



HAL
open science

Optimal enzyme profiles in unbranched metabolic pathways

Elad Noor, Wolfram Liebermeister

► **To cite this version:**

Elad Noor, Wolfram Liebermeister. Optimal enzyme profiles in unbranched metabolic pathways. Interface Focus, 2024, 14, pp.20230029. 10.1098/rsfs.2023.0029 . hal-04448194

HAL Id: hal-04448194

<https://hal.inrae.fr/hal-04448194>

Submitted on 9 Feb 2024

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

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



Distributed under a Creative Commons Attribution 4.0 International License



Research

Cite this article: Noor E, Liebermeister W. 2024 Optimal enzyme profiles in unbranched metabolic pathways. *Interface Focus* **14**: 20230029.
<https://doi.org/10.1098/rsfs.2023.0029>

Received: 30 June 2023
Accepted: 19 October 2023

One contribution of 4 to a theme issue '50 Years of metabolic control analysis'.

Subject Areas:

systems biology, computational biology, biotechnology

Keywords:

metabolic pathway, enzymatic rate law, enzyme demand, optimization, protein allocation, bacterial growth law

Authors for correspondence:

Elad Noor
e-mail: elad.noor@weizmann.ac.il
Wolfram Liebermeister
e-mail: wolfram.liebermeister@inrae.fr

[†]Corresponding authors contributed equally to this study.

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.6976367>.

Optimal enzyme profiles in unbranched metabolic pathways

Elad Noor^{1,†} and Wolfram Liebermeister^{2,†}

¹Department of Plant and Environmental Sciences, Weizmann Institute of Science, 76100 Rehovot, Israel
²Université Paris-Saclay, INRAE, MaIAGE, 78350 Jouy-en-Josas, France

EN, 0000-0001-8776-4799; WL, 0000-0002-2568-2381

How to optimize the allocation of enzymes in metabolic pathways has been a topic of study for many decades. Although the general problem is complex and nonlinear, we have previously shown that it can be solved by convex optimization. In this paper, we focus on unbranched metabolic pathways with simplified enzymatic rate laws and derive analytic solutions to the optimization problem. We revisit existing solutions based on the limit of mass-action rate laws and present new solutions for other rate laws. Furthermore, we revisit a known relationship between flux control coefficients and enzyme abundances in optimal metabolic states. We generalize this relationship to models with density constraints on enzymes and metabolites, and present a new local relationship between optimal reaction elasticities and enzyme amounts. Finally, we apply our theory to derive simple kinetics-based formulae for protein allocation during bacterial growth.

1. Introduction

The idea that living beings show optimal shapes or behaviour has a very long history. A process like evolution, which combines random mutations with a selection for favourable properties, could potentially lead to optimization, but the question of *if* and/or *when* should we expect living beings to function optimally has been widely debated and is far from solved. In practice, it can be useful to invoke optimality principles to seek insights and design principles that might be relevant in naturally evolved systems [1]. Specifically, cell metabolism has often been studied using this approach [2,3], thanks to the powerful mathematical models that we have to describe it. But although natural selection has been the main inspiration for this study, the evolutionary aspects of pathway optimization are not discussed here, and are rather left for the reader to reflect upon.

Within cells, protein is arguably the most important and central resource, both in terms of contributing to fitness, but also since protein synthesis requires large amounts of energy, metabolic precursors, and ribosomes, and the proteins themselves occupy a significant portion of cellular space. Therefore, a cell should generally save protein wherever it can. This notion, specifically for enzymes, has been mathematically applied in genome-scale metabolic models [4,5], models of core metabolism [6], and in direct comparisons between pathways [7,8]. Since the proteome is a limiting resource, its allocation to different sectors (metabolism, translation, etc.) is a topic of high interest. In some works, people assumed optimality. Other papers assumed a general rule based on linear growth rate control. Interestingly, even a very simple linear chain model, with two reactions representing metabolism and protein synthesis and with a bound on the total protein budget, has been successful in explaining bacterial growth laws and overflow metabolism [9,10]. However, these bacterial growth law models did not consider enzyme kinetics.

Here, we focus on a special case of this cost/benefit analysis: the efficient use of metabolic enzymes in unbranched pathways operating at steady state,

giving priority to scenarios that can be solved analytically. We explore several types of kinetic rate laws and introduce the idea of bounding the total metabolite concentration (which is required in some cases for meaningful results). To define states of maximal enzyme efficiency, we can consider two equivalent optimality problems: maximizing a production flux at a given enzyme budget or minimizing protein usage at a given required production flux. In both cases, we maximize the production flux per enzyme usage within the given constraints. If the product of the pathway is directly tied to biomass, the overall enzyme efficiency, called ‘biomass/enzyme efficiency’, can serve as a proxy for cell growth [6]. Furthermore, this optimality problem is also relevant in other contexts, such as metabolic engineering of synthetic pathways using a set of existing and/or new-to-nature enzymes with known kinetic parameters [11,12].

Another perspective often used to analyse metabolic systems is through their control, e.g. the effect of changes in a level of an enzyme on pathway flux [13]. In optimal metabolic states with a bound on the sum of enzyme levels, each enzyme effectively carries an opportunity cost. This cost must be balanced by a marginal benefit, given by the flux control coefficient—as defined in metabolic control analysis (MCA) [13–15]. Hence, for systems in optimal states, there are simple relationships between enzyme abundance and flux control [16–18]. We will recapitulate these results below, generalize them to models with a density constraint on enzyme and metabolite levels, and illustrate them using some of our analytic solutions. In addition, we present a simple rule that links optimal enzyme investments around a given metabolite to the reaction elasticities of this metabolite.

Finally, we show how the analytic solutions derived here might be useful for modelling phenomena on the level of entire cells such as the Monod curve (i.e. the relationship between the concentration of a limiting nutrient and the growth rate of bacteria [19]). We use this model to demonstrate how each kinetic parameter should affect the growth rate under different conditions.

In summary, this paper revisits the question of optimal enzyme allocation and adds to previous results. We focus on unbranched metabolic pathways, extend the optimality problem, discuss new optimality conditions, and present analytic solutions that directly show how different factors determine optimal enzyme levels and fluxes. We discuss general principles, in particular how optimal enzyme levels reflect flux control and local reaction elasticities, and use our theory to derive formulae for kinetics-based bacterial growth.

2. Results

2.1. Optimal states of unbranched pathways

One of the first attempts at analytically solving the enzyme allocation problem was published by Waley [20], who studied short pathways of two to three reactions while assuming that the concentrations of metabolites (which were denoted *linking intermediates*) are much below the enzymes’ K_M values, and therefore affect the flux linearly. Given the total amount of catalytically active protein (bounded by ϵ_{tot}), the relative enzyme concentrations should be such that they maximize rate (figure 1). Based on these assumptions, one can derive simple formulae for the optimal enzyme levels and maximal pathway flux. Later studies repeated this result and

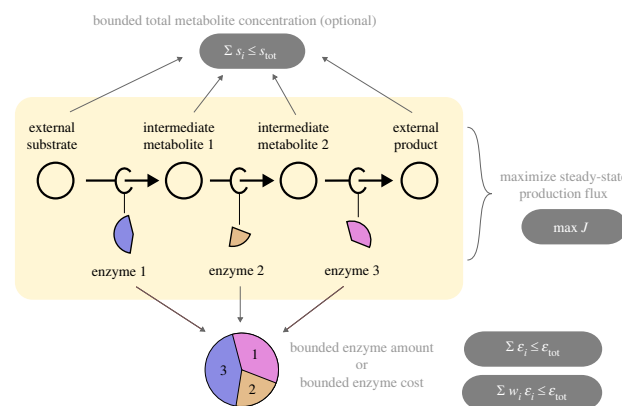
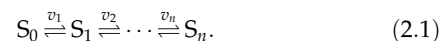


Figure 1. Optimal enzyme levels in unbranched metabolic pathways. In the basic optimality problem in this paper, we consider a chain of reactions and ask how a given protein budget should be spent on metabolic enzymes to achieve a maximal steady-state flux. If the efficiencies (i.e. the reaction rates per catalysing enzyme) of all the enzymes were known, the steady-state flux would determine the enzyme levels, and there would be nothing to optimize. Here we assume that the enzyme efficiencies can be adjusted by choosing the metabolite concentrations (not shown), and we search for the optimal metabolite and enzyme profile. The aim is to maximize the flux at a given total enzyme amount (bottom) and possibly under a constraint on metabolite concentrations (top). The flux ratios between reactions are predefined, for example, assuming equal fluxes in all the reactions (right).

generalized it to linear pathways of any size [17,18,21]. Here, we will revisit this general solution and extend it to other types of rate laws beyond the one considered by Waley [20] (which, from now on, we will refer to as *mass-action*).

Consider the following unbranched pathway [22] (figure 1):



In a kinetic model, each variable s_i represents the concentration of a metabolite i and each variable ϵ_i represents the level (molar concentration or mass concentration) of the enzyme catalysing reaction i . Imagine that the total enzyme level in the pathway is bounded by ϵ_{tot} , i.e.

$$\sum_i \epsilon_i \leq \epsilon_{\text{tot}}. \quad (2.2)$$

What would be the optimal strategy for distributing this resource between the reactions in order to maximize the steady-state flux? To answer this question, we need to know how the rate of each reaction depends on the levels of enzyme, substrate, and product, as well as on kinetic parameters. This is described by rate laws. For some rate laws, we can solve the optimization problem and obtain an analytic solution that describes exactly how much of each enzyme should be allocated. Below, we will also consider a variant of this problem with an extra bound on the sum of metabolite levels or with fixed initial and final metabolite concentrations (s_0 and s_n).

Since single analytic solutions are rare but instructive, we explore them in this article. We consider four different rate laws (summarized in figure 2): the general Haldane rate law (saturable and reversible) which has no analytic solution, and three solvable approximations derived from it. As explained above, the *enzyme levels* ϵ_i may refer to either molar concentrations or mass concentrations (depending on the modeller’s preference). In the case of molar concentrations,

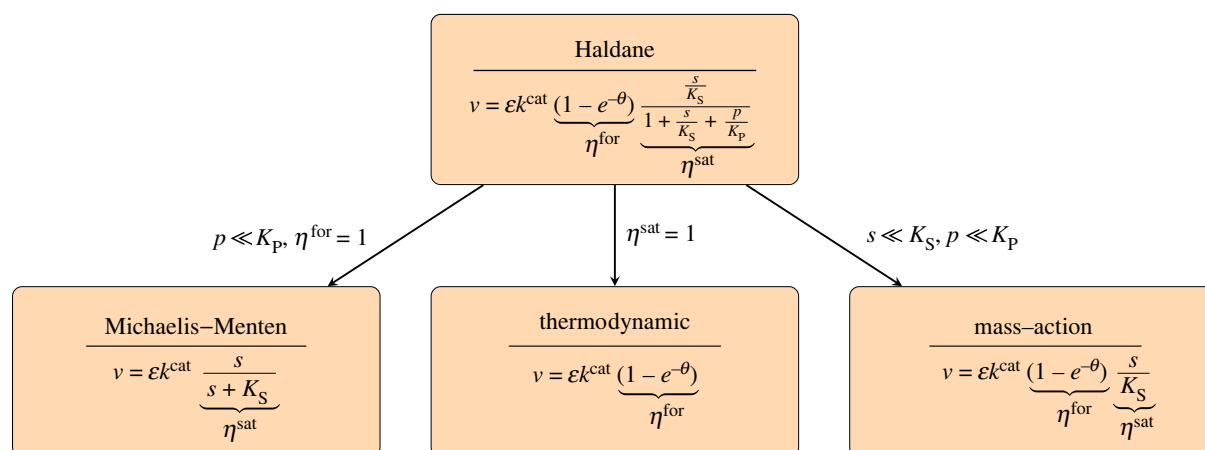


Figure 2. The Haldane rate law and some simplified rate laws. Simplified rate laws are obtained as limiting cases by setting efficiency terms η^{for} or η^{sat} to 1 or another constant value (figure 3) or by assuming that reactant concentrations are small (in the case of the mass-action rate law). For the enzyme optimality problem in unbranched pathways, we do not know of any analytic solutions for the Haldane rate law. We report here solutions for the other, simplified rate laws (where the solution for the thermodynamic rate law contains an unknown auxiliary parameter). Symbols are explained in the text.

the k^{cat} values are catalytic constants (e.g. in units of s^{-1}). In the case of mass concentrations, the k^{cat} values are specific activities (e.g. in units of $\mu\text{mol} \times \text{min}^{-1} \times \text{mg}^{-1}$).

