

RESEARCH ARTICLE

Maintaining maximal metabolic flux by gene expression control

Robert Planqué^{1*}, Josephus Hulshof¹, Bas Teusink², Johannes C. Hendriks², Frank J. Bruggeman²

1 Department of Mathematics, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands, **2** Systems Bioinformatics, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

* r.planque@vu.nl



OPEN ACCESS

Citation: Planqué R, Hulshof J, Teusink B, Hendriks JC, Bruggeman FJ (2018) Maintaining maximal metabolic flux by gene expression control. *PLoS Comput Biol* 14(9): e1006412. <https://doi.org/10.1371/journal.pcbi.1006412>

Editor: Christopher V. Rao, University of Illinois at Urbana-Champaign, UNITED STATES

Received: August 22, 2017

Accepted: August 1, 2018

Published: September 20, 2018

Copyright: © 2018 Planqué et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All software files are available at <http://www.few.vu.nl/~rplanque/Research/qORAC/index.html>.

Funding: FJB acknowledges funding of Nederlandse Organisatie NWO-VICI project No. 864-11-011; BT acknowledges funding of Nederlandse Organisatie NWO-VICI project 865.14.005. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

One of the marvels of biology is the phenotypic plasticity of microorganisms. It allows them to maintain high growth rates across conditions. Studies suggest that cells can express metabolic enzymes at tuned concentrations through adjustment of gene expression. The associated transcription factors are often regulated by intracellular metabolites. Here we study metabolite-mediated regulation of metabolic-gene expression that maximises metabolic fluxes across conditions. We developed an adaptive control theory, *q*ORAC (for ‘Specific Flux (*q*) Optimization by Robust Adaptive Control’), and illustrate it with several examples of metabolic pathways. The key feature of the theory is that it does not require knowledge of the regulatory network, only of the metabolic part. We derive that maximal metabolic flux can be maintained in the face of varying *N* environmental parameters only if the number of transcription-factor binding metabolites is at least equal to *N*. The controlling circuits appear to require simple biochemical kinetics. We conclude that microorganisms likely can achieve maximal rates in metabolic pathways, in the face of environmental changes.

Author summary

To attain high growth rates, microorganisms need to sustain high activities of metabolic reactions. Since the catalysing enzymes are in finite supply, cells need to carefully tune their concentrations. When conditions change, cells need to adjust those concentrations. How cells maintain high metabolism rates across conditions by way of gene regulatory mechanisms and whether they can maximise metabolic activity is far from clear. Here we present a general theory that solves this metabolic control problem, which we have called *q*ORAC for specific flux (*q*) Optimisation by Robust Adaptive Control. It considers that external changes are sensed by internal “sensor” metabolites that bind to transcription factors in order to regulate enzyme-synthesis rates. We show that such a combined system of metabolism and its gene network can self-optimize its metabolic activity across conditions. We present the mathematical conditions for the required adaptive control for robust system-steering to optimal states across conditions. We provide explicit examples of such self-optimising coupled metabolism and gene network systems. We prove that a cell can be robust to changes in *K* parameters, e.g. external conditions, if at least *K* internal

metabolite concentrations act transcription-factor binding sensors. We find that the optimal relation of the enzyme synthesis rates of self-optimising systems and the concentration of the sensor metabolites can generally be implemented by basic biochemistry. Our results indicate how cells are able to maintain maximal reaction rates, even in changing conditions.

Introduction

Microbes need to grow fast to outcompete others. They therefore have to maintain high growth rates in changing environments. To achieve this specific metabolic fluxes (metabolic rates per unit of expended enzyme) need to be kept as high as possible. Since metabolic enzymes are a limited resource, cells should behave economically: synthesise the right enzymes in the right amounts, and adapt their levels when conditions change. In this paper we show how cells can achieve this in the case when the growth rate itself is fixed, but a limited protein pool needs to be optimally distributed over metabolic pathway reactions to maximise its steady-state rate.

Experimental evidence is mounting that cells are indeed able to tune enzyme levels to maximise the growth rate (Fig 1; [1, 2, 3, 4, 5, 6, 7, 8, 10]). Efficient enzyme allocation has also recently been shown explain measured flux values [11], and to underlie a surprising number of other general physiological phenomena [12, 13], such as the bacterial growth laws [14, 15, 16], overflow metabolism (the Crabtree or Warburg effect; [13, 17]), and catabolite repression [18]. Except perhaps for the case of optimal ribosomal synthesis [15, 16], it is not clear in any of these examples how cells can find the optimal protein expression state out of all possible ones.

Finding optimal states is difficult for microorganisms. They generally do not have sensor proteins in their membranes to alert them of the presence or absence of nutrients or stresses, because their membrane space is limited. It needs to be filled with transporters and respiratory proteins that directly contribute to fitness. Thus cells have to decide how to allocate their resources from internal cues only. Cells are evidently able to accomplish this feat, but that raises the question how they are able to achieve such “blind optimisation”.

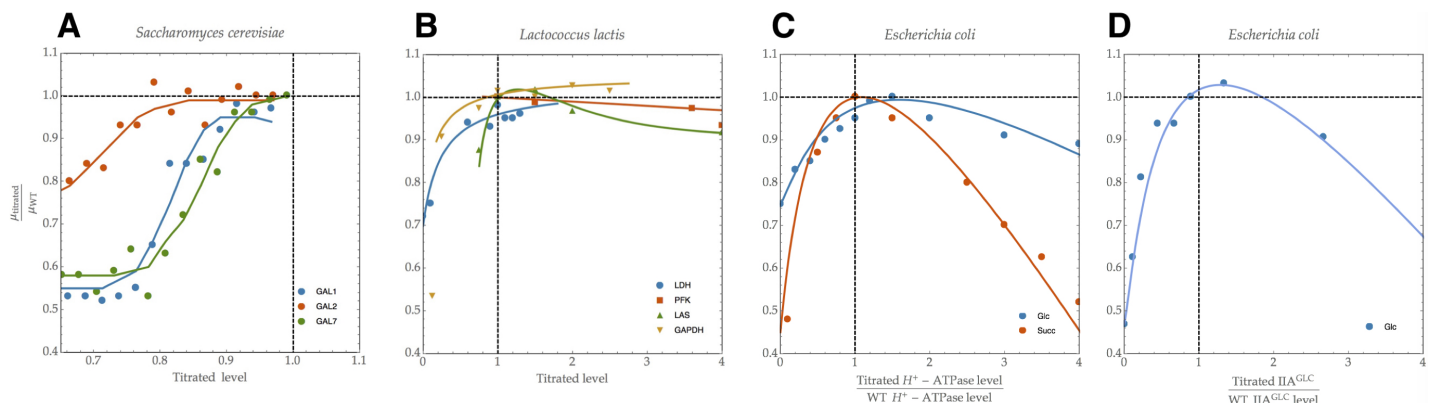


Fig 1. Experimental evidence indicating that microbes tune their enzyme levels to maximise growth rate. In each example, the wild type (WT) is shown to express enzyme concentrations at which the growth rate μ is approximately maximal. Axes show enzyme concentrations relative to wild type (WT) levels (abscissa) and growth rates relative to WT. Data adapted from: A, [9]; B, [4, 5, 6], C, [3]; D, [2]. Abbreviations: GAL1, galactokinase; GAL2, Galactose permease; GAL7, Galactose-1-phosphate uridyl transferase; LDH, lactate dehydrogenase; PFK, phosphofructokinase; LAS, las operon; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; Glc, glucose; Succ, succinate. In [9] there are many other examples, including several proteins that do not show levels at which growth rate is optimised.

<https://doi.org/10.1371/journal.pcbi.1006412.g001>

Gene expression regulation is largely achieved by transcription factors that are either affected by direct binding of metabolites, or signal transduction cascades, as readouts of environmental and cellular states. Even though transcription factor binding by sensor metabolites is widely accepted in the field [19, 20], the identity of the sensors is only known in a handful of cases (Fig 2). In *E. coli*, fructose-1,6-bisphosphate (FBP), a glycolytic intermediate, binds to the transcription factor Cra to regulate genes involved in glycolysis [18, 21]; in yeast, the galactose catabolic pathway is induced by intracellular galactose [22]; in *E. coli*, uncharged-tRNAs induce synthesis of ppGpp when amino acids are limited, leading to the adjustment of ribosome expression [15, 16]; like most amino acid pathways, the amino acid L-tryptophan regulates the transcription of several enzymes involved in its own biosynthetic pathway [23]; perhaps the best known example is the lactose operon, which is induced by allolactose, an intermediate of the pathway [24]. There is even very recent experimental evidence that *E. coli*'s central metabolism is in fact controlled by just three such sensor metabolites (cyclic AMP (cAMP), FBP and fructose-1-phosphate (F1P); [25]).

