

REVIEW

Open Access



# Engineering microbial chemical factories using metabolic models

Debolina Sarkar and Costas D. Maranas\*

## Abstract

Living organisms in analogy with chemical factories use simple molecules such as sugars to produce a variety of compounds which are necessary for sustaining life and some of which are also commercially valuable. The metabolisms of simple (such as bacteria) and higher organisms (such as plants) alike can be exploited to convert low value inputs into high value outputs. Unlike conventional chemical factories, microbial production chassis are not necessarily tuned for a single product overproduction. Despite the same end goal, metabolic and industrial engineers rely on different techniques for achieving productivity goals. Metabolic engineers cannot affect reaction rates by manipulating pressure and temperature, instead they have at their disposal a range of enzymes and transcriptional and translational processes to optimize accordingly. In this review, we first highlight how various analytical approaches used in metabolic engineering and synthetic biology are related to concepts developed in systems and control engineering. Specifically, how algorithmic concepts derived in operations research can help explain the structure and organization of metabolic networks. Finally, we consider the future directions and challenges faced by the field of metabolic network modeling and the possible contributions of concepts drawn from the classical fields of chemical and control engineering. The aim of the review is to offer a current perspective of metabolic engineering and all that it entails without requiring specialized knowledge of bioinformatics or systems biology.

**Keywords:** Systems biology, Metabolic engineering, Flux balance analysis, Control theory

## Introduction

Chemical engineering draws from a wide range of disciplines such as physics, chemistry, computer science, mathematics, operations research, and the life sciences. Almost seven decades ago during World War II the rising demands for penicillin was met by co-opting chemical engineering principles to carry out large-scale fermentation. A quadruple rise in production was achieved by refining the original mold species, developing an effective growth medium, and changing the fermentation process from rudimentary surface culture to submersion in tanks [1].

These early efforts were the vanguards of current mechanistic descriptions of biological processes. Cellular metabolism is a temporally-varying process that gives rise to a wide variety of dynamical phenomena such as multiple steady states and temporal oscillations. The elucidation, and subsequent prediction of the behavior of

metabolic systems is one of the major challenges of the postgenomic era [2–4]. To this end, significant strides have been made in recent years to construct and investigate detailed models of cellular processes [5–8]. Such a model can be treated as a “virtual laboratory” that allows one to build a characteristic description of the system and elucidate an understanding of the design principles of cellular functions, robustness, adaptability, and optimality. The insights gleaned can then be translated into the rational engineering of microbes to serve as miniature chemical factories to produce products of interest. Microbial fermentation is a popular production mode for many biofuels and biochemicals as it generally (1) relies on a sustainable feedstock (i.e., usually sugars), (2) involves a reduced environmental footprint, (3) is easily scalable, and (4) bypasses the need for complex separations.

The goal of this article is to review how chemical engineering is playing a germane role in the study of metabolic networks. We first describe the use of principles

\* Correspondence: [costas@psu.edu](mailto:costas@psu.edu)

Department of Chemical Engineering, Pennsylvania State University, University Park, PA, USA



such as reaction kinetics and linear programming in modeling metabolic networks. This is followed by a description of the tools used to identify the processes that control and limit flux in biological systems. Finally, we discuss the challenges and opportunities associated with the successful establishment of microbial chemical factories.

### Steady-state analysis of metabolic networks

A cell's metabolism is described by the gamut of biochemical conversions occurring within it that work together to support life. Cells intake carbohydrates, proteins, lipids, and many small molecules and ions. These species, called metabolites, act as building materials and fuel for the cell as it grows, exports and imports materials from its environment, and replicates its genome in order to divide and multiply. A metabolic network can be used to model these reactions using principles derived from chemical kinetics. A basic premise is conservation of mass – if  $X_i(t)$  is the mass of a chemical species  $i$  at time  $t$ , then accordingly

$$X_i(t + \Delta t) - X_i(t) = (v_{i,in} - v_{i,out})\Delta t$$

where  $v_{i,in}$  and  $v_{i,out}$  are the flux rates at which species  $i$  is created and consumed per unit time  $\Delta t$ , respectively. Thus, as  $\Delta t \rightarrow 0$ , the above equation can be written as

$$\frac{dX_i}{dt} = v_{i,in} - v_{i,out}$$

The entire set of metabolic reactions can be represented as a stoichiometric matrix  $\mathbf{S}$  whose entries are the stoichiometric coefficients of every metabolite  $i$  in every reaction  $j$ . A metabolic quasi-steady state is assumed, based on the fact that metabolic reactions are typically much faster compared to the time-scale of cellular growth and environmental changes. Thus, all metabolic fluxes that lead to the production and degradation of metabolites must balance, leading to the flux balance eq. [9, 10]:

$$\mathbf{S}\mathbf{v} = \mathbf{b}$$

where  $\mathbf{v}$  is the vector of metabolic fluxes,  $\mathbf{S}$  the stoichiometric matrix, and  $\mathbf{b}$  is a vector containing the net metabolite uptake/secretion rates. The resulting system is typically under-determined (due to metabolites participating in multiple reactions) and an objective function is introduced as a teleological driver of cellular metabolism. If  $c(\mathbf{v})$  is the objective function (usually maximizing organism growth), the resultant linear programming model is

$$\max\{c(\mathbf{v}) : \mathbf{S}\mathbf{v} = \mathbf{0}, \mathbf{LB} \leq \mathbf{v} \leq \mathbf{UB}\}$$

where  $\mathbf{LB}$  and  $\mathbf{UB}$  are vectors representing the lower

and upper bounds on the reaction fluxes. The above is the most common example of flux balance analysis (FBA) [11]. To construct such a mathematical formulation, two major inputs are required – 1) information of all the metabolic enzymes existing in an organism, as this will inform the possible metabolic conversions, and 2) demands placed on the system (see Fig. 1 for an overview of the reconstruction process). This includes processes such as growth (modelled as flux through a biomass synthesis reaction), maintenance requirements, or secretion of a product of interest.

A genome-scale metabolic model (GSM) includes internal metabolic conversions as well as the reactions transporting metabolites in and out of cells. Thus, reactions can be limited by the resources available in the environment, resulting in a flexible network structure that can adapt to different ecological niches and perturbations. For instance, an *E. coli* cell can survive with or without oxygen, investigations of which using FBA captured the varied routes of energy production and protein biosynthesis employed under the two regimes [12]. The same metabolic model can be used in both cases by simply adjusting the bound of the reaction that transports oxygen into the cell.

FBA has been used to make significant contributions to understanding biochemical networks and metabolic engineering. Its primary goal is to design engineered organisms that can achieve higher efficiencies in metabolite overproduction through alterations in the flux distribution. This has been realized in numerous cases such as lycopene [13] and butanol [14] production in *E. coli*, and lysine production in *Corynebacterium glutamicum* [15]. FBA can also be used to improve productivity such as by optimizing process parameters and rationally designing the cell culture medium. Optimal uptake rates were first computed using FBA and then used to alter the nutrient feed composition in a hybridoma culture which reduced waste accumulation by several folds [16]. A metabolic model of *E. coli* was used to determine culture conditions that increased strain stability by optimizing the glucose to oxygen supply ratio [17].

FBA predictions have been found to achieve biological fidelity -- its ability to identify essential genes (i.e., genes whose deletion negates biomass synthesis) exceeds 90% in many metabolic models [18]. Thus, such analyses can be used to glean insights into an organism's physiology by examining its metabolism quantitatively. For instance, input vs output trade-offs can be computed that describe the impact of nutrient supply rates on by-product secretion and/or growth. Such a study in *E. coli* predicted growth rates to increase with the nutrient supply, as is expected, but at higher growth rates the secretion of metabolites such as acetate was also predicted [19]. This is because the oxygen utilization

capacity is reached at higher growth rates and the cell thus possesses surplus reductive potential. This leads to a redistribution of metabolic fluxes and by-products (such as acetate) are secreted so as to eliminate the surplus redox potential [20–22].