— *Reversible saturable rate law (Haldane).* As a general rate law for a reaction $S \leftrightarrow P$, we consider the reversible saturable rate law

$$v = \varepsilon \frac{k_+^{\text{cat}} s/K_S - k_-^{\text{cat}} p/K_P}{1 + s/K_S + p/K_P}, \quad (2.3)$$

with Michaelis–Menten constants K_S and K_P , which can be factorized into

$$v = \varepsilon k_+^{\text{cat}} \underbrace{(1 - e^{-\theta})}_{\eta^{\text{for}}(\theta(s,p))} \underbrace{\frac{s/K_S}{1 + s/K_S + p/K_P}}_{\eta^{\text{sat}}(s,p)}, \quad (2.4)$$

with the thermodynamic driving force $\theta(s, p) = \ln(K^{\text{eq}} s/p)$, the thermodynamic force efficiency η^{for} and the saturation efficiency η^{sat} (figure 3). Note that the driving force can also be defined as $-\Delta G'$ which is equal to $RT \ln(K^{\text{eq}} s/p)$, but here we drop the gas constant R and temperature T to have a unitless (positive) variable. The two efficiency terms can only assume values between 0 and 1. This factorized formulation of the Haldane rate law is equivalent to the one in equation (2.3), based on the constraint that $K^{\text{eq}} = k_+^{\text{cat}}/k_-^{\text{cat}} \cdot K_P/K_S$, which is commonly known as the *Haldane relationship* (see [23] for more details). This rate law is the most realistic one that we discuss in this work, and deriving it requires only a few assumptions. However, it is also the most mathematically complex and therefore most questions we raise below do not have analytic solutions. So, in addition, we consider three simplified rate laws as limiting cases.

— *Mass-action rate law.* A very common approximation for enzymatic rate laws (the one also made by Waley [20]) is based on the limit of low metabolite concentrations ($s \ll K_S$ and $p \ll K_P$). In this case, the concentration-dependent terms in the denominator (i.e. $s/K_S + p/K_P$) can be neglected and we get

$$v = \varepsilon (k_+^{\text{cat}} s/K_S - k_-^{\text{cat}} p/K_P). \quad (2.5)$$

There are many equivalent ways to write down this rate law. For instance, we can apply the same approximation to the factorized form in equation (2.4) to get $v = \varepsilon k_+^{\text{cat}} (1 - e^{-\theta}) s/K_S$, and we can further replace θ with its explicit definition based on reactant concentrations and write $v = \varepsilon k_+^{\text{cat}}/K_S (s - p/K^{\text{eq}})$. Another common form for this rate law is based on the first-order rate constants $k_+ \equiv k_+^{\text{cat}}/K_S$ and $k_- \equiv k_-^{\text{cat}}/K_P$. Using them in equation (2.5) looks like this: $v = \varepsilon (k_+ s - k_- p)$. As the rate law resembles mass-action kinetics for non-enzymatic reactions, we will refer to this as the ‘mass-action’ rate law although here the enzyme level appears as a prefactor. Throughout this paper, we will switch between these four different notations based on convenience.

— *Thermodynamic rate law.* If the saturation efficiency η^{sat} is approximated by 1 (e.g. in the limit $s \gg K_S$ and $p \ll K_P$), we obtain

$$v = \varepsilon k_+^{\text{cat}} \left(1 - \frac{p}{s} \frac{1}{K^{\text{eq}}}\right) = \varepsilon k_+^{\text{cat}} (1 - e^{-\theta}). \quad (2.6)$$

We will also consider a special case of this rate law where $\theta \rightarrow 0$ and therefore $p/s \approx K^{\text{eq}}$, i.e. the reaction is close to equilibrium. In this case, the rate law becomes $v = \varepsilon k_+^{\text{cat}} \theta$.

— *Irreversible saturable rate law (Michaelis–Menten kinetics).* We next assume that both $p \ll s \approx K^{\text{eq}}$ (which means that $\theta \rightarrow \infty$ and therefore the thermodynamic force efficiency η^{for} can be approximated by 1) and also $p \ll K_P$ (so we can drop p/K_P from the denominator in η^{sat}). In this case, we obtain the Michaelis–Menten rate law

$$v = \varepsilon k_+^{\text{cat}} \frac{s}{s + K_S}, \quad (2.7)$$

which depends on the substrate, but not on the product concentration. Originally, Michaelis & Menten [24] developed this irreversible rate law by assuming that the rate of enzyme–substrate binding is very fast compared to catalysis, and that the catalytic step is irreversible. The assumptions made here lead to the same result but are less stringent.

In the approximations, instead of setting the thermodynamic or saturation efficiencies to their maximal value of 1, we

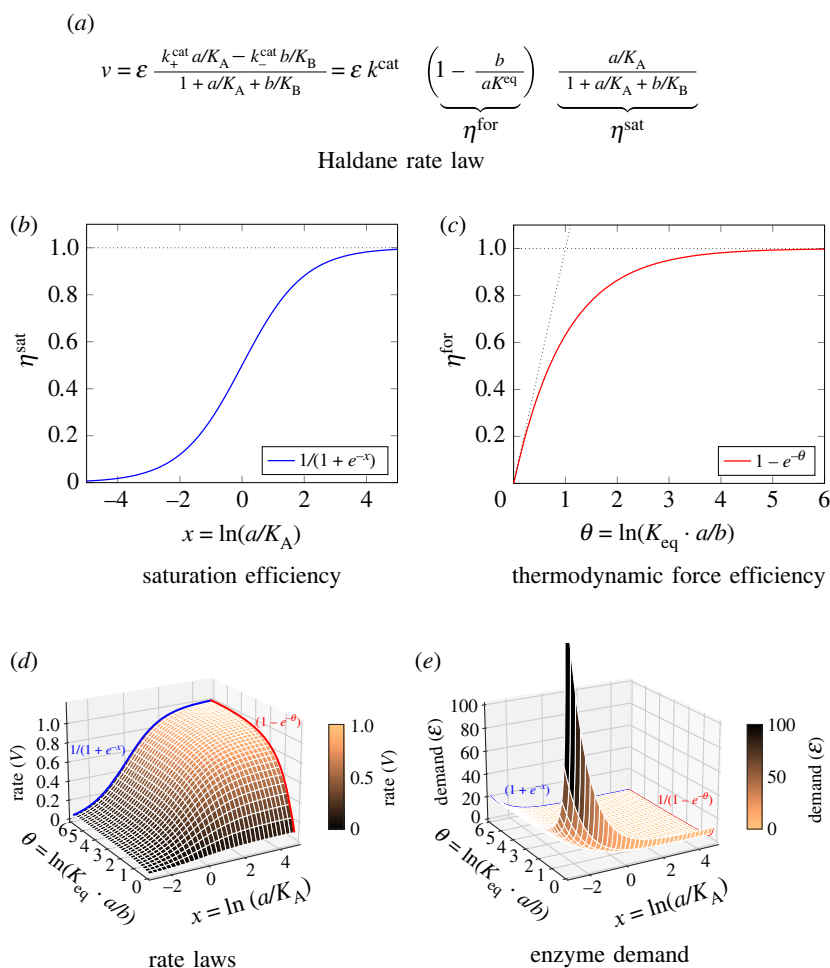


Figure 3. The Haldane rate law and its factorization into efficiency terms. (a) The Haldane rate law. In the factorized form (right), thermodynamic and saturation effects are described by separate efficiency factors. (b) The saturation efficiency as a function of the substrate concentration (in log-scale, relative to the K_A), assuming $b \ll K_B$. (c) The thermodynamic force efficiency as a function of the driving force θ . (d) A surface plot showing the rate v as a function of θ and the substrate concentration (in log-scale, relative to the K_A). The parameters are $K^{\text{eq}} = K_A = k^{\text{cat}} = \varepsilon = 1$ and $K_B = 10$. (e) A surface plot showing the enzyme demand for a given rate ($v = 1$). All kinetic parameters are the same as in (d).

may also approximate them by constant numbers smaller than 1; we obtain exactly the same rate laws, but instead of the k^{cat} value we obtain a smaller apparent value in the approximated rate law.

2.2. Analytic formulae for unbranched metabolic pathways

How can we characterize metabolic states by using analytic formulae? Here we consider chains of uni–uni reactions, or reactions in which all reactant concentrations except for one substrate and one product are given as model parameters. For such an unbranched pathway with a given type of rate law (e.g. mass-action or Michaelis–Menten rate laws), we are interested in formulae for a number of quantities:

1. *Metabolic steady state.* Given the enzyme levels, external metabolite concentrations and kinetic constants, we can directly compute the stationary fluxes and internal metabolite concentrations. For general metabolic networks, no explicit formulae are known, but for unbranched pathways with some simplified rate laws, explicit formulae exist. Incidentally, this also shows that in these models the steady-state concentrations are unique.
2. *Stability of steady state.* If the Jacobian matrix in a steady state has positive eigenvalues, the state is asymptotically

unstable and is not able to persist under the inevitable chemical noise in the cell. Stability is also a prerequisite for metabolic control coefficients being defined. A sufficient (but not necessary) condition for stable steady states in unbranched metabolic pathways is given in electronic supplementary material, section 5.8.

3. *Metabolic control.* The metabolic response coefficients R^X are defined by MCA as the derivatives between steady-state concentrations or fluxes and model parameters (e.g. the enzyme levels). If two model parameters act (exclusively) on the same reaction, all their response coefficients will be the same except for a proportional scaling. Taking this into account, one can define the metabolic control coefficients C^X , which describe the same type of derivatives, but for a set of hypothetical, reaction-specific parameters. In practice, if reaction rates are proportional to enzyme levels and if each reaction is catalysed by a single specific enzyme, we can write the enzyme response coefficients as $R_{\varepsilon_i}^X = C_i^X \cdot v_i / \varepsilon_i$. Importantly, throughout this paper we use elasticities, response coefficients, and control coefficients in their unscaled form, for instance unscaled enzyme response coefficients $R_{\varepsilon_i}^J = \partial J / \partial \varepsilon_i$ instead of the common scaled form $(\varepsilon_i / J)(\partial J / \partial \varepsilon_i) = \partial \ln J / \partial \ln \varepsilon_i$.

The metabolic control coefficients can be computed from the stoichiometric matrix and the elasticity matrix,

and they satisfy summation and connectivity theorems. They can be computed in two ways: if an analytic formula for the metabolic steady state is known (as is the case in some unbranched pathway models studied below), we may differentiate symbolically by the enzyme levels; otherwise, we may compute the elasticity coefficient matrices by differentiating the rate laws and then compute the control coefficient matrices from them using a known formula; however, since this formula involves a matrix inversion, writing this down as an analytic formulae may be extremely complicated, and control coefficients are usually computed numerically. Symbolic expressions for metabolic control coefficients can be conveniently handled by the PYSCESTOOLBOX [25]. Another way to compute control coefficients, which follows from the enzyme-control rule and works only in optimal states, is described below.

