The value structure of metabolic states

Wolfram Liebermeister[1,2]

[1] INRA, UR1404, MaIAGE, Université Paris-Saclay, France
[2] Institut für Biochemie, Charité – Universitätsmedizin Berlin, Germany

Abstract

To improve their metabolic performance, cells have to realise good compromises between large production fluxes, low enzyme investments, and well-adapted metabolite levels. In models, this idea can be formulated in the form of optimality principles that trade a high metabolic benefit against low enzyme cost. However, different modelling approaches are often incompatible. I propose a unified theory, called metabolic economics, to bridge the gap between different optimality-based cell models that exploits hidden equivalences between these approaches. Metabolic economics introduces new variables on the network, called economic variables, which represent the cost and benefit of metabolites, fluxes, and enzymes, and can be defined by Lagrange multipliers, auxiliary variables that are commonly used to handle constraints in optimality problems. Metabolic economics translates optimality conditions into local balance equations between these variables. The economic potentials and loads describe the value of metabolite production and metabolite concentrations. As "proxy variables", they can describe indirect fitness effects, arising elsewhere in the network, as if they arose locally in a reaction of interest. Here I derive these variables and their balance equations for three types of optimality problems: for direct optimisation of enzyme levels in kinetic models; for flux cost minimisation (FCM), a minimisation of enzyme cost, with flux and metabolite profiles as the variables to be optimised; and for optimal protein allocation in whole-cell models, where growth rate or other whole-cell objectives are maximised. The economic balance equations add a new layer of description to mechanistic models, a description in terms of beneficial cell functions and associated costs, and can seen as economic laws of metabolism. Metabolic economics provides concepts for comparing and combining metabolic optimality problems, employing different modelling paradigms or different levels of detail, which can be useful for semi-automatic or modular modelling.

Keywords: Metabolic network, kinetic model, optimality condition, Lagrange multiplier, resource allocation

1 Introduction

A living cell can show various metabolic states, with fluxes, metabolite profiles, and enzyme profiles that provide different selective advantages. The choice of these flux states, or "metabolic strategies", raises many questions: Which metabolic pathways should a cell use in different situations, and why? How should cells distribute their protein budget within and across pathways, and how should the protein levels change if cells are perturbed? For example, if an enzyme is inhibited, should this enzyme (and other enzymes in the same pathway) be overexpressed to compensate for the perturbation, or should the entire pathway be switched off (because it works inefficiently)? If the perturbation increases, at which point should the cell switch from the first strategy to the second one? And, finally, how fast can cells grow under optimal circumstances and what limits their growth? These questions do not primarily concern the actual behaviour shown by cells, but the behaviour that would be most profitable. Thus, studying mechanisms – e.g. how enzyme levels are regulated biochemically – is not sufficient. Instead, one needs to ask an economical question: what behaviour – e.g. what enzyme levels, and what ensuing metabolite levels

and fluxes – would be most advantageous in a given situation, and thus likely to emerge through evolutionary adaptation?

The metabolic behaviour of cells may be understood, at least partially, through economic principles. In this view, the cell pursues some general task, for example, to produce biomass. This global objective induces a direct incentive for a high reaction flux in the biomass reaction; but it also induces an incentive for high fluxes, metabolite levels, and enzyme levels elsewhere in the network. This choice involves trade-offs. Higher fluxes, for example, allow cells to produce biomass faster, but the catalysing enzymes would occupy valuable space [1] and their production and maintenance would consume more resources such as amino acids, ribosomes, and energy. In bacteria, protein costs have been quantified by measuring growth defects after the forced expression of proteins [2, 3]). In view of these costs, cells must prioritise the usage of metabolic pathways and must adjust their fluxes to changing supplies and demands: the resulting metabolic fluxes will reflect a compromise between metabolic tasks – e.g. production of biomass components – and a limited enzyme capacity. Notably, by repressing some enzymes, cells do not only save protein cost – they can also disrupt futile substrate cycles, which would waste valuable metabolites [4], and can choose between pathways, and thus between different metabolic strategies, e.g. maximising either cell growth or product yield [5, 6]. In models, such cost-benefit considerations can be described by optimality problems. For example, to find the enzyme levels in a kinetic model that lead to an optimal state (i.e. flux distributions and metabolite levels) [7, 8, 9, 2, 6], enzyme profiles are scored by benefit terms – usually functions of the steady-state fluxes and metabolite levels – and, at the same time, penalised by cost functions [7, 8, 2] or constrained by fixing the sum of enzyme levels [8]. As an important result, relationships between the costs and control coefficients of individual enzymes have been predicted.

Optimality problems in cell models can be solved numerically on a case-by-case basis. However, this is not the aim of this article. To answer these questions, we do not need more simulations, but a general theory. To this aim, I will ask how different optimality problems are related and how we can abstract from their particular formulations to obtain intuitive, general, local laws for optimal metabolic states. In the first part of this article [10], optimality conditions for many optimality problems were formulated as balances between cost and benefit gradients. Here, as a next step, I describe a framework that connects and unifies these problems. The framework, called metabolic economics, employs new, "economic" variables that satisfy balance equations on the metabolic network. I will associate the single network elements – enzymes, reactions, and metabolites – with new economic variables that describe their importance – or "economic value" – for the overall cell objective. With their help, the balances can be rewritten as local balance equations which may concern the economic values of single enzymes and their catalysed reactions, the economic values of single reactions and their direct substrates and products, or the economic values of single metabolites and their adjacent reactions. Even if the balances are formulated locally, they capture all the complexity of the surrounding network because the economic variables implicity describe global effects. Mathematically, the economic values are Lagrange multipliers, dual variables that automatically arise in the optimality problems and describe the economic value of fluxes, enzymes, and metabolites for cells in the metabolic state in question. Economic variables and the economic balance equations provide a general framework for optimal metabolic states.

2 Metabolic value theory

In [10], I described a number of metabolic optimality problems as variants of a general problem: an optimisation on the set of feasible metabolic states (characterised by fluxes, metabolite levels, and enzyme levels, and cconstrained by stationarity, rate laws, physiological bounds, and possibly fixed flux benefits). I assumed that there are trade-offs between metabolic objectives (e.g., between flux benefit, metabolite cost, and enzyme cost). I showed that such trade-offs, no matter how they are modelled, lead to a simple optimality condition: a weighted sum of gradients, for the different metabolic objectives, must cancel out. This holds not only if the overall objective is
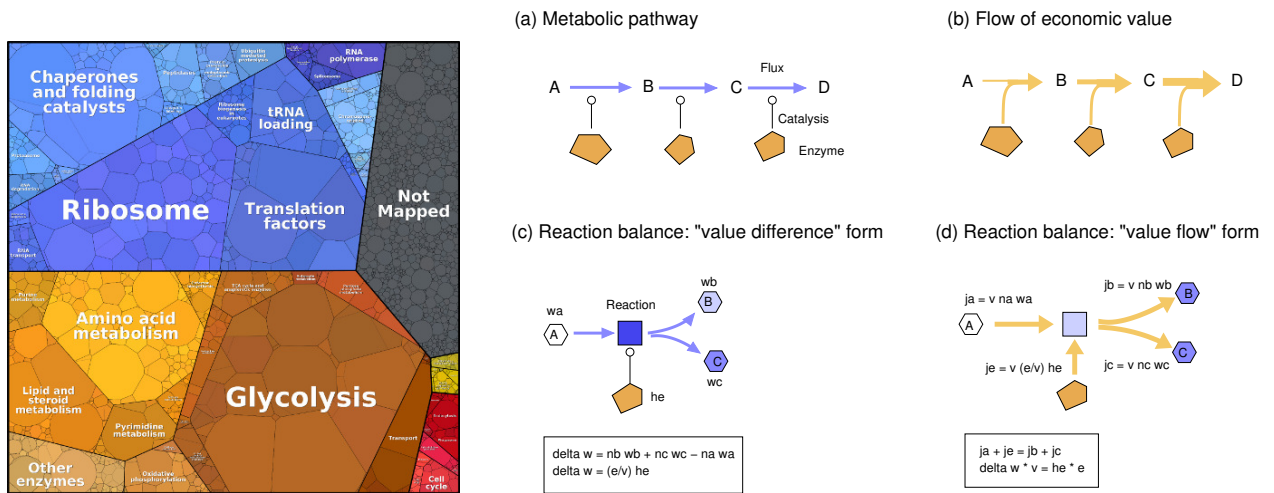
Figure 1: Protein investment and economic balance equations. (a) Proteome of yeast cells, shown as a proteomap. Metabolic enzyme level (brown polygons) can be interpreted as enzyme investments, which need to provide a benefit. The relation between cost and benefit, for individual enzymes, is described by metabolic economics.

a weighted sum of the individual objectives, but also if some objectives are used as constraints, and in multi-objective (Pareto) optimisation. This fact – that various formulations of the problem lead to the same optimality conditions – is convenient. However, some of the gradients (e.g., the gradient of flux benefit function in the space of enzyme levels) may be hard to compute and will depend on various details of the kinetic model. This is no surprise: in order to understand how the change of an enzyme level changes the global flux benefit, we need to trace the effects of this change through the network.