### Dynamic models of metabolism

FBA considers metabolism using only reaction stoichiometry and reversibility subject to a steady-state condition. Thus, such a framework performs well when predicting the long-term response of the system to environment or genetic perturbations [23] but is unable to explain transient behavior or even how a specific metabolite state is achieved [24, 25]. In the absence of any regulatory or mechanistic information, the in silico solution space available to constraint-based models is much larger than the biologically feasible sample space [14, 15]. Thus, attempts to engineer microbial systems as in cell factories to over-produce metabolites must consider the kinetics associated with their production pathways along with the interaction of the designated pathways with the entire metabolic network.

Efforts to construct genome-scale dynamic models have been made by extending constraint-based models -- starting from stationary fluxes and introducing pseudo-kinetic behavior, such as in dynamic FBA (DFBA) [26, 27] and structural kinetic modeling [28, 29]. DFBA works by first discretizing the time period into intervals and then solving the following optimization problem at the beginning of each interval to obtain fluxes at that instant:

$$\begin{aligned} & \max_{\mathbf{v}(t)} \sum w_j v_j(t) \\ \text{s.t.} & \\ & \mathbf{X}(t + \Delta T) \geq 0 \\ & \mathbf{v}(t) \geq 0 \\ & c(\mathbf{v}(t), \mathbf{X}(t)) \leq 0, \forall t \in [t_0, t_f] \\ & |\mathbf{v}(t) - \mathbf{v}(t - \Delta T)| \leq \mathbf{v}_{max} \Delta T, \forall t \in [t_0, t_f] \\ & \mathbf{X}(t + \Delta T) = \mathbf{X}(t) + \mathbf{S}\mathbf{v}\Delta T \\ & X_{biomass}(t + \Delta T) = X_{biomass}(t) + \mu X_{biomass}(t)\Delta T \end{aligned}$$

Where  $\mathbf{X}$  is the vector of metabolite concentrations ( $X_{biomass}$  represents the concentration of the biomass metabolite),  $\mu$  is the growth rate, and  $w_j$  is the vector of weights associated with the objective function in the current time interval  $\Delta T$ .  $t_0$  and  $t_f$  represent the initial and final time points. The non-linearity in this formulation arises from the kinetic expressions for flux rates contained in the vector  $c(\mathbf{v}(t), \mathbf{X}(t))$ , usually used to constrain input fluxes. The flux values so obtained are

considered to be constant over the current time interval and are then used to solve a dynamic model describing metabolite time profiles. Mahadevan et al. [26] used DFBA to describe the biphasic growth of *E.coli* on glucose and acetate. They could successfully predict the onset of acetate production and sequential substrate utilization where *E. coli* preferred glucose followed by acetate instead of simultaneous utilization. Furthermore, they could also qualitatively match the predicted rates of metabolite consumption to those measured experimentally. Since then DFBA has been used to predict the accumulation of high-value storage compounds in microalgae under varying light and nutrient conditions [30], determine optimal aerobic and anaerobic culture times and thus scale-up a batch culture of ethanol production in *S. cerevisiae* by 5-folds [31], and optimize glucose and ethanol production in *E. coli* by calculating the optimal flux profile for reactions controlled by genes under genetic perturbation [32].

Thus, DFBA as an extension of classical FBA can indeed be used to analyze the dynamic reprogramming of a network [33], especially in response to external perturbations. However, it still hinges on the inherent assumption that the time constants related to intracellular dynamics are much smaller than those describing changes in external concentrations. This is not always true for biological systems as they exhibit control on various levels and thus a direct kinetic description of metabolism that incorporates regulatory mechanisms would likely lead to a higher biological fidelity. For instance, analysis of a hybrid kinetic-FBA model of *S. cerevisiae* demonstrated that the inclusion of a relatively small number of enzyme kinetic expressions substantially improves the predictive accuracy of FBA, especially if these are used to describe initial reactions in the metabolism of exogenous substrates and reactions at crucial metabolic branch points [34].

Kinetic models of metabolism take the next step in this direction by using mechanistic enzyme kinetics to model network fluxes which is subject to thermodynamic and regulatory constraints and the underlying network stoichiometry. Kinetic models can thus account for changes in metabolite concentrations while capturing the non-linearities inherently present in the system [35, 36]. A set of ordinary differential equations are used to model the temporal concentration of each metabolite -

$$\frac{d\mathbf{X}}{dt} = \mathbf{S}\mathbf{v}(\mathbf{E}, \mathbf{X}, \mathbf{k}), \mathbf{X}(0) = \mathbf{X}_0$$

Where the reaction flux  $\mathbf{v}$  is a function of metabolite concentrations  $\mathbf{X}$ , kinetic parameters  $\mathbf{k}$ , and enzyme concentrations  $\mathbf{E}$ , and  $\mathbf{X}_0$  represents the initial metabolite concentrations. Because many of the enzyme kinetic

parameters are unknown, approximating kinetic mechanisms is a way to improve the tractability of these models when applied to large networks [37]. These include substitutes such as power-law, lin-log and log-lin kinetics, and modular rate laws (reviews of the various rate laws and their differences can be found in [38, 39]). Due to a lack of experimentally-measured kinetic parameters, many times in vitro kinetic data is used to fit and approximate in vivo enzyme kinetics [40, 41]. The thermodynamic feasibility of kinetic models can be enforced using the generalized Wegscheider condition:  $B^T \log K_{eq} = 0$ , where  $B$  is the right null space of the stoichiometric matrix  $S$  [42, 43]. This enforces thermodynamic feasibility by constraining the values of  $K_{eq}$  for every reaction such that a reaction flux can be non-zero only if the corresponding change in the reaction's Gibbs free energy is negative.

To counter the paucity of data, a top-down modeling approach is used, wherein model parameters are adjusted iteratively using an optimization routine until the model predicted flux distribution matches the one observed. This method uses Monte Carlo modeling to reduce parameter uncertainty by replacing the use of a single set of parameters by multiple parameter sets obtained through random sampling [44–46]. The feasibility of these methods is improved by defining the parameter space on the basis of known in vivo information before sampling commences. Model validation in this approach involves comparing against data from a different physiological state, usually nutrient stress or a mutant strain [47]. There are three primary frameworks in existence for Monte Carlo modeling in genome-scale reconstructions – ORACLE [48–50], Jacobian-based structural kinetic models (SKMs) [51] and Ensemble Modeling [52–55]. The ensemble modeling approach was recently used to construct a genome-scale kinetic model of *E. coli* called k-ecoli457, consisting of 457 reactions, 337 metabolites, and 295 substrate-level regulatory interactions [52]. Model parameterization was done via a genetic algorithm where all available fluxomic data was simultaneously imposed. The k-ecoli457 model [52] was able to capture a wide array of perturbations, with a Pearson correlation coefficient of 0.84 between the experimental data and predicted product yields for 320 engineered strains spanning 24 product metabolites. More recently, the K-FIT decomposition based parameterization approach was introduced [56] which offers orders of magnitude improvement in parameterization times allowing for detailed *a posteriori* local sensitivity analyses. Despite their obvious virtues, constructing detailed kinetic models remains challenging. For instance, the latest *E. coli* constraint-based model contains 2719 reactions involving 1192 metabolites and spans 1515 genes (accounting for ~ 34% of the genome) [57].

The overarching goal is to be able to capture the hierarchical organization seen in biological systems, where the overall phenotype is a function of the inherent cooperativity between layers such as the transcriptome, metabolome, and proteome. Whole-cell computational models are a step in that direction and are able to predict a wide range of cellular behavior by incorporating the function of each gene, gene product, and metabolite [58, 59]. Cellular functions are split into independent modules describing processes such as DNA replication, segregation, and repair, RNA transcription, protein folding, ribosome assembly, and biochemical conversions modelled via metabolic networks. These modules are then integrated and the overall model performance validated against known properties such as organism doubling time, cellular chemical composition, and gene expression. Thus, whole cell models herald a new age of biological discovery driven by in silico modeling, but the data- and computationally-intensive inputs are a major deterrent to their construction and accessibility [60, 61].