4. *Optimal metabolic states.* In our basic metabolic optimality problem, we define optimal states as states in which a given enzyme budget (a fixed sum of enzyme levels) is allocated to maximize a production flux. Kinetic constants and external metabolites are given, and we compute the optimal metabolite profile, the optimal enzyme profile and the optimal flux. If the flux distribution is known (e.g. a steady-state flux in an unbranched pathway) and increases or decreases proportionally with the enzyme levels, this problem is equivalent to the problem of minimizing the enzyme demand at a given (unit) flux. This convex problem can be solved numerically [26], but analytic solutions were known only for very few cases. Below we present some new analytic solutions. Formulae for optimal enzyme levels and the optimal achievable flux are shown in table 1. We also consider a related problem, maximizing the flux under a constraint on the total enzyme plus metabolite amount.
5. *Metabolic control in optimal states.* The control coefficients tell us how a metabolic system responds to perturbations (resulting in a new state that is stationary, but probably non-optimal). For optimal states (with a constraint on the sum of enzyme levels), the enzyme-control rule states that enzyme levels and flux control coefficients must be proportional. Since we know (from the summation theorem) that the control coefficients must sum to 1, we can conclude: whenever there is an analytic formula for optimal enzyme levels, we also obtain a formula for the control coefficients. Below we extend this to models with general density constraints on enzymes and metabolites, and derive an additional rule that relates enzyme levels to metabolite elasticities around each metabolite.

2.3. Optimal metabolic states: analytic solutions

What are the general principles behind optimal enzyme allocation? One important principle, valid in optimal metabolic states, has been shown for pathways with mass-action rate laws [21] and later been confirmed for general rate laws [17]: in optimal states, where the metabolic flux has been maximized at a fixed enzyme budget, enzyme levels and flux control coefficients must be proportional:

$$\varepsilon_i^* \propto C_i^* \quad (2.8)$$

Here the star * denotes variables in the optimal state.

To derive this rule, we note that in an optimal state each enzyme has a marginal cost (contribution to the enzyme budget, per mole of enzyme) which must be balanced by the same marginal benefit (contribution to the production flux, per mole of enzyme). If all enzymes contribute to the enzyme budget with equal weights, their marginal costs are the same. This means that also their marginal benefit must be the same, which are given by the (unscaled) flux response coefficients $R_{e_i}^J = \partial v^{\text{st}} / \partial e_i$. This, finally, means that their flux control coefficients $C_i^J = (\partial v^{\text{st}} / \partial e_i) / (v^{\text{st}} / e_i)$ must be proportional to the enzyme levels e_i .

Together with the summation theorem for flux control coefficients [14,27], $\sum_i C_i^J = 1$, we obtain the conversion formulae

$$C_i^J = \frac{\varepsilon_i^*}{\sum_j \varepsilon_j^*} \quad \text{and} \quad \varepsilon_i^* = \varepsilon_{\text{tot}}^* C_i^J \quad (2.9)$$

The enzyme-control rule (2.8) provides a condition for metabolic states, independent of the type of rate laws considered. Importantly, it holds only in states of maximal flux, given a fixed total enzyme amount and no other constraints. More realistic models employ also bounds on metabolite levels [5,22,28] for different reasons. First, metabolite levels in real cells are bounded: while metabolite molecules may be small, their concentration is high, and they contribute much more than macromolecules to osmotic pressure. Second, as we will see below, some models without metabolite bounds lead to paradoxical results. We will therefore present a generalized version of the enzyme-control rule that takes metabolite bounds into account.

But what are the general shapes of optimal enzyme profiles, i.e. how do enzyme levels vary across the network? And on what factors (kinetic, thermodynamic or cost factors) does this depend? To answer this, the enzyme-control rule alone would not be enough (because no kinetic details appear in the rule). Also numerical studies would not be enough (because they apply to single cases and yield no general laws). Hence, to study this it would be good to consider analytic solutions. Unfortunately, analytic solutions are not known for general metabolic models, but rather only for special cases of unbranched pathways (such as ones with mass-action rate laws [20]).

We now present analytic formulae for optimal metabolic states with different types of rate laws. We discuss the rate laws in increasing order of difficulty. An overview of all analytic solutions is given in table 1.

2.3.1. Michaelis–Menten rate law

We first consider the Michaelis–Menten rate law (i.e. irreversible reactions with simple saturation kinetics):

$$v_i = \varepsilon_i \cdot k_i^{\text{cat}} \frac{S_{i-1}}{S_{i-1} + K_{M,i}} \quad (2.10)$$

In an unbranched pathway at steady state, all rates must be equal. To describe them, we introduce a new variable J called the pathway flux, and require that $\forall i v_i = J$. Now we can use equation (2.10) to find a relationship between substrate and enzyme levels:

$$\varepsilon_i = \frac{J}{k_i^{\text{cat}}} \cdot \left(1 + \frac{K_{M,i}}{S_{i-1}} \right) = J \left(\alpha_i + \frac{\beta_i}{S_{i-1}} \right) \quad (2.11)$$

Table 1. A summary of all the rate laws considered in this paper, along with their solutions for optimal enzyme allocation and the resulting minimum pathway cost per flux. The rate laws are roughly ordered by increasing level of complexity. The ‘trivial’ rate law does not have its own section in the text, but is shown here as the limit of the Michaelis–Menten rate law when $s_{j-1} \gg K_{Mj}$, as well as the thermodynamic rate law when $k_{\text{tot}}^{\text{eq}} s_0 / s_n \rightarrow \infty$. The ‘mass-action’ rate law shown in this table is equivalent to the form used in previous studies [16,17,20]. Here it is shown in a modified form for the purpose of using the same set of kinetic parameters as the other rate laws. In order to calculate the absolute optimal enzyme levels from the relative ones (given in the column titled ‘optimal enzyme allocation’), one can use $\varepsilon_j^* = \varepsilon_{\text{tot}} (x_j / \sum_j x_j)$ where x_j are the relative values.

name (assumptions)	enzymatic rate law $v_j / \varepsilon_j =$	comments/definitions	optimal enzyme allocation: $\varepsilon_j^* \propto$	minimum pathway cost per flux: $\varepsilon_{\text{tot}} / j^*$
trivial (irreversible, saturated)	k_j^{cat}	$\alpha_j \equiv 1 / k_j^{\text{cat}}$	$\ \alpha\ _1$	
Michaelis–Menten (irreversible, saturable)	$k_j^{\text{cat}} \cdot \frac{s_{j-1}}{s_{j-1} + K_{Mj}}$	$\beta_j \equiv K_{Mj} / k_j^{\text{cat}} \sum_j s_j \leq s_{\text{tot}}$	$\alpha_j + \sqrt{\beta_j \ \beta\ _{1/2} / s_{\text{tot}}}$	$\ \alpha\ _1 + \ \beta\ _1 / s_{\text{tot}}$
thermodynamic (reversible, unsaturable)	$k_j^{\text{cat}} \cdot (1 - \frac{s_j}{s_{j-1} k_j^{\text{eq}}})$	find Ψ s.t. $\prod_i (\sqrt{\Psi} \alpha_i + \sqrt{1 + \Psi \alpha_i}) = \sqrt{\frac{s_0 k_{\text{tot}}^{\text{eq}}}{s_n}}$	$\alpha_j \left(\frac{1}{2} + \frac{1}{2} \sqrt{1 + (\Psi \alpha_j)^{-1}} \right)$	$\sum_j \alpha_j \left(\frac{1}{2} + \frac{1}{2} \sqrt{1 + (\Psi \alpha_j)^{-1}} \right)$
thermodynamic (approximate solution)	$k_j^{\text{cat}} \cdot (1 - \frac{s_j}{s_{j-1} k_j^{\text{eq}}})$	same as above	no explicit formula	$\ \alpha\ _1 \cdot (1 - \exp(-\ \alpha\ _1 / \ \alpha\ _{1/2} \theta_{\text{tot}}))^{-1}$
mass-action [20] (reversible, unsaturable)	$k_j^{\text{cat}} \cdot (1 - \frac{s_j}{s_{j-1} k_j^{\text{eq}}}) \cdot \frac{s_{j-1}}{K_{Mj}}$	$s_0 = \text{const}, s_n = \text{const}, \gamma_j \equiv \beta_j \cdot \prod_{l=j}^n k_l^{\text{eq}}$	$\sqrt{\gamma_j}$	$\ \gamma\ _1 / \frac{1}{2} \left(s_0 - \frac{s_n}{k_{\text{tot}}^{\text{eq}}} \right)$
Haldane (reversible, saturable)	$k_j^{\text{cat}} \cdot (1 - \frac{s_j}{s_{j-1} k_j^{\text{eq}}}) \cdot \frac{s_{j-1} / K_{Sj}}{1 + \frac{s_{j-1}}{K_{Sj}} + \frac{s_j}{K_{Pj}}}$	no known solution, use convex optimization solver		

where we defined $\alpha_i \equiv 1/k_i^{\text{cat}}$ and $\beta_i \equiv K_{M,i}/k_i^{\text{cat}}$ for convenience.

Combining this equation with an upper bound on total enzyme from equation (2.2) we get

$$\begin{aligned} \varepsilon_{\text{tot}} &\geq \sum_i \varepsilon_i = \sum_i J \left(\alpha_i + \frac{\beta_i}{s_{i-1}} \right) \\ &= J \left(\sum_i \alpha_i + \sum_i \frac{\beta_i}{s_{i-1}} \right), \end{aligned} \quad (2.12)$$

and by rearranging we obtain

$$J \leq \frac{\varepsilon_{\text{tot}}}{\sum_i \alpha_i + \sum_i \frac{\beta_i}{s_{i-1}}}. \quad (2.13)$$

Maximizing J would mean that we reach the upper bound, and therefore we can also treat this as an equality. Since ε_{tot} is constant and the only free variables are the metabolite concentrations, the maximal flux is reached when the denominator on the right-hand side is minimized. The problem is that it is a monotonically decreasing function in s_i (for each i) and since metabolite concentrations are unbounded, the optimum would be at $s_i \rightarrow \infty$, which formally is not a mathematically defined state. But even if we consider this as a limit, the resulting state would not be controllable, but instead very fragile against any changes of enzyme levels (see electronic supplementary material, section 3.1). In reality, of course, the range of physiological osmotic pressures does impose some constraint on the concentrations of small molecules. As a proxy for this effect, we can add another constraint to bound the sum of all metabolite concentrations, $\sum_i s_i \leq s_{\text{tot}}$. Now, one can show (see electronic supplementary material, section 5.1) that the optimal allocation of enzymes will obey

$$\varepsilon_i^* \propto \alpha_i + \sqrt{\beta_i} \cdot \frac{\sqrt{\|\boldsymbol{\beta}\|_{1/2}}}{s_{\text{tot}}}, \quad (2.14)$$

where $\boldsymbol{\beta}$ is the vector of all β_i and $\|\mathbf{x}\|_{1/2} = (\sum_i \sqrt{x_i})^2$ is the $l_{1/2}$ norm (i.e. $\|\boldsymbol{\beta}\|_{1/2} \equiv (\sum_i \sqrt{\beta_i})^2$). In this case, the maximal flux would be

$$J^* = \varepsilon_{\text{tot}} \cdot (\|\boldsymbol{\alpha}\|_1 + \|\boldsymbol{\beta}\|_{1/2}/s_{\text{tot}})^{-1}. \quad (2.15)$$

Note that the solution looks essentially the same even if we constrain the first metabolite (s_0) to have a fixed concentration (see electronic supplementary material, section 5.1.4). We will revisit this case in §2.5 in the context of a course-grained model of a growing cell.