When we write the optimality condition for specific problems (e.g., with metabolite levels or enzyme levels as variables to be optimised), we obtain balance equations that connect the physical (concentrations, fluxes, enzyme levels) and economical (value and price) variables of different network elements. But the problem remains: in the equations, the variables are coupled non-locally through control coefficients. This makes sense, since the incentive for an enzyme adjustment in one place can depend on the incentives for flux or concentration changes in other places. However, these non-local dependencies are hard to handle and pose epistemological problems: if the economical variables in a pathway depend on variables anywhere in the network, how can we model this pathway alone? To obtain a self-contained, realistic pathway model, we need to represent all external fitness effects by proxy variables at the model boundary. To do this, we need a local theory – a theory in which global effects can be replaced by local variables. Here, I present such a theory. I start from metabolic optimality problems as in [10], write them down in an expanded form, with explicit local constraints, treat the constraints with Lagrange mulitpliers, and obtain general optimality conditions with a simple local form. These local optimality conditions are the laws of metabolic economics. The approach holds for a wide range of modelling paradigms and applies to systems both inside and outside metabolism.

How can we understand patterns in measured proteomes, or, metaphorically speaking, the enzyme investments in cells? Here I make the premise that all enzyme investments must be justified by positive benefits. Implicitly, the same premise underlies all variants of Flux Balance Analysis that suppress all fluxes that do not contribute – at least indirectly – to biomass production. In metabolic economics, this balance between cost and benefit is explicitly assumed. It is justified by an optimality assumption, and benefit is taken to arise from a local production of value, i.e. conversion of low-value into high-value metabolites. The economic values of metabolites are mathematically defined and complementary to our physical variables (such as metabolite levels and fluxes). Let me already state some results of metabolic economics (which will be derived and explained further down in the article). For kinetic metabolic models, we introduce economic variables called economic potentials w_{r_l}, which are coupled to enzyme
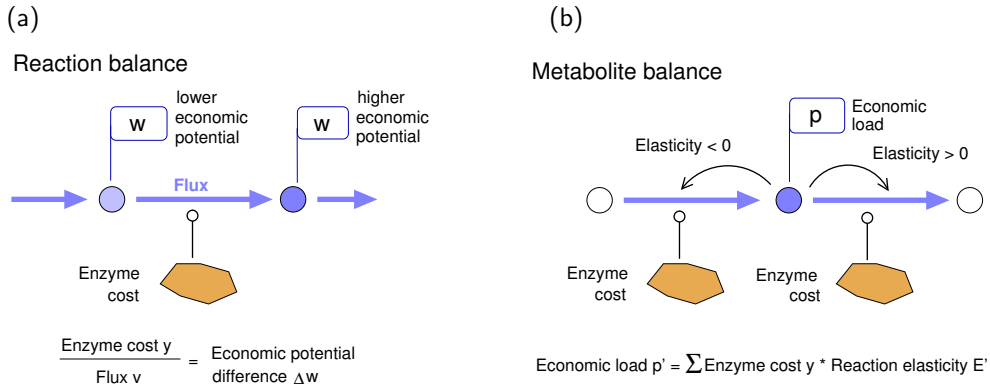
(a)
Reaction balance

(b)
Metabolite balance

Figure 2: Economic balance equations. The reaction balance (a) and the metabolite balance (b) represent optimality conditions in biochemical networks. Different formulations of the optimality problem lead to the same economic balance equations. The balance equations provide a unified, local view on metabolic optimality problems. They apply to constraint-based or kinetic models and pathway or whole-cell models.

costs by two types of balance equations (see Figure 2):

$$[\Delta w_{r_l} + b_{v_l}] \, v_l \;=\; h_{e_l} \, e_l \qquad \text{Reaction balance} \tag{1}$$

$$\sum_l h_{e_l} \, e_l \, \mathcal{E}_{c_i}^{v_l} \;=\; y_{c_i}^{\text{int}} \, c_i \qquad \text{Metabolite balance.} \tag{2}$$

The equations relate the physical variables in the model (fluxes v_l, metabolite levels c_i, and enzyme levels e_l) to economic variables (economic potentials w_{r_l}, the enzyme prices h_{e_l}, and metabolite loads $y_{c_i}^{\text{int}}$) describing the "values" network elements (or, more precisely, of some physical constraints associated with these elements). The symbol Δ denotes the difference between the products and substrates in a reaction (and is closely related to the stoichiometric coefficients n_{il}), and $\mathcal{E}_{c_i}^{v_l}$ is the scaled elasticity coefficient, known from Metabolic Control Analysis. After dividing by the physical variables, we can write the balance equations, as

$$\Delta w_{r_l} + b_{v_l} \;=\; a_{v_l} \qquad \text{Reaction balance}$$

$$\sum_l a_{v_l} \, E_{c_i}^{v_l} \;=\; y_{c_i}^{\text{int}} \qquad \text{Metabolite balance.} \tag{3}$$

with flux prices $a_{v_l} = h_{e_l} \, e_l / v_l$ and unscaled elasticities $E_{c_i}^{v_l}$.

The first equation holds for kinetic and stoichiometric models with optimality assumptions, and follows directly from mass balance constraints in these models. The second equation is related to enzyme kinetics, which appear in the equation in the form of reaction elasticities. Together, the balance equations (1) and (2) describe how optimal enzyme investments (or, as proxies, the enzyme amounts in proteome data) reflect the structure, kinetics, and flux state of the metabolic system. The definitions of enzyme cost and metabolic objectives – from which the economic potentials are derived – depend on the type of model. The equations give some immediate insights. The reaction balance (1) states that the cost of an enzyme in optimal states is given by the economic potential difference in the catalysed reaction, multiplied by the flux (or, in other words, by the rate of value production in the reaction). This implies, for example: the total enzyme cost in a metabolic pathway (or in any set of reactions) is given by the total rate of value production $\Delta \mathbf{w}_{r^{\text{int}}} = \Delta \mathbf{w}_r^{\text{int}} \cdot \mathbf{v}$ (from Eq. (1)). It also implies that all fluxes, to justify the enzyme investments, must run from lower to higher potentials. This resembles a rule in thermodnamics, the rule that fluxes must run from higher to lower chemical potentials to dissipate Gibbs free energy. According to the balance equations, a metabolite embodies, in its economic potential, the cost of transporters or enzymes that were needed to import or produce this metabolite. Therefore, knowledge about quantitative enzyme costs and benefits can be easily propagated across the network. If we apply Eq. (2) to a

4

chain fo reactions without any bounds on metabolite levels, the costs of the enzymes before and after a metabolite are inversely proportional to their reaction elasticities. In a model with all kinetic parameters set to 1, the enzyme point costs would gradually decrease along the chain[1].

So far, I postulated that economic variables exist and that they satisfy economic balance equations. But how can these variables be defined and how can the balance equations be proven? To do so, we consider metabolic optimality problems and analyse their optimality conditions. We first unify the optimality problems by writing all relationships between variables as explicit constraints. The constraints can be handled with Lagrange multipliers, and the Lagrange multipliers can be seen as economic variables. Lagrange multipliers are auxiliary variables used in constrained optimisation, and therefore also in physical theories based on variational principles (such as Lagrangian mechanics and classical thermodynamics). In optimality problems with inequality constraints, optimal states need to satisfy the Karush-Kuhn-Tucker conditions (SI S3.1): each constraint is described by a Lagrange multiplier, and the necessary optimality conditions are equations between these Lagrange multipliers and the derivatives of objective and constraint functions. The signs of a Lagrange multiplier depends on the type of bounds (upper or lower) and whether a bound is active or not. In maximisation problems, lower bounds correspond to positive signs and upper bounds correspond to negative signs (see SI section S3.1), and inactive constraints go with zero values.