### Metabolic control analysis

Reaction properties such as rate and efficiency can be tweaked by altering metabolite concentrations and the catalytic and binding properties of enzymes. Enzymes enable high metabolic flux rates by acting as primary 'levers' governing a cell's metabolism constrained by the thermodynamics and stoichiometry of the system. Drug design was one of the first paradigms under which modification of metabolism was attempted: by identifying and therapeutically targeting 'essential metabolic pathways' in a pathogen. This rational drug design was based on Hans Krebs' proposition that once the exact description of a metabolic pathway has been given, the 'pacemaker enzyme' (i.e., the rate-limiting step) can be identified and this controls flux through the entire pathway.

However, the flux through most biological pathways has been seen to be controlled by more than one enzyme. Glycolysis, which is responsible for breaking down glucose to release energy for cellular processes and provide precursors for downstream metabolic pathways, has been found to be controlled by three enzymes – hexokinase (HK), phosphofructokinase (PFK), and pyruvate kinase (PK) [62]. These enzymes are the slowest in the pathway ( $V_{max}$  values lower than the rest by at least an order of magnitude) and catalyze reactions that operate far away from equilibrium (ratio of mass-action ratio to equilibrium constant  $\sim 10^{-3}$  to  $10^{-4}$ ). Hence, it becomes necessary to quantitatively determine the degree of control that a given enzyme exerts on the flux through the entire pathway and/or the final rate of product formation so as to be able to rationally reengineer metabolism.

Metabolic control analysis (MCA) quantifies the degree of control that a given enzyme in a pathway exerts

on the flux and on the concentration of metabolites. This is done by determining the dependence of system variables (such as the rate of a reaction or the concentration of a metabolite) on enzymatic parameters such as the activity or concentration of an enzyme. Control coefficients  $C_{E_k}^{v_j}$  and  $C_{E_k}^{X_i}$  can be calculated which indicate the steady-state change in the concentration of metabolite  $X_i$  or flux  $v_j$ , respectively, in response to a perturbation in the enzyme concentration or activity ( $E_k$ ). Thus, the flux control coefficient ( $C_{E_k}^{v_j}$ ) and concentration control coefficient ( $C_{E_k}^{X_i}$ ) are defined as:

$$C_{E_k}^{v_j} = \left( \frac{\partial v_j}{\partial E_k} \right) \left( \frac{E_{k,o}}{v_{j,o}} \right)$$

$$C_{E_k}^{X_i} = \left( \frac{\partial X_i}{\partial E_k} \right) \left( \frac{E_{k,o}}{X_{i,o}} \right)$$

Here,  $\left( \frac{E_{k,o}}{v_{j,o}} \right)$  and  $\left( \frac{E_{k,o}}{X_{i,o}} \right)$  are scaling factors used to obtain dimensionless and normalized values of the flux coefficients. They represent the ratio between the initial values from which the slopes  $\left( \frac{\partial v_j}{\partial E_k} \right)$  and  $\left( \frac{\partial X_i}{\partial E_k} \right)$  are calculated. Two summation theorems have been shown to hold for control coefficients [63]:

$$\sum_{k=1}^K C_{E_k}^{v_j} = 1$$

$$\sum_{k=1}^K C_{E_k}^{X_i} = 0$$

If a small change in  $E_k$  causes a significant variation in flux  $v_j$ , then the enzyme exerts high flux control. Similarly, if a minor change in  $v_j$  is observed when  $E_k$  is varied a lot, then the enzyme does not exert significant flux control.

MCA has indeed been used to analyze control in a metabolic pathway and rationally identify potential enzymatic intervention targets, such as in the parasite *Entamoeba histolytica* which causes amebiasis in humans. *E. histolytica* does not have a functional mitochondrion, TCA cycle enzymes, or the oxidative phosphorylation pathway. Thus, glycolysis is the only route of generating ATP, the primary energy 'currency' of the cell, and thus inhibiting this pathway will potentially cripple this parasite. Saavedra et al. [64] constructed a kinetic model of this organism's glycolytic pathway in order to determine its distribution of control. They found that the majority of control is distributed between HK and phosphoglycerate mutase (PGAM). The amoeba model was also used for evaluating the effect of inhibiting individual enzymes and its effect on flux through the entire pathway. Model predictions indicate that in order to reduce glycolytic flux and ATP levels by 50%, HK and PGAM should be inhibited by 24 and 55%, respectively, or both enzymes

by 18%. In contrast, if other enzymes such as phosphofructokinase and pyruvate phosphate dikinase are targeted, their activities have to be inhibited by over 70% to achieve similar glycolytic flux reductions. These predictions were later experimentally verified [65] using reconstitution experiments which expressed glycolytic enzymes in vitro at near physiological conditions of pH, temperature, and enzymatic activity. PGAM was found to be the main flux controlling enzyme for lower glycolysis while upper glycolysis was controlled by HK, and to a much smaller extent, PFK, and ALD.

However, considerations have to be made before applying MCA to biological systems due to its dependence on the reaction flux  $v_j$  being a homogeneous function of  $E_k$ , which is not necessarily always true. The summation law for flux control coefficients can be written as

$$E_1 \left( \frac{\partial v_j}{\partial E_1} \right) + E_2 \left( \frac{\partial v_j}{\partial E_2} \right) + \dots + E_K \left( \frac{\partial v_j}{\partial E_K} \right) = v_j$$

Using Euler's theorem reveals that the above implies  $v_j$  to be homogeneous with degree 1:

$$v_j(tE_1, \dots, tE_K) = tv_j(E_1, \dots, E_K)$$

However, this homogeneity is hard to realize as a change, say, in the concentration of enzyme  $E_k$  will also change components such as the concentration of free substrate which will, in turn, affect the reaction rate  $v_j$ . Only for systems where enzymes are substrate saturated, i.e., where the concentrations of metabolites exceed those of the enzymes by two or three orders of magnitude does the above serve as a fair approximation (such as 70% of the enzymes measured in *E. coli*, yeast, and a mammalian cell line [66]). This is exemplified in the case of Rubisco, the most abundant enzyme in the world, as efforts to engineer it by modifying its catalytic efficiency or by overexpressing it, have been largely unsuccessful. Its concentration was measured to be 3–4 mM in the stroma of spinach chloroplasts [67], whereas the concentration of its substrate ribulose 1,5 bisphosphate is often lower than 0.5 mM. As expected, the kinetic rate law of the corresponding carbon-fixing reaction was found to be non-homogeneous w.r.t. Rubisco [67].

### Thermodynamic analysis to predict pathway bottlenecks

MCA uses a given steady state to describe how changes in enzyme abundance (or activity) affect the pathway flux. This requires the careful calculation of enzyme kinetic properties which are experimentally laborious to measure and also differ between organisms and isozymes, which further confounds the ability to compare across pathways and/or multiple organisms. Using thermodynamics and stoichiometry to calculate a single

metric called Max-min Driving Force (MDF) [68] alleviates these problems. MDF identifies reactions within a pathway whose rates are constrained by a low thermodynamic driving force and thus would need the catalyzing enzymes to be present in higher concentrations and/or be engineered to achieve higher turnovers. Furthermore, MDF can also predict if a pathway is likely to support high flux under cellular conditions, thus allowing one to select among alternative pathways as intervention targets [69, 70].