If the metabolite density constraint is not very tight (i.e. s_{tot} is large enough), we can ignore the second term in equation (2.14), which would be equivalent to assuming that all enzymes are substrate-saturated. In this case, the optimal allocation of enzymes will be proportional to α_i (or inversely proportional to k_i^{cat}) and therefore

$$\lim_{s_{\text{tot}} \rightarrow \infty} J^* = \varepsilon_{\text{tot}} \cdot (\|\boldsymbol{\alpha}\|_1)^{-1}. \quad (2.16)$$

Interestingly, $\|\boldsymbol{\alpha}\|_1$, which is equal to the sum of k^{cat} reciprocals, is in fact the inverse of the *pathway specific activity* (PSA) as originally defined by Bar-Even *et al.* [11]. Indeed, the idealized scenario considered in that study (where all enzymes were irreversible and saturated) provides an upper

bound on the maximal flux achievable. Interestingly, if we see the $\alpha_i = 1/k_i^{\text{cat}}$ as minimal enzymatic turnover times and $\alpha_{\text{PW}} = \varepsilon_{\text{tot}}/J$ is the turnover time of the pathway, we can simply rephrase equation (2.16) by saying that, like in the case of the PSA, turnover times along a pathway are additive. In fact, this holds more generally for any rate law $v_i = e_i k_i(s)$, if the enzyme turnover times are defined as $\tau_i = 1/k_i$, and resembles the fact that in physics, electric resistances in a series of resistors are additive.

One of the reasons the irreversible Michaelis–Menten model is unrealistic is that it does not capture reactions that are close to equilibrium and therefore suffer from a counter-productive reverse flux. This is yet another reason why it might be impossible for some metabolites to reach very high concentrations: they might be products of unfavourable reactions. In most metabolic networks, about half of the reactions are reversible and therefore it would be more realistic to use a reversible rate law such as Haldane's. However, using the Haldane equation would typically create a system of equations for which no analytic solution is known. So, first, we will consider rate laws that account for the reverse flux but, for simplicity, ignore substrate saturation.

2.3.2. Thermodynamic rate law

Irreversible rate laws like the Michaelis–Menten kinetics depend only on the substrate concentration; in the formula there is no product-dependent reverse term that would decrease the total rate or could make it become negative. However, according to thermodynamics, such laws can only be approximations: thermodynamically feasible rate laws must contain a reverse term and must depend on the thermodynamic imbalance of substrate and product concentrations expressed by the thermodynamic driving force. Some rate laws can even be written as functions of the thermodynamic force alone. Here we describe such a 'thermodynamic' rate law, where v is proportional to the thermodynamic force efficiency $1 - e^{-\theta}$, while the saturation efficiency η^{sat} is assumed to be constant and given by 1:

$$v_i = \varepsilon_i k_i^{\text{cat}} (1 - e^{-\theta_i}) = \varepsilon_i k_i^{\text{cat}} \left(1 - \frac{s_i}{s_{i-1}} \frac{1}{K_i^{\text{eq}}} \right). \quad (2.17)$$

This type of kinetics approximates cases where all reactions are saturated ($s_i \gg K_{M,i}$) and therefore the thermodynamic force efficiency η^{sat} in equation (2.4) becomes 1. The parameters here are the turnover numbers k_i^{cat} and the equilibrium constants K_i^{eq} . However, as we will learn soon, the individual equilibrium constants K_i^{eq} do not change the result, and only the overall equilibrium constant ($K_{\text{tot}}^{\text{eq}} = \prod_i K_i^{\text{eq}}$) matters.

As before, we can define the steady-state pathway flux J and apply an upper bound on the total enzyme to get

$$\left. \begin{aligned} \varepsilon_{\text{tot}} &\geq \sum_i \varepsilon_i = \sum_i \frac{J}{k_i^{\text{cat}} (1 - e^{-\theta_i})} = J \cdot \left(\sum_i \frac{1}{k_i^{\text{cat}} \cdot \frac{1}{1 - e^{-\theta_i}}} \right) \\ \text{and} \quad J &\leq \varepsilon_{\text{tot}} \cdot \left(\sum_i \frac{1}{k_i^{\text{cat}} \cdot \frac{1}{1 - e^{-\theta_i}}} \right)^{-1} \end{aligned} \right\} \quad (2.18)$$

Unfortunately, there is no closed-form analytic solution for the maximal rate of this thermodynamic rate law. But

we do have something very close which requires rather simple computations. First, we need to invert the functional relationship between the overall driving force (θ_{tot}) and an auxiliary variable Ψ which is defined by the inverse function of

$$\ln\left(\frac{s_0}{s_n} K_{\text{tot}}^{\text{eq}}\right) = \theta_{\text{tot}} = 2 \sum_i \ln\left(\sqrt{\frac{\Psi}{k_i^{\text{cat}}}} + \sqrt{1 + \frac{\Psi}{k_i^{\text{cat}}}}\right). \quad (2.19)$$

The expression on the right-hand side is analytic and strictly increasing in the range $\Psi \in [0, \infty)$, so there is a unique solution that can be found by simple numerical methods. Then, we can use that value to directly calculate the optimal driving forces, enzyme levels, and pathway flux:

$$\left. \begin{aligned} \theta_i^* &= 2 \ln\left(\frac{1}{k_i^{\text{cat}}}\left(1 + \sqrt{1 + \frac{k_i^{\text{cat}}}{\Psi}}\right)\right) + \ln(\Psi k_i^{\text{cat}}), \\ \varepsilon_i^* &= \frac{\varepsilon_{\text{tot}}}{2} \cdot \frac{1}{k_i^{\text{cat}}}\left(1 + \sqrt{1 + \frac{k_i^{\text{cat}}}{\Psi}}\right) \end{aligned} \right\} \text{and } J^* = 2\varepsilon_{\text{tot}} \cdot \left(\sum_i \frac{1}{k_i^{\text{cat}}}\left(1 + \sqrt{1 + \frac{k_i^{\text{cat}}}{\Psi}}\right)\right)^{-1}. \quad (2.20)$$

The full derivation of this solution can be found in electronic supplementary material, section 5.2.2. An example of what the relationship between the driving force and the optimal flux looks like is illustrated in figure 4 for a pathway with two enzymes. Furthermore, a comparison between our exact formula presented here and a numerical solution based on convex optimization [8] shows no difference at all.

These formulae cannot be directly evaluated because of the unknown parameter Ψ . To obtain solutions that do not depend on this parameter, we now consider two limiting cases in which Ψ is either infinitely high (very high driving force) or infinitely low (very low driving force).

When the driving forces are very high (i.e. $\theta_{\text{tot}} \rightarrow \infty$), the solution for Ψ in equation (2.19) will approach infinity and the optimal flux becomes

$$\begin{aligned} \lim_{\theta_{\text{tot}} \rightarrow \infty} J^* &= \lim_{\Psi \rightarrow \infty} \frac{2 \varepsilon_{\text{tot}}}{\sum_i 1/k_i^{\text{cat}} \cdot (1 + \sqrt{1 + k_i^{\text{cat}}/\Psi})} \\ &= \frac{\varepsilon_{\text{tot}}}{\sum_i 1/k_i^{\text{cat}}} = \varepsilon_{\text{tot}} \cdot (\|\alpha\|_1)^{-1}, \end{aligned} \quad (2.21)$$

with parameters $\alpha_i = 1/k_i^{\text{cat}}$ defined as before. This solution indeed makes sense as it is equivalent to the fully saturated limit in the Michaelis–Menten case (see §2.3.1, equation (2.16)).

However, when the driving force in the pathway is limited (Ψ has a finite positive value), this driving force will be split between reactions according to their k_i^{cat} values: enzymes with a higher k_i^{cat} value will have to pay a higher penalty due to their driving force being closer to 0. On the other hand, slow enzymes will obtain more driving force, which will help them by having a smaller fraction of reverse flux. Notably, the distribution does *not* depend on the reaction equilibrium constants (only on the overall $K_{\text{tot}}^{\text{eq}}$). Perhaps this is not that surprising if we consider the fact that the concentrations of intermediate substrates and products are unconstrained and therefore we have enough degrees of freedom to adjust to any value of K_i^{eq} as long as the total driving force stays the same.

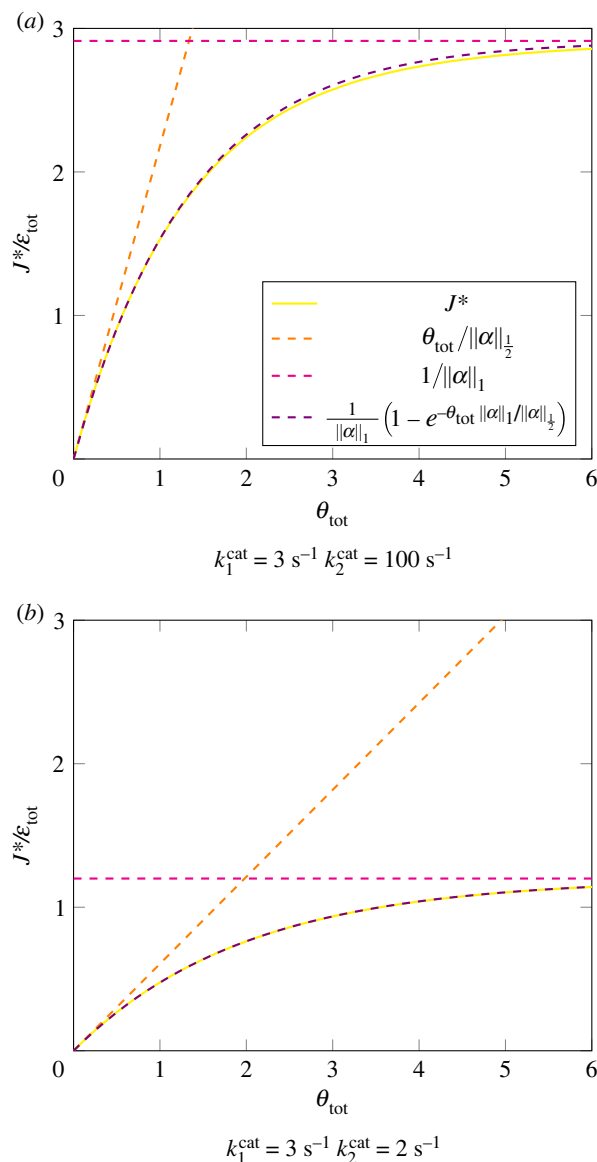


Figure 4. The relationship between the maximal flux per enzyme and the overall driving force for the thermodynamic rate law. Even though we do not have a closed-form solution for J^* as a function of θ_{tot} , we can still plot $J^*(\Psi)$ against $\theta_{\text{tot}}(\Psi)$ for varying values of Ψ using equations (2.19) and (2.20). Here, we show this relationship for a pathway with 2 steps (yellow). The parameters are: $k_i^{\text{eq}} = 1$, $k_1^{\text{cat}} = 3 \text{ s}^{-1}$, and the k^{cat} of the second enzyme is either (a) $k_2^{\text{cat}} = 100 \text{ s}^{-1}$ or (b) $k_2^{\text{cat}} = 2 \text{ s}^{-1}$. The orange dashed line represents the approximations for very small driving forces based on equation (2.23), and similarly in pink for very large driving forces as in equation (2.21). The approximation (dashed purple line) using equation (2.24) is based on effective parameters chosen to match the slope at $\theta_{\text{tot}} = 0$ and the limiting maximal flux at $\theta_{\text{tot}} \rightarrow \infty$.

One can also consider the other extreme where all reactions are close to equilibrium, which means that θ_{tot} is close to 0 (and therefore also each $\theta_i \rightarrow 0$). In this case, we can approximate equation (2.18) by

$$J \leq \varepsilon_{\text{tot}} \cdot \left(\sum_i \frac{1}{k_i^{\text{cat}}} \cdot \frac{1}{1 - e^{-\theta_i}}\right)^{-1} \approx \varepsilon_{\text{tot}} \cdot \left(\sum_i \frac{1}{k_i^{\text{cat}} \theta_i}\right)^{-1}. \quad (2.22)$$

Maximizing J under the constraint $\sum_i \theta_i = \theta_{\text{tot}}$ yields the following solution (see the full derivation in electronic

supplementary material, section 5.2.3):

$$\left. \begin{aligned} \theta_i^* &\approx \theta_{\text{tot}} \frac{\sqrt{1/k_i^{\text{cat}}}}{\sum_i \sqrt{1/k_i^{\text{cat}}}} \\ \text{and } J^* &\approx \frac{\theta_{\text{tot}} \varepsilon_{\text{tot}}}{\left(\sum_i \sqrt{1/k_i^{\text{cat}}}\right)^2} = \varepsilon_{\text{tot}} \cdot \frac{\theta_{\text{tot}}}{\|\alpha\|_{1/2}} \end{aligned} \right\} \quad (2.23)$$

Using the two limiting cases ($\theta_{\text{tot}} \rightarrow \infty$ and $\theta_{\text{tot}} \approx 0$), we can approximate the solution from equation (2.20) by a much simpler formula, which has the same shape as the thermodynamic rate law for a single reaction:

$$J^* \approx \frac{\varepsilon_{\text{tot}}}{\|\alpha\|_1} \left(1 - \exp\left(-\frac{\|\alpha\|_1}{\|\alpha\|_{1/2}} \theta_{\text{tot}}\right)\right). \quad (2.24)$$

Figure 4 compares between the precise and approximate solution for two example cases (a more detailed comparison for metabolic pathways of different lengths and parameter choices is shown in electronic supplementary material, section 5.3). One can appreciate that the two curves are nearly indistinguishable, illustrating the good quality of the approximation.

