The view of microbes as energy converters illustrates the trade-off between growth rate and yield

St. Elmo Wilken[1], Victor Vera Frazão[1], Nima P. Saadat[1], Oliver Ebenhöh[1,2,*]

April 2021

1. Institute of Quantitative and Theoretical Biology, Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany.
2. Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine-Universität Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany.
* Author to whom correspondence should be addressed.

S.E.W.: wilkenst@hhu.de
V.V.F.: viver101@hhu.de
N.P.S.: nima.saadat@uni-duesseldorf.de
O.E.: oliver.ebenhoeh@hhu.de

Perspective

1. Microbial growth is the foundation of many biotechnological applications. The key to optimizing microbial growth lies in thermodynamics, similar to how classical thermodynamics helped optimize steam engines in the 19th century.

2. Genome-scale metabolic models have become widely available, and are used to predict microbial growth. These predictions often fail because these models do not distinguish between growth rate and yield.

3. Classical black box models present a sound thermodynamic theory, by viewing microbes as energy converters. Incorporating such concepts into genome-scale metabolic models has the promise to advance our fundamental understanding of microbial growth, and thus to improve the predictive power of these models.

Abstract

The application of thermodynamics to microbial growth has a long tradition that originated in the middle of the 20th century. This approach reflects the view that self-replication is a thermodynamic process that is not fundamentally different from mechanical thermodynamics. The key distinction is that a free energy gradient is not converted into mechanical (or any other form of) energy, but rather into new biomass. As such, microbes can be viewed as energy converters that convert a part of the energy contained in environmental nutrients into chemical energy that drives self-replication. Before the advent of high-throughput sequencing technologies, only

the most central metabolic pathways were known. However, precise measurement techniques allowed for the quantification of exchanged extracellular nutrients and heat of growing microbes with their environment. These data, together with the absence of knowledge of metabolic details, drove the development of so-called black box models, which only consider the observable interactions of a cell with its environment and neglect all details of how exactly inputs are converted into outputs. Now, genome sequencing and genome-scale metabolic models provide us with unprecedented detail about metabolic processes inside the cell. However, the derived modelling approaches make surprisingly little use of thermodynamic concepts. Here, we review classical black box models and modern approaches that integrate thermodynamics into genome-scale metabolic models. We also illustrate how the description of microbial growth as an energy converter can help to understand and quantify the trade-off between microbial growth rate and yield.

1   Introduction

Thermodynamics has its origins in the optimization of steam engines in the 19$^{\text{th}}$ century. During this era, two of the fundamental laws of thermodynamics were postulated: the first law holds that energy is conserved in an isolated system; and the second law holds that no process is possible that only transfers heat from a lower temperature to a higher temperature[1] [1]. While the resultant theory mostly dealt with systems at equilibrium, in the middle of the 20$^{\text{th}}$ century strides were made to extend it to irreversible, non-equilibrium regimes, allowing for the direct analysis of less idealized systems [2, 3]. Today, thermodynamics is integral to understanding efficiency limitations of natural and engineered systems, and has been widely applied in biochemistry and biotechnology [4, 5, 6, 7, 8].

Efficiency considerations guide researchers in their quest to rationalize the design of living systems in light of evolutionary optimization. The importance of understanding efficiency limitations is further stressed by the fact that the economic competitiveness of bio-processes is often seen as a major hurdle in the development of a sustainable bio-economy [9]. Because many biotechnological processes are based on exploiting microbes as bio-based factories [10, 11, 12], the continued development and application of thermodynamic concepts and tools to study and optimize microbial growth processes is necessary. Heterotrophs obtain energy by metabolizing higher energy nutrients into lower energy components through a process known as catabolism. Since a part of this energy gradient is used to build new biomass (anabolism), heterotrophic growth can be seen as a thermodynamic energy converter [13], where ATP is the central coupling metabolite, which facilitates the conversion of energy released in catabolic processes to drive biomass formation. This thermodynamic view of microbial growth opens the door to a rich array of tools that can be used to analyze these systems [14].

A complete understanding and quantitative description of heterotrophic growth will depend on the knowledge of the detailed mechanisms of how this energy converter functions. Historically, a black box description was used to investigate the overall conversion [15], but this approach did not reveal how the coupling is performed in detail. For example, it is unknown how many molecules of ATP are being produced per carbon consumed in catabolism, and how many are needed to incorporate one carbon into biomass in anabolism. Genome-scale metabolic models (GEMs) reveal a view into the black box by mapping pathways and reactions to specific genes, and are increasingly used in systems biology applications [16], but often do not adequately reflect thermodynamic constraints [17]. It is apparent that a mechanistically sound description of the microbial energy conversion process must obey the laws of thermodynamics. Incorporating concepts from thermodynamic black box models into GEMs therefore seems to be a promising way forward to develop a thermodynamically consistent, and

---

[1]In the words of Clausius who originally formulated the second law, "No process is possible whose sole result is the transfer of heat from a body of lower temperature to a body of higher temperature."

mechanistically detailed, description of metabolism.