What remains unexplained is why certain sensor metabolites bind to transcription factors and others do not. How many sensors can we expect to be functioning? When do cells rely on just a few sensors? What are the design criteria for regulating circuits that maintain optimal metabolism in fluctuating environments? Does this regulation require complex, hard to evolve,

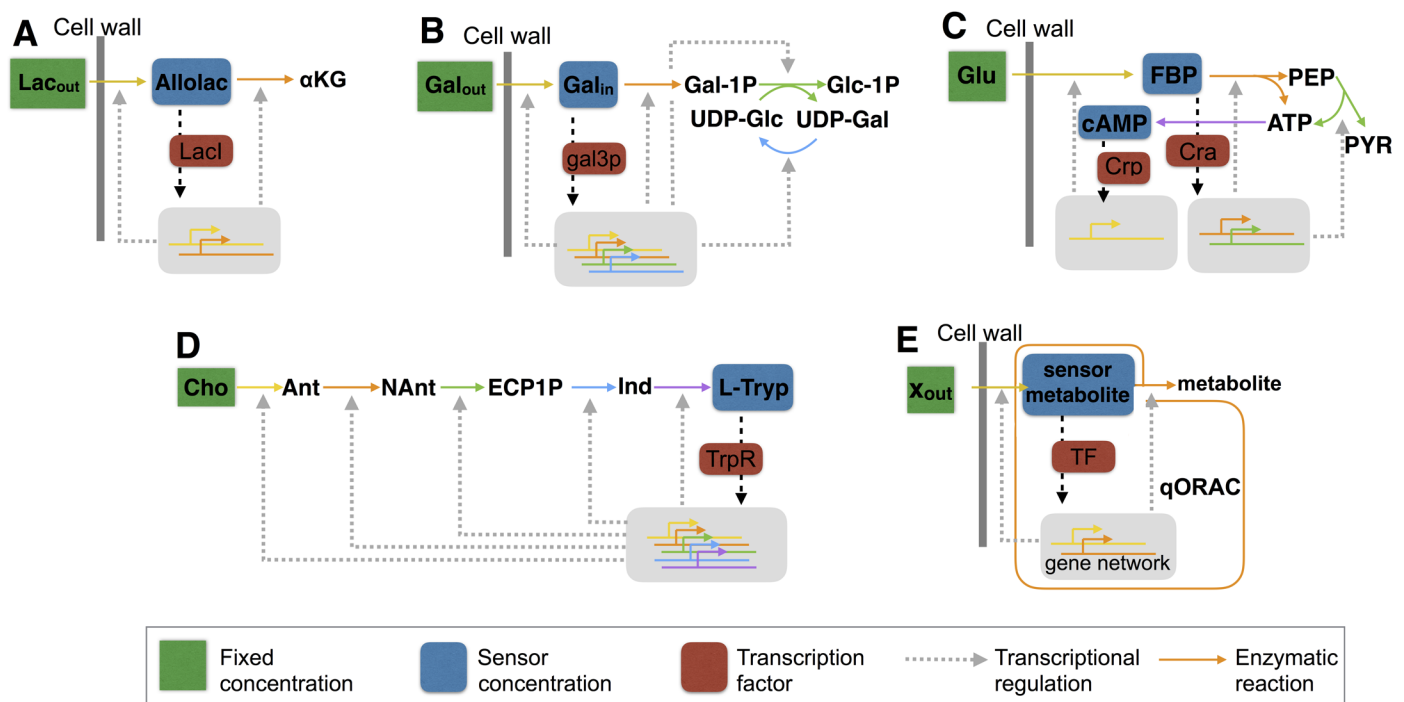


Fig 2. Biological examples of qORAC. Four well-characterised metabolic pathways in which a metabolite binds to a transcription factor (TF) to influence gene expression. The qORAC framework applies to each of them: in each case, the qORAC formalism gives rise to the enzyme synthesis rates that steer the metabolic pathway to maximal metabolic rates that are robust to changes in the external concentration (external with respect to the pathway). (A) The lac operon in *E. coli*, with sensor Allolactose binding to Lacl; (B) The galactose uptake system in yeast, with sensor internal galactose binding to gal3p; (C) The control of glycolytic enzymes via sensors FBP (binding to Cra), and cAMP (binding to Crp); (D) The control of the L-Tryptophan biosynthesis pathway by the amino acid binding to TrpR; (E) The general scheme of a qORAC-steered pathway. Abbreviations: LaC_{out}, external lactose; Allolac, allolactose; αKG, α-ketoglutarate; Gal_{out}, external galactose; Gal_{in}, internal galactose; Gal-1P, galactose-1-phosphate; Glc-1P, glucose-1-phosphate; UDP-Glc, uridine-diphosphate-glucose; UDP-Gal, uridine-diphosphate-galactose; Glu, glucose; FBP, fructose-1,6-bisphosphate; PEP, phosphoenolpyruvate; PYR, pyruvate; cAMP, cyclic AMP; ATP, adenosine-triphosphate; Cho, chorismate; Ant, Anthranilate; NAnt, N-(5'-phosphoribosyl)-anthranilate; ECP1P, Enol-1-0-carboxy-phenylamino-1-deoxyribulose phosphate; Ind, Indole-3-glycerol-P; L-Tryp, L-tryptophan.

<https://doi.org/10.1371/journal.pcbi.1006412.g002>

biochemistry or it is almost gratuit? We derive a universal theory, called *q*ORAC (for Specific Flux (*q*) Optimisation by Robust Adaptive Control), that gives answers to these questions. Understanding how growth rate itself is maximised is beyond the scope of this paper. Instead we focus on the important case of maximising specific rates of metabolic subnetworks at fixed growth rate.

In order to achieve maximal metabolic rates without *direct* knowledge of those external conditions and how they change, a controlling gene regulatory network must work as follows. At each point in time, internal sensor metabolites must influence a gene regulatory network, causing changes in gene expression. The strength of this signal depends on the concentration of the sensor metabolite. The crucial ingredient is that the gene network must be made in such a way that it expresses proteins at optimal steady state rates with respect to the current sensor metabolite concentration. The network thus ‘assumes’ a steady-state optimum at each point in time. As long as there is a mismatch between the enzyme synthesis rates and the external concentration, so that the metabolic system is not in an optimal state, the system will display dynamics. The sensor metabolite concentration will therefore continue to change and the enzyme synthesis rates will change with it. However, when the enzyme rates are optimal, given the current external environment, a steady state should be reached, which is then necessarily also optimal. In this way, the cell has achieved an optimal state without direct information about it. Its allocation of limited biosynthetic resources for protein synthesis will then be optimal. A gene network, informed by sensor metabolites, that causes optimal steady state enzyme levels in different conditions therefore must necessarily implement some form of *q*ORAC control (Fig 3).

The experimental evidence presented in Fig 1 indicates that a *q*ORAC-like control mechanism is active in cells. If cells are able to reach maximal growth rates (and hence maximal metabolic rates to attain this) in different conditions, at different optimal enzyme concentrations, then the gene regulatory network responsible must necessarily cause the correct enzyme synthesis rates (or must approximate these to a good degree). If this gene network works on the basis of internal metabolic information (rather than on information from signalling pathways, for instance), the control is adaptive, and indeed a form of *q*ORAC control.

Remarkably, the *q*ORAC theory we present here shows that a metabolic system, coupled to its controlling gene network, has a unique optimal steady state, no matter what the environmental conditions are—even though that optimum *changes* with those conditions. We prove that the dynamics of enzyme synthesis that is required for attaining optimal metabolic states can be inferred from the kinetic rate laws from metabolic enzymes alone. This is in direct agreement with a celebrated engineering principle, the internal model principle [26]. Our results also suggest that the optimising enzyme dynamics of a gene circuits circuit can be achieved with basic biochemistry.

The *q*ORAC theory predicts which metabolites may act as sensors. A fundamental insight is that maintenance of optimal metabolism in the face of *N* parameters requires *N* sensor metabolites. The *q*ORAC theory indicates that recent findings, such as the pervasive optimisation of enzyme levels in yeast [9], or the small number of sensor metabolites found in *E. coli*'s central metabolism [25], should necessarily be seen as surprising.

The phenotypic plasticity of microorganisms is a marvel of evolution. What would be even more remarkable is that cells can maximise their performance in changing conditions, without direct information about those changes. This appears almost impossible in view of the bewildering biochemical complexity of the cell. Part of what we achieve in this paper is to show that this skepticism is most likely unfounded: cells can do this. The insight can explain the robustness to human interventions in metabolic engineering and medicine, and provide opportunities to circuit design in synthetic biology.