2.3.3. Mass-action rate law

Next, we consider the same unbranched pathway but assuming mass-action rate laws:

$$v_i = \varepsilon_i k_i^{\text{cat}} / K_{M,i} \left(s_{i-1} - \frac{s_i}{K_i^{\text{eq}}}\right) = \varepsilon_i \beta_i^{-1} \left(s_{i-1} - \frac{s_i}{K_i^{\text{eq}}}\right). \quad (2.25)$$

As before, we defined $\beta_i \equiv K_{M,i}/k_i^{\text{cat}}$. Note that instead of equation (2.5) we use a form of the mass-action rate law that does not require the turnover rate and K_M of the product (and instead uses K_i^{eq}), to avoid confusing indexation of forward and backward parameters.

Like with the previous rate laws, we define the pathway flux J and apply an upper bound on the total enzyme levels:

$$\left. \begin{aligned} \varepsilon_{\text{tot}} &\geq \sum_i \varepsilon_i = \sum_i \frac{J \beta_i}{s_{i-1} - s_i / K_i^{\text{eq}}} = J \sum_i \frac{\beta_i}{s_{i-1} - s_i / K_i^{\text{eq}}} \\ \text{and } J &\leq \varepsilon_{\text{tot}} \cdot \left(\sum_i \frac{\beta_i}{s_{i-1} - s_i / K_i^{\text{eq}}}\right)^{-1} \end{aligned} \right\} \quad (2.26)$$

Again, we maximize the flux at a constrained total enzyme level. Optimization with Lagrange multipliers yields the following expressions for the optimal individual enzyme levels (ε_j^*) and the maximal flux (J^*) (see full derivation in electronic supplementary material, section 5.4):

$$\left. \begin{aligned} \varepsilon_j^* &= \varepsilon_{\text{tot}} \sqrt{\frac{\gamma_j}{\|\gamma\|_{1/2}}} \\ \text{and } J^* &= \varepsilon_{\text{tot}} \cdot \frac{s_0 K_{\text{tot}}^{\text{eq}} - s_n}{\|\gamma\|_{1/2}} \end{aligned} \right\} \quad (2.27)$$

where γ_j is defined as

$$\gamma_i \equiv \beta_i \prod_{j=i}^n K_j^{\text{eq}}. \quad (2.28)$$

Of course, the exact value of this maximum depends on all the different system parameters. However, it is interesting to consider a naive assumption where all the γ_j parameters are identical. In such a case, the flux in the pathway would decrease quadratically with the number of steps [13]. Of course, we know that in real metabolic pathways the equilibrium constants are typically not close to 1, and therefore this approximation might not have many applications in biology.

What would happen if a mutation improved the catalytic rate of only one of the enzymes ε_i by a factor of a (i.e. β_i decreases by a factor of a , but the equilibrium constant K_i^{eq} remains the same)? In this case, γ_i would be divided by a but the optimal enzyme concentration for this reaction ε_i^* would only decrease by a factor of \sqrt{a} . This saving would then be distributed proportionally among all the n enzymes and would contribute to an increase in the pathway flux J . On the flip side, if for some reason the activity of one enzyme is decreased by a multiplicative factor b , it would need to ‘pay’ (increase the enzyme’s concentration) only by a factor \sqrt{b} , and this increase would be ‘funded’ by all of the enzymes together in order to keep the same ε_{tot} , thus lowering J .

2.3.4. Haldane rate law

So far we analysed three cases where all enzymes could be described by the same kinetic rate law: mass-action, Michaelis–Menten, or thermodynamic. Although these rate laws can reliably describe some enzymes in specific conditions, it is very unlikely that such an approximation would apply to *all* the reactions in a single pathway (except for the trivial case of a 1-reaction pathway). A more realistic model would allow all reactions to be reversible and saturable. Here, we will analyse such a case based on the factorized rate law with one substrate and one product. Note that although it is equivalent to equation (2.3), we prefer this formulation because it is easier to separate thermodynamics from saturation effects.

$$v_i = \varepsilon_i \cdot k_i^{\text{cat}} \cdot (1 - e^{-\theta_i}) \cdot \frac{s_{i-1}/K_{S,i}}{1 + s_{i-1}/K_{S,i} + s_i/K_{P,i}}. \quad (2.29)$$

As always, we can assume that all fluxes are equal to J , and use the total enzyme budget to get an upper bound:

$$\left. \begin{aligned} \varepsilon_{\text{tot}} &\geq \sum_i \varepsilon_i = \sum_i \frac{J}{k_i^{\text{cat}} \cdot (1 - e^{-\theta_i}) \cdot \frac{s_{i-1}/K_{S,i}}{1 + s_{i-1}/K_{S,i} + s_i/K_{P,i}}} \\ &= J \cdot \left(\sum_i \frac{1}{k_i^{\text{cat}} (1 - e^{-\theta_i})^{-1} \left(1 + \frac{K_{S,i}}{s_{i-1}} + \frac{s_i K_{S,i}}{s_{i-1} K_{P,i}}\right)}\right) \\ \text{and } J &\leq \varepsilon_{\text{tot}} \cdot \left(\sum_i \frac{1}{k_i^{\text{cat}} (1 - e^{-\theta_i})^{-1} \left(1 + \frac{K_{S,i}}{s_{i-1}} + \frac{s_i K_{S,i}}{s_{i-1} K_{P,i}}\right)}\right)^{-1} \end{aligned} \right\} \quad (2.30)$$

where we can now appreciate how this is a generalization of both equations (2.13) (Michaelis–Menten) and (2.18) (thermodynamic). We can see that the maximal pathway flux would be realized when the term in parentheses is minimized with respect to the s_i , i.e.

$$\text{minimize}_s \sum_i \frac{1}{k_i^{\text{cat}} (1 - e^{-\theta_i(s)})^{-1} \left(1 + \frac{K_{S,i}}{s_{i-1}} + \frac{s_i K_{S,i}}{s_{i-1} K_{P,i}}\right)}. \quad (2.31)$$

Unfortunately, as we already mentioned before, no analytic solution to this problem is known in the general case.

2.4. Optimal metabolic states: insights from metabolic control analysis

2.4.1. Enzyme-control rule and enzyme-elasticity rule

We now continue with some observations about the enzyme-control rule. The enzyme-control rule equation (2.8) states that enzyme levels and flux control coefficients in optimal states are proportional: $\varepsilon_i^* \propto C_i^{J^*}$. When the aim is to compute control coefficients in optimal states (at a given enzyme budget ε_{tot}), this rule comes in handy. The summation theorem tells us that the flux control coefficients must sum to 1; therefore, the optimal enzyme levels are given by the flux control coefficients, multiplied by the (predefined) total enzyme level, and we obtain the simple conversion $C_i^{J^*} = \varepsilon_i^*/\varepsilon_{\text{tot}}$ and $\varepsilon_i^* = \varepsilon_{\text{tot}} C_i^{J^*}$. In the basic form of the rule, we put a constraint on the enzyme mass and assume that enzyme levels are mass densities. If the enzyme levels are molar concentrations, and differently weighted in the constraint, the weights can be taken into account by modifying the rule.

To what types of optimality problems does the rule apply? If the metabolic system is not an unbranched pathway but a general network with one target flux, the rule will still hold (where the flux control coefficients refer to this target flux). This also holds for models with metabolite dilution. However, there are some limitations. Obviously, the rule cannot hold in states in which control coefficients are not defined. This condition may seem unproblematic, but it is actually violated in the pathway with Michaelis–Menten rate laws, because in the limit $s_i \rightarrow \infty$ required by optimization, any variations of single enzyme levels would break the steady state. We discuss this in electronic supplementary material, section 3.1.

One way to avoid this problem is to add a bound on metabolite levels. For example, we may consider a generalized density constraint on enzyme levels and metabolite concentrations

$$\text{Maximize } \mathbf{z} \cdot \mathbf{v} \quad \text{s.t. } \mathbf{a} \cdot \boldsymbol{\varepsilon} + \mathbf{b} \cdot \mathbf{s} \leq \rho,$$

with a linear flux objective $\mathbf{z} \cdot \mathbf{v}$ instead of a single target flux, and enzyme weights a_i , metabolite weights b_i , and an upper bound ρ on the molecule density. Problems with this constraint lead to a generalized form of the enzyme-control rule. For an unbranched pathway with equal weights a for enzymes and equal weights b for metabolites (e.g. molecular masses), we obtain the new enzyme-control rule

$$\varepsilon_i^* = \varepsilon_{\text{tot}}^* C_i^{J^*} - J^* \frac{b}{a} \sum_i C_i^{s_i^*}, \quad (2.32)$$

where the symbols $C_i^{J^*}$ and $C_i^{s_i^*}$ denote, respectively, the flux and control coefficients in the optimal state. Likewise, $\varepsilon_{\text{tot}}^*$ is the (non-fixed) sum of enzyme levels emerging in the optimal state. For derivation and details, see electronic supplementary material, section 3.2.

The original enzyme-control rule has a direct and useful consequence. Since enzyme levels (in optimal states) are proportional to flux control coefficients, they need to satisfy a connectivity theorem. Connectivity theorems relate the elasticities of a given metabolite i to control coefficients of the reactions around this metabolite. In the case of flux control coefficients, the right-hand side of the theorem is zero, and we obtain a simple equation for the optimal enzyme levels around a metabolite i :

$$\sum_j \varepsilon_j \cdot E_{j,i} = 0, \quad (2.33)$$

where $E_{j,i} = \partial v_j / \partial s_i$ denotes the unscaled elasticity between metabolite i and reaction j .

We call this formula the enzyme-elasticity rule (see electronic supplementary material, section 4). A similar rule for small adaptations of enzyme levels, instead of the enzyme levels themselves, had previously been shown in [29]. For a linear pathway, the enzyme-elasticity rule yields a simple result: in an optimal state, for each metabolite i and its producing reaction $j = i - 1$ and its consuming reaction $j = i$, the ratio of enzyme levels $\varepsilon_i/\varepsilon_j$ must be equal to the absolute inverse ratio of the elasticities $|E_{i,i}^j/E_{i,i}^{j+1}|$. An example of this, with the mass-action rate law, is derived in electronic supplementary material, section 4. In contrast to the enzyme-control rule (which yields only one equation for the entire pathway), the enzyme-elasticity rule yields an equation for every internal metabolite, which determines the ratio of the enzyme levels around this metabolite. Together with the known sum of enzyme levels, these rules determine the enzyme profile completely (given all the elasticities in the optimal state).

2.4.2. Analytic formulae for metabolic control

We learned that enzyme levels in optimal states (with a bound on total enzyme, and maximizing the steady-state flux) satisfy two general laws, the enzyme-control rule and the enzyme-elasticity rule. Hence, if we have a formula for the optimal enzyme levels, and if the enzyme-control rule applies, we can compute the flux control coefficients. Moreover, the enzyme-elasticity rule (which comes from the connectivity theorem) relates the optimal enzyme levels to elasticities. These formulae can be trusted, but for didactic reasons we set out to compute the control coefficients for some of the rate laws analytically (also demonstrating different ways to compute them) and compared them to optimal enzyme levels. Here we summarize this briefly; details can be found in electronic supplementary material, section 5.

