

We assume that we know the dependence of the growth rate on an external parameter, for example, the concentration s of a limiting nutrient S. Here we achieve this using the Monod relation [54], $\mu(s) = \mu_{max}(s/s+K_s)$. We desire the metabolic pathway performance to be optimal with respect to the concentration of S. Thus, if s changes as a function of time, the cell should adapt its protein expression to again achieve maximal performance.

We also assume also that the cell has a fixed content $Y_{P/X}$ (mol product/gram biomass) of the product of the biosynthesis pathway, for example, of the amino acid; in other words, its molar content per gram cellular dry weight (gram drw) is fixed. Now we can calculate the biosynthetic rate of the product as function of the limiting nutrient

concentration, i.e.,

$$\frac{J\,(s)}{\text{gram drw}} = \frac{\text{mol product}}{\text{hr } \times \text{gram drw}} = \frac{\text{mol prod}}{\text{gram drw}} \times \frac{\text{gram drw}}{\text{gram drw} \times \text{hr}} = Y_{P/X} \times \mu\,(s)\,.$$

Next, we assume that the cell has a constant protein content (ρ = mol total protein/gram drw). Then

$$\frac{J\,(s)}{\text{mol total protein}} = \frac{Y_{P/X} \times \mu\,(s)}{\rho}.$$

This shows that maximization of the growth rate coincides with the maximization of the product biosynthesis rate.

From the last relation, we can determine the desired production flux per unit cellular protein that is required to grow at rate μ(s), given a concentration of the limiting nutrient S [36]. We assume that the cell aims to achieve this flux by minimizing the required concentrations of the biosynthetic proteins. In this way, the protein amount that the cell has left over is maximal (i.e., mol total protein-mol biosynthetic protein is maximal) and can be allocated to different cellular tasks associated with a competitive fitness such as stress tolerance, new-nutrient adaptation, etc [3,12].

All of this makes sense from the premise that a cell minimizes the protein amount allocated to each task that is currently needed for its fitness and that it maximizes its fitness in this manner [40,50]. This assumes that task performance is a rising function of the associated protein investment and that genetic circuits exist for each task that maximize the task performance by optimal protein expression [3].

An important aspect to note is that the substrate of the biosynthetic pathway is not the limiting nutrient S. It is an intracellular compound, an intermediate in cellular metabolism whose steady-state concentration either depends on the concentration of S or not. Here we consider that it depends on this concentration in a hyperbolic manner. Other cases we do not consider here.

The optimization problem associated with this scenario requires that we set the flux to a value J and the starting metabolite concentration to s and subsequently minimize the total biosynthetic protein needed to achieve this flux [36]. This leads to optimal concentrations of all biosynthetic enzymes that sum to this minimal total. We repeat this minimization then over the entire range of considered s values ($0.1K_s \leq s \leq 10K_s$) (step 2 in Figure 3A).

The optimization of the metabolic network leads to sets of optimal concentration of metabolic network enzymes ($e_1(s), \ldots, e_4(s)$) and metabolites, as a function of the concentration of S. This allows for a plot of the dependence of the protein synthesis rates of these enzymes, equaling $\mu(s)e_i(s)$ (with $i = 1, \ldots, 4$), as a function of the metabolite that binds to the components of the genetic circuit [36]; we choose X_3 as the metabolite playing this role, and plot $x_3(s)$ on the x-axis (Figure 3B). One can think of this relation as the steady-state input-output (I/O) relation of the genetic circuit, with the concentration of X_3 as its input and the optimal enzyme synthesis rates as its outputs. Any genetic network that generates this I/O relation is capable of sustaining optimal steady-state performance of the metabolic network as function of the concentration of S [36,38].

Next, we need to find a genetic network which displays these optimal I/O relations, by choosing protein synthesis functions that agree with the deduced ones in Figure 3B and that can be 'made' from biochemical interactions occurring in genetic networks or from those available in a library of synthetic-biology components. An example of a feasible network and its fit is shown in Figure 3B,C. We emphasize that different network designs can often fit one and the same I/O relation, which provides the synthetic biologist some flexibility in network design.

To illustrate the performance of the genetic network in controlling the optimal performance of the metabolic network, we show simulation results of a mathematical model of the coupled genetic and metabolic network in Figure 4. The genetic circuit design we consider is shown in Figure 4A, the range of concentrations of S in Figure 4B and an illustration of the dynamics leading to robust optimization is shown in Figure 4C,D. More complicated examples, with multiple metabolic intermediates as inputs to the I/O relation, may be found in [36,38].