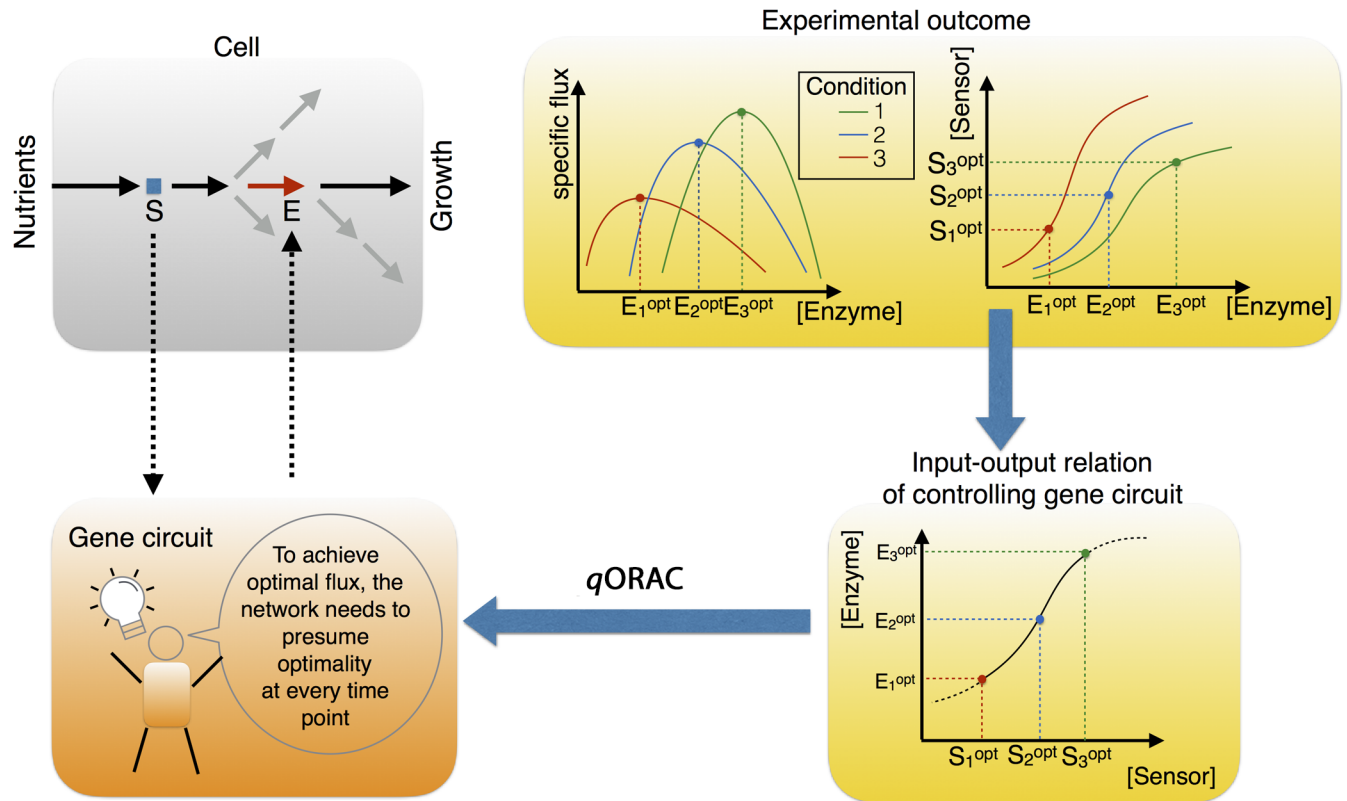


Fig 3. Integrating *q*ORAC with experimental evidence. Top, left: We consider a cell which takes up glucose (Glu) and converts it into biomass using a metabolic pathway. A sensor metabolite (*S*) influences gene expression and hence enzyme levels. Let *E* be one such enzyme in the active pathway. Top, right: The concentration of *E* is titrated experimentally under different glucose conditions, Low, Mid and High. In each condition, the maximal growth rate is measured, at different levels of titrated enzyme levels. In the same experiment, the sensor concentration is monitored. Bottom, right: Plotting the optimal enzyme levels at different conditions together with the measured sensor concentrations indicates the input-output relation of the gene network necessary to achieve maximal growth rates. Any gene network that implements such an input-output relation automatically expresses optimal enzyme levels in each condition. Bottom, left: To ensure that the steady state of the combined metabolic-gene system is always optimal, the gene network must presume optimality of the sensor at each time point. If the sensor is not optimal, it will change (and so will the enzyme levels); if it is optimal and stationary, the whole pathway will achieve maximal rates. *q*ORAC also describes the input-output relation in other conditions than the cell may have experienced (dotted lines in graph bottom right).

<https://doi.org/10.1371/journal.pcbi.1006412.g003>

Motivating *q*ORAC with an example

We will first introduce the control problem that a cell faces. We consider a well-understood example: the regulation of galactose metabolism in yeast (Figs 1A and 2B). We aim to characterise the dynamics of a controlling gene circuit that always maximises the steady-state flux per unit invested enzyme in this pathway (the *specific flux*) upon an environmental change, such as in the extracellular galactose concentration. The controlling gene network has to distribute a finite amount of biosynthetic resources for enzyme synthesis over the four pathway enzymes to maximise the steady state pathway flux.

Depending on the external galactose concentration, less or more enzymatic resources should be invested in the galactose import reaction. This leaves a correspondingly smaller or larger pool of enzymatic resources for the remaining pathway reactions. An increase in $[Gal_{out}]$ will cause an increase in $[Gal_{in}]$, which is therefore indicative for the external change. Gal_{in} can thus act as a signal for the adjustment of enzyme concentrations of the pathway: the transporter concentration should decrease and the others should increase.

In yeast, Gal_{in} plays the role of metabolic sensor [27]. It relays information to the GAL operon by binding to gal3p, a regulatory protein that can activate transcription factors, such as gal80 and gal4. The key question is how the concentration of Gal_{in} should influence the gene network in order to steer the galactose pathway to maximal specific flux.

We refer to the relation between the steady-state concentrations of the metabolic sensor ([Gal_{in}]) and the metabolic enzymes as the input-output relation of the gene circuit. *q*ORAC specifies this relation for robust maximisation of specific pathway flux. Whether a gene circuit with realistic biochemical kinetics can be found that can implement this input-output relation then still needs to be determined. Since the gene network for the galactose pathway in yeast is known, the optimal input-output relation may be found by fitting parameters in this network, which we achieved in an earlier paper [28]. In the current paper, however, we show that the problem of finding optimal input-output relations for a given metabolic pathway has a general solution, applicable to all examples shown in Fig 2A–2D. This indicates that cells can implement *q*ORAC using simple regulating circuits.

Methods

The *q*ORAC theory starts with the dynamics of the intracellular metabolite concentrations $\mathbf{x}_I = (x_1, \dots, x_n)$ of a metabolic network,

$$\dot{\mathbf{x}}_I = N\mathbf{v}(\mathbf{x}_I; \mathbf{x}_E) - \mu\mathbf{x}_I. \quad (1)$$

Here, N is the stoichiometry matrix, $\mathbf{v}(\mathbf{x}_I; \mathbf{x}_E)$ is the vector of reaction rates, \mathbf{x}_E are fixed external concentrations, and μ is the cellular growth rate. It is generally assumed that the dilution rate of concentrations by growth, $-\mu\mathbf{x}_I$, is negligible for metabolism. We take the same view here, and consider

$$\dot{\mathbf{x}}_I = N\mathbf{v}(\mathbf{x}_I; \mathbf{x}_E). \quad (2)$$

The *q*ORAC framework couples this metabolic pathway to enzyme dynamics, by choosing

$$\dot{\mathbf{e}} = \mathbf{E}(\mathbf{x}_S) - \mu\mathbf{e}. \quad (3)$$

Since enzyme dynamics occur at time scales of similar order as the growth rate, the dilution by growth cannot be neglected this time. Throughout the paper, the growth rate is a predefined parameter, and not part of the optimisation problem (see the Discussion for more information). $\mathbf{E}(\mathbf{x}_S)$ denote the enzyme synthesis rates for all the different enzymes involved in the pathway. These functions may only depend on internal sensor metabolite concentrations, as explained in the Introduction. The task is to define these functions in such a way that the combined dynamical metabolic-enzyme system converges to a steady state in which flux through the pathway is maximal.

Results

As explained in the Introduction, *q*ORAC relies on allocating resources on the basis of sensor metabolite information alone. The optimal allocation must therefore be uniquely defined for each set of sensor concentrations. By considering the optimisation problem in detail, we show that this requires several steps:

1. We must restrict the original pathway to a minimal set of metabolic reactions, connecting sources to sinks, called an Elementary Flux Mode (EFM). EFMs have been shown to arise naturally within the context of this optimisation problem [29, 30], as we will see.

2. Having restricted the pathway to an EFM, the optimisation problem of maximal steady state specific flux must have, for each choice of external conditions, a unique optimal allocation of enzymes and internal metabolites. We show that for practically all enzyme kinetics rate laws, this is the case.
3. We show under what conditions a set of chosen sensor concentrations may be used in a *q*ORAC control. One of the conditions is that the number of sensors must be equal to the number of varying external concentrations.
4. The metabolic pathway, coupled to *q*ORAC enzyme dynamics, must have a unique steady state, which is necessarily optimal. We show this is true for a very large class of pathways.

We now consider these four steps in detail.

Step 1: Restricting to minimal pathways

We aim to maximise a steady-state specific flux v_r/e_T through the network where v_r is some chosen output flux (e.g. in mM/hr) and e_T (e.g. in grams) is total amount of invested enzyme. The optimisation problem we study is

$$\max_{x_i, e} \left\{ \frac{v_r}{e_T} \mid N\mathbf{v} = 0, \sum_j e_j = e_T \right\}, \tag{4}$$

with e_j as the concentration of enzyme j . Thus, we wish to maximise a given output flux v_r per unit of total invested enzyme e_T of a metabolic network at steady state.

The optimisation problem stated in Eq (4) is equivalent to minimising the amount of enzyme necessary to sustain a given steady-state flux v_r at rate V_r ,

$$\min_{x_i, e} \left\{ \frac{\sum_j e_j}{v_r} \mid N\mathbf{v} = 0, v_r = V_r \right\}. \tag{5}$$

A crucial observation is now that since reaction functions generally are of the form $v_j = e_j f_j(\mathbf{x}_i; \mathbf{x}_E)$ [31], we may prescribe $v_r = 1$. After all, if we can solve that problem then we can solve it for $v_r = V_r$ as well by multiplying all the enzyme concentrations by V_r , because the specific flux v_r/e_T remains the same. Hence, we simplify (5) to

$$\min_{x_i, e} \left\{ \sum_j e_j \mid N\mathbf{v} = 0, v_r = 1 \right\}. \tag{6}$$

The relation $v_j = e_j f_j(\mathbf{x}_i; \mathbf{x}_E)$ may also be used to write $e_j = v_j / f_j(\mathbf{x}_i; \mathbf{x}_E)$ and rewrite (6) to

$$\min_{x_i} \left\{ \sum_j \frac{v_j}{f_j(\mathbf{x}_i; \mathbf{x}_E)} \mid N\mathbf{v} = 0, v_r = 1 \right\}. \tag{7}$$

Observe that the enzyme concentration vector \mathbf{e} has disappeared from the problem. (Note also that this optimisation is not a stoichiometric-model optimisation, such as flux balance analysis [32]. The *q*ORAC method takes into account the kinetics of the metabolic enzymes and the metabolite concentrations are the variables in this approach. The outcome of *q*ORAC is the definition of a self-optimising dynamical system; this has nothing to do with the optimisation associated with stoichiometric modelling.)