3 From optimality problems to economic balance equations

In this section, we will see how the economic balance equations can be derived. As example cases, we consider some metabolic optimality problems (see Figure 3), including the ones mentioned in [10]: optimality problems in flux, metabolite, and enzyme space (as possible projections of the general state manifold into specific spaces). To make the optimality problems comparable, I will write them as *maximisation* problems. This means that lower bounds (which resemble maximisation objectives) lead to positive Lagrange multipliers, while upper bounds (which resemble minimisation objectives) lead to negative Lagrange multipliers (see SI S3.1). Detailed descriptions are given in SI section S1.

3.1 Flux analysis models

As a first type of optimality problems, we consider flux cost minimisation, the problem of choosing a metabolic flow $\mathbf{v}$ that realises a given linear benefit[2] $b(\mathbf{v}) = \mathbf{b}_\mathrm{v} \cdot \mathbf{v}$ at a minimal cost $a(\mathbf{v})$, (see Figure 3 (b) and (c) as well as SI section S2.1).

$$
\begin{array}{lll}
\text{Minimise flux cost } a(\mathbf{v}) & \text{with respect to fluxes } \mathbf{v} & \\
\text{subject to} & \mathbf{N}\,\mathbf{v} = 0 & \text{Stationarity} \\
& \mathbf{b}_\mathrm{v} \cdot \mathbf{v} \geq b' & \text{Benefit bound.} \quad\quad (4)
\end{array}
$$

A typical example is FBA with a principle of minimal fluxes [12]. The stoichiometric matrix $\mathbf{N}$ refers to internal (i.e. balanced) metabolites. The benefit function b may describe some production rate (e.g. of a pathway product or biomass). The cost function may represent, for example, a weighted sum of absolute fluxes, but we may also consider non-linear flux cost functions; the main condition for a meaningful optimality problem is that the cost $a(\mathbf{v})$ increases with each flux (mathematical condition $\mathrm{sign}(\partial a/\partial v_l) = \mathrm{sign}(v_l)$). In (4), we impose an inequality

---

[1]In linear pathways with equal driving forces and K_M values in all reactions, the substrate elasticity in each reaction is larger than the product elasticity. According to the compound balance equation, the enzyme cost in a producing reaction l should therefore be larger than in consuming reaction $l + 1$. Since this holds for all metabolites, the enzyme costs must decrease along the chain. This confirms a known result for flux optimisation at fixed total enzyme levels [11].

[2]A non-linear benefit function with gradient $b_\mathbf{v} = \partial b/\partial \mathbf{v}$ would lead to the same optimality condition.
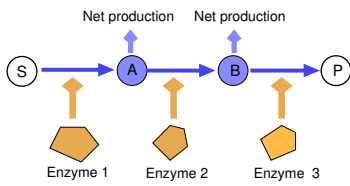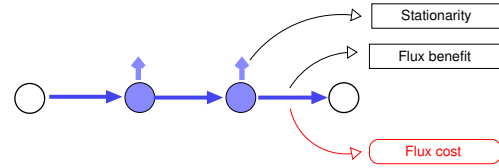
Figure 3: Optimality problems considered in this paper. (a) Network elements. In a schematic example pathway, network elements (metabolites, reactions, and enzymes) are associated with constraints and optimisation objectives. (b) Flux cost minimisation. (c) Flux cost minimisation with kinetic flux cost functions. (d) Enzyme optimisation. (e) Kinetic cost optimisation in metabolite space ("Enzyme cost minimisation")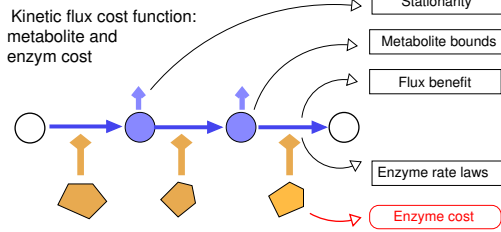. (f) Maximisation of cell growth rate. (g) Resource balance analysis with linear objective, scoring fluxes and compound concentrations. The compounds (blue) comprise metabolites and enzymes. Following the usual terminology, stationarity, rate laws, and concentration bounds are replaced by mass-balance constraints, capacity constraints, and density constraints. (h) Resource balance analysis with growth rate maximisation. For clarity reasons, only some of the arrows are shown in the graphics: for example, in (b) stationarity is assumed for all internal metabolites, but the arrow is shown only for metabolite B.
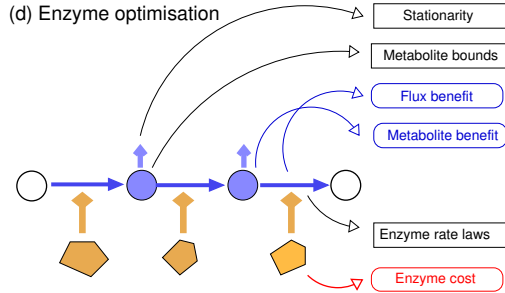
constraint: i.e. we require a benefit $b(\mathbf{v})$ at least as high as the desired value b'. In practice, this has the same effect as an equality constraint $\mathbf{b}_{\mathrm{v}} \cdot \mathbf{v} = b'$, but it keeps the problem general and leads to some further insights. To solve the problem (4), we rewrite it as a maximisation problem, introduce Lagrange multipliers (a vector $\boldsymbol{\omega}_{\varphi}$

for stationarity, and a scalar ω_b for the flux benefit constraint), and obtain the formulation

$$\text{Maximise } \mathcal{L}(\mathbf{v}) = -a(\mathbf{v}) + \boldsymbol{\omega}_{r\text{-int}}^{\top} \mathbf{N}\, \mathbf{v} + \omega_b\, (\mathbf{b}_v \cdot \mathbf{v} - b') \quad \text{with respect to fluxes } \mathbf{v} \tag{5}$$

with the Lagrangian $\mathcal{L}$, where the constraints from (4) still need to be satisfied. The necessary optimality condition (Karush-Kuhn-tucker condition) reads

$$0 = \nabla_{\mathbf{v}}\, \mathcal{L} = \nabla_{\mathbf{v}} a - \mathbf{N}^{\top} \boldsymbol{\omega}_{r\text{-int}} - \omega_b\, \mathbf{b}_v. \tag{6}$$

For a flux distribution $\mathbf{v}$ to be optimal, there must be a vector $\boldsymbol{\omega}_{r\text{-int}}$ and a scalar ω_b that solve this equation. Since ω_b refers to a lower bound and since (5) is a *maximisation problem*, ω_b must be positive (see SI section S3.1). To write equation (6) in a standard form, comparable across optimality problems, we change the variable names. With the symbols $w_{r_i}^{\text{int}} = \omega_i$ ("economic potential"), $b_{v_l}^{\star} = \omega_b\, \mathbf{b}_v$, ("direct flux gain") and $a_{v_l} = \frac{\partial a}{\partial v_l}$ ("flux price"), we obtain the economic reaction balance

$$\Delta w_{r_l}^{\text{int}} + b_{v_l}^{\star} = a_{v_l} \tag{7}$$

which must hold for each active reaction l (and where a_{v_l} must have the same sign as v_l because of our assumptions about meaningful flux cost functions). Eq. (7) states that the flux price $a_{v_l} = \partial a/\partial v_l$ is balanced by the flux gain $b_{v_l} = \partial b/\partial v_l$ and by the difference of economic potentials $\Delta w_{r_l}^{\text{int}}$. . This equation holds for all flux cost minimisation problems. With a specific choice of our cost function, the balance equation can be used to express thermodynamic constraints: after formulating the principle of minimal entropy production as a flux cost minimization problem, the balance equations tell us that fluxes run from higher to lower chemical potentials.