- *Michaelis–Menten rate law.* In metabolic pathways with irreversible rate laws, the first reaction has usually full flux control and therefore a flux control coefficient of 1, while the remaining reactions have no flux control and therefore flux control coefficients of 0. But under an optimization, this changes. All metabolite levels go to infinity and all elasticities go to zero. Mathematically, the optimal state does not exist (because infinite concentrations are not real numbers), but even if we assume that enzymes could be saturated completely, any enzyme variation would break the steady state. This means that the control coefficients are not defined, and so the enzyme-control rule does not apply (see electronic supplementary material, section 3.1). To obtain meaningful results, we therefore studied an optimality problem with additional metabolite constraints. In this case, metabolite levels remain finite and the control coefficients remain defined.
- *Thermodynamic rate law.* For pathway models with the thermodynamic rate law, we derived a simple formula for the flux control coefficients that contains only the enzyme levels, the k^{cat} values, and the flux. Even if we do not have a closed formula for the flux as a function of enzyme levels, we can obtain a closed formula for the flux control coefficients by using a trick. From the rate law, we first obtain a relationship between enzyme levels,

external metabolite levels, kinetic constants and flux:

$$\prod_{i=1}^n \left(1 - \frac{J}{\varepsilon_i k_i^{\text{cat}}}\right) = \frac{s_n}{s_0} \frac{1}{K_{\text{tot}}^{\text{eq}}} \quad (2.34)$$

(see equation (S31) in electronic supplementary material, section 5.2.1, for the derivation). While we cannot solve this for the flux J directly, we can obtain the response coefficients $R_{\varepsilon_i}^J = \partial J / \partial \varepsilon_i$ by implicit differentiation. This yields the flux control coefficients (derivation in electronic supplementary material, section 5.3.2):

$$C_i^J = \frac{(\varepsilon_i k_i^{\text{cat}} - J)^{-1}}{\sum_{i=1}^n (\varepsilon_i k_i^{\text{cat}} - J)^{-1}}. \quad (2.35)$$

The control coefficients are proportional to $(\varepsilon_i k_i^{\text{cat}} - J)^{-1}$ and normalized to a sum of 1 as required by the summation theorem. For didactical reasons, we checked the connectivity theorem (electronic supplementary material, section 5.3.1) and the enzyme-control rule (electronic supplementary material, section 5.3.4).

- *Mass-action rate law.* With the reversible mass-action rate law, we can analytically compute the flux control coefficients and verify that they satisfy summation and connectivity theorems. To verify that the enzyme-control rule is satisfied in the optimal state, we use the explicit formula (2.26) for the steady-state flux and take derivatives to obtain the flux control coefficients:

$$C_i^J = \frac{\gamma_i / \varepsilon_i}{\sum_j \gamma_j / \varepsilon_j}, \quad (2.36)$$

with γ_i defined as above (derivation in electronic supplementary material, section 5.4.2). The coefficients sum to 1 as required by the summation theorem. Again, we verified the connectivity theorem (electronic supplementary material, section 5.4.4) and the enzyme-control rule (electronic supplementary material, section 5.4.5) for didactical reasons.

2.5. A cell model with enzyme kinetics and metabolite constraints

As an illustrative example, we now apply our formulae to a simple model of growing cells, describing transport processes, metabolism and macromolecule synthesis, respectively, by lumped reactions. Obviously, none of the rate laws considered throughout this work can fully capture the dynamics of growing cells. For example, having only one single reaction representing metabolism is a gross oversimplification (and likewise for transport and translation). Nevertheless, we might still be able to draw insights from this model if we make the right assumptions. This approach has been successful in the past: by considering simple cell models and assuming that enzyme efficiencies are constant (i.e. completely independent of growth rate and metabolite concentrations), Basan *et al.* [10] were able to show how overflow metabolism in *E. coli* corresponds to a proteome-efficient allocation of enzymes. Here we ask how the predictions would change by accounting for enzyme kinetics. Instead of assuming that enzyme efficiencies are constant and given, we consider a model in which protein allocation and metabolite concentrations are always optimized in order to maximize the growth rate.

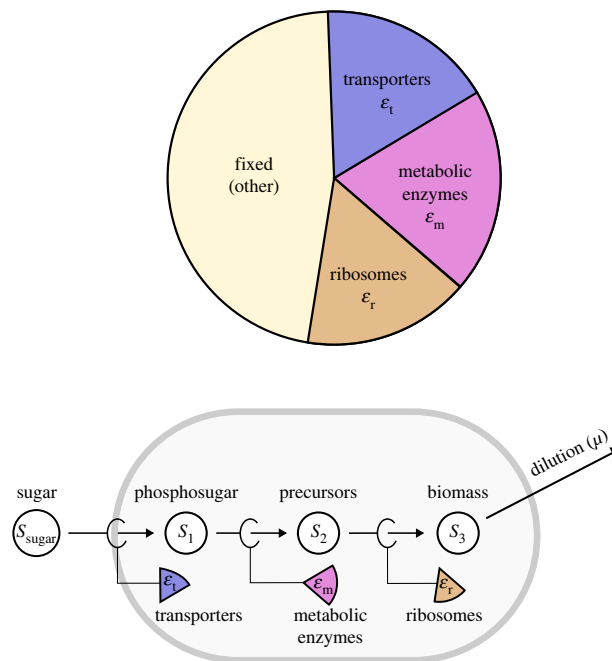


Figure 5. A growing cell described as a chain of reactions. Top: Inspired by the cell model in [9] explaining ‘bacterial growth laws’, we assume a fixed protein budget consisting of a fixed protein fraction and variable fractions for transporters, metabolic enzymes, and ribosomal proteins, but with a fixed sum. The proportions of the variable fractions are chosen to maximize growth. Bottom: The growing cell is described as a chain of reactions, each catalysed by one of the protein fractions. For a constant biomass production, the three reactions must be in steady state.

In our simple cell model (figure 5), production processes are represented by an unbranched chain of reactions, describing overall transport, overall metabolism, and protein synthesis. The steady-state fluxes are proportional to the cell growth rate and optimized under a constraint on the sum of enzyme levels (ε_{tot}) and another one on metabolite levels (s_{tot}). For convenience, in this section we will replace the pathway flux J^* with the standard symbol μ for growth rate. It should be noted that dilution of the intermediate compounds is ignored even though in reality some of them can have a high total concentration (e.g. the precursors in S_2) and thus a considerable growth dilution rate which affects protein allocation strategies.

For our model, we chose to use the Michaelis–Menten approximation, which gives the most realistic results (although we also discuss the two other analytical solutions at the end of this section). Metabolite concentrations (as extra variables) can be adjusted and become part of the optimization problem. We then apply the formulae derived in the previous sections of this paper to find the optimal allocation of enzymes and thereby maximize the growth rate of the cell. Importantly, all calculations are completely based on analytic expressions.

In §2.3.1, we derived a formula for the optimal allocation and maximal flux in the case where all metabolite levels are free variables (including s_0 , which is denoted here by s_{sugar}). However, since in this cell model s_{sugar} represents the concentration of an external nutrient that is not subject to optimization, we would like to treat it as a constant system parameter (and later show how the optimum responds to changes in it). Fortunately, adding this assumption changes the optimal solution only slightly, as described in electronic

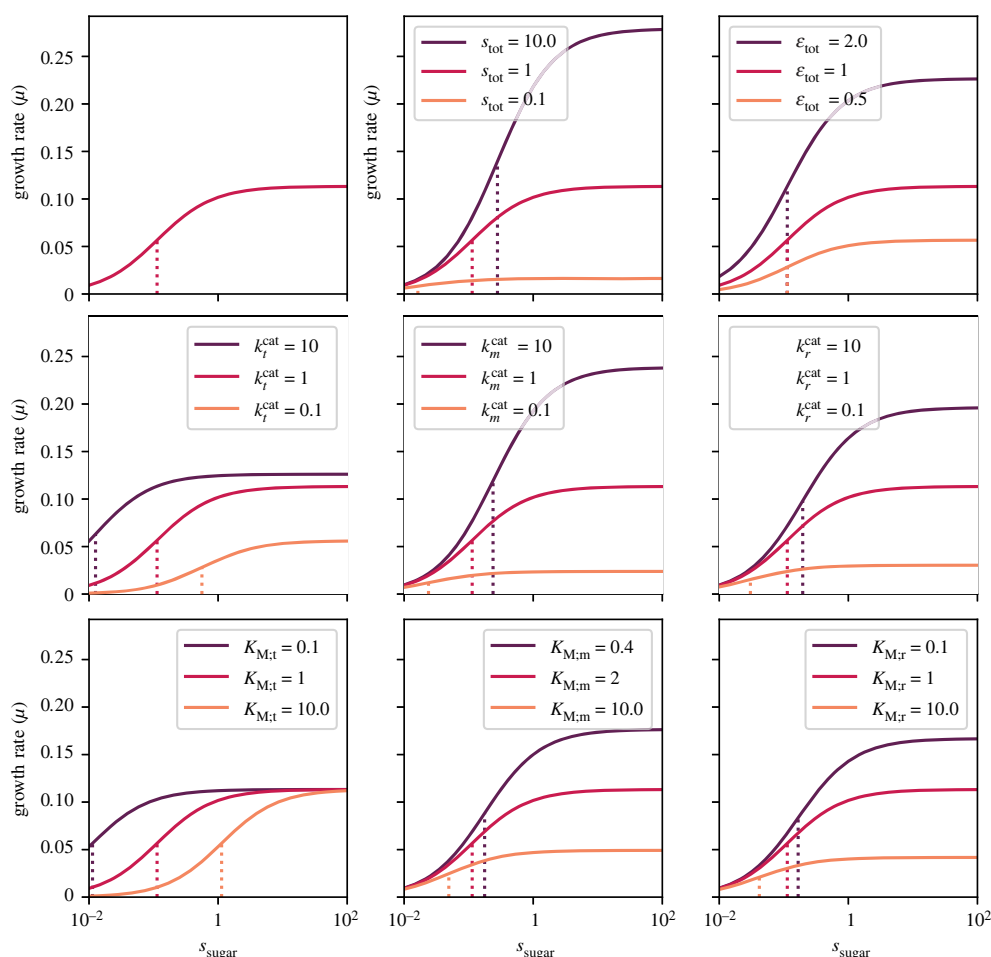


Figure 6. The optimal achievable flux in a cell model—i.e. the growth rate μ —for a given total enzyme ε_{tot} . The model contains three irreversible steps based on Michaelis–Menten kinetics. The default parameters are $k_t^{\text{cat}} = k_m^{\text{cat}} = k_r^{\text{cat}} = K_{M,t} = K_{M,r} = 1$, $K_{M,m} = 2$, and the upper bounds are set to $s_{\text{tot}} = 1$, $\varepsilon_{\text{tot}} = 1$. The panel on the upper-left corner shows the maximal flux as a function of s_{sugar} for the default parameters. In each other panel, only one of the parameters is changed. The vertical dotted lines mark the K_{Monod} (i.e. the concentration of s_{sugar} where the flux is 50% of the maximum). Note that the x -axis shows values on log-scale, and therefore the curves have a sigmoidal shape (rather than the familiar Monod-style). See electronic supplementary material, figure S1, for plots of the enzyme demands.

supplementary material, section 5.1.4. The optimal growth rate as a function of s_{sugar} can be written in the following simple form:

$$\mu = \mu^{\text{max}} \cdot \frac{s_{\text{sugar}}}{s_{\text{sugar}} + K_{\text{Monod}}}, \quad (2.37)$$

where we define

$$\left. \begin{aligned} \mu^{\text{max}} &\equiv \frac{\varepsilon_{\text{tot}}}{1/k_t^{\text{cat}} + 1/k_m^{\text{cat}} + 1/k_r^{\text{cat}} + \left(\sqrt{K_{M,m}/k_m^{\text{cat}}} + \sqrt{K_{M,r}/k_r^{\text{cat}}} \right)^2 / s_{\text{tot}}} \\ \text{and } K_{\text{Monod}} &\equiv \frac{K_{M,t}/k_t^{\text{cat}}}{1/k_t^{\text{cat}} + 1/k_m^{\text{cat}} + 1/k_r^{\text{cat}} + \left(\sqrt{K_{M,m}/k_m^{\text{cat}}} + \sqrt{K_{M,r}/k_r^{\text{cat}}} \right)^2 / s_{\text{tot}}} \end{aligned} \right\} \quad (2.38)$$

As implied by the symbol for K_{Monod} , this form corresponds to empirical observations by Monod [19] (and further followed up by others [30]) who stated that growth rate increases with the nutrient concentration until reaching a saturation level where growth is fastest. Interestingly, the value of K_{Monod} (e.g. the level of s_{sugar} for which growth rate is half of its maximum) is not determined solely by the kinetic parameters of the transporter.