It has recently been shown that the flux profiles that solve (7) (and therefore also the equivalent original problem (4)) are always subnetworks with a particularly simple structure, called Elementary Flux Modes (EFMs; [30, 29]). Such EFMs are one-degree-of-freedom flux vectors

satisfying $Nv = 0$ that cannot be simplified further by deleting reactions without violating the steady state assumption [33, 34]. A given EFM is thus characterised by $\lambda(V_1, \dots, V_m)$, where λ is a free parameter and the flux vector (V_1, \dots, V_m) is fixed.

Step 2: Unique allocation within a given EFM

If we want to optimise specific flux *within a given EFM* with flux vector (V_1, \dots, V_m) , we still need to find a vector x_I for

$$\min_{x_I} \left\{ \sum_j \frac{\lambda V_j}{f_j(x_I; x_E)} \right\}. \tag{8}$$

This motivates the introduction of the *objective function*

$$O(x_I) := \sum_j \frac{\lambda V_j}{f_j(x_I; x_E)}, \tag{9}$$

which is to be minimised, for given external concentrations x_E , by suitably choosing internal concentrations x_I . This function is convex for pathways with many kinds of reaction kinetics [11], and in the Supporting Information (SI) we show that it is in fact strictly convex, for an even larger class of rate laws. Hence, the optimum is uniquely specified by the external concentrations x_E .

Note that the objective function has a lower value if the values of $f_j(x_I; x_E)$ are higher. Maximising specific flux may thus be reinterpreted as maximising the values of all f_j 's simultaneously. These f_j are closely associated to the saturation levels of enzyme j with its reactants (and effectors). This optimisation can be done by making as little enzyme as possible, so that the enzymes are used at their maximal capacity.

If we find the vector x_I^o which minimises $O(x_I)$, then we can infer the corresponding optimal enzyme concentrations e^o by setting

$$e_j^o = \frac{\lambda V_j}{f_j(x_I^o; x_E)}. \tag{10}$$

It is clear that we may choose $\lambda = 1$ in $O(x_I)$: having found the minimiser of $O(x_I)$ for $\lambda = 1$, we have found it for all λ : the corresponding enzyme levels e_j^o just scale with λ . In hindsight, we may also for instance normalise the enzyme concentrations such that they sum to a given total concentration e_T .

Step 3: Implementing qORAC: Choosing the right (number of) sensors

At this stage, the optimal enzyme concentrations that maximise the specific flux at steady state are still defined in terms of external concentrations x_E : for each choice of x_E , the objective function (9) needs to be minimised to find x_I^o , and subsequently e^o needs to be calculated. In order to characterise gene regulatory networks that produce the right concentrations of enzymes in steady state, robustly with respect to changes in external concentrations but without direct knowledge of those changes, we need to understand the defining characteristics of optimal solutions.

Steady-state optimisers x_I^o are minima of $O(x_I)$, and are dependent on (i.e., parameterised by) x_E . So, x_I^o is a (in fact, the) critical point of $O(x_I) = O(x_1, \dots, x_n)$, satisfying the *optimality*

relations

$$0 = \frac{\partial O}{\partial x_i} = \frac{\partial}{\partial x_i} \sum_j \frac{V_j}{f_j(\mathbf{x}_I; \mathbf{x}_E)}, \quad i = 1, \dots, n. \quad (11)$$

So instead of minimizing $O(\mathbf{x})$ for given external conditions \mathbf{x}_E , we could solve (11) by prescribing \mathbf{x}_E and solving for the remaining variables, the internal concentrations \mathbf{x}_I . However, the gene network does not have access to \mathbf{x}_E . Eq (11) should be solved with knowledge of the current sensor concentrations only. We therefore solve (11) by prescribing a subset of the internal metabolite concentrations, sensor values \mathbf{x}_S , and solving for all remaining concentrations, namely all other internal concentrations, but now also the (unknown) external concentrations. The solution is denoted by $\xi = (\xi_I, \xi_E)$, and is the estimated optimal concentration vector, under the assumption of steady state and optimality of the sensor values. In short, we call ξ the optimum as predicted by the sensors. Here, ξ_E are the external concentrations for which the current sensor values would have been optimal if the pathway had been in steady state. The part of ξ_I corresponding to sensor metabolites, ξ_S , of course coincides with the real concentrations \mathbf{x}_S , by construction. Since ξ is defined by \mathbf{x}_S , we denote it by $\xi(\mathbf{x}_S)$.

To solution of $\partial O(\mathbf{x})/\partial x_i = 0$ for different sensor values is well-defined mathematically if the Implicit Function Theorem (IFT) holds (see SI for a more detailed exposition). In essence, this means that it is then possible to calculate the optimal allocation by varying the sensors appropriately. The sensors are able to “track” the optima. Any choice of sensor metabolites for which the IFT holds is a candidate for the proposed adaptive control. An immediate consequence of the IFT is that the number of sensor metabolite concentrations must equal the number of changing external metabolite concentrations to which the system needs to be robust. This makes intuitive sense: to track changes (and hence achieve robustness) in N parameters, the gene network should be influenced by (at least) N (independent) internal sensors. Examples of parameters are environmental nutrient concentrations, temperature, pH and toxin concentrations.

Step 4: The *q*ORAC pathway has a unique steady state, the optimum

With $\xi(\mathbf{x}_S)$, we can define corresponding predicted optimal enzyme levels, analogous to (10), by setting

$$e_j^o = \frac{V_j}{f_j(\xi(\mathbf{x}_S))}. \quad (12)$$

At these enzyme concentrations, the pathway is either in steady state or not. If not, the metabolic concentrations are still changing, including the sensor concentrations. Hence, the predicted optimal enzyme levels also change. This argument indicates that the only steady state of the metabolic network steered in this fashion is the optimal one.

In the SI we prove that an EFM metabolic pathway with added *q*ORAC control has a unique steady state, the optimum. The proof is fully worked out for linear chains of enzymatic reactions (Theorem 3 in SI), but the techniques of the proof extend to a much larger class of pathways. All one needs to require is that for each choice of enzyme concentrations, the metabolic pathway has a unique steady state (a common enough assumption), and that the sensors are a few reaction steps away from the external concentrations (which makes intuitive sense). This result therefore ensures that when the *q*ORAC-controlled pathway has reached a steady state, it necessarily must be optimal.

Putting it all together

We now finish by implementing the enzyme synthesis rate functions E_j in

$$\dot{e}_j = E_j - \mu e_j.$$

By setting

$$E_j = \mu \frac{V_j}{f_j(\xi(\mathbf{x}_s))}, \tag{13}$$

we have ensured that at steady state the enzyme levels are optimal. The complete construction is termed *qORAC*, and is summarised in Definition 1. A fully-worked out example for the small pathway shown in Figs 3 and S2 is specified in Example 1.

Definition 1 (*qORAC*): The following differential-algebraic system of equations implements Specific Flux (*q*) Optimisation by Robust Adaptive Control (*qORAC*) through an EFM with flux vector (V_1, \dots, V_m) in a cell culture growing at fixed growth rate μ . Let I be the index set of internal metabolite concentrations, E the index set of external concentrations, and S the index set of sensor concentrations. Let furthermore $O(\mathbf{x}_I) = \sum_{j=1}^m V_j/f_j(\mathbf{x}_I; \mathbf{x}_E)$ be the objective function. Then we consider for $i \in I$, and $j = 1, \dots, m$,

$$\dot{x}_i = \sum_{j=1}^m N_{ij} v_j = \sum_{j=1}^m N_{ij} e_j f_j(\mathbf{x}_I; \mathbf{x}_E), \tag{14}$$

$$\dot{e}_j = E_j(\mathbf{x}_s) - \mu e_j, \tag{15}$$

$$E_j(\mathbf{x}_s) = \mu \frac{V_j/f_j(\xi(\mathbf{x}_s))}{\sum_{i=1}^m V_i/f_i(\xi(\mathbf{x}_s))}, \tag{16}$$

where $\xi(\mathbf{x}_s) = (\xi_I(\mathbf{x}_s), \xi_E(\mathbf{x}_s))$ is the predicted optimum, and is the (time-dependent) solution of

$$\xi_s = \mathbf{x}_s, \tag{17}$$

$$\frac{\partial O}{\partial \xi_i}(\xi) = 0. \tag{18}$$

The rescaling of $E_j(\mathbf{x}_s)$ in (16) by the sum of all the inverses of $1/f_j$ implies that total enzyme concentration is chosen to be equal to 1. Other rescalings give identical results, up to the chosen scaling factor. The choice above, however, is particularly useful, since it ensures positive synthesis rates both for positive and negative metabolic rates through the pathway, and it ensures that it is well-defined also at thermodynamic equilibrium (see SI for details).