Metabolic fluxes are catalysed by enzymes. If this is what makes them costly, how can we compute the cost of a flux from the underlying enzyme cost? If we multiply both sides of Eq. (7) by the fluxes v_l, we obtain the reaction balance in "point cost form" (see Figure **??**)

$$[\Delta w_{r_l} + b_{v_l}]\, v_l = a_{v_l}\, v_l \tag{8}$$

The term $a_{v_l}\, v_l = \partial a/\partial \ln v_l$ on the right, called *flux point cost* $\underline{a}_{v_l}$, is directly derived from our flux cost function and has the units of fitness (Darwins). To make a connection to enzyme levels, we now consider a flux cost function that effectively describes enzyme cost. To do so, we choose a kinetic model which defines a relationship between enzyme levels, metabolite concentrations, and fluxes. The *enzymatic flux cost* of a flux distribution $\mathbf{v}$ is the minimal enzyme cost at which $\mathbf{v}$ can be realised in this model, and with optimally adjusted enzyme levels (see SI sections S1 and S2.1). With enzyme levels e_l and an enzyme cost function $h(\mathbf{e})$, we obtain the economic reaction balance

$$[\Delta w_{r_l}^{\text{int}} + b_{v_l}^{\star}]\, v_l = h_{e_l}\, e_l, \tag{9}$$

which relates economic potentials, direct flux gain, and fluxes to the *enzyme point costs* $\underline{h}_{e_l} = h_{e_l}\, e_l$ (see Figure 1 a). Thus, in flux cost minimisation with an enzymatic flux cost function $a(\mathbf{v})$, a comparison of Eqs (8) and (9) shows that $\frac{\partial a}{\partial v_l} = \frac{\partial h}{\partial e_l} \frac{e_l}{v_l}$ (proof in SI S3.2). This means that enzymatic flux costs, flux point cost and enzyme point cost are identical: $a_{v_l}\, v_l = h_{e_l}\, e_l$.

3.2 Kinetic models

We have seen that flux optimisation leads to Lagrange multipliers that can be interpreted as economic variables, and that the optimality conditions can be written in the form of economic balance equations. Let us see this for

some other optimality problems. Our next case, which also leads us to the second balance equation (2), concerns kinetic models with optimally chosen enzyme levels. In *enzyme cost-benefit optimisation* (see Figure **??**, Figure 3 (d), and SI section S2.2), we determine enzyme levels that maximise the metabolic fitness

$$f(\mathbf{v}, \mathbf{c}, \mathbf{e}) = b(\mathbf{v}) - q(\mathbf{c}) - h(\mathbf{e}), \tag{10}$$

the difference between flux benefit, metabolite and enzyme cost. If we restrict ourselves to steady states, determined by enzyme levels[3], the optimality problem reads

$$\text{Maximise the fitness } f^{\text{st}}(\mathbf{e}) = b(\mathbf{v}^{\text{st}}(\mathbf{e})) - q(\mathbf{c}^{\text{st}}(\mathbf{e})) - h(\mathbf{e}) \text{ with respect to } \mathbf{e}, \tag{11}$$

where steady-state fluxes $\mathbf{v}^{\text{st}}(\mathbf{e})$ and metabolite levels $\mathbf{c}^{\text{st}}(\mathbf{e})$ are functions of the enzyme levels. These functions, however, are only defined implicitly: to satisfy stationarity, rate laws, and conserved moiety constraints, they must obey the equations

$$\mathbf{N}_{\text{int}}\,\mathbf{v}^{\text{st}}(\mathbf{e}) = 0, \quad \mathbf{v}(\mathbf{c}^{\text{st}}(\mathbf{e}), \mathbf{e}) = \mathbf{v}^{\text{st}}(\mathbf{e}), \quad \mathbf{G}\,\mathbf{c}^{\text{st}}(\mathbf{e}) = \mathbf{c}'_{\text{cm}} \tag{12}$$

with the internal stoichiometric matrix $\mathbf{N}_{\text{int}}$, rate laws $k_l(\cdot)$, moiety conservation matrix $\mathbf{G}$, and the conserved moiety concentrations in a vector $\mathbf{c}'_{\text{cm}}$. To derive our optimality conditions, we do not use the functions $\mathbf{c}^{\text{st}}(\mathbf{e})$ and $\mathbf{v}^{\text{st}}(\mathbf{e})$, but choose an "expanded" formulation in which $\mathbf{e}, \mathbf{c}$, and $\mathbf{v}$ are free variables, and the relations Eq. (12) are used as explicit constraints. We obtain the optimality problem

$$\begin{aligned}
\text{Maximise the fitness } q(\mathbf{c}, \mathbf{v}) - h(\mathbf{e}) \quad &\text{with respect to } \mathbf{v}, \mathbf{c}, \mathbf{e} \\
\text{subject to} \quad &\mathbf{N}_{\text{int}}\,\mathbf{v} = 0 &&\text{Stationarity} \\
&\mathbf{v}(\mathbf{c}, \mathbf{e}) = \mathbf{v} &&\text{Rate laws} \\
&\mathbf{G}\,\mathbf{c} = \mathbf{c}_{\text{cm}} &&\text{Conserved moieties} \\
&\mathbf{c}_{\text{min}} \leq \mathbf{c} \leq \mathbf{c}_{\text{max}} &&\text{Physiological bounds}
\end{aligned} \tag{13}$$

For a full derivation, see SI section S1 and appendix S2.2. As before, the optimality conditions can be written with Lagrange multipliers, and after introducing new variable names, we obtain the simple economic rules

$$\begin{aligned}
\mathbf{w}_{\text{v}} &= \Delta\,\mathbf{w}_{\text{r}}^{\text{int}} + \mathbf{b}_{\text{v}}^{\star} &&\text{Reaction rule} \\
\mathbf{w}_{\text{c}} &= \mathbf{E}_{\text{c}}^{\top}\,\mathbf{w}_{\text{v}} - \mathbf{q}_{\text{c}^{\text{int}}} &&\text{Metabolite rule} \\
\mathbf{f}_{\text{e}} &= \mathbf{E}_{\text{e}}^{\top}\,\mathbf{w}_{\text{v}} - \mathbf{h}_{e} &&\text{Enzyme rule.}
\end{aligned} \tag{14}$$

Each rule links economic variables (the fitness values of fluxes, metabolite levels, or enzyme levels) between neighbouring network elements. The economic variables have different meanings: the flux and metabolite gains $\mathbf{b}_{\text{v}}^{\star}$ and $\mathbf{b}_{\text{c}}$ are derivatives of the benefit function , while the enzyme prices in $\mathbf{h}_{e}$ are derivatives of the enzyme cost function h. The vector $\mathbf{f}_{\text{e}} = \mathbf{g}_{\text{e}} - \mathbf{h}_{e}$ contains the enzyme derivatives of the fitness function. The economic potentials (in $\mathbf{w}_{\text{r}}^{\text{int}}$), the flux demands (in $\mathbf{w}_{\text{v}}$), and the concentration demands (in $\mathbf{w}_{\text{c}}$) are derived from Lagrange multipliers for stationarity constraints, rate law constraints, and moiety conservation constraints, respectively. The concentration value vector is defined as $\mathbf{w}_{\text{c}} = \mathbf{G}^{\top}\,\mathbf{y}_{\text{cm}}$, where $\mathbf{y}_{\text{cm}}$ contains the Lagrange multipliers for the conserved-moiety constraints. If we assume an enzyme-optimal state (where $\mathbf{f}_{\text{e}} = 0$),

---

[3]In some models, there are conserved chemical moieties (e.g. [ATP]+[ADP]=const, if ATP and ADP appear only as a cofactor pair). In such models, one may predefine moiety concentrations and use them as an additional constraint.

the economic rules (14) lead to the equalities

$$
\begin{aligned}
\mathbf{E}_{\mathrm{e}}^{\top}\,\mathbf{w}_{\mathrm{v}} &= \mathbf{h}_e && \text{(*) from enzyme rule} \\
\mathbf{E}_{\mathrm{e}}^{\top}\,(\Delta\,\mathbf{w}_{\mathrm{r}}^{\mathrm{int}} + \mathbf{b}_{\mathrm{v}}^{\star}) &= \mathbf{h}_e && \text{from (*) and reaction rule} \\
\mathbf{E}_{\mathrm{c}}^{\top}\,\mathbf{E}_{\mathrm{e}}^{-1}\,\mathbf{h}_e &= \underbrace{\mathbf{w}_{\mathrm{c}} + \mathbf{q}_{\mathrm{c^{int}}}}_{\mathbf{y}_{\mathrm{c}}^{\mathrm{int}}} && \text{from (*) and metabolite rule,}
\end{aligned}
\tag{15}
$$

which directly yield our balance equations (1) and (2) (for derivations, see SI section S1). The partial enzyme costs $\underline{h}_{e_l}$ are defined as before as $\underline{h}_{e_l} = \frac{\partial h}{\partial e_l}\,e_l$, and $\mathcal{E}_{li} = \frac{c_i}{v_l}\frac{\partial v_l}{\partial c_i}$ is the scaled elasticity between reaction l and compound i. The reaction balance (1) has the same form as the reaction balance Eq. (8) that we obtained previously from flux cost minimisation. If we perform flux cost optimisation with kinetic flux cost functions, then even the numerical values will be identical.