Discussion

Adaptation to new stresses and nutrients is quicker and growth is faster when the associated stress, adaptation and growth-associated proteins are expressed at higher levels. The performance of cellular tasks is, therefore, a rising function of protein investment. Since cells have finite biosynthetic resources, protein investment in cellular tasks cannot all be high simultaneously. Therefore, cellular tasks trade off, task performance varies with conditions, and shifts occur between tasks (and metabolic strategies). Balancing of cellular tasks is the outcome of environmental sensing and induction of adaptive responses by cellular regulatory circuitry [2,3,15,50].

We currently lack a comprehensive understanding of the regulatory circuits of microbial cells [2,15]. Such an understanding would involve understanding the objective that those circuits aim to achieve and how this contributes to

Essays in Biochemistry (2024) **68** 41–51
https://doi.org/10.1042/EBC20230045

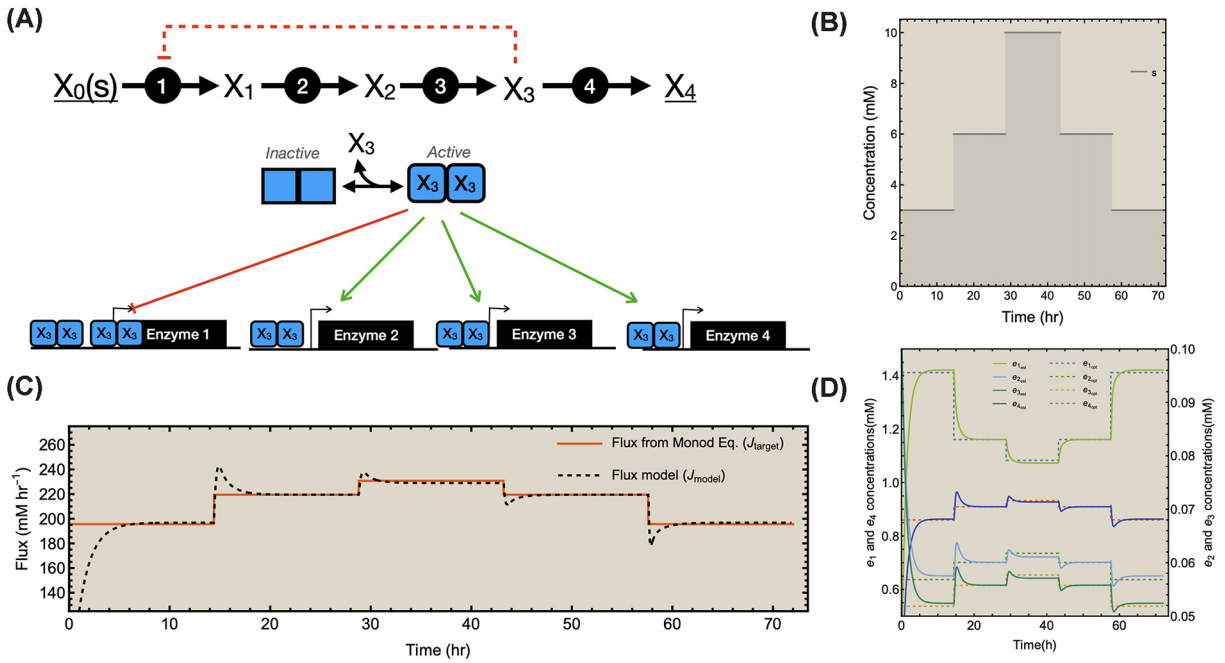

Figure 4. Illustration of optimal environmental tracking of the steady-state metabolic flux by the optimal genetic network

(**A**) The optimizing genetic network inferred in Figure 3. (**B**) Stepwise changes in the environmental concentration of the nutrient *S*. (**C**) Illustration of optimal environmental tracking: comparison of the flux of the model with the desired optimal behavior. (**D**) Overview of the dynamics of the enzyme concentration during the dynamics of (C). Model details and parameters may be found in the Appendix.

whole-cell fitness. We lack this understanding because most experimental studies are focused on steady-state growing cultures and their associated steady-state metabolism, while regulation is an inherently dynamic concept, requiring quantitative monitoring of dynamic responses.