The rate of a reaction is constrained by its thermodynamic driving force, with near-equilibrium reactions requiring exponentially more enzyme to sustain the same rate as reactions operating far from equilibrium. Due to this interdependence between thermodynamic potential and flux, pathways operating near equilibrium will incur a kinetic penalty due to backward flux. Thus, MDF computes reactant concentrations that maximize the minimum driving force  $B$  associated with all reactions in a pathway:

$$\max_{\mathbf{x}, B} B$$

s.t.

$$-(\mathbf{G}^\circ + RT\mathbf{S}^T \mathbf{x}) \geq B$$

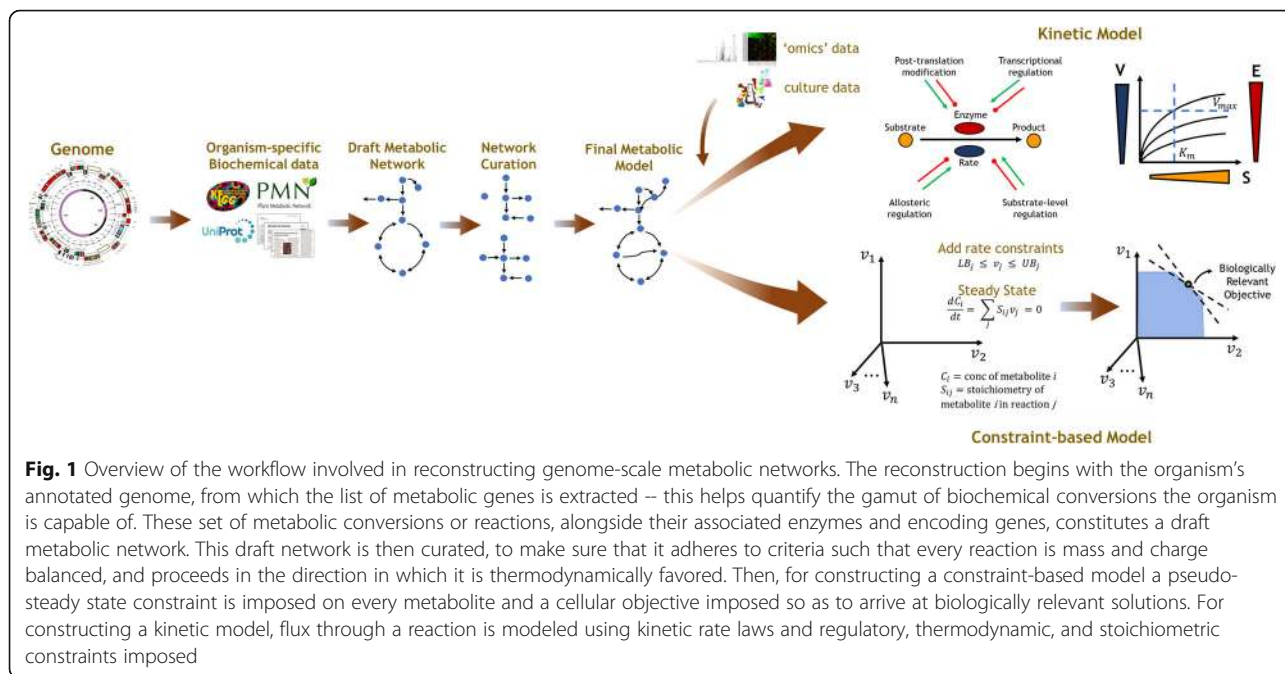
$$\ln(C_{min}) \leq \mathbf{x} \leq \ln(C_{max})$$

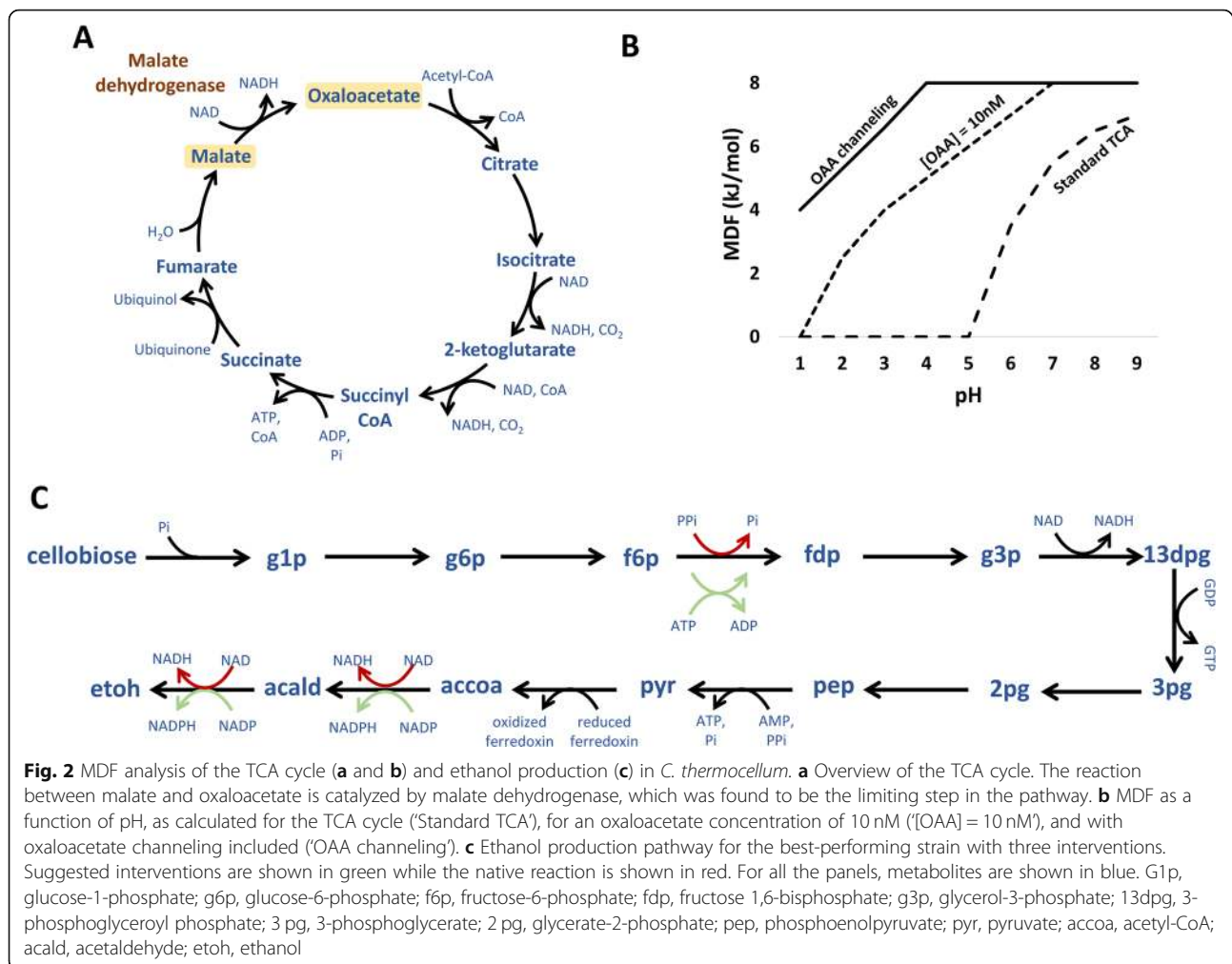
where  $\mathbf{S}$  represents the stoichiometric matrix (rows corresponding to compounds and columns to reactions),  $\mathbf{G}^\circ$  is a column vector containing the standard change in Gibbs energy for a reaction  $j$ ,  $\mathbf{x}$  is a column vector of the

log-concentrations of metabolites,  $C_{max}$  and  $C_{min}$  the maximum and minimum allowed metabolite concentrations, respectively.

MDF has been used to glean biological insights into pathways such as the TCA cycle and EMP glycolysis. An examination of the TCA pathway revealed that flux through it is constrained by malate dehydrogenase, which has a large positive standard Gibbs energy of  $> 30$  kJ/mol at  $\text{pH} \leq 7$  (Fig. 2a and b). However, the TCA cycle is known to sustain high fluxes which can also vary according to the organism's need [71]. This apparent contradiction is explained by the low physiological concentrations of oxaloacetate, the product of malate dehydrogenase, as it is channeled between it and the following enzyme in the pathway -- citrate synthase. Thus, malate dehydrogenase and citrate synthase can be treated as a single enzyme which makes the Gibbs energy drop down to below  $-20$  kJ/mol and removes the thermodynamic bottleneck.