Our third case also concerns kinetic models, but we consider an optimisation of enzyme and metabolite levels at given fluxes (see Figure 3 (e)). Given the desired flow $\mathbf{v}$, each metabolite profile $\mathbf{c}$ implies an enzyme profile $\mathbf{e}$, given by $e_l(v_l, \mathbf{c}) = \frac{v_l}{k_l(\mathbf{c})}$ (and from the rate laws $v_l(\mathbf{c}, \mathbf{e}) = e_l\,k(\mathbf{c})$). Thus, our kinetic model defines an enzyme demand function $\mathbf{e}(\mathbf{v}, \mathbf{c})$, which can be easily computed: using metabolite levels as the free variables, we combine enzyme cost $h(\mathbf{e})$ and metabolite cost $q(\mathbf{c})$ can be combined into an apparent metabolite cost $q^{\mathrm{app}}(\mathbf{c}) = q(\mathbf{c}) + h(\mathbf{e}(\mathbf{v}, \mathbf{c}))$ (where the second term is an "overhead cost"). If metabolite levels can be constrained by physiological bounds and by predefining the concentrations of conserved moieties, we obtain the optimality problem[4]

$$
\begin{aligned}
&\text{Minimise effective metabolite cost } q(\mathbf{c}) + h(\mathbf{e}(\mathbf{v}, \mathbf{c})) && \text{with respect to } \mathbf{c} \\
&\qquad\qquad \text{subject to} \quad \mathbf{G}\,\mathbf{c} = \mathbf{c}_{\mathrm{cm}} && \text{Moiety conservation} \\
&\qquad\qquad\qquad\qquad\quad \mathbf{c}_{\min} \le \mathbf{c} \le \mathbf{c}_{\max} && \text{Physiological bounds}
\end{aligned}
\tag{16}
$$

So again, we expand our set of variables and to write all dependencies as explicit constraints. Using metabolite levels $\mathbf{c}$ and enzyme levels $\mathbf{e}$ as free variables, we obtain the problem

$$
\begin{aligned}
&\text{Minimise metabolite and enzyme costs } q(\mathbf{c}) + h(\mathbf{e}) && \text{with respect to } \mathbf{c}, \mathbf{e} \\
&\qquad\qquad \text{subject to} \quad \mathbf{v}(\mathbf{c}, \mathbf{e}) = \mathbf{v} && \text{Rate laws} \\
&\qquad\qquad\qquad\qquad\quad \mathbf{G}\,\mathbf{c} = \mathbf{c}_{\mathrm{cm}} && \text{Moiety conservation} \\
&\qquad\qquad\qquad\qquad\quad \mathbf{c}_{\min} \le \mathbf{c} \le \mathbf{c}_{\max} && \text{Physiological bounds}
\end{aligned}
\tag{17}
$$

From the optimality conditions with Lagrange multipliers (see appendix S1 and SI section S2.3), we obtain the metabolite balance (2). This time, there is no reaction balance: since all fluxes are predefined, there is no need to enforce stationarity by mass-balance constraints[5].

3.3   Models of growing cells

Once we can model a metabolic pathway, we may try to model an entire cell. To model growing cells, we may first extend our pathway models to non-metabolic processes (e.g., protein production and maintenance). To account for cell growth, we describe its effect on concentrations by hypothetical dilution reactions, which consume each compound at a rate proportional to the compound concentration. Since the same enzyme levels,

---

[4]To optimise the fitness at given fluxes, we just need to find a metabolite profile that minimises this cost.

[5]To obtain a reaction balance equation, we would have to modify our optimality problem. Instead of imposing predefined fluxes, we treat the fluxes as free variables, require stationarity, and add a flux benefit as a side objective. After converting the metabolite costs to metabolite benefits, we reobtain our optimality problem (13) as well as reaction balance and metabolite balance equations (see SI section S2.3).

at a higher dilution rate, will require higher production rates, the enzyme cost function for our metabolic pathway models (where enzyme levels are treated as tunable model parameters) must be dilution-dependent. Finally, we employ constraints that define a viable cell, e.g. a lower bound on ATP level or DNA content (forcing our cell to continuously regenerate ATP and DNA). In the optimality problem, we can treat the dilution rate – i.e., the cell growth rate – as a maximisation objective while requiring a steady, viable cell state. No other optimality criteria are needed, because all other requirements (e.g. the need for metabolic fluxes or enzyme production) will follow, indirectly, from growth and from the viability constraints. Whole-cell models built in this way can either be kinetic and simple (containing only a few variables, and aimed at an optimal choice of kinetic constants that regulate protein expression) [6], or constraint-based and complex (with a direct optimisation of genome-scale fluxes and protein levels) as in resource balance analysis [13].

Thus, between metabolic pathway models and models of growing cells, there are two main differences. First, to model growing cells, we need to consider how dilution affects the metabolic states. If a cell dilutes its compounds (including metabolites, enzymes, proteins, ribosomes, DNA, and any other molecules), they need to be reproduced to keep them at given concentrations. If the model variables are concentrations (not amounts), dilution can be described by hypothetical dilution fluxes proportional to the compound concentrations, which must be balanced by production. Instead of $\mathbf{N}\,\mathbf{v} = 0$, the stationary condition reads $\mathbf{N}\,\mathbf{v} = \lambda\,\mathbf{c}$, describing a balance between protein production and dilution of each compound. The dilution term has an impact on flow patterns, dynamics, and optimality requirements. Second, to formulate a growth maximisation problem, we assume that the cell growth rate λ itself is the objective to be maximised, while all cell variables must stay in physiological ranges.

Surprisingly, growth optimisation problems lead to the same economic balance equations that we encountered before. Let us see this for a pathway model with dilution, where enzyme levels and dilution rate serve as control parameters, and then for a whole-cell model.

In models of growing cells, three different effects of dilution need to be considered: an apparent "consumption" of internal metabolites, described by the right-hand side $\lambda\,\mathbf{c}$ of the stationarity condition; possible changes of metabolite concentrations at the pathway boundary (which depend on adapations outside the pathway, and which we neglect for simplicity); and a dilution of enzymes, which needs to be compensated by higher production rates, and which can be described, effectively, by assuming a growth-rate dependent enzyme cost function. In the resulting optimality problem, we maximise the dilution rate λ for pathway metabolites (or a benefit function depending on λ), and subtract a term for enzyme cost. Maintaining a given protein level at a faster cell growth requires a higher protein production rate. If enzyme cost is determined by the overall protein production, the dilution rate affects enzyme cost. In the model (see Figure 3 (f)), the concentration vector $\mathbf{c}$ contains metabolites only, while the enzyme levels appear as control parameters. Our optimality problem reads

$$
\begin{aligned}
\text{Maximise pathway fitness } b(\lambda) - h(\mathbf{e}, \lambda) \quad &\text{with respect to } \lambda \text{ and } \mathbf{e} \\
\text{subject to} \quad \mathbf{N}_{\text{int}}\,\mathbf{v} - \lambda\,\mathbf{c} = 0 \quad &\text{Mass balance with dilution} \\
\mathbf{v} = \mathbf{v}(\mathbf{c}) \quad &\text{Rate laws} \\
\mathbf{c}_{\min} \leq \mathbf{c} \leq \mathbf{c}_{\max} \quad &\text{Physiological bounds}
\end{aligned}
\tag{18}
$$

Using Lagrange multipliers, the optimality conditions read

$$
\begin{aligned}
[\Delta\,\mathbf{w}_{\text{r}} + \mathbf{b}_{\text{v}}] \circ \mathbf{v} &= \mathbf{h}_e \circ \mathbf{e} \\
h_\lambda + \mathbf{w}_{\text{r}}^{\text{int}} \cdot \mathbf{c} &= b_\lambda.
\end{aligned}
\tag{19}
$$

Aside from the well-known reaction balance (first line), we obtain a new balance equation for the cost h_λ and benefit b_λ of dilution (or, equivalently, of cell growth).