In this review, we focused on three examples of regulation of metabolic pathways, that are based on [36–39]. We introduced the maximization of the flux per unit invested protein as the fitness measure for a subnetwork [34,35,55], which can be directly related to whole-cell fitness when it is defined as the instantaneous growth rate. In this context, the objective of the regulatory circuit is to minimise the protein expression needed to achieve a certain metabolic flux. This minimization can be done with a numerical algorithm [34,35], given a dynamic model of metabolism in terms of enzyme kinetics, and the results can be used to either design a novel genetic circuit [36], or understand an existing one from the perspective of optimality [37]. Since maximizing instantaneous growth rate is the fitness objective in environments with relative long periods of constant conditions [3], the approach we propose has broad applicability.

Our analysis of how *E. coli* regulates ribosome expression using ppGpp to sense ribosome substrate saturation, and thus balances biosynthesis and ribosomal gene expression, indicates that *E. coli* aims to maximize its instantaneous growth rate by controlling ribosomal protein levels efficiently. This is in agreement with other studies [3,40,50]. Current understanding suggests that growth-rate optimization applies to intermediate and high growth rate, but not at slow growth [3]. A corollary of instantaneous growth rate maximization is that none of the metabolic systems in the cell have protein over-or under-expression and that protein expression is, therefore, exactly matching the demand. This makes sense given that metabolic rates are proportional to the total protein investment and cells have finite biosynthetic resources. Thus, the translation of the cellular fitness measure (instantaneous growth rate) to a metabolic subnetwork implies that its gene regulation has maximization of the flux per unit protein (or, equivalently, minimization of the protein needed to obtain a desired flux) as objective. We illustrated that this allows for the computational design of genetic circuits. The example from amino-acid biosynthesis in *E. coli* suggests that such genetic circuits are actually simple in structure [36,38].

The strategy to design input-output relations to optimize flux per unit invested protein can be extended to much larger metabolic networks, indeed to any network in which each enzymatic reaction is essential to sustain the flux through the network (so-called Elementary Flux Modes [34,35,38,56]). These I/O relations may be fit using simple

gene regulatory circuits (see [36] for another example). It may even be shown mathematically that the gene circuits controlling the metabolic network generally steer the network to an optimal steady state [57]. Real biotechnological applications thus seem within reach.

Summary

- We study how high fitness in single-celled organisms is regulated through gene expression.
- We show how to infer and design gene regulatory networks that steer metabolic pathways to states of maximal flux.
- Examples include ribosome regulation and *E. coli* and amino acid biosynthesis.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Acknowledgements

F.J.B. and M.R. acknowledge funding by NWO-XL grant OCENW.XL21.XL21.007 'Taking Control of Metabolism in Microbial Cell Factories by Applying Non-canonical Redox Cofactors'. L.S. acknowledges funding by CONACYT. R.P.'s work was funded by NWO grant 613.009.131 'Control of maximal growth rate by single-celled organisms'. A.B.'s work was funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 713669.

Abbreviation

qORAC, q (denoting a specific flux with unit $mol(gr\ drw)^{-1}hr^{-1}$) Optimization by Robust Adaptive Control; ppGpp, Guanosine tetraphosphate; TCA, tricarboxylic acid.