Dash et al. [72] used MDF to examine the thermodynamic bottlenecks associated with ethanol production in *C. thermocellum*. They found five reactions belonging to central carbon metabolism as being limiting under high external ethanol concentrations. They further evaluated the effects of imposing a minimal set of genetic perturbations on pathway thermodynamics and energy production. In doing so they found that modifications involving ATP-linked phosphofructokinase (PFK-ATP) and NADPH linked alcohol dehydrogenase (ADH-NADPH) with NADPH linked aldehyde dehydrogenase (ALDH-NADPH) had the highest performances. The inclusion of ATP-PFK yields a higher MDF at the expense





of ATP, while the ADH-NADPH reaction decouples ethanol production flux from those reactions involving NADH (Fig. 2c). ALDH-NADPH is required for ensuring NADPH production and also ensure redox balance. Interestingly, studies involving high ethanol-yielding *C. thermocellum* strains have shown that the cofactor specificity of ADH changes to NADPH from NADH [73].

### Minimal protein utilization drives cellular metabolism

MDF exploits the fact that the thermodynamic driving force behind a reaction dictates its rate, where higher forces correspond to high forward and low backward fluxes. This translates into efficient enzyme usage by decreasing the amount of enzyme needed per unit of metabolic flux. However, computing enzyme demand from metabolic fluxes is not trivial as enzymes tend to not function at maximal capacity. This is mainly due to metabolites causing incomplete substrate saturation and acting as allosteric regulators (affecting enzyme turnover

by binding to sites other than the active site). This becomes a cyclic inference problem as steady-state metabolite levels depend on enzyme profiles. Thus, in order to arrive at a single solution, one can look for the enzyme profile with the least cost which is needed to realize a certain flux distribution. This is well justified in biological systems, where metabolic enzymes are a limited resource and thus cells economize by synthesizing the right enzymes in the right amounts, and adapting their levels when conditions change.

A reaction rate  $v = Er(c)$  depends on enzyme level  $E$  and metabolite concentrations  $c$  through the enzymatic rate law  $r(c)$ . As metabolite levels are often unknown and also vary between experimental conditions, the enzyme demand cannot simply be computed as  $E = v/r(c)$ . This leads to the definition of an enzyme cost function and choosing the enzyme profile with the lowest cost, while imposing thermodynamic constraints and constraining the metabolite levels to physiological ranges. Using Michaelis-Menten kinetics, a reversible rate law can be written as

$$v = Ek_{cat}^+ \frac{s/K_s \left(1 - \frac{k_{cat}^- p/K_p}{k_{cat}^+ s/K_s}\right)}{1 + s/K_s + p/K_p}$$

$$v = Ek_{cat}^+ \left(1 - \frac{k_{cat}^- p/K_p}{k_{cat}^+ s/K_s}\right) \left(\frac{s/K_s}{1 + s/K_s + p/K_p}\right)$$

$$v = Ek_{cat}^+ \eta^{rev}(\mathbf{c}) \eta^{kin}(\mathbf{c})$$

Where  $E$  is the enzyme level,  $k_{cat}^+$  is the forward catalytic constant,  $\eta^{rev}$  is the driving force (defined as the ratio between forward and backward reaction fluxes), and  $\eta^{kin}$  is the reduction in flux due to kinetic effects (such as substrate saturation or allosteric regulation). Thus, the enzyme demand of a single reaction  $j$  can be written as:

$$E_j(\mathbf{c}, v_j) = \frac{v_j}{k_{cat}^+ \eta^{rev}(\mathbf{c}) \eta^{kin}(\mathbf{c})}$$

A burden  $h_{E_j}$  can be defined for each enzyme representing its molecular mass, post-translation modifications, or effects of misfolding and non-specific catalysis. To determine the demand for an entire pathway, all the reactions are summed over and the final cost function to be minimized is –

$$q(\mathbf{x}, \mathbf{v}) = \sum_j h_{E_j} E_j(\mathbf{c}, v_j) = \sum_j h_{E_j} \frac{v_j}{k_{cat}^+ \eta^{rev}(\mathbf{c}) \eta^{kin}(\mathbf{c})}$$

This function  $q(\mathbf{x}, \mathbf{v})$  represents the trade-off between the fluxes which can be realized and the enzyme levels required to sustain that. Noor et al. [74] used enzyme-cost minimization (ECM) to predict enzyme levels and metabolite concentrations in *E. coli* using fluxes found by 13-C MFA [75]. They found that the prediction fidelity increases monotonically as more complex cost functions are used. The root-mean square error varied from 1.35 (when enzyme levels are considered to be proportional to reaction fluxes) to 0.42 (when the modular rate laws [76] are used and the form of  $\eta^{kin}(\mathbf{c})$  is determined using the reaction mechanism and order of enzyme-substrate binding). However, the caveat of ECM is the a priori knowledge of reaction fluxes, which is difficult to realize on a genome-scale. Although it is true that metabolic states with a maximal specific rate constitute an elementary flux mode [77], but their enumeration is computationally intensive [78]. Furthermore, ECM relies on the assumption that the metabolic states of a cell are optimized for enzyme levels, which isn't always true. Cells often function at sub-optimal levels for robustness or maintaining the metabolic flexibility needed for tidying over future perturbations [79].

## Summary and perspectives

Metabolic engineering has been used for the analysis, design, and optimization of metabolic pathways with significant successes [13, 14, 80–82]. In this review we discussed metabolic engineering tools (using flux balance analysis) that enable the formulation of a cell's metabolism as a resource allocation problem driven by biological objectives such as maximizing growth rate or energy production. The construction of genome-scale models of metabolism require, as input, the set of all known metabolic conversions (or reactions) occurring inside the organism, and the thermodynamic favorability of each. Although such constraint-based models of metabolism have found wide use and adaptation, their primary drawback is an inability to capture the dynamic behavior evinced by biological systems. To this end, conventional FBA has been augmented such as by incorporating pseudo-kinetic reaction descriptions for a subset of reaction fluxes (dynamic FBA). Kinetic models take the next step in this direction by modeling reaction fluxes as a function of metabolic concentrations, enzyme kinetic parameters, and enzyme levels themselves. Such models are able to predict the dynamic behavior of metabolic networks but at the expense of intensive data-driven or computationally expensive parameterization. Nevertheless, a kinetic description of reaction mechanisms can be used to identify major flux-controlling steps [83] and identify pathway bottlenecks (MCA and MDF). Modeling regulations at different levels of metabolism, such as enzymatic or gene expression regulation, draws heavily from control theoretic approaches and can be further extended using classical concepts such as proportional and integral control. This will enable the study of cellular processes such as robust adaptation to environmental perturbations within the well-established fields of control systems for both steady and transient states.

In nature, organisms rarely exist in isolation but interact with others in a variety of biological and ecological niches. Microbial modeling allows us to explore the co-product potential of such communities by modeling the dynamics of inter-species interactions. Microbes can interact with each other and their hosts via processes such as metabolite cross-feeding, which can link disparate pathways from individual species to give rise to novel emergent metabolic functions [84]. By intelligent design of the growth media [85], standalone growth can be negated and thus co-culture growth and product secretion can be made to be an obligatory outcome of microbial biomass synthesis. The composition of a synthetic consortia can be further tuned by using genome-scale metabolic models to scan potential members and subsequently determine the ability of the culture to synthesize desired compounds.



## Conclusions

Thus, a thorough and mechanistic understanding of an organism's cellular processes would revolutionize our abilities to repair or even guide metabolism. Synthetic biology offers the promise of replacing traditional, high carbon footprint processes fed by unsustainable feedstocks with tunable microbial reactors. Using rational approaches derived from metabolic engineering, designing clean processes that use renewable feedstocks as the raw material can also help provide feasible solutions to current problems of global warming and fossil fuel exhaustion. Indeed, a number of instances where metabolic engineering has helped sustainably improve the economy and efficiency of production processes are already available. Engineered bacteria being used to produce energy from sunlight, water, and organic wastes; synthetic molecules produced by biocatalysts being used as novel drugs and vaccines; and increasing the productivity of existing crop systems by implementing an optimal set of genetic interventions – these are but a few of the possible applications of metabolic engineering [30, 86, 87].