In our next model, we describe a cell as a whole, including macromolecule production and dilution. We treat

enzymes (and other macromolecules) as "metabolites" (acting as catalysts through the rate laws), include their production in the biochemical network, and maximise the growth rate. Since costly effects arise automatically from constraints, i.e. from the fact that enzyme production puts a burden on metabolism, and that an enzyme level can only be increased at the expense of others, no further cost terms needed. In the resulting model (SI section S2.4), we consider a cell during steady growth (corresponding to the exponential phase in experiments), describe all compounds by concentrations , and describe dilution by dilution fluxes. In expanded form (with the cell growth rate λ, concentration vector $\mathbf{c}$, and flux vector $\mathbf{v}$ as free variables, and describing all relationships by explicit constraints), we obtain the optimality problem

$$
\begin{array}{llll}
\text{Maximise growth rate } \lambda & \text{with respect to } \lambda, \mathbf{v} \text{ and } \mathbf{c} & \\
\text{subject to} & \mathbf{N}_{\text{int}}\,\mathbf{v} - \lambda\,\mathbf{c} = 0 & \text{Mass balance with dilution} \\
& \mathbf{v} = \mathbf{v}(\mathbf{c}) & \text{Rate laws} \\
& \mathbf{c}_{\text{min}} \leq \mathbf{c} \leq \mathbf{c}_{\text{max}} & \text{Physiological bounds} & (20)
\end{array}
$$

As a variant of this problem for large, detailed cell models, we consider linear Resource Balance Analysis (RBA) (see SI section S2.7). RBA models have a complicated structure, but if we describe all cellular compounds (including metabolites, macromolecules, and molecular complexes) by a single concentration vector $\mathbf{c}$ and all reactions and processes by a single flux vector $\mathbf{v}$, we can write the problem in a simple general form (see Figure 3 (g) and (h)). Compared to the previous problem, there are two changes. First, all reaction rates are now linearised and described by apparent forward and reverse k^{cat} values in matrices $\mathbf{E}_{\text{for}}$ and $\mathbf{E}_{\text{rev}}$. Second, we do not assume physiological bounds on single compounds, but general density constraints, i.e. bounds on some weighted sums of compound concentrations (describing, e.g., the total volume demand in diferent cell compartments); the weighted sums are defined by a non-negative matrix $\mathbf{D}$. The optimality problem reads:

$$
\begin{array}{llll}
\text{Maximise growth rate } \lambda & \text{with respect to } \lambda, \mathbf{v} \text{ and } \mathbf{c} & \\
\text{subject to} & \mathbf{N}_{\text{int}}\,\mathbf{v} - \lambda\,\mathbf{c} = 0 & \text{Mass-balance constraints} \\
& -\mathbf{E}_{\text{rev}}\,\mathbf{c} \leq \mathbf{v} \leq \mathbf{E}_{\text{for}}\,\mathbf{c} & \text{Capacity constraints} \\
& \mathbf{c}_{\text{min}} \leq \mathbf{D}\,\mathbf{c} \leq \mathbf{c}_{\text{max}} & \text{Density constraints} & (21)
\end{array}
$$

For small metabolites, the term $\lambda\,\mathbf{c}$ in the mass-balance equation can be neglected (as we can tell from their fast turnover times). From the optimality conditions, written with Lagrange multipliers, we obtain the economic balance equations (see Eq. (23) in appendix)

$$
\begin{array}{lcll}
\mathbf{N}_{\text{int}}^{\top}\,\mathbf{w}_{\text{r}}^{\text{int}} & = & \mathbf{a}_{\text{v}} & \text{Reaction balance} \\
\mathbf{E}_{\text{cap}}^{\top}\,\mathbf{a}_{\text{v}} & = & \mathbf{D}^{\top}\,\mathbf{p}_{\text{d}} + \mathbf{w}_{\text{r}}^{\text{int}}\,\lambda & \text{Metabolite balance} \\
\mathbf{c}\cdot\mathbf{w}_{\text{r}}^{\text{int}} & = & 1 & \text{Scaling condition (growth balance).} & (22)
\end{array}
$$

Here, $\mathbf{w}_{\text{r}}^{\text{int}}$, $\mathbf{a}_{\text{v}}$, and $\mathbf{p}_{\text{d}}$ are the Lagrange multipliers for mass-balance constraints, capacity constraints, and density constraints, respectively. The matrix $\mathbf{E}_{\text{cap}}$ contains "active" apparent k^{cat} values (i.e. the forward or reverse apparent k^{cat} values, depending on the flux direction in the optimal state). the signs in $\mathbf{a}_{\text{v}}$ must match the signs of the flux vector (to ensure positive enzyme prices), and the signs in $\mathbf{p}_{\text{d}}$ must be positive (unless some density constraint is inactive; then $p_{\text{d}_{\text{j}}}$ vanishes).

As a variant of RBA, we can also consider our model at a fixed growth rate, and with a linear optimisation

Figure 4: Economic balances in a cell model with growth maximisation. (a) In a schematic model, biomass is produced from glucose via precursor molecules. Metabolic reactions are catalysed by enzymes, which are produced by ribosomes. (b) Economic potentials and flux values, our economic variables, must satisfy two types of relations: (i) Reaction law: in each reaction, the flux value (squares) matches the difference of economic potentials (circles). For instance, in catabolism (dark blue): $w^{\mathrm{v}} = w^{\mathrm{c}}_{\mathrm{Energy}} + w^{\mathrm{c}}_{\mathrm{Precursor}} - w^{\mathrm{c}}_{\mathrm{Glucose}} = 1 + 1 - 0 = 2$. (ii) Compound law: for each compound, the flux values in the adjacent reactions, weighted with the elasticities, must vanish (e.g. ribosome (dark yellow): $w^{v}_{\mathrm{RibSynth}} + w^{v}_{\mathrm{EnzSynth}} + w^{v}_{\mathrm{RibDeg}} = 1 + 2 - 3 = 0$). The numbers follow directly from our choice of reaction elasticities. The economic potential for external glucose ($w_{\mathrm{glucose}} = 0$), the biomass value ($z^{\mathrm{c}}_{\mathrm{biomass}} = 4$), and the cell growth rate λ are fixed and given. From the reaction and metabolite rules, and with elasticities chosen (see SI S3.3), we obtain internal economic potentials $\mathbf{w}^{\mathrm{int}}_{\mathrm{r}} = (1, 1, 4, 3)^{\top}$ and flux values $\mathbf{w}_{\mathrm{v}} = (2, 2, 2, 1)^{\top}$.

objective $F = \mathbf{b}_{\mathrm{v}} \cdot \mathbf{v} - \mathbf{q}_{\mathrm{c}} \cdot \mathbf{c}$. Again, we obtain similar economic balance equations (see appendix S1)

$$
\begin{aligned}
\mathbf{N}^{\top}_{\mathrm{int}} \, \mathbf{w}^{\mathrm{int}}_{\mathrm{r}} + \mathbf{b}_{\mathrm{v}} &= \mathbf{a}_{\mathrm{v}} & \text{Reaction balance} \\
\mathbf{E}_{\mathrm{cap}}{}^{\top} \mathbf{a}_{\mathrm{v}} &= \mathbf{d}^{\mathrm{app}}_{\mathrm{c}} & \text{Metabolite balance}
\end{aligned}
\tag{23}
$$

with the apparent compound price $\mathbf{d}^{\mathrm{app}}_{\mathrm{c}} = \mathbf{q}_{\mathrm{c}} + \mathbf{D}^{\top} \mathbf{p}_{\mathrm{d}} + \lambda \, \mathbf{w}^{\mathrm{int}}_{\mathrm{r}}$.

3.4 Solving the economic balance equations

In some cases, the economic balance equations determine the economic variables uniquely and can be directly solved. An example case is discussed in the appendix, below Eq. (23).

Models of growing cells, as in [6], explain how cells should distribute the protein resources between metabolism and enzyme production in order to maximise growth. We can study such models with metabolic economics: instead of numerically searching for optimal states, we analyse the economic values of different elements (e.g., metabolites, enzymes, and ribosomes) and how they must be related in optimal states. The model illustrates two things: one, that metabolic economics is not restricted to metabolic pathways, but also applies to whole-cell models including protein and ribosome production; and two, how enzyme costs can effectively arise from a need to reproduce proteins, thus consuming metabolic resources. Figure 4 shows the model structure: biomass is produced from external glucose through a catabolic reaction, which produces precursors and energy (high-energy phosphate groups in ATP), and an anabolic reaction that converts precursors into biomass. Enzymes are not treated as control variables as in the models before, but as compounds with a catalytic activity. Enzyme production requires energy, precursors, and ribosomes, and the same holds for ribosome production. Finally, as a cell is growing, all cell compounds are diluted. The metabolic objective scores the biomass concentration at a given dilution rate. This is equivalent to maximising the dilution rate while keeping the biomass concentration above some predefined threshold value. Based on the reaction balance and metabolite balance equations, optimal enzyme patterns can

even be computed by hand (see SI S3.3).