Example 1: *qORAC* for a simple pathway The example *qORAC*-controlled metabolic pathway from Figs 4 and S2 is specified by the following set of equations for the metabolite concentrations $\mathbf{x} = (x_1, \dots, x_4) = ([C], [C'], [N'], [C_3N_2])$. Note that $x_1 = [C]$ is an external concentration which may change value periodically, as shown in Fig 4.

$$\begin{aligned} \dot{x}_1 &= 0, \\ \dot{x}_2 &= v_1 - v_3, \\ \dot{x}_3 &= v_2 - v_3, \\ \dot{x}_4 &= v_3 - v_4, \end{aligned}$$

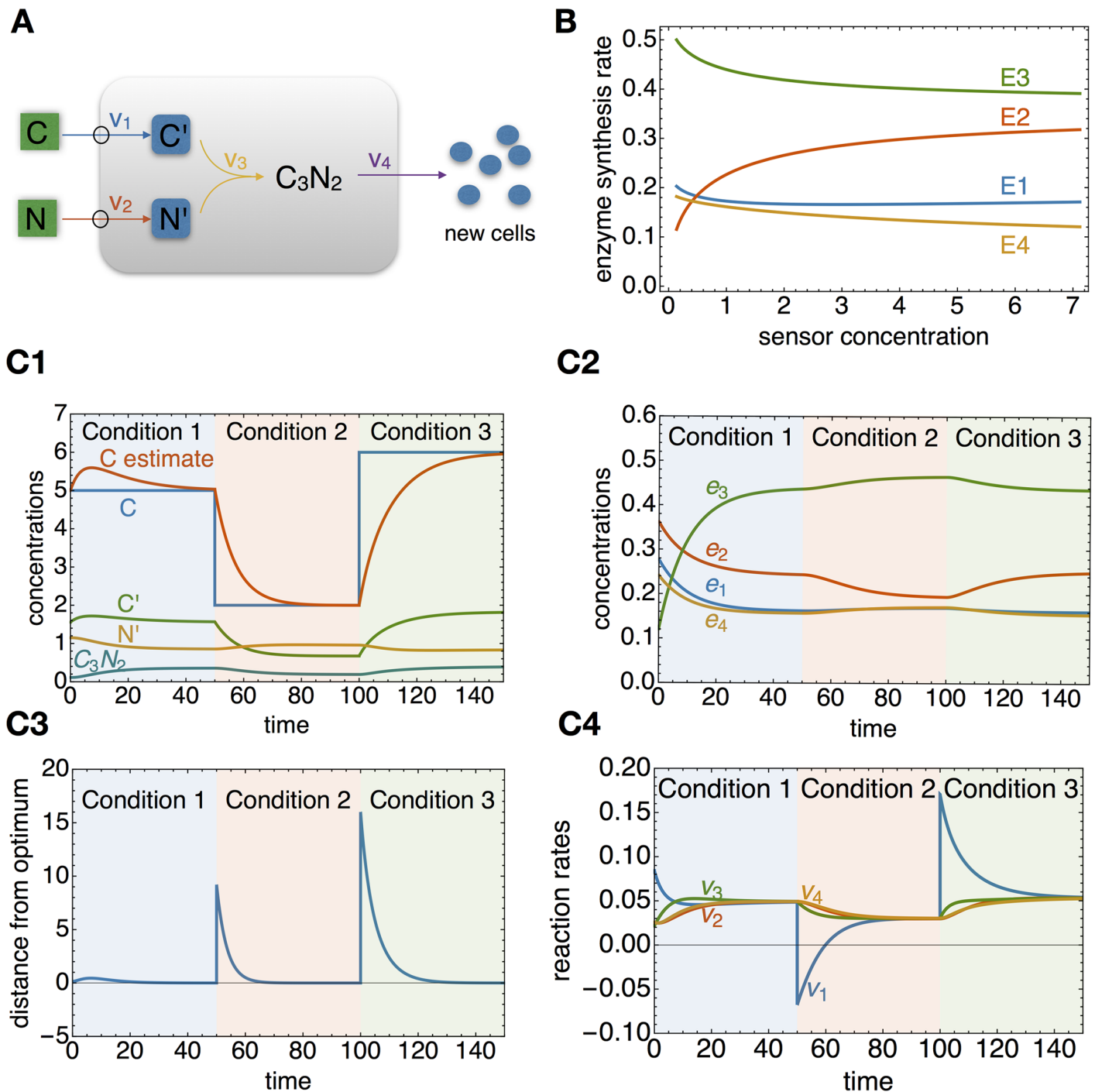


Fig 4. Example qORAC dynamics. The dynamics are illustrated for the network shown in A. The green box depicts a varying external concentration, the blue box denotes the sensor concentration. B: the optimal input-output relations, showing enzyme synthesis rates as a function of changing sensor concentration C' . In plots C1 to C4, the external C concentration is changed after 50 time units, and again after 100 time units. C1: The optimal C concentration predicted by the sensor (red line) converges to the real external C concentration (blue), even when the external concentration changes at $t = 50$ and $t = 100$. C2: enzyme dynamics equilibrate after each change in external conditions, and reach their optimal levels. C3: the steered metabolic pathway reaches the optimum after each external change, as the distance to the (periodically changing) optimum reaches zero after some time. C4: flux dynamics equilibrate, showing that the pathway has reached steady state each time the external conditions change. Full equations are given in Box 2, code is given in the SI.

<https://doi.org/10.1371/journal.pcbi.1006412.g004>

where $v_i = e_i f_i(\mathbf{x})$, $i = 1, \dots, 4$, and the kinetics functions $f_i(\mathbf{x})$ are defined by

$$\begin{aligned} f_1(\mathbf{x}) &= \frac{0.6x_1 - 0.75x_2}{(0.2x_1 + 1.0)(0.33x_2 + 1.0)}, \\ f_2(\mathbf{x}) &= \frac{[N] - 3.0x_3}{(0.5x_3 + 1)([N] + 1)}, \\ f_3(\mathbf{x}) &= \frac{0.2x_2x_3 - 0.17x_4}{(0.2x_3/5 + 1.0)(0.33x_2 + 0.17x_4 + 1.0)}, \\ f_4(\mathbf{x}) &= \frac{x_4 - 0.0025}{0.33x_4 + 1.0025}. \end{aligned}$$

The objective function is given by $O(\mathbf{x}) = \frac{1}{f_1(\mathbf{x})} + \dots + \frac{1}{f_4(\mathbf{x})}$. The enzyme dynamics are given $\dot{e}_j = E_j(x_2) - e_j$, $j = 1, \dots, 4$, where

$$E_j(x_2) = \frac{1/f_j(\xi(x_2))}{\sum_k 1/f_k(\xi(x_2))}, \quad j = 1, \dots, 4$$

and the predicted optimum $\xi(x_2)$ is defined by

$$\begin{cases} \xi_2 &= x_2, \\ \frac{\partial O}{\partial \xi_2}(\xi_1, \dots, \xi_4) &= \frac{\partial O}{\partial \xi_3}(\xi_1, \dots, \xi_4) = \frac{\partial O}{\partial \xi_4}(\xi_1, \dots, \xi_4) = 0. \end{cases}$$

Illustrations of the q ORAC framework

A toy metabolic network, with two external parameters and one output flux, is shown in Fig 4 (see Box 2 for the mathematical implementation). In this example, only the external [C] concentration is allowed to vary, so one internal sensor metabolite is required. Upon changes in this external concentration, the sensor concentration changes, causing changes in enzyme synthesis, which finally result in adaptation to the new optimum. The optimal enzyme synthesis relations of the gene network are also shown. They are simple curves, suggesting that small gene circuits are sufficient for optimal steering of this pathway.

To illustrate the general applicability of q ORAC, consider the complicated branched example network in Fig 5. It has two inputs and two outputs and two allosteric interactions; by employing four sensors, it can be made robust to changes in all four external concentrations.

The q ORAC framework is able to start from nearly any initial condition. As an extreme example, with no enzymes present, and only the sensor concentration and no other internal metabolite, the q ORAC-controlled pathway still steers to optimum (S1 Fig). Similarly, if the sensor concentrations are ‘wrong’, such that they predict a metabolic flow in the opposite direction to the one dictated by external concentrations, the combined controlled system nevertheless converges to the correct optimum (S2 Fig).

The q ORAC control does not guarantee that a metabolic pathway is actually steered towards the optimum. In an example in which one of the periodically changing parameters is a K_m parameter of a rate law, the choice of sensors matters critically (Figs 6 and S3). With one choice, the system robustly steers to the optimal specific flux steady state, but with another choice it does not. In both cases, the technical requirements to use the internal metabolites as sensors are met.