## Abbreviations

ADH: Alcohol Dehydrogenase; ALDH: Aldehyde Dehydrogenase; DFBA: Dynamic Flux Balance Analysis; ECM: Enzyme-Cost Minimization; FBA: Flux Balance Analysis; GSM: Genome-Scale Model; HK: Hexokinase; MCA: Metabolic Control Analysis; MDF: Max-min Driving Force; MFA: Metabolic Flux Analysis; PFK: Phosphofructokinase; PGAM: Phosphoglycerate Mutase; PK: Pyruvate Kinase

## Acknowledgements

Not applicable.

## Authors' contributions

C.D.M. and D.S. generated the ideas for the manuscript, and wrote the manuscript, and edited it. Both authors reviewed and approved the final manuscript.

## Funding

This work is funded by the Center for Bioenergy Innovation (CBI) (DE-AC05-000R22725) and NSF (DE-AC05-000R22725). This funding source had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

## Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

Received: 22 June 2019 Accepted: 30 September 2019

Published online: 01 November 2019

## References

- Quinn R. Rethinking antibiotic research and development: world war II and the penicillin collaborative. *Am J Public Health*. 2013;103(3):426–34. <https://doi.org/10.2105/AJPH.2012.300693>

- Fernie AR, Trethewey RN, Krotzky AJ, Willmitzer L. Metabolite profiling: from diagnostics to systems biology. *Nat Rev Mol Cell Biol*. 2004. <https://doi.org/10.1038/nrm1451>
- Kell DB. Metabolomics and systems biology: making sense of the soup. *Curr Opin Microbiol*. 2004;7(3):296–307. <https://doi.org/10.1016/j.mib.2004.04.012>
- Palsson B. The challenges of in silico biology. *Nat Biotechnol*. 2000;18:1147–50. <https://doi.org/10.1038/81125>
- Saha R, Liu D, Hoynes-O'Connor A, Liberton M, Yu J, Bhattacharyya-Pakrasi M, et al. Diurnal Regulation of Cellular Processes in the Cyanobacterium *Synechocystis* sp. Strain PCC 6803: Insights from Transcriptomic, Fluxomic, and Physiological Analyses. *MBio*. 2016;7(3):e00464–16. <https://doi.org/10.1128/mBio.00464-16>
- Chowdhury R, Chowdhury A, Maranas C. Using Gene Essentiality and Synthetic Lethality Information to Correct Yeast and CHO Cell Genome-Scale Models. *Metabolites*. 2015;5(4):536–70. <https://doi.org/10.3390/metabo5040536>
- Simons M, Saha R, Amior N, Kumar A, Guillard L, Clément G, et al. Assessing the Metabolic Impact of Nitrogen Availability Using a Compartmentalized Maize Leaf Genome-Scale Model. *PLANT Physiol*. 2014;166(3):1659–74. <https://doi.org/10.1104/pp.114.245787>
- Sarkar D, Mueller TJ, Liu D, Pakrasi HB, Maranas CD. A diurnal flux balance model of *Synechocystis* sp. PCC 6803 metabolism. *PLoS Comput Biol*. 2019. <https://doi.org/10.1371/journal.pcbi.1006692>
- Fell DA, Small JR. Fat synthesis in adipose tissue. An examination of stoichiometric constraints. *Biochem J*. 1986.
- Savinell JM, Palsson BO. Network analysis of intermediary metabolism using linear optimization. I. Development of mathematical formalism. *J Theor Biol*. 1992;154:421–54. [https://doi.org/10.1016/S0022-5193\(05\)80161-4](https://doi.org/10.1016/S0022-5193(05)80161-4)
- Orth JD, Thiele I, Palsson BØ. What is flux balance analysis? *Nat Biotechnol*. 2010;28(3):245–8. <https://doi.org/10.1038/nbt.1614>
- Edwards JS, Palsson BO. Metabolic flux balance analysis and the in silico analysis of *Escherichia coli* K-12 gene deletions. *BMC Bioinformatics*. 2000. <https://doi.org/10.1186/1471-2105-1-1>
- Alper H, Jin YS, Moxley JF, Stephanopoulos G. Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metab Eng*. 2005;7:155–64. <https://doi.org/10.1016/j.jymben.2004.12.003>
- Atsumi S, Cann AF, Connor MR, Shen CR, Smith KM, Brynildsen MP, et al. Metabolic engineering of *Escherichia coli* for 1-butanol production. *Metab Eng*. 2008. <https://doi.org/10.1016/j.jymben.2007.08.003>
- Vallino JJ, Stephanopoulos G. Metabolic flux distributions in *Corynebacterium glutamicum* during growth and lysine overproduction. *Biotechnol Bioeng*. 2000;41:633–46. [https://doi.org/10.1002/\(SICI\)1097-0290\(20000320\)67:6<872::AID-BIT21>3.0.CO;2-X](https://doi.org/10.1002/(SICI)1097-0290(20000320)67:6<872::AID-BIT21>3.0.CO;2-X)
- Xie L, Wang DIC. Fed-batch cultivation of animal cells using different medium design concepts and feeding strategies. *Biotechnol Bioeng*. 2006;95:270–84. <https://doi.org/10.1002/bit.21160>
- Varma A, Palsson BO. Predictions for oxygen supply control to enhance population stability of engineered production strains. *Biotechnol Bioeng*. 1994;43:275–85. <https://doi.org/10.1002/bit.260430403>
- Fong SS, Palsson B. Metabolic gene-deletion strains of *Escherichia coli* evolve to computationally predicted growth phenotypes. *Nat Genet*. 2004;36:1056–8. <https://doi.org/10.1038/ng1432>
- Varma A, Boesch BW, Palsson BO. Stoichiometric interpretation of *Escherichia coli* glucose catabolism under various oxygenation rates. *Appl Environ Microbiol*. 1993.
- Bajpai R. Control of bacterial fermentations. *Ann N Y Acad Sci*. 1987;506:446–58. <https://doi.org/10.1111/j.1749-6632.1987.tb23840.x>
- Majewski RA, Domach MM. Simple constrained-optimization view of acetate overflow in *E. coli*. *Biotechnol Bioeng*. 1990;35:732–8. <https://doi.org/10.1002/bit.260350711>
- Varma A, Palsson BO. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Appl Environ Microbiol*. 1994.
- Edwards JS, Ibarra RU, Palsson BO. In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat Biotechnol*. 2001;19:125–30. <https://doi.org/10.1038/84379>
- Machado D, Costa RS, Ferreira EC, Rocha I, Tidor B. Exploring the gap between dynamic and constraint-based models of metabolism. *Metab Eng*. 2012;14:112–9. <https://doi.org/10.1016/j.jymben.2012.01.003>
- Smallbone K, Simeonidis E, Broomhead DS, Kell DB. Something from nothing - Bridging the gap between constraint-based and kinetic