4 Discussion

Economics is typically concerned with opposing needs and limited resources. Applying an economic perspective to cells may improve our understanding of enzyme investments and pathway usage in cells. How should cells allocate limited or costly enzyme resources to metabolic pathways, and how should the usage of pathways depend on costs and efficiencies of pathway enzymes? To answer such questions, we can study optimality problems for fluxes or enzyme levels. If we solve these problems numerically, we observe that certain cycle fluxes are suppressed, that active enzymes have a positive effect on the metabolic metabolic performance, and that these positive effects are balanced with enzyme costs. Metabolic economics explains these observations by general laws and by notions such as economic potentials, economical flows, or futile flux modes. Why would such general laws be useful? Numerical results are obtained on a case-by-case basis, for specific models with specific parameter choices. The laws of metabolic economics, in contrast, apply generally and can provide insights, even if kinetic details or parts of network are unknown. They show that different models of a pathway have something in common (e.g. the "values" of metabolites), which directly determines their metabolic behaviour (e.g. the exclusion of certain futile cycles). They can also show that different optimality-based models of the same system must yield the same general result (e.g. that the difference in metabolic potentials along a pathway, multiplied by the flux, is given by the sum of partial enzyme costs), which can then be tested *even without* a detailed quantitative model. However, all these predictions concern enzyme-optimal states, and to test them empirically, evolution experiments would be needed.

4.1 Concepts of metabolic economics

The main idea of metabolic economics is to replace global fitness objectives by local economic variables, such as economic potentials, which refer to individual metabolites and are governed by local balance equations. After solving optimality problems in abstract flux, enzyme, or metabolite spaces, this brings us right back to the metabolic network (see Figure 5). In a way, metabolic economics closes a circle: to understand the functioning of cells, we start from single biochemical elements, e.g. enzymes and their reactants, and describe their direct interactions, e.g. how enzyme and reactant levels determine reaction rates and how these reaction rates change the metabolite levels. The biochemical entities form pathways and networks, and using dynamic models, we can trace how perturbations propagate through the network and can simulate metabolic dynamics and fluxes in the network. To screen the different states of such systems, and to find optimal states, we need more abstract descriptions: for example, seeing flux distributions as points in an abstract flux space and enzyme cost as a function on this space. Such theoretical frameworks are useful for optimality problems, and they capture global dependencies and compromises well, but they are less intuitive than our original "realistic" picture, which captured biochemical elements and physical interactions occurring between them. It is a dilemma: the benefit of enzymes in cells is not determined by local effects, but by their global effects on the cell state. It is surprising that we can describe this locally. And by doing so, metabolic economics brings us back to a network picture of the cell. Using metabolic economics, we can understand how the optimal global state is reflected in local enzyme investments.

Metabolic pathways are tightly coupled to the rest of the cell – most notably, to macromolecule synthesis, which relies on precursors and energy from metabolism and which produces the enzymes that support metabolic performance. All pathways in a cell need to be orchestrated, and their demands for material and energy need to be balanced to optimise cell fitness (e.g., to maximise cell growth). As a simplification, we focused here on metabolism: certain metabolic states afford a high cell fitness, while others don't, and we assume that we can express this fitness contribution by an effective, metabolic fitness function, describing a pathway of interest.
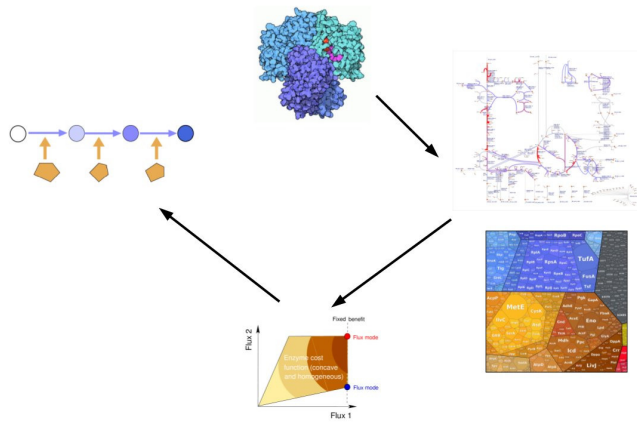
Figure 5: From biochemical molecule interactions to optimality problems and back. The dynamics of biomolecule concentrations (top) can be visualised on cellular networks and be compared to omics data (right). To explain patterns in these data, we analyse them in abstract spaces (e.g. flux space, shown at the bottom). Metabolic economics (left) brings us back to the metabolic network: it represents global fitness demands by local economic values which are assigne to metabolites, enzymes, and reactions and which are interdependent between neighbouring network elements.

This fitness function scores all variables that (potentially) interact with the rest of the cell – namely fluxes (or metabolite production rates), metabolite levels, and enzyme levels. It can be seen as a proxy of the whole-cell fitness, but expressed in the local variables of our pathway. Obviously, such a pathway fitness function cannot be chosen once and forever, because it reflects the cell's demands (e.g., for energy or pathway products), capabilities (e.g., to provide protein resources), and environment (e.g., the availability of nutrients). In metabolic economics, we can apply the same logic to all subsystems of our metabolic system, down to single reactions, metabolites, and enzymes. Given a whole-cell model, we may zoom in on a pathway. For consistency reasons, the pathway model should lead to the same balance equations as the whole-cell model. This is always possible: we simply assume that economic potentials at the pathway boundary are given by the corresponding economic potentials in the whole-cell model with growth optimisation. These potentials define a metabolic pathway objective. If we solve the pathway optimality problem with this objective, we obtain the same metabolic state, and the same internal economic potentials! In this sense, a pathway objective can perfectly mimic, as for as the pathway variables are concerned, the economic requirements posed by maximal cell growth. A global objective, such as fast growth, can thus be used to justify and interpret our pathway objectives.

4.2 Advantages of the metabolic value theory

If metabolic optimality problems can be solved numerically, why do we need a general theory of optimal metabolic states? Why should we interpret the Lagrange multipliers as "economic values", and why should we care about their balances in the network? In the first part of this article [10], I showed that various metabolic optimality problems share the same optimality conditions, but these conditions wer not local – they implied knowledge about the entire system; that is, when modelling a single pathway, they implied knowledge about the (non-modelled) rest of the cell surrounding this pathway . Here, in the second part of the articel, we saw that these conditions can be framed as local balance equations, referring to for individual metabolites and reactions, and linking them with their direct neighbours. This local formulation is a striking result: general economic laws allow us to describe optimal metabolic states as such, "as they are", irrespective of our specific formulation of the optimality problems. Such general concepts can broaden our view of metabolism:

1. When translating our pictures of cells into mathematical optimality problems, we cannot avoid some arbitrariness. For example, a cost-benefit trade-off can be modelled by minimising enzyme cost at a fixed flux, by maximising fluxes at a given enzyme cost, by maximising the flux/enzyme cost ratio, or by employing multi-objective optimality. All these different formulations refer to a single biochemical model and to a single biological trade-off. Accordingly, these problems will lead to the same solutions, provided that the values of the constraints – the fixed flux, or the fixed enzyme cost – are properly chosen. Which formulation of the problem we choose is a matter of convenience only. By focusing on the economic balance equations, we get

rid of this arbitrariness and become independent of the formulation of the optimality problem.