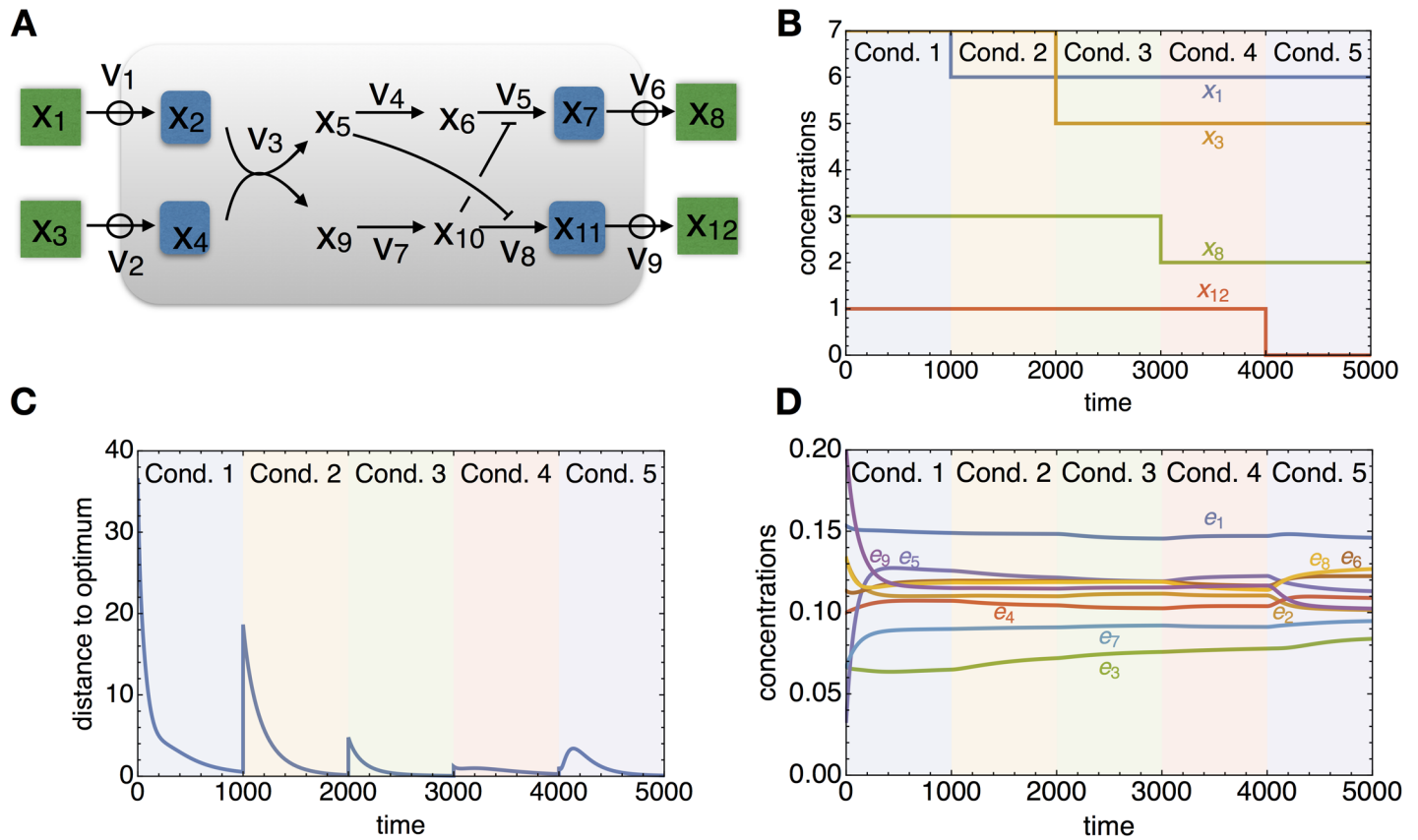


Fig 5. Example dynamics for a more complicated pathway. A: metabolic network with two inputs and two outputs, and with allosteric cross-inhibition. This pathway is robust to changes in both input and output concentration (in green), which requires four sensors (in blue). B: Each of the external concentrations is changed once, and the system adapts accordingly. C: the metabolite concentrations converge to the (periodically changing) predicted optimum over time. D: enzyme concentration dynamics. See SI text for details of the pathway, and the matlab file `daes_double_branched.m` for the code.

<https://doi.org/10.1371/journal.pcbi.1006412.g005>

Biological examples

In each of the pathways shown in Fig 2A–2D, the sensor metabolite(s) and transcription factor(s) have been identified. Specifying the kinetics for each enzymatic step in the pathway

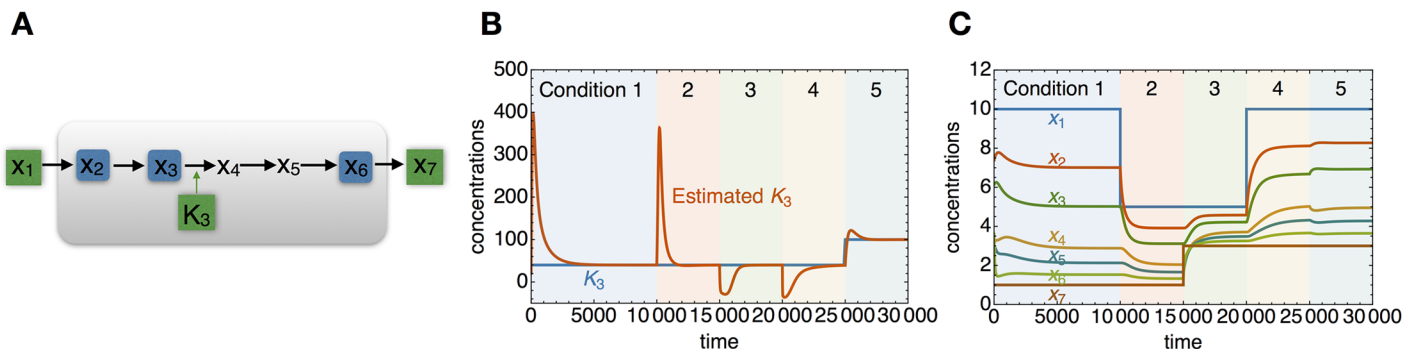


Fig 6. qORAC for an internal parameter. In this example qORAC is illustrated for a K_m parameter in the third reaction, K_3 . In A the same pathway is drawn, with sensors in blue. B: metabolite dynamics in which first external concentrations are varied (green) and at the end also K_3 is varied. C: K_3 (in green) is varied at time $t = 2500$, and the predicted optimal value (in orange) subsequently converges, illustrating robust adaptive control. An example in which the same pathway is controlled using a different set of sensors, resulting in lack of convergence to the optimum, is found in the SI, S4 Fig.

<https://doi.org/10.1371/journal.pcbi.1006412.g006>

now directly gives the corresponding objective function (9) and the q ORAC framework can be set up. The case of galactose uptake (Fig 2B) in yeast has been studied theoretically in detail by [28], including fitting the parameters of the well-characterised GAL gene network to approximate optimal input-output relations. Recent experimental evidence moreover shows that yeast cells are indeed able to tune the levels of these enzymes to optimise growth rate ([9]; Fig 1A).

Discussion

Experimental evidence is accumulating that suggests that cells can tune their enzyme resources to maximise growth rate [1, 2, 3, 4, 5, 6, 7, 8, 10]. We addressed whether cells growing at a fixed rate can tune limited enzyme resources to steer metabolism to optimal flux states, given only limited information about the current metabolic state of the cell in the form of sensor-metabolite concentrations. We demanded robustness of optimality in the face of environmental changes. We logically derived the q ORAC framework, which implements such control for Elementary Flux Modes, the minimal steady state pathways that maximise specific flux [29, 30]. Maximisation of specific fluxes is a requirement for maximisation of the specific growth rate of cells.

We use the term Specific Flux (q) Optimisation by Robust Adaptive Control (q ORAC) to describe the regulatory mechanism that we study. ‘Robust’ signifies that attaining optimal states is independent of (environmental) parameter values—the system is robust to them. ‘Adaptive’ means that the control system steers the metabolic system to optimality without direct knowledge of external changes, contrary to the more widely studied problem of ‘optimal control’, in which the steering mechanism works using external changes as inputs to the controller [35].

It is important to note that the growth rate itself is not optimised in our approach. Maximising steady state growth rate rather than specific flux requires a fundamentally different approach. The modelling framework should be extended to Metabolite-Enzyme models in which enzymes are made from precursors [36, 37]. In such models, the growth rate features quadratically rather than linearly, in the resulting steady state and optimality equations. EFMs therefore no longer apply, and the objective function $O(x)$ is also absent. Our approach is therefore more suitable to isolated pathways than to all of metabolism. For such smaller pathways, it is more reasonable to assume that there is a fixed amount of enzyme resources to distribute, and that the cellular growth rate is considered constant. Recent work does suggest, however, that the objective function $O(x)$ studied here in fact matters to cells also on a more global metabolic level [11].

An important finding of our work is that the number of sensor metabolites must be (at least) equal to the number of parameters for which the metabolic pathway is robustly optimal. In other words, if the metabolic pathway always achieves states of maximal specific flux, regardless of the values of three (independently changing) environmental parameters, such as, for example, osmolarity, temperature and some nutrient concentration, then the number of sensors is expected to be three. This is a general result that follows from the associated mathematics of this control problem. Finding the sensors experimentally is difficult, and the number of known sensors is still quite small. However, it is telling that the whole of central carbon metabolism in *E. coli* seems to be controlled by just three sensors, FBP, cAMP and F1P [25].

The identity of suitable sensors does not follow immediately from the optimisation problem. In general, one needs to make sure that the Implicit Function Theorem applies to the optimum Eq (11), and this is not a trivial matter. However, a different argument shows that sensors near the beginning or ends of the pathway would work in most cases. The reason is that for all metabolites in between a set of fixed concentrations, their optimal value is uniquely

determined by minimising the corresponding optimisation problem (i.e. finding the minimum of a suitable objective function $O(\mathbf{x}; \mathbf{x}_S)$ with \mathbf{x} the set of metabolites between the sensors \mathbf{x}_S). The remaining variables, including the external concentrations, then need to be determined using the optimum Eq (11). This is easiest (it involves the smallest number of equations and unknowns to solve for) when sensors are close to the external metabolites. Also from a biological standpoint this makes sense: such sensors obviously provide the most information of any change in external concentrations.

An important question is whether the adaptive control can be achieved by molecular circuits, given our understanding of biochemical kinetics and molecular interactions. The explicit example from galactose metabolism in yeast [28] gives hope that this might be true in general. If the necessary gene network is small, then the optimal circuit is likely also evolvable. We cannot give definite answers about this, but the computational analyses of different networks, of which some are shown in this paper, indicate that *q*ORAC-controlled networks show remarkably simple dynamics and input-output relations. One would expect that biochemical systems are capable of evolving those, and that synthetic biologists are capable of designing them.

The parameterisation of the optimising circuit is completely determined by the kinetics and the wiring of the metabolic pathway that it controls, since the objective function (9) contains only this information. This interdependence between the controller and the controlled is sometimes called the ‘internal model principle’ in engineering [26] which roughly states that the control system should have knowledge of the dynamic behaviour of the system in order to be able to control it. Additional control mechanisms may then prevent for instance undesired oscillations or slow responses.

The internal model principle, applied to metabolic pathway control, suggests a new perspective on the larger problem of understanding metabolic regulation. The theory presented here indicates that knowledge of the metabolic pathway, including properties of catalysing enzymes, is sufficient to understand how this pathway needs to be controlled to maximise flux. It is not necessary to know the controlling regulatory pathway in advance. This offers hope for situations in which this circuit has not been characterised yet, or for which it needs to be designed synthetically.