To better understand how changes in the model parameters affect cell growth, we considered a toy example where all constants are set to a default value of 1 (except for $K_{M,m} = 2$). In figure 6, we plot the value of μ as a function of s_{sugar} (based on equations (2.37)–(2.38)), each time

changing one of the parameters. In almost all cases, we find a trade-off between growth and binding affinity, namely that improving the kinetics (or relaxing a constraint) improves the maximal growth rate while also increasing K_{Monod} (i.e. making it worse since higher K_{Monod} means higher a concentration of sugar is required to reach the same growth rate). The three exceptions are the response to changes in ε_{tot} which only affects μ^{max} , in K_t which only affects the K_{Monod} , and in k_t^{cat} which can improve both growth parameters at the same time. Indeed, by observing the formulae in equation (2.38), one can see that μ^{max} and K_{Monod} are proportional, with a ratio given by $\varepsilon_{\text{tot}}/(K_{M,t}/k_t^{\text{cat}})$. Hence, all parameters that are only in the denominator should affect μ^{max} and K_{Monod} in the same fashion (which leads to a trade-off), while the three parameters in the numerators have each their own unique effect.

This result might have implications for the design of optimal enzyme regulation. As the rate of sugar transport is not affected by any of the metabolite concentrations inside the cytoplasm, the transport reaction acts as an information buffer in the sense that any changes in the external sugar concentration do not affect the optimal allocation of enzymes (and metabolites) within the cytoplasm. The only thing that changes is the relative abundance of the transporter versus all other (cytoplasmic) enzymes. This rule becomes clear when plotting the optimal enzyme allocations as a function

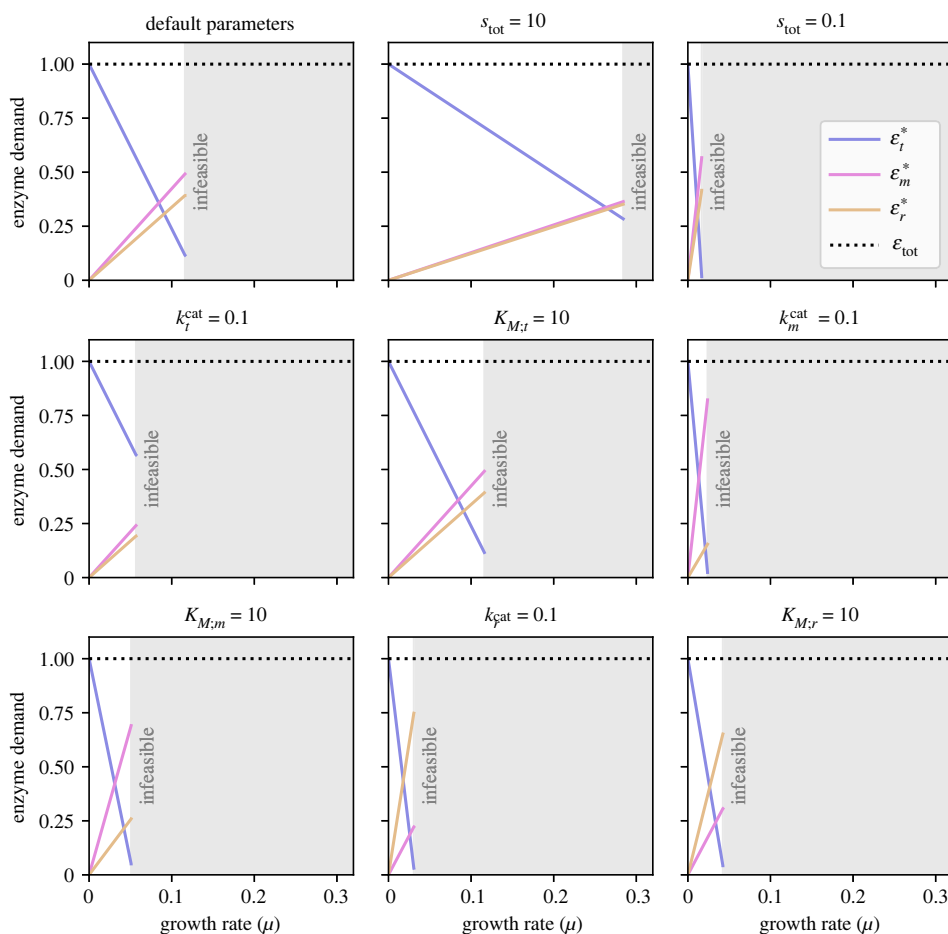


Figure 7. The sizes of the three enzyme sectors (ε_t^* , ε_m^* and ε_r^*) as a function of the flux (J^*) when the allocation is optimal. The default parameters are $k_t^{\text{cat}} = k_m^{\text{cat}} = k_r^{\text{cat}} = 1$, $K_{M,t} = K_{M,r} = 1$, $K_{M,m} = 2$ (we chose another value in order prevent the curves for ε_m and ε_r from overlapping), and the upper bounds are set to $s_{\text{tot}} = 1$, $\varepsilon_{\text{tot}} = 1$. In each panel, only one of these parameters is changed. The grey shaded area marks growth rates that are higher than the maximum, μ^{max} . The linear protein allocation curves are caused by the irreversible transport step. In the optimization, this leads to constant optimal metabolite concentrations, and the protein fractions ε_m^* and ε_r^* must change proportionally with the flux. The line slopes are given by the enzyme efficiencies at the optimal metabolite concentrations, and the transporter fraction—the residual of the protein budget—must be a falling straight line.

of the pathway flux (i.e. the growth rate), as shown in figure 7. The two cytoplasmic fractions (metabolic enzymes and ribosomes) change proportionally and appear as straight lines that cross the axis origin. Since ε_{tot} is a constant, the remainder (i.e. the level of the transporter) also decreases linearly with growth rate. These three straight lines extend only up to a point (where $\mu = \mu^{\text{max}}$). This limit represents the fully saturated transporter, which has an efficiency of k_m^{cat} and therefore an optimal abundance of $\mu^{\text{max}}/k_m^{\text{cat}}$. Higher concentrations of s_{sugar} would not generate any further benefits.

In fact, the principle described above can be shown to be true in a much more general case. As long as one section of the metabolic network is ‘isolated’ from the rest (i.e. connected only by irreversible steps), changes in the upstream parameters will affect the incoming flux, but would not change the optimal allocation of enzymes (and also not the metabolite concentrations). Curiously, in such cases our results predict, based on an underlying kinetic model, that the optimal protein efficiencies are independent of the external glucose concentration and, as a consequence, protein fractions vary linearly. Thus, our model predictions are very similar to ones from previous proteome allocation models that are based on empirical observations.

A series of irreversible steps is only one option for a simple model of cell growth. Alternatively, one might consider cells

living in an environment that provides high metabolite concentrations, albeit a limited overall thermodynamic driving force. In this case, applying our thermodynamic rate law with the approximation from equation (S46) in the electronic supplementary material, section 5.3, yields another simple growth formula:

$$\left. \begin{aligned} \mu(s_{\text{sugar}}) &= \mu^{\text{max}} \left(1 - \left(\frac{A}{s_{\text{sugar}}} \right)^B \right) \\ \text{and } \mu^{\text{max}} &\equiv \varepsilon_{\text{tot}} \left(\sum_i 1/k_i^{\text{cat}} \right)^{-1}, \end{aligned} \right\} \quad (2.39)$$

where $A \equiv s_n K_{\text{tot}}^{\text{eq}}$ and $B \equiv (\sum_i 1/k_i^{\text{cat}}) / (\sum_i 1/\sqrt{k_i^{\text{cat}}})^2$. Again, the curve parameters depend on all the k_i^{cat} values. In contrast to the Michaelis–Menten based growth formula in equation (2.38), this formula is also sensitive to product concentrations (s_n). This may become important close to chemical equilibrium, where product accumulation can decrease growth in a way that cannot be bypassed by any regulatory mechanisms. In this case, our cells would profit from the presence of a ‘cleaner’ strain that removes the accumulating product, which has been proposed as a general mechanism that would encourage symbiosis [31].

Finally, the last option to consider for an analytical cell growth model is based on the mass-action rate law. However, from equation (2.27) we can immediately see that in this case

the optimal enzyme fractions are direct functions of the kinetic parameters and do not depend on s_0 or s_n . Therefore, when changing the sugar concentration, enzyme allocation will remain constant. Furthermore, the growth rate will be a linear function of s_{sugar} without any saturation behaviour. This means there is no way to define μ^{max} and K_{Monod} for the mass-action case.

3. Discussion

In this study, we addressed the enzyme allocation problem in unbranched pathways analytically by considering a number of different rate-law approximations. Along the way, we learned that each approximation comes with its own idiosyncrasies. For instance, a model with Michaelis–Menten kinetics requires an extra metabolite constraint to obtain a valid optimal state, and the ‘thermodynamic’ rate law requires inverting a function in order to find the value of the Lagrange multiplier. Fortunately, there is a very good approximate solution that works well across the entire range of driving forces.

For historical reasons, we solved the optimality problem under the assumption that there is a bound on the sum of enzyme abundances— $\sum_i \varepsilon_i \leq \varepsilon_{\text{tot}}$ —without explicitly specifying the units. The common interpretation would be that ε_i are molar concentrations (where the symbol k_i^{cat} represents the turnover numbers in units of 1 over time). However, the total mass concentration of all pathway enzymes may be a better proxy for cost, since crowding effects in the cytoplasm are often limiting the total mass concentration of all soluble proteins, including most enzymes. Furthermore, fast-growing cells are often limited by the rate of protein elongation, and the molecular mass is nearly proportional to the gene length. Therefore, we usually think of ε_i as mass concentrations (e.g. in $\text{g} \times \text{m}^{-3}$) and of k_i^{cat} as specific activities (e.g. in $\text{mol} \times \text{s}^{-1} \times \text{g}^{-1}$). Nevertheless, the analytic derivation and solutions provided here are agnostic to the choice of units and are equally valid for both these interpretations. Moreover, one could imagine a completely different set of linear weights for the enzyme cost function (i.e. $\sum_i w_i \varepsilon_i \leq \varepsilon_{\text{tot}}$, as in figure 1). If one thinks of the weights as scaling factors for the k^{cat} values, the provided solutions will still hold (while making use of the new ‘effective’ k^{cat} values).

Previous studies focused on the ‘mass-action’ rate law, justifying it by saying that the general reversible form derived by Haldane can be approximated at the limit of low concentrations. However, it is quite rare to have all the substrates and products of an enzyme at concentrations that are way below their K_M values [32]. Furthermore, the limit of all reactant concentrations going to zero is not very meaningful because, in the first place, Haldane derived his rate law assuming enzymes are much less abundant than metabolites [33]. Interestingly, just being close to equilibrium is not enough for this approximation, since the product can still affect the rate nonlinearly via η^{sat} (figure 2). Here we tried to consider more comprehensively all the different approximations that yield an analytic solution to optimal enzyme allocation in unbranched pathways (table 1).