References

1 Jacob, F. and Monod, J. (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318–356, https://doi.org/10.1016/S0022-2836(61)80072-7
2 Chubukov, V., Gerosa, L., Kochanowski, K. and Sauer, U. (2014) Coordination of microbial metabolism. *Nat. Rev. Microbiol.* **12**, 327–340, https://doi.org/10.1038/nrmicro3238
3 Bruggeman, F.J., Teusink, B. and Steuer, R. (2023) Trade-offs between the instantaneous growth rate and long-term fitness: consequences for microbial physiology and predictive computational models. *Bioessays* **45**, 2300015, https://doi.org/10.1002/bies.202300015
4 van Hoek, P.I.M., van Dijken, J.P. and Pronk, J.T. (1998) Effect of specific growth rate on fermentative capacity of Baker's yeast. *Appl. Environ. Microbiol.* **64**, 4226–4233, https://doi.org/10.1128/AEM.64.11.4226-4233.1998
5 Elsemman, I.E., Prado, A.R., Grigaitis, P., Albornoz, M.G., Harman, V., Holman, S.W. et al. (2022) Whole-cell modeling in yeast predicts compartment-specific proteome constraints that drive metabolic strategies. *Nat. Commun.* **13**, 801, https://doi.org/10.1038/s41467-022-28467-6
6 Basan, M., Hui, S., Okano, H., Zhang, Z., Shen, Y., Williamson, J.R. et al. (2015) Overflow metabolism in Escherichia coli results from efficient proteome allocation. *Nature* **528**, 99–104, https://doi.org/10.1038/nature15765
7 Lewis, K. (2010) Persister cells. *Annu. Rev. Microbiol.* **64**, 357–372, https://doi.org/10.1146/annurev.micro.112408.134306
8 Balaban, N.Q., Merrin, J., Chait, R., Kowalik, L. and Leibler, S. (2004) Bacterial persistence as a phenotypic switch. *Science* **305**, 1622–1625, https://doi.org/10.1126/science.1099390
9 Errington, J. (2003) Regulation of endospore formation in Bacillus subtilis. *Nat. Rev. Microbiol.* **1**, 117–126, https://doi.org/10.1038/nrmicro750
10 Orr, H.A. (2009) Fitness and its role in evolutionary genetics. *Nat. Rev. Genet.* **10**, 531–5391, https://doi.org/10.1038/nrg2603
11 Scott, M., Gunderson, C.W., Mateescu, E.M., Zhang, Z. and Hwa, T. (2010) Interdependence of cell growth and gene expression: origins and consequences. *Science* **330**, 1099–1102, https://doi.org/10.1126/science.1192588
12 Nyström, T. (2004) Microreview: growth versus maintenance: a trade-off dictated by RNA polymerase availability and sigma factor competition? *Mol. Microbiol.* **54**, 855–862, https://doi.org/10.1111/j.1365-2958.2004.04342.x
13 Snoep, J.L., Yomano, L.P., Westerhoff, H.V. and Ingram, L.O. (1995) Protein burden in Zymomonas mobilis: negative flux and growth control due to overproduction of glycolytic enzymes. *Microbiology* **141**, 2329–2337, https://doi.org/10.1099/13500872-141-9-2329
14 Dekel, E. and Alon, U. (2005) Optimality and evolutionary tuning of the expression level of a protein. *Nature* **436**, 588–592, https://doi.org/10.1038/nature03842
15 Scott, M. and Hwa, T. (2023) Shaping bacterial gene expression by physiological and proteome allocation constraints. *Nat. Rev. Microbiol.* **21**, 327–342, https://doi.org/10.1038/s41579-022-00818-6