- modelling. *FEBS J.* 2007;274:5576–85. <https://doi.org/10.1111/j.1742-4658.2007.06076.x>
26. Mahadevan R, Edwards JS, Doyle FJ 3rd. Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys J.* 2002;83:1331–40. [https://doi.org/10.1016/S0006-3495\(02\)73903-9](https://doi.org/10.1016/S0006-3495(02)73903-9)
27. Vargas FA, Pizarro F, Pérez-Correa JR, Agosin E. Expanding a dynamic flux balance model of yeast fermentation to genome-scale. *BMC Syst Biol.* 2011; 5:75. <https://doi.org/10.1186/1752-0509-5-75>
28. Steuer R, Gross T, Selbig J, Blasius B. Structural kinetic modeling of metabolic networks. *Proc Natl Acad Sci.* 2006;103:11868–73. <https://doi.org/10.1073/pnas.0600013103>
29. Grimbs S, Selbig J, Bulik S, Holzhütter HG, Steuer R. The stability and robustness of metabolic states: identifying stabilizing sites in metabolic networks. *Mol Syst Biol.* 2007;3:146. <https://doi.org/10.1038/msb4100186>
30. Flassig RJ, Fachet M, Höffner K, Barton PI, Sundmacher K. Dynamic flux balance modeling to increase the production of high-value compounds in green microalgae. *Biotechnol Biofuels.* 2016;9:165. <https://doi.org/10.1186/s13068-016-0556-4>
31. Hanly TJ, Tiernan AR, Henson MA. Validation and optimization of a yeast dynamic flux balance model using a parallel bioreactor system. In: *IFAC Proceedings Volumes (IFAC-PapersOnline)*; 2013, p. 113–8. <https://doi.org/10.3182/20131216-3-IN-2044.00002>
32. Gadkar KG, Doyle FJ, Edwards JS, Mahadevan R. Estimating optimal profiles of genetic alterations using constraint-based models. *Biotechnol Bioeng.* 2005;89:243–51. <https://doi.org/10.1002/bit.20349>
33. Serrano-Bermúdez L, Barrios AFG, Montoya D. Clostridium butyricum population balance model: predicting dynamic metabolic flux distributions using an objective function related to extracellular glycerol content. *PLoS One.* 2018;13. <https://doi.org/10.1371/journal.pone.0209447>
34. Pozo C, Miró A, Guillén-Gosálbez G, Sorribas A, Alves R, Jiménez L. Global optimization of hybrid kinetic/FBA models via outer-approximation. *Comput Chem Eng.* 2015;72:325–33. <https://doi.org/10.1016/j.compchemeng.2014.06.011>
35. Voit EO. Design principles and operating principles: the yin and yang of optimal functioning. *Math Biosci.* 2003;182:81–92. [https://doi.org/10.1016/S0025-5564\(02\)00162-1](https://doi.org/10.1016/S0025-5564(02)00162-1)
36. Chowdhury A, Khodayari A, Maranas CD. Improving prediction fidelity of cellular metabolism with kinetic descriptions. *Curr Opin Biotechnol.* 2015;36: 57–64. <https://doi.org/10.1016/j.copbio.2015.08.011>
37. Costa RS, Machado D, Rocha I, Ferreira EC. Hybrid dynamic modeling of *Escherichia coli* central metabolic network combining Michaelis–Menten and approximate kinetic equations. *Biosystems.* 2010;100(2):150–7. <https://doi.org/10.1016/j.biosystems.2010.03.001>
38. Tummiler K, Lubitz T, Schelker M, Klipp E. New types of experimental data shape the use of enzyme kinetics for dynamic network modeling. *FEBS Journal.* 2014;281:549–71. <https://doi.org/10.1111/febs.12525>
39. Heijnen JJ. Approximate kinetic formats used in metabolic network modeling. *Biotechnol Bioeng.* 2005;91:534–45. <https://doi.org/10.1002/bit.20558>
40. Beard DA. Simulation of cellular biochemical system kinetics. *Wiley Interdiscip Rev Syst Biol Med.* 2011;3:136–46. <https://doi.org/10.1002/wsbm.116>
41. Teusink B, Passarge J, Reijenga CA, Esgalhado E, Van Der Weijden CC, Schepper M, et al. Can yeast glycolysis be understood terms of vitro kinetics of the constituent enzymes? Testing biochemistry. *Eur J Biochem.* 2000;267: 5313–29. <https://doi.org/10.1046/j.1432-1327.2000.01527.x>
42. Liebermeister W, Klipp E. Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theor Biol Med Model.* 2006;3:41. <https://doi.org/10.1186/1742-4682-3-41>
43. Stanford NJ, Lubitz T, Smallbone K, Klipp E, Mendes P, Liebermeister W. Systematic Construction of Kinetic Models from Genome-Scale Metabolic Networks. *PLoS One.* 2013;8(11):e79195. <https://doi.org/10.1371/journal.pone.0079195>
44. Murabito E, Smallbone K, Swinton J, Westerhoff HV, Steuer R. A probabilistic approach to identify putative drug targets in biochemical networks. *J R Soc Interface.* 2011;8:880–95. <https://doi.org/10.1098/rsif.2010.0540>
45. Mišković L, Hatzimanikatis V. Modeling of uncertainties in biochemical reactions. *Biotechnol Bioeng.* 2011;108:413–23. <https://doi.org/10.1002/bit.22932>
46. Janasch M, Asplund-Samuelsson J, Steuer R, Hudson EP. Kinetic modeling of the Calvin cycle identifies flux control and stable metabolomes in *Synechocystis carbon* fixation. *J Exp Bot.* 2019;70(3):973–83. <https://doi.org/10.1093/jxb/ery382>
47. Rohwer JM. Kinetic modelling of plant metabolic pathways. *J Exp Bot.* 2012; 63(6):2275–92. <https://doi.org/10.1093/jxb/ers080>
48. Wang L, Birol I, Hatzimanikatis V. Metabolic Control Analysis under Uncertainty: Framework Development and Case Studies. *Biophys J.* 2004; 87(6):3750–63. <https://doi.org/10.1529/biophysj.104.048090>
49. Hameri T, Fengos G, Ataman M, Miskovic L, Hatzimanikatis V. Kinetic models of metabolism that consider alternative steady-state solutions of intracellular fluxes and concentrations. *Metab Eng.* 2019;52:29–41. <https://doi.org/10.1016/j.ymben.2018.10.005>
50. Chakrabarti A, Miskovic L, Soh KC, Hatzimanikatis V. Towards kinetic modeling of genome-scale metabolic networks without sacrificing stoichiometric, thermodynamic and physiological constraints. *Biotechnol J.* 2013;8(9):1043–57. <https://doi.org/10.1002/biot.201300091>
51. Steuer R, Gross T, Selbig J, Blasius B. Structural kinetic modeling of metabolic networks. *Proc Natl Acad Sci U S A.* 2006;103(32):11868–73. <https://doi.org/10.1073/pnas.0600013103>
52. Khodayari A, Maranas CD. A genome-scale *Escherichia coli* kinetic metabolic model k-ecoli457 satisfying flux data for multiple mutant strains. *Nat Commun.* 2016.
53. Tan Y, Lafontaine Rivera JG, Contador CA, Asenjo JA, Liao JC. Reducing the allowable kinetic space by constructing ensemble of dynamic models with the same steady-state flux. *Metab Eng.* 2011;13(1):60–75. <https://doi.org/10.1016/j.ymben.2010.11.001>
54. Zomorodi AR, Lafontaine Rivera JG, Liao JC, Maranas CD. Optimization-driven identification of genetic perturbations accelerates the convergence of model parameters in ensemble modeling of metabolic networks. *Biotechnol J.* 2013;8(9):1090–104. <https://doi.org/10.1002/biot.201200270>
55. Tran LM, Rizk ML, Liao JC. Ensemble Modeling of Metabolic Networks. *Biophys J.* 2008;95(12):5606–17. <https://doi.org/10.1529/biophysj.108.135442>
56. Gopalakrishnan S, Dash S, Maranas C. K-FIT: An accelerated kinetic parameterization algorithm using steady-state fluxomic data. *bioRxiv.* 2019. <https://doi.org/10.1101/612994>
57. Monk JM, Lloyd CJ, Brunk E, Mih N, Sastry A, King Z, et al. iML1515, a knowledgebase that computes *Escherichia coli* traits. *Nat Biotechnol.* 2017; 35:904–8. <https://doi.org/10.1038/nbt.3956>
58. Atlas JC, Nikolaev EV, Browning ST, Shuler ML. Incorporating genome-wide DNA sequence information into a dynamic whole-cell model of *Escherichia coli*: application to DNA replication. *IET Syst Biol.* 2008, p. 369–82. <https://doi.org/10.1049/iet-syb:20070079>
59. Karr JR, Sanghvi JC, Macklin DN, Gutschow MV, Jacobs JM, Bolival B, et al. A whole-cell computational model predicts phenotype from genotype. *Cell.* 2012. <https://doi.org/10.1016/j.cell.2012.05.044>
60. Szigeti B, Roth YD, Sekar JAP, Goldberg AP, Pochiraju SC, Karr JR. A blueprint for human whole-cell modeling. *Curr Opin Syst Biol.* 2018;7:8–15. <https://doi.org/10.1016/j.coisb.2017.10.005>
61. Goldberg AP, Chew YH, Karr JR. Toward scalable whole-cell modeling of human cells. In: *SIGSIM-PADS 2016 - Proceedings of the 2016 Annual ACM Conference on Principles of Advanced Discrete Simulation*; 2016. <https://doi.org/10.1145/2901378.2901402>
62. Moreno-Sánchez R, Saavedra E, Rodríguez-Enríquez S, Olín-Sandoval V. Metabolic Control Analysis: A tool for designing strategies to manipulate metabolic pathways. *J Biomed Biotechnol.* 2008;1–30. <https://doi.org/10.1155/2008/597913>
63. Kacser H, Burns JA, Kacser H, Fell DA. The control of flux: 21 years on. *Biochem Soc Trans.* 1995. <https://doi.org/10.1042/bst0230341>
64. Saavedra E, Marín-Hernández A, Encalada R, Olivos A, Mendoza-Hernández G, Moreno-Sánchez R. Kinetic modeling can describe in vivo glycolysis in *Entamoeba histolytica*. *FEBS J.* 2007;274:4922–40. <https://doi.org/10.1111/j.1742-4658.2007.06012.x>
65. Saavedra E, Encalada R, Pineda E, Jasso-Chávez R, Moreno-Sánchez R. Glycolysis in *Entamoeba histolytica*: biochemical characterization of recombinant glycolytic enzymes and flux control analysis. *FEBS J.* 2005;272: 1767–83. <https://doi.org/10.1111/j.1742-4658.2005.04610.x>
66. Park JO, Rubin SA, Xu YF, Amador-Noguez D, Fan J, Shlomi T, et al. Metabolite concentrations, fluxes and free energies imply efficient enzyme usage. *Nat Chem Biol.* 2016;12:482–9. <https://doi.org/10.1038/nchembio.2077>
67. Farquhar GD. Models describing the kinetics of ribulose biphosphate carboxylase-oxygenase. *Arch Biochem Biophys.* 1979;193:456–68. [https://doi.org/10.1016/0003-9861\(79\)90052-3](https://doi.org/10.1016/0003-9861(79)90052-3)
68. Noor E, Bar-Even A, Flamholz A, Reznik E, Liebermeister W, Milo R. Pathway thermodynamics highlights kinetic obstacles in central metabolism. *PLoS Comput Biol.* 2014;10:e1003483. <https://doi.org/10.1371/journal.pcbi.1003483>