Technological advances have spurred recent interest in studying control properties of gene regulatory networks in cellular metabolism. One line of work involves characterising a particular gene control system and studying its theoretical properties. Examples are the perfect adaptation in the chemotaxis network in *E. coli* [38, 39], the robustness properties of the heat-shock response system [40] and of the circadian clock [41]. Several authors have considered dynamic optimisation of resources in pathways from a mostly computational perspective, e.g. to minimise the time of adaptive response [42], *de*FBA [43], and for other objectives than maximal specific flux, such as detecting equilibrium regimes of pathways [44], robustness to flux perturbations [45], and noise propagation [46]. In many studies, the control is not adaptive, but optimal; the objective is then usually to maximise the long term production of biomass [47, 42, 48, e.g.].

The approach taken here differs principally from most previous works in the following respect. The objective (maximal specific flux) is defined in advance, and the optimal input-output relations are characterised later. The framework is also analytic rather than computational: the input-output relations are obtained by solving the optimum equations (11) for the pathway, rather than by using a numerical optimisation routine. The latter is impossible, since this would require knowing the external concentrations.

A few recent papers have used adaptive controls similar to ours. So-called Flux Control Regulation (FCR; [49]) comes closest, and uses the same type of adaptive control as proposed in *q*ORAC. FCR also explicitly relies on making estimates at each time point under the

assumption of steady state. When the system is in fact in steady state, it has reached the desired objective. The principle difference between FCR and q ORAC lies in the objective. The input-output relations in FCR come from measurements and ensure steady state properties only. q ORAC, however, solves a steady state optimisation problem, and constructs input-output relations directly from the kinetic rate laws of the metabolic pathway itself. Another recent example of a coarse-grained model of cellular physiology including gene expression control can be found in [50]. Two other examples using adaptive control are from the context of optimal ribosomal allocation to maximise the growth rate in *E. coli*. The free amino acid concentration acts as a sensor to ppGpp, which downstream influences gene expression. Two models have been proposed that are based on optimal synthesis of ribosomes so as to maximise growth rate [15, 16]. The input-output relations used in these models are not derived from kinetic properties as in q ORAC, but are designed by hand to approximate maximal growth rates in different conditions.

The choice of sensors sometimes matters for the control to steer the pathway to optimum (Figs 6 and S4). This example already indicates that, although the q ORAC control follows logically from the design objective, it is not easy to decide which intermediate metabolites make it controllable. We cannot expect completely general mathematical theorems. Apparently, some choices of sensors do work, and others do not, for the same pathway, using the same initial conditions. A second, mathematical reason why one cannot expect convergence to optimal states is that if time would be reversed, the control would remain the same, but dynamics would be reversed. The control is based on steady state properties of the system, and these do not change upon time reversal.

q ORAC has direct applications in synthetic biology. To achieve maximal production rates in a biotechnological-product producing pathway requires a controller that q ORAC provides. The only ingredient to design such a controller are the enzymatic rate laws in the pathway. q ORAC then immediately makes predictions about the optimal enzyme synthesis rates, as a function of one or more intermediate metabolites. As the synthetic biology field advances, synthetic circuits with the required input-output relationships for the constituent enzymes of the pathway can be designed and built. q ORAC therefore does not only contribute to the general understanding of steering mechanisms to optimal states, but provides direct operational relevance for microbiology, synthetic biology and biotechnological applications.

Supporting information

S1 Text. Supporting information text in which we prove that the optimisation problem (8) has a unique solution for a large class of reaction kinetics. We also give a detailed explanation which and how many sensor metabolites may be used in q ORAC. We prove that many pathways with q ORAC control only have one steady state, the actual optimum. We also give additional illustrations of the q ORAC formalism, give details on the numerical integration of q ORAC-controlled pathways, and fully describe the kinetics of the pathways considered in this paper.

(PDF)

S1 Fig. Extreme robustness: A simulation for the same pathway as in Fig 3 in the main text, but now with minimal initial conditions: At the start, enzymes are completely absent, and all internal metabolites except the sensor are absent. The pathway is still steered to the optimal specific flux steady state. A: the pathway; B: metabolite concentrations over time (all except orange), and predicted optimal external metabolite concentration (orange); C: reaction fluxes over time; D: enzyme concentrations over time. See Box 2 in main text for details of the

pathway, and the matlab file `daes_CN_minimal_ICs.m` for the code.
(TIF)

S2 Fig. With initial conditions for the sensor concentration such that they actually predict an optimal flow from end to beginning rather than the reverse, the predicted optimum needs to ‘straighten out’, and move through a singular point: Thermodynamic equilibrium. Although the requirements for sensor control are not upheld in this point, the predicted optimum moves smoothly through this singular points and the system adapts as it should. A: linear chain pathway, with external conditions such that flow is initially from x_1 to x_7 ; B: dynamics for the predicted optimal metabolite concentrations (ξ). The intersection point of all the curves is thermodynamic equilibrium; C: reaction flux dynamics. Note that the fluxes do not pass through $v_1 = \dots = v_6 = 0$. Three do, and the others do not, in this example. See SI text for details of the pathway, and the matlab file `daes_linearchain_reversal.m` for the code.
(TIF)

S3 Fig. An example of qORAC steering for an changing internal parameter, showing lack of convergence to the optimum. A: The pathway, which is identical to that in Fig 5 in the main text—only the choice of sensors (in blue) is different. Sensor x_3 is swapped with x_4 . B/C: The dynamics of metabolites (B) and predicted K_3 values (C) do start to change. However, the dynamics converge to a singular point, and the dynamical system can not continue. This second choice of sensors does not yield a gene expression control system which steers the pathway to optimal specific flux. See matlab code `daes_extra_param_wrong.m` for the code.
(TIF)

S4 Fig. King-Altman patterns of an ordered $M - K$ mechanism. The left one shows the master pattern, the middle and right figure show two alternative patterns that yield the constant term in the denominator of f .
(TIF)

Author Contributions

Conceptualization: Robert Planqué, Bas Teusink, Frank J. Bruggeman.

Formal analysis: Robert Planqué, Bas Teusink, Johannes C. Hendriks, Frank J. Bruggeman.

Funding acquisition: Bas Teusink.

Investigation: Robert Planqué, Josephus Hulshof, Bas Teusink, Johannes C. Hendriks, Frank J. Bruggeman.

Software: Robert Planqué.

Supervision: Robert Planqué.

Visualization: Robert Planqué.

Writing – original draft: Robert Planqué, Josephus Hulshof, Bas Teusink, Johannes C. Hendriks, Frank J. Bruggeman.

Writing – review & editing: Robert Planqué, Josephus Hulshof, Bas Teusink, Frank J. Bruggeman.

References

1. Walsh K, Koshland DE. Characterization of rate-controlling steps in vivo by use of an adjustable expression vector. Proc Nat Acad Sciences USA. 1985; 82:3577–3581. <https://doi.org/10.1073/pnas.82.11.3577>