16 Ihssen, J. and Egli, T. (2004) Specific growth rate and not cell density controls the general stress response in *Escherichia coli*. *Microbiology* **150**, 1637–1648, https://doi.org/10.1099/mic.0.26849-0
17 Meyers, L.A. and Bull, J.J. (2002) Fighting change with change: adaptive variation in an uncertain world. *Trends Ecol. Evol.* **17**, 551–5571, https://doi.org/10.1016/S0169-5347(02)02633-2
18 Basan, M., Honda, T., Christodoulou, D., Hörl, M., Chang, Y.-F., Leoncini, E. et al. (2020) A universal trade-off between growth and lag in fluctuating environments. *Nature* **584**, 470–474, https://doi.org/10.1038/s41586-020-2505-4
19 Sander, T., Farke, N., Diehl, C., Kuntz, M., Glatter, T. and Link, H. (2019) Allosteric feedback inhibition enables robust amino acid biosynthesis in *E. coli* by enforcing enzyme overabundance. *Cell Syst.* **8**, 66–75, https://doi.org/10.1016/j.cels.2018.12.005
20 Hofmeyr, J.-H.S. and Cornish-Bowden, A. (2000) Regulating the cellular economy of supply and demand. *FEBS Lett.* **476**, 47–51, https://doi.org/10.1016/S0014-5793(00)01668-9
21 Xu, Y.-F., Zhao, X., Glass, D.S., Absalan, F., Perlman, D.H., Broach, J.R. et al. (2012) Regulation of yeast pyruvate kinase by ultrasensitive allostery independent of phosphorylation. *Mol. Cell* **48**, 52–62, https://doi.org/10.1016/j.molcel.2012.07.013
22 Feldman-Salit, A., Hering, S., Messiha, H.L., Veith, N., Cojocaru, V., Sieg, A. et al. (2013) Regulation of the activity of lactate dehydrogenases from four lactic acid bacteria. *J. Biol. Chem.* **288**, 21295–21306, https://doi.org/10.1074/jbc.M113.458265
23 LaPorte, D.C., Walsh, K. and Koshland, D.E. (1984) The branch point effect ultrasensitivity and subsensitivity to metabolic control. *J. Biol. Chem.* **259**, 14068–14075, https://doi.org/10.1016/S0021-9258(18)89857-X
24 Gey, U., Czupalla, C., Hoflack, B., Rodel, G. and Krause-Buchholz, U. (2008) Yeast pyruvate dehydrogenase complex is regulated by a concerted activity of two kinases and two phosphatases. *J. Biol. Chem.* **283**, 9759–9767, https://doi.org/10.1074/jbc.M708779200
25 Snoep, J.L., Joost, M., de Mattos, T. and Neijssel, O.M. (1991) Effect of the energy source on the nadh/nad ratio and on pyruvate catabolism in anaerobic chemostat cultures of enterococcus faecalis nctc775. *FEMS Microbiol. Lett.* **81**, 63–66, https://doi.org/10.1111/j.1574-6968.1991.tb04713.x
26 Falgueras, M.L.S., Viñuela, E., Salas, M. and Sols, A. (1965) Citrate inhibition of phosphofructokinase and the pasteur effect. *Biochem. Biophys. Res. Commun.* **19**, 371–376, https://doi.org/10.1016/0006-291X(65)90471-7
27 Hofmeyr, J.-H.S. (1995) Metabolic regulation: a control analytic perspective. *J. Bioenerg. Biomembr.* **27**, 479–490, https://doi.org/10.1007/BF02110188
28 Atkinson, D.E. (1969) Regulation of enzyme function. *Annu. Rev. Microbiol.* **23**, 47–68, https://doi.org/10.1146/annurev.mi.23.100169.000403
29 Orth, J.D., Thiele, I. and Palsson, B.Ø. (2010) What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–248, https://doi.org/10.1038/nbt.1614
30 Teusink, B., Passarge, J., Reijenga, C.A., Esgalhado, E., van der Weijden, C.C., Schepper, M. et al. (2000) Can yeast glycolysis be understood in terms of *in vitro* kinetics of the constituent enzymes? Testing biochemistry. *Eur. J. Biochem.* **267**, 5313–5329, https://doi.org/10.1046/j.1432-1327.2000.01527.x
31 Vasilakou, E., van Loosdrecht, M.C.M. and Wahl, S.A. (2020) *Escherichia coli* metabolism under short-term repetitive substrate dynamics: adaptation and trade-offs. *Microbial Cell Factories* **19**, 1–19, https://doi.org/10.1186/s12934-020-01379-0
32 Erickson, D.W., Schink, S.J., Patsalo, V., Williamson, J.R., Gerland, U. et al. (2017) A global resource allocation strategy governs growth transition kinetics of *Escherichia coli*. *Nature* **551**, 119–123, https://doi.org/10.1038/nature24299
33 van Heerden, J.H., Wortel, M.T., Bruggeman, F.J., Heijnen, J.J., Bollen, Y.J.M., Planqué, R. et al. (2014) Lost in transition: start-up of glycolysis yields subpopulations of nongrowing cells. *Science* **343**, 1245114, https://doi.org/10.1126/science.1245114
34 Wortel, M.T., Peters, H., Hulshof, J., Teusink, B. and Bruggeman, F.J. (2014) Metabolic states with maximal specific rate carry flux through an elementary flux mode. *FEBS J.* **281**, 1547–1555, https://doi.org/10.1111/febs.12722
35 Noor, E., Flamholz, A., Bar-Even, A., Davidi, D., Milo, R. and Liebermeister, W. (2016) The protein cost of metabolic fluxes: prediction from enzymatic rate laws and cost minimization. *PLoS Comput. Biol.* **12**, e1005167, https://doi.org/10.1371/journal.pcbi.1005167
36 Berkhout, J., Bosdriesz, E., Nikerel, E., Molenaar, D., de Ridder, D., Teusink, B. et al. (2013) How biochemical constraints of cellular growth shape evolutionary adaptations in metabolism. *Genetics* **194**, 505–512, https://doi.org/10.1534/genetics.113.150631
37 Bosdriesz, E., Molenaar, D., Teusink, B. and Bruggeman, F.J. (2015) How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization. *FEBS J.* **282**, 2029–2044, https://doi.org/10.1111/febs.13258
38 Planqué, R., Hulshof, J., Hendriks, J.C., Teusink, B. and Bruggeman, F.J. (2018) Maintaining maximal metabolic rate using gene expression control. *PLoS Comput. Biol.* **14**, e1006412, https://doi.org/10.1371/journal.pcbi.1006412
39 Planqué, R. and Hulshof, J. (2019) Metabolic pathways and optimisation. In *Complexity Science* (Peletier, M.A., Santen, van and Steur, E., eds), pp. 375–392, https://doi.org/10.1142/9789813239609˙0012
40 Scott, M., Klumpp, S., Mateescu, E.M. and Hwa, T. (2014) Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Mol. Syst. Biol.* **10**, 747, https://doi.org/10.15252/msb.20145379
41 Towbin, B.D., Korem, Y., Bren, A., Doron, S., Sorek, R. and Alon, U. (2017) Optimality and sub-optimality in a bacterial growth law. *Nat. Commun.* **8**, 14123, https://doi.org/10.1038/ncomms14123
42 Chure, G. and Cremer, J. (2023) An optimal regulation of fluxes dictates microbial growth in and out of steady state. *Elife* **12**, e84878, https://doi.org/10.7554/eLife.84878
43 Schaechter, M., Maaløe, O. and Kjeldgaard, N.O. (1958) Dependency on medium and temperature of cell size and chemical composition during balanced growth of salmonella typhimurium. *Microbiology* **19**, 592–606, https://doi.org/10.1099/00221287-19-3-592
44 Bremer, H. and Dennis, P.P. (2008) Modulation of chemical composition and other parameters of the cell at different exponential growth rates. *EcoSal Plus* **3**, 10–1128, https://doi.org/10.1128/ecosal.5.2.3
45 Cornish-Bowden, A. (2004) *Fundamentals of Enzyme Kinetics*, 4th edition, Wiley-Blackwell