69. Yang X, Yuan Q, Luo H, Li F, Mao Y, Zhao X, et al. Systematic design and in vitro validation of novel one-carbon assimilation pathways. *Metab Eng.* 2019;56:142–53. <https://doi.org/10.1016/j.ymben.2019.09.001>
70. Jacobson TB, Adamczyk PA, Stevenson DM, Regner M, Ralph J, Reed JL, et al. 2H and 13C metabolic flux analysis elucidates in vivo thermodynamics of the ED pathway in *Zymomonas mobilis*. *Metab Eng.* 2019;54:301–16. <https://doi.org/10.1016/j.ymben.2019.05.006>
71. Heyland J, Fu J, Blank LM. Correlation between TCA cycle flux and glucose uptake rate during respiro-fermentative growth of *Saccharomyces cerevisiae*. *Microbiology.* 2009;155:3827–37. <https://doi.org/10.1099/mic.0.030213-0>
72. Dash S, Olson DG, Joshua SH, Amador-Noguez D, Lynd LR, Maranas CD. Thermodynamic analysis of the pathway for ethanol production from cellobiose in *Clostridium thermocellum*. *Metab Eng.* 2019;55:161–9.
73. Zheng T, Olson DG, Tian L, Bomble YJ, Himmel ME, Lo J, et al. Cofactor specificity of the Bifunctional alcohol and aldehyde dehydrogenase (AdhE) in wild-type and mutant *Clostridium thermocellum* and *Thermoanaerobacterium saccharolyticum*. *J Bacteriol.* 2015;2610–2619:197. <https://doi.org/10.1128/jb.00232-15>
74. Noor E, Flamholz A, Bar-Even A, Davidi D, Milo R, Liebermeister W. The Protein Cost of Metabolic Fluxes: Prediction from Enzymatic Rate Laws and Cost Minimization. *PLoS Comput Biol.* 2016;12:e1005167. <https://doi.org/10.1371/JOURNAL.PCBI.1005167>
75. Haverkorn Van Rijsewijk BRB, Nanchen A, Nallet S, Kleijn RJ, Sauer U. Large-scale 13C-flux analysis reveals distinct transcriptional control of respiratory and fermentative metabolism in *Escherichia coli*. *Mol Syst Biol.* 2011. <https://doi.org/10.1038/msb.2011.9>
76. Liebermeister W, Uhlendorf J, Klipp E. Modular rate laws for enzymatic reactions: thermodynamics, elasticities and implementation. *Bioinformatics.* 2010;26:1528–34. <https://doi.org/10.1093/bioinformatics/btq141>
77. Wortel MT, Peters H, Hulshof J, Teusink B, Bruggeman FJ. Metabolic states with maximal specific rate carry flux through an elementary flux mode. *FEBS J.* 2014; 281:1547–55. <https://doi.org/10.1111/febs.12722>
78. Klamt S, Stelling J. Combinatorial complexity of pathway analysis in metabolic networks. *Mol Biol Rep.* 2002. <https://doi.org/10.1023/A:1020390132244>
79. Schuetz R, Zamboni N, Zampieri M, Heinemann M, Sauer U. Multidimensional optimality of microbial metabolism. *Science.* 2012;336: 601–4. <https://doi.org/10.1126/science.1216882>
80. Yim H, Haselbeck R, Niu W, Pujol-Baxley C, Burgard A, Boldt J, et al. Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nat Chem Biol.* 2011;7:445–52. <https://doi.org/10.1038/nchembio.580>
81. Vu TT, Hill EA, Kucek LA, Konopka AE, Beliaev AS, Reed JL. Computational evaluation of *Synechococcus* sp. PCC 7002 metabolism for chemical production. *Biotechnol J.* 2013;8(5):619–30. <https://doi.org/10.1002/biot.201200315>
82. Suástegui M, Yu Ng C, Chowdhury A, Sun W, Cao M, House E, et al. Multilevel engineering of the upstream module of aromatic amino acid biosynthesis in *Saccharomyces cerevisiae* for high production of polymer and drug precursors. *Metab Eng.* 2017;42:134–44. <https://doi.org/10.1016/j.ymben.2017.06.008>
83. Dai Z, Locasale JW. Thermodynamic constraints on the regulation of metabolic fluxes. *J Biol Chem.* 2018;293(51):19725–39. <https://doi.org/10.1074/jbc.RA118.004372>
84. Chiu HC, Levy R, Borenstein E. Emergent Biosynthetic Capacity in Simple Microbial Communities. *PLoS Comput Biol.* 2014;10:e1003695. <https://doi.org/10.1371/journal.pcbi.1003695>
85. Klitgord N, Segrè D. Environments that induce synthetic microbial ecosystems. *PLoS Comput Biol.* 2010;6:e1001002. <https://doi.org/10.1371/journal.pcbi.1001002>
86. Cases I, De Lorenzo V. Genetically modified organisms for the environment: stories of success and failure and what we have learned from them. In: *International Microbiology*; 2005.
87. Twyman RM, Schillberg S, Fischer R. Transgenic plants in the biopharmaceutical market. *Expert Opin Emerg Drugs.* 2005.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