2. Combining models from different modelling paradigms (constraint-based or kinetic models, pathway or whole-cell models) is usually difficult, and transferring results from one model to another one can be difficult as well (e.g. computing a flux distribution by FBA and then realising this flux distribution by an enzyme-optimal state of a kinetic model) . Variables related to fitness, such as flux cost weights, are usually meaningful only within a specific model. To be transferable across models, and to have a, "absolute", model-independent meaning, they must be measured on well-defined, comparable scales and the optimality problems must be compatible too, but this is rarely the case. Constraint-based models, for example, implement the same constraints as kinetic models in steady state and can be seen as a simplification of kinetic models, but in practice they are not derived from kinetic models but built directly from heuristic assumptions. By matching the optimality conditions between kinetic and stoichiometric models, metabolic economics can thus justify constraint-based optimality problems. A similar link can be made between whole-cell models to metabolic models; metabolic fitness functions are usually heuristic (e.g. fitness functions comparing ATP production and the enzyme demand of a pathway, which are measured on different physical scales). If such fitness functions are derived from a common objective, e.g. from the growth rate in a whole-cell problem, their meaning becomes much more transparent.

Aside from new notions, metabolic economics also provides some practical advantages for modelling. Similar to thermodynamics, it adds a new level of description to dynamical cell models, which makes models comparable across modelling paradigms. This can be used in several ways.

1. The economic balance equations can be used as constraints in flux balance analysis, similar to the laws of thermodynamics. By imposing these constraints, we can systematically detect and remove futile flux cycles (which would preclude a consistent choice of economic potentials). The remaining flows are *economical*. Economical flows, and only those, satisfy a principle of minimal linear flux cost [12]. This means: a flux mode can be realised by enzyme-optimal kinetic models *if and only if* this flux mode minimises some weighted sum of fluxes ("Nocturno principle").

2. The fact that metabolic economics applies "locally", to parts of the cell, makes it useful for modular models. Our statements about a single reaction – e.g., the reaction balance – do not depend on the reaction's environment – the local economic variables provide all knowledge we need about the surrounding pathway. Likewise, what we can state about a single pathway does not depend on surrounding (metabolic or non-metabolic) pathways – *given* the economic variables on the pathway boundary, which incorporate all what we need to know about the pathway's environment. This allows us to model a pathway in isolation, but describes its behaviour *as if* the pathway were embedded in a cell, which also behaves optimally. Defining suitable pathway objectives – by simply considering all direct effects of the pathway, plus all indirect effects represented by economic potentials on the pathway boundary – is all that needs to be done.

Metabolic economics allows us to make a number of testable predictions. First, compounds that embody a large cost (e.g. costs of transporters and enzymes embodied in small molecules, or costs of biosynthesis machinery embodied in macromolecules) should be used efficiently. Second, the *enzyme values* in a pathway (e.g. enzyme prices divided by the enzyme efficiencies) should match the differences of economic potentials. This prediction could be experimentally tested on metabolic pathways under strong selection pressures: the economic potentials of metabolites could be measured by finding the break-even point for transporter costs: i.e. the cost at which a biosynthesis pathway and a transporter will be equally cost-efficient. Third, predicted economic imbalances, as predicted selection pressures, could be compared to the measured rates of evolutionary change of enzyme levels, e.g. during evolution in the lab. All these tests would require precise measurements of enzyme costs, and since optimal states are considered, an evolution in the lab prior to any measurements [2, 14]. Enzyme prices could

be measured, for example, by assessing artificial protein expression [2, 3] or the effects of increased protein costs (e.g. by nitrogen depletion or inhibitors of translation). Generally, cell models would greatly profit from precise measurements of enzyme levels, catalytic constants, and enzyme costs at different growth rates.

4.3 Metabolic economics beyond metabolic steady states

Metabolic economics, as described here, can serve as a paradigm for other economic problems in cells. First, metabolic economics can be extended to non-metabolic processes, e.g. macromolecule production, by formally treating compounds as "metabolites". In whole-cell models, for example, any mass-balanced compound, including lipids, proteins, or DNA can have an economic potential defined by the compound's involvement in cell fitness. In such models, no explicit enzyme cost function is needed, because protein cost and ribosome cost arise automatically from the mechanistic model, e.g. through the consumption of energy and precursors, the redirection of fluxes to produce them, and the ensuing constraints on achieving fast growth. The benefit principle – the assumption that the positive (point) costs of an enzyme needs to be justified by a positive (point) benefit – remains valid outside metabolism, e.g. in protein turnover, usage of structural proteins and membranes, or cell signalling. We can still use the same approach: the marginal benefits of compounds, or of other cell quantities, are defined by control coefficients, and economic balance equations are derived from thought experiments about compensated variations. Given a metabolic economics model of an entire cell, we may "zoom in" on individual pathways, accounting for the connections to the surrounding networks.

In this artice, I developed metabolic economics for metabolic steady states. However, metabolic economics can also be extended to other phenomena, including dynamical behaviour in cells (caused by temporally varying enzyme activities), the spatial distributions of compounds (e.g. spatial reaction-diffusion models based on partial differential equations), to the management of uncertainty and variability in cellular variables, and to temporal profiles of signalling substances and the information they carry [15]. Economic balance equations could be applied to models involving other dynamic equations, optimal static or temporal enzyme adaptation, and even optimal behaviour of cells in communities (considering game theory). Other fitness objectives matter for cells in fluctuating environments, for example robustness, optimal dynamic responses to perturbations, and anticipation of predictable or unpredictable changes. To describe this mathematically, we need to consider optimal behaviour under uncertainty or optimal choices of probability distributions that determine stochastic cell behaviour. If cells employ anticipatory behaviour or probabilistic strategies [16, 15], this could explain "non-optimal" preemptive enzyme expression observed in experiments (where enzymes seem to be "overexpressed" and allosterically inhibited until they are suddenly needed), or the production of storage compounds such as starch or glycogen. Testing such predictions experimentally would require evolution experiments in temporally changing environments (e.g. in serial diluation experiments with varying supplies of carbon sources).

Acknowledgements

References

[1] K. Zhuang, G.N. Vemuri, and R. Mahadevan. Economics of membrane occupancy and respiro-fermentation. *MSB*, 7:500, 2011.

[2] E. Dekel and U. Alon. Optimality and evolutionary tuning of the expression level of a protein. *Nature*, 436:588–692, 2005.

[3] I. Shachrai, A. Zaslaver, U. Alon, and E. Dekel. Cost of unneeded proteins in E. coli is reduced after several generations in exponential growth. *Molecular Cell*, 38:1–10, 2010.

[4] M.V. Kritz, M.T. dos Santos, S. Urrutia, and J.-M. Schwartz. Organising metabolic networks: cycles in flux distributions. *J. Theor. Biol.*, 265:250–260, 2010.

[5] S. Schuster, T. Pfeiffer, and D. Fell. Is maximization of molar yield in metabolic networks favoured by evolution? *J. Theor. Biol.*, 252:497–504, 2008.

[6] D. Molenaar, R. van Berlo, D. de Ridder, and B. Teusink. Shifts in growth strategies reflect tradeoffs in cellular economics. *Molecular Systems Biology*, 5:323, 2009.

[7] J.G. Reich. Zur Ökonomie im Proteinhaushalt der lebenden Zelle. *Biomed. Biochim. Acta*, 42(7/8):839–848, 1983.

[8] E. Klipp and R. Heinrich. Competition for enzymes in metabolic pathways: implications for optimal distributions of enzyme concentrations and for the distribution of flux control. *BioSystems*, 54:1–14, 1999.

[9] W. Liebermeister, E. Klipp, S. Schuster, and R. Heinrich. A theory of optimal differential gene expression. *BioSystems*, 76:261–278, 2004.

[10] W. Liebermeister. Metabolic economics: the search for optimal metabolic states. *In preparation*, 2018.

[11] R. Heinrich and E. Klipp. Control analysis of unbranched enzymatic chains in states of maximal activity. *J. Theor. Biol.*, 182(3):243–252, 1996.

[12] H.-G. Holzhütter. The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks. *Eur. J. Biochem.*, 271(14):2905–2922, 2004.

[13] A. Goelzer and V. Fromion. Bacterial growth rate reflects a bottleneck in resource allocation. *Biochim Biophys Acta*, 1810(10):978–988, 2011.

[14] H. Bachmann, M. Fischlechner, I. Rabbers, N. Barfa, F. Branco dos Santos, D. Molenaar, and B. Teusink. Availability of public goods shapes the evolution of competing metabolic strategies. *PNAS*, 110(35):14302–14307, 2013.

[15] O. Rivoire and S. Leibler. The value of information for populations in varying environments. *Journal of Statistical Physics*, 142(6):1124–1166, 2011.

[16] E. Kussell and S. Leibler. Phenotypic diversity, population growth, and information in fluctuating environments. *Science*, 309(5743):2075 – 2078, 2005.