2. van der Vlag J, van Dam K, Postma PW. Quantification of the regulation of glycerol and maltose metabolism by IAGlc of the phosphoenolpyruvate-dependent glucose phosphotransferase system in *Salmonella typhimurium*. *J Bact.* 1994; 176(12):3518–3526. <https://doi.org/10.1128/jb.176.12.3518-3526>. 1994 PMID: 8206828
3. Jensen PR, Michelsen O, Westerhoff HV. Experimental determination of control by the H⁺-ATPase in *Escherichia coli*. *J Bioenerg Biomem.* 1995; 27:543–554. <https://doi.org/10.1007/BF02111653>
4. Andersen HW, Pedersen MB, Hammer K, Jensen PR. Lactate dehydrogenase has no control on lactate production but has a strong negative control on formate production in *Lactococcus lactis*. *FEBS Journal.* 2001; 268(24):6379–6389.
5. Solem C, Koebmann BJ, Jensen PR. Glyceraldehyde-3-Phosphate Dehydrogenase Has No Control over Glycolytic Flux in *Lactococcus lactis* MG1363. *J Bact.* 2003; 185(5):1564–1571. <https://doi.org/10.1128/JB.185.5.1564-1571.2003> PMID: 12591873
6. Koebmann B, Solem C, Jensen PR. Control analysis as a tool to understand the formation of the *las* operon in *Lactococcus lactis*. *FEBS Journal.* 2005; 272(9):2292–2303. <https://doi.org/10.1111/j.1742-4658.2005.04656.x> PMID: 15853813
7. Dekel E, Alon U. Optimality and evolutionary tuning of the expression level of a protein. *Nature.* 2005; 436:588–592. <https://doi.org/10.1038/nature03842> PMID: 16049495
8. Li GW, Burkhardt D, Gross C, Weissman JS. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell.* 2014; 157(3):624–635. <https://doi.org/10.1016/j.cell.2014.02.033> PMID: 24766808
9. Keren L, Haussler J, Lotan-Pompan M, Slutskin IV, Alisar H, Kaminski S, et al. Massively parallel interrogation of the effects of gene expression levels on fitness. *Cell.* 2016; 166(5):1282–1294.e18. <https://doi.org/10.1016/j.cell.2016.07.024> PMID: 27545349
10. Towbin BD, Korem Y, Bren A, Doron S, Sorek R, Alon U. Optimality and sub-optimality in a bacterial growth law. *Nature Comm.* 2017; 8:14123. <https://doi.org/10.1038/ncomms14123>
11. Noor E, Flamholz A, Bar-Even A, Davidi D, Milo R, Liebermeister W. The Protein Cost of Metabolic Fluxes: Prediction from Enzymatic Rate Laws and Cost Minimization. *PLoS Comp Biol.* 2016; 12(11): e1005167. <https://doi.org/10.1371/journal.pcbi.1005167>
12. Weiße A, Oyarzún DA, Danos V, Swain PS. Mechanistic links between cellular trade-offs, gene expression, and growth. *Proc Nat Acad Sciences USA.* 2015; p. E1038–E1047.
13. Molenaar D, van Berlo R, de Ridder D, Teusink B. Shifts in growth strategies reflect tradeoffs in cellular economics. *Mol Syst Biol.* 2009; 5:323. <https://doi.org/10.1038/msb.2009.82> PMID: 19888218
14. Scott M, Gunderson CW, Mateescu EM, Zhang Z, Hwa T. Interdependence of Cell Growth and Gene Expression: Origins and Consequences. *Science.* 2010; 330:1099–1102. <https://doi.org/10.1126/science.1192588> PMID: 21097934
15. Scott M, Klumpp S, Mateescu EM, Hwa T. Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Mol Syst Biol.* 2014; 10:747. <https://doi.org/10.15252/msb.20145379> PMID: 25149558
16. Bosdriesz E, Molenaar D, Teusink B, Bruggeman FJ. How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximisation. *FEBS Journal.* 2015; 282(10):2029–44. <https://doi.org/10.1111/febs.13258> PMID: 25754869
17. Basan M, Hui S, Okano H, Zhang Z, Shen Y, Williamson JR, et al. Overflow metabolism in *E. coli* results from efficient proteome allocation. *Nature.* 2015; 528:99–104. <https://doi.org/10.1038/nature15765> PMID: 26632588
18. You C, Okano H, Hui S, Zhang Z, Kim M, Gunderson CW, et al. Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature.* 2013; 500:301–306. <https://doi.org/10.1038/nature12446> PMID: 23925119
19. Kotte O, Zaugg JB, Heinemann M. Bacterial adaptation through distributed sensing of metabolic fluxes. *Mol Syst Biol.* 2010; 6:355. <https://doi.org/10.1038/msb.2010.10> PMID: 20212527
20. Kochanowski K, Volkmer B, Gerosa L, van Rijsewijk BRH, Schmidt A, Heinemann M. Functioning of a metabolic flux sensor in *Escherichia coli*. *Proc Nat Acad Sciences USA.* 2013; 110(3):1130–1135. <https://doi.org/10.1073/pnas.1202582110>
21. Kotte O, Volkmer B, Radzikowski JL, Heinemann M. Phenotypic bistability in *Escherichia coli*'s central carbon metabolism. *Mol Syst Biol.* 2014; 10(7):736–736. <https://doi.org/10.15252/msb.20135022> PMID: 24987115
22. Sellick CA, Campbell RN, Reece RJ. Galactose Metabolism in Yeast—Structure and Regulation of the Leloir Pathway Enzymes and the Genes Encoding Them. *Int Rev Cell Mol Biol.* 2008; 269:111–150. [https://doi.org/10.1016/S1937-6448\(08\)01003-4](https://doi.org/10.1016/S1937-6448(08)01003-4) PMID: 18779058

23. Gollnick P, Babitzke P, Antson A, Yanofsky C. Complexity in regulation of tryptophan biosynthesis in *Bacillus subtilis*. *Ann Rev Genetics*. 2005; 39:47–68. <https://doi.org/10.1146/annurev.genet.39.073003.093745>
24. Gilbert W, Müller-Hill B. Isolation of the lac repressor. *Proc Nat Acad Sciences USA*. 1966; 56(6):1891–1898. <https://doi.org/10.1073/pnas.56.6.1891>
25. Kochanowski K, Gerosa L, Brunner SF, Christodoulou D, Nikolaev YV, Sauer U. Few regulatory metabolites coordinate expression of central metabolic genes in *Escherichia coli*. *Mol Syst Biol*. 2017; 13:903. <https://doi.org/10.15252/msb.20167402> PMID: 28049137
26. Francis BA, Wonham WM. The internal model principle of control theory. *Automatica*. 1976; 12(5):457–465. [https://doi.org/10.1016/0005-1098\(76\)90006-6](https://doi.org/10.1016/0005-1098(76)90006-6)
27. Lohr D, Venkov P, Zlatanova J. Transcriptional regulation in the yeast GAL gene family: a complex genetic network. *FASEB Journal*. 1995; 9(9):777–787. <https://doi.org/10.1096/fasebj.9.9.7601342> PMID: 7601342
28. Berkhout J, Teusink B, Bruggeman FJ. Gene network requirements for regulation of metabolic gene expression to a desired state. *Sci Rep*. 2013; 3:1417. <https://doi.org/10.1038/srep01417> PMID: 23475326
29. Müller S, Regensburger G, Steuer R. Enzyme allocation problems in kinetic metabolic networks: Optimal solutions are elementary flux modes. *J Theor Biology*. 2014; 347:182–190. <https://doi.org/10.1016/j.jtbi.2013.11.015>
30. Wortel MT, Peters H, Hulshof J, Teusink B, Bruggeman FJ. Metabolic states with maximal specific rate carry flux through an elementary flux mode. *FEBS Journal*. 2014; 281:1547–1555. <https://doi.org/10.1111/febs.12722> PMID: 24460934
31. Cornish-Bowden A. *Fundamentals of Enzyme Kinetics*. Portland Press, London; 1995.
32. Orth JD, Thiele I, Palsson BØ. What is flux balance analysis? *Nature Biotechnol*. 2010; 28(3):245–248. <https://doi.org/10.1038/nbt.1614>
33. Schuster S, Hilgetag C. On elementary flux modes in biochemical reaction systems at steady state. *J Biol Systems*. 1994; 2:165–185. <https://doi.org/10.1142/S0218339094000131>
34. Schuster S, Hilgetag C, Woods JH, Fell DA. Reaction routes in biochemical reaction systems: algebraic properties, validated calculation procedure and example from nucleotide metabolism. *J Math Biology*. 2002; 45:153–181. <https://doi.org/10.1007/s002850200143>
35. Krstic M, Kanellakopoulos I, Kokotovic P. *Nonlinear and Adaptive Control Design*. Wiley, New York; 1995.
36. Goelzer A, Fromion V, Scorletti G. Cell design in bacteria as a convex optimisation problem. *Automatica*. 2011; 47:1210–1218. <https://doi.org/10.1016/j.automatica.2011.02.038>
37. Lerman JA, Hyduke DR, Latif H, Portnoy VA, Lewis NE, Orth JD, et al. In silico method for modelling metabolism and gene product expression at genome scale. *Nature Comm*. 2012; 3:929. <https://doi.org/10.1038/ncomms1928>
38. Barkai N, Leibler S. Robustness in simple biochemical networks. *Nature*. 1997; 387:913–917. <https://doi.org/10.1038/43199> PMID: 9202124
39. Yi TM, Huang Y, Simon MI, Doyle J. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc Nat Acad Sciences USA*. 2000; 97(9):4649–4653. <https://doi.org/10.1073/pnas.97.9.4649>
40. El-Shamad H, Kurata H, Doyle JC, Gross CA, Khammash M. Surviving heat shock: Control strategies for robustness and performance. *Proc Nat Acad Sciences USA*. 2005; 102(8):2736–2741. <https://doi.org/10.1073/pnas.0403510102>
41. Stelling J, Gilles ED, III FJD. Robustness properties of circadian clock architectures. *Proc Nat Acad Sciences USA*. 2004; 101(36):13210–13215. <https://doi.org/10.1073/pnas.0401463101>
42. Pavlov MY, Ehrenberg M. Optimal control of gene expression for fast proteome adaptation to environmental change. *Proc Nat Acad Sciences USA*. 2013; 110(51):20527–20532. <https://doi.org/10.1073/pnas.1309356110>
43. Waldherr S, Oyarzún D, Bockmayr A. Dynamic optimization of metabolic networks coupled with gene expression. *J Theor Biology*. 2015; 365:469–485. <https://doi.org/10.1016/j.jtbi.2014.10.035>
44. Oyarzún DA, Chaves M, Hoff-Hoffmeyer-Zlotnik M. Multistability and oscillations in genetic control of metabolism. *J Theor Biology*. 2012; 295:139–153. <https://doi.org/10.1016/j.jtbi.2011.11.017>
45. Oyarzún DA, Stan GBV. Synthetic gene circuits for metabolic control: design trade-offs and constraints. *J Roy Soc Interface*. 2012; 10(78).
46. Oyarzún DA, Lugagne JB, Stan GBV. Noise Propagation in Synthetic Gene Circuits for Metabolic Control. *ACS Synth Biol*. 2015; 4:116–125. <https://doi.org/10.1021/sb400126a> PMID: 24735052

47. van den Berg HA, Kiselev YN, Kooijman SALM, Orlov MV. Optimal allocation between nutrient uptake and growth in a microbial trichome. *J Math Biology*. 1998; 37:28–48. <https://doi.org/10.1007/s002850050118>
48. Giordano N, Mairet F, Gouzé JL, Geiselmann J, de Jong H. Dynamical Allocation of Cellular Resources as an Optimal Control Problem: Novel Insights into Microbial Growth Strategies. *PLoS Comp Biol*. 2016; 12(3):e1004802. <https://doi.org/10.1371/journal.pcbi.1004802>
49. Erickson DW, Schink SJ, Patsalo V, Williamson JR, Gerland U, Hwa T. A global resource allocation strategy governs growth transition kinetics of *Escherichia coli*. *Nature*. 2017; 551:119–123. <https://doi.org/10.1038/nature24299> PMID: 29072300
50. Liao C, Blanchard AE, Lu T. An integrative circuit–host modelling framework for predicting synthetic gene network behaviours. *Nature Microbiol*. 2017; 2:1658–1666. <https://doi.org/10.1038/s41564-017-0022-5>