46 Potrykus, K., Murphy, H., Philippe, N. and Cashel, M. (2011) ppGpp is the major source of growth rate control in *E. coli*. *Environ. Microbiol.* **13**, 563–575, https://doi.org/10.1111/j.1462-2920.2010.02357.x
47 Steinchen, W. and Bange, G. (2016) The magic dance of the alarmones (p)ppGpp. *Mol. Microbiol.* **101**, 531–544, https://doi.org/10.1111/mmi.13412
48 Agirrezabala, X., Fernández, I.S., Kelley, A.C., Cartón, D.G., Ramakrishnan, V. and Valle, M. (2013) The ribosome triggers the stringent response by RelA via a highly distorted tRNA. *EMBO Rep.* **14**, 811–816, https://doi.org/10.1038/embor.2013.106
49 Paul, B.J., Barker, M.M., Ross, W., Schneider, D.A., Webb, C., Foster, J.W. et al. (2004) DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell* **118**, 311–322, https://doi.org/10.1016/j.cell.2004.07.009
50 Bruggeman, F.J., Planqué, R., Molenaar, D. and Teusink, B. (2020) Searching for principles of microbial physiology. *FEMS Microbiol. Rev.* **44**, 821–844, https://doi.org/10.1093/femsre/fuaa034
51 Ecker, R.E. and Schaechter, M. (1963) Ribosome content and the rate of growth of *Salmonella typhimurium*. *Biochim. Biophys. Acta* **76**, 275–279, https://doi.org/10.1016/0926-6550(63)90040-9
52 Wu, C., Balakrishnan, R., Braniff, N., Mori, M., Manzanarez, G., Zhang, Z. et al. (2022) Cellular perception of growth rate and the mechanistic origin of bacterial growth law. *Proc. Natl. Acad. Sci.* **119**, e2201585119, https://doi.org/10.1073/pnas.2201585119
53 Young, R. and Bremer, H. (1976) Polypeptide-chain-elongation rate in escherichia coli b/r as a function of growth rate. *Biochem. J.* **160**, 185–194, https://doi.org/10.1042/bj1600185
54 Monod, J. (1949) The growth of bacterial cultures. *Annu. Rev. Microbiol.* **3**, 371–394, https://doi.org/10.1146/annurev.mi.03.100149.002103
55 Müller, S., Regensburger, G. and Steuer, R. (2014) Enzyme allocation problems in kinetic metabolic networks: optimal solutions are elementary flux modes. *J. Theor. Biol.* **347**, 182–190, https://doi.org/10.1016/j.jtbi.2013.11.015
56 Schuster, S. and Hilgetag, C. (1994) On elementary flux modes in biochemical reaction systems at steady state. *J. Biol. Systems* **2**, 165–185, https://doi.org/10.1142/S0218339094000131
57 Planqué, R., Hulhof, J. and Bruggeman, Frank J. (2024) Sensing cellular growth rate facilitates robust optimal adaptation to changing conditions. in preparation