One of these approximations is the Michaelis–Menten rate law, which is widely used in enzymatic assays and metabolic modelling of irreversible reactions. Curiously, using it for the simple optimality problem with linear chains leads to paradoxical results: the metabolite concentrations go to infinity, the

elasticities vanish and flux control coefficients are not defined. For solving this problem, we introduced a new upper bound on the total metabolite concentration in order to obtain realistic results and derive analytic solutions based on this rate law. Although it is very reasonable to assume that concentrations of small molecules (and not just enzymes) are restricted in cells, this fact is often ignored in metabolic models. One of the rare cases where this constraint was taken into account is the work of Dourado *et al.* [34], who found empirical evidence to the fact that there is a balance between enzymes and substrates when minimizing the total mass concentration. In addition to an analytic solution, we also found a new enzyme-control rule for models with a constraint on enzyme and metabolite concentrations: in this case, the enzyme amounts do not only reflect the flux control coefficients, but a sum of flux and concentration control coefficients (see electronic supplementary material, section 3.2).

Besides ‘mass-action’ and Michaelis–Menten kinetics, we discussed a solution for one other rate law which we call *thermodynamic*, as the rate is only affected by metabolite levels through the thermodynamic driving force (i.e. ignoring any saturation effects). One advantage of this rate law is that it does not require knowing the K_M values (which are often difficult to come by). The thermodynamic-only approximation was also used for justifying the max–min driving force (MDF) method (as described in electronic supplementary material, section 5.3.5), which similarly aims to quantify the efficiency of metabolic pathways [35]. But, unlike the solution presented here, MDF does not explicitly optimize a simplified version of a kinetic rate law, but rather applies a heuristic based on the assumption that the lowest driving forces are mainly responsible for increased enzyme demands, and the total driving force should therefore be as evenly distributed as possible. On the other hand, the advantage of the MDF method is that it takes metabolite concentration bounds into consideration.

All throughout this paper, we only considered pathways with uni–uni reactions (one substrate, one product, with stoichiometric coefficients of 1). In reality, many reactions involve co-factor pairs or other substrates or products. Instead of deriving results for this general case, we assumed that these extra reactants may exist, but with known concentrations. In this case, the rate laws contain extra terms, but these terms can be rearranged to yield the same simple formulae as in the uni–uni case, using *effective* kinetic constants. We demonstrate this for the case of two substrates and two products following convenience kinetics [36] in electronic supplementary material, section 5.6. Notably, this logic also applies to reactions with more than two substrates and products as well as to enzyme activation or inhibition with constant activator or inhibitor levels.

This paper can be seen as an exercise in solving the enzyme allocation problem and describing the optimal states analytically using MCA. Although some might argue that the required approximations are not realistic, they do represent a step forward compared to the very common approach of assuming that metabolites have no effect on enzyme efficiency at all. On the other hand, adding metabolite concentrations as extra variables greatly increases the complexity of models and typically renders them unsolvable. Therefore, the solutions provided here might be handy, as the assumption of metabolite steady state combined with the optimality argument give us analytic expressions that only depend on the initial and final metabolites in the pathway. We demonstrated this result using a toy example for a course-grained model of cell

growth, and showed how the analytic solutions provide valuable insights about the effects of changes in each parameter—all this without the need to simulate the metabolic network or use nonlinear solvers. One could imagine a scenario where metabolic engineers, while trying to improve a biochemical pathway, have to decide in which enzyme to invest their time. Although MCA can perfectly serve this purpose while covering more scenarios, it requires tools and language that can seem opaque to people lacking the proper mathematical background. The solutions provided here should be more approachable and serve as simple guidelines while capturing a wider range of cases than previously explored in the literature. We hope that future studies will apply and extend this approach to other, more complex models.

Ethics. This work did not require ethical approval from a human subject or animal welfare committee.

References

- Rosen R. 1967 *Optimality principles in biology*. Berlin, Germany: Springer.
- Dekel E, Alon U. 2005 Optimality and evolutionary tuning of the expression level of a protein. *Nature* **436**, 588–692. (doi:10.1038/nature03842)
- Schütz R, Zamboni N, Zampieri M, Heinemann M, Sauer U. 2012 Multidimensional optimality of microbial metabolism. *Science* **336**, 601–604. (doi:10.1126/science.1216882)
- Beg QK, Vazquez A, Ernst J, de Menezes MA, Bar-Joseph Z, Barabási A-L, Oltvai ZN. 2007 Intracellular crowding defines the mode and sequence of substrate uptake by *Escherichia coli* and constrains its metabolic activity. *Proc. Natl Acad. Sci. USA* **104**, 12 663–12 668. (doi:10.1073/pnas.0609845104)
- Tepper N, Noor E, Amador-Noguez D, Haraldsdóttir HS, Milo R, Rabinowitz J, Liebermeister W, Shlomi T. 2013 Steady-State metabolite concentrations reflect a balance between maximizing enzyme efficiency and minimizing total metabolite load. *PLoS ONE* **8**, e75370. (doi:10.1371/journal.pone.0075370)
- Wortel MT, Noor E, Ferris M, Bruggeman FJ, Liebermeister W. 2018 Metabolic enzyme cost explains variable trade-offs between microbial growth rate and yield. *PLoS Comput. Biol.* **14**, e1006010. (doi:10.1371/journal.pcbi.1006010)
- Flamholz A, Noor E, Bar-Even A, Liebermeister W, Milo R. 2013 Glycolytic strategy as a tradeoff between energy yield and protein cost. *Proc. Natl Acad. Sci. USA* **110**, 10 039–10 044. (doi:10.1073/pnas.1215283110)
- Noor E, Flamholz A, Bar-Even A, Davidi D, Milo R, Liebermeister W. 2016 The protein cost of metabolic fluxes: prediction from enzymatic rate laws and cost minimization. *PLoS Comput. Biol.* **12**, e1005167. (doi:10.1371/journal.pcbi.1005167)
- Scott M, Gunderson CW, Mateescu EM, Zhang Z, Hwa T. 2010 Interdependence of cell growth and gene expression: origins and consequences. *Science* **330**, 1099. (doi:10.1126/science.1192588)
- Basan M, Hui S, Okano H, Zhang Z, Shen Y, Williamson JR, Hwa T. 2015 Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature* **528**, 99. (doi:10.1038/nature15765)
- Bar-Even A, Noor E, Lewis NE, Milo R. 2010 Design and analysis of synthetic carbon fixation pathways. *Proc. Natl Acad. Sci. USA* **107**, 8889–8894. (doi:10.1073/pnas.0907176107)
- Löwe H, Kremling A. 2021 In-depth computational analysis of natural and artificial carbon fixation pathways. *BioDesign Res.* **2021**, 9898316. (doi:10.34133/2021/9898316)
- Heinrich R, Schuster S. 2012 *The regulation of cellular systems*. New York, NY: Springer.
- Kacser H, Burns JA. 1973 The control of flux. *Symp. Soc. Exp. Biol.* **27**, 65–104.
- Heinrich R, Rapoport TA. 1974 A linear steady-state treatment of enzymatic chains. General properties, control and effector strength. *Eur. J. Biochem.* **42**, 89–95. (doi:10.1111/j.1432-1033.1974.tb03318.x)
- Brown GC. 1991 Total cell protein concentration as an evolutionary constraint on the metabolic control distribution in cells. *J. Theor. Biol.* **153**, 195–203. (doi:10.1016/S0022-5193(05)80422-9)
- Heinrich R, Klipp E. 1996 Control analysis of unbranched enzymatic chains in states of maximal activity. *J. Theor. Biol.* **182**, 243–252. (doi:10.1006/jtbi.1996.0161)
- Klipp E, Heinrich R. 1999 Competition for enzymes in metabolic pathways: implications for optimal distributions of enzyme concentrations and for the distribution of flux control. *BioSystems* **54**, 1–14. (doi:10.1016/S0303-2647(99)00059-3)
- Monod J. 1949 The growth of bacterial cultures. *Annu. Rev. Microbiol.* **3**, 371–394. (doi:10.1146/annurev.mi.03.100149.002103)
- Waley SG. 1964 A note on the kinetics of multi-enzyme systems. *Biochem. J.* **91**, 514–517. (doi:10.1042/bj0910514)
- Burns JA. 1971 Studies on complex enzyme systems. PhD thesis, University of Edinburgh, UK.
- Schuster S, Heinrich R. 1991 Minimization of intermediate concentrations as a suggested optimality principle for biochemical networks. *J. Math. Biol.* **29**, 425–442. (doi:10.1007/BF00160471)
- Noor E, Flamholz A, Liebermeister W, Bar-Even A, Milo R. 2013 A note on the kinetics of enzyme action: a decomposition that highlights thermodynamic effects. *FEBS Lett.* **587**, 2772–2777. (doi:10.1016/j.febslet.2013.07.028)
- Michaelis L, Menten ML. 1913 Die Kinetik der Invertinwirkung. *Biochem. Z.* **49**, 352.
- Christensen CD, Hofmeyr J-HS, Rohwer JM. 2018 PySCeSToolbox: a collection of metabolic pathway analysis tools. *Bioinformatics* **34**, 124–125. (doi:10.1093/bioinformatics/btx567)
- Liebermeister W, Noor E. 2015 The enzyme cost of given metabolic flux distributions, as a function of logarithmic metabolite levels, is convex. (<http://arxiv.org/abs/1501.02454>)
- Heinrich R, Rapoport TA. 1973 Linear theory of enzymatic chains: its application for the analysis of the crossover theorem and of the glycolysis of human erythrocytes. *Acta Biol. Med. Germ.* **31**, 479–494.
- Dourado H, Mori M, Hwa T, Lercher MJ. 2021 On the optimality of the enzyme-substrate relationship in bacteria. *PLoS Biol.* **19**, e3001416. (doi:10.1371/journal.pbio.3001416)
- Liebermeister W, Klipp E, Schuster S, Heinrich R. 2004 A theory of optimal differential gene expression. *BioSystems* **76**, 261–278. (doi:10.1016/j.biosystems.2004.05.022)
- Shehata TE, Marr AG. 1971 Effect of nutrient concentration on the growth of *Escherichia coli*. *J. Bacteriol.* **107**, 210–216. (doi:10.1128/jb.107.1.210-216.1971)
- Großkopf T, Soyer OS. 2016 Microbial diversity arising from thermodynamic constraints. *ISME J.* **10**, 2725–2733. (doi:10.1038/ismej.2016.49)

32. Bennett BD, Kimball EH, Gao M, Osterhout R, Van Dien SJ, Rabinowitz JD. 2009 Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol.* **5**, 593–599. (doi:10.1038/nchembio.186)
33. Haldane SJB. 1930 *Enzymes*, ch. V, pp. 74–92. London, UK: Longmans, Green and Co.
34. Dourado H, Maurino VG, Lercher MJ. 2017 Enzymes and substrates are balanced at minimal combined mass concentration *in vivo*. *bioRxiv*. (doi:10.1101/128009)
35. Noor E, Bar-Even A, Flamholz A, Reznik E, Liebermeister W, Milo R. 2014 Pathway thermodynamics highlights kinetic obstacles in central metabolism. *PLoS Comput. Biol.* **10**, e1003483. (doi:10.1371/journal.pcbi.1003483)
36. Liebermeister W, Klipp E. 2006 Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theor. Biol. Med. Model.* **3**, 41. (doi:10.1186/1742-4682-3-41)
37. Noor E, Liebermeister W. 2023 Optimal enzyme profiles in unbranched metabolic pathways: Jupyter notebooks. *Zenodo*. (doi:10.5281/zenodo.10069509)
38. Noor E, Liebermeister W. 2023 Optimal enzyme profiles in unbranched metabolic pathways. *Figshare*. (doi:10.6084/m9.figshare.c.6976367)