In this review, we first discuss the basis for all further thermodynamic calculations used to describe biological systems (Section 2). This basis is given by the concept of Gibbs free energies of reactions. We discuss its importance, its connection to biochemistry, existing methods to calculate it, and how it can be applied to glean insight into single reactions and complete metabolic pathways. We then review the black box approaches used to investigate microbial thermodynamics (Section 3), before we review and discuss more recent techniques that make use of GEMs for this purpose (Section 4). In Section 5 we discuss how these approaches may be connected. Viewing microbial growth as a thermodynamic energy converter provides access to key quantities, such as thermodynamic efficiency. This, in turn, immediately raises fundamental questions regarding the trade-off between growth rate (kinetic efficiency) and yield (stoichiometric efficiency). This trade-off is a recurring theme in this review, as it is encountered on various levels in different models describing microbial growth as a thermodynamic energy conversion process.

2 Gibbs free energies - the universal driving force of reactions and pathways

The laws of thermodynamics can be used to deduce properties of reactions. Typically, chemical reactions take place at constant temperature and pressure, leading to the use of Gibbs free energy, G, as the primary thermodynamic potential of interest. As a consequence of the second law of thermodynamics, the change in Gibbs free energy of a spontaneous reaction, $\Delta_r G$, can only be negative, constraining the direction of net flux for a given reaction. This potential difference can also be used to find the equilibrium conditions of a reaction through the relation $\Delta_r G^\circ = -RT \ln(K)$, where K is the equilibrium reaction quotient, T the thermodynamic temperature, and $\Delta_r G^\circ$ the standard[2] Gibbs free energy change of the reaction. Standard Gibbs free energies can be found in thermodynamic databases [1].

The practical importance of Gibbs free energy has led to the compilation of numerous databases storing thermodynamic data for biologically relevant species and reactions. In particular, two current examples of such databases include the Thermodynamics of Enzyme-Catalyzed Reactions Database [18], and eQuilibrator [19]. The latter database offers the highest coverage of biochemical reactions to date [20], while being thermodynamically consistent [21]. Importantly, these databases deviate from conventional chemical reaction thermodynamics by their use of *apparent* equilibrium constants, K', as well as *apparent* standard Gibbs free energies of formation, $\Delta_f G'^\circ$, and reaction, $\Delta_r G'^\circ$, to describe the thermodynamic potential of biochemical reactions. In essence, the qualifier "apparent" denotes that pH and ionic strength are held constant when calculating thermodynamic potentials. This deviation is necessitated by practical challenges associated with biochemistry, e.g. measurements of enzymatic reactions are typically made in buffered solutions where H^+ is not conserved, and protonation states of metabolites are difficult to measure independently [22]. The most important consequence of this is that chemical and biochemical thermodynamic databases are not directly compatible with each other, however, the fundamentally important relations between G and properties of interest remain unchanged, as shown in equations (1) – (3), where Q' is the apparent reaction quotient, ν_i is the stoichiometric coefficient associated with species i in a reaction with N' different chemical species (excluding protons).

$$\Delta_r G'^\circ = \sum_i^{N'} \nu_i \Delta_f G_i'^\circ = -RT \ln(K') \tag{1}$$

---

[2]Concentrations of all species 1 M, temperature 298.15 K, and pressure 1 bar.

$$\Delta_{\mathrm{r}} G' = \Delta_{\mathrm{r}} G'^{\circ} + RT \ln(Q') = RT \ln \frac{Q'}{K'} \tag{2}$$

$$\text{spontaneous reaction} \iff \Delta_{\mathrm{r}} G' \leq 0 \text{ (with equality at equilibrium)} \tag{3}$$

Gibbs free energy can be used to identify energetic bottlenecks in native and engineered microbial metabolism. However, the use of thermodynamic considerations in flux analysis has not been without controversy. The primary point of contention is the relationship between flux (or rate) and thermodynamic driving force. The sequence of correspondence [23, 24, 25, 26, 27] highlights the common misconception that flux is independent of the thermodynamic driving force. In contrast, under certain circumstances the flux, J, through an enzymatic reaction *is* proportional to the associated thermodynamic driving force $(-\Delta_{\mathrm{r}} G)$. The flux through a reaction can be decomposed into kinetic (F_{K}) and thermodynamic (F_{T}) factors, i.e. $J = F_{\mathrm{K}} F_{\mathrm{T}}$ [28, 29, 30, 31, 32, 33]. The thermodynamic factor can be written in the form $F_{\mathrm{T}} = 1 - \mathrm{e}^{\frac{\Delta_{\mathrm{r}} G'}{RT}}$ [34, 33]. If the thermodynamic driving force of a reaction is large $(\Delta_{\mathrm{r}} G \ll 0)$ the thermodynamic component approaches unity, approximately reducing the rate expression to the kinetic component only. It should be noted that even in this case the kinetic parameters of an enzymatic reaction are constrained by thermodynamics through the Haldane relation [35]. However, in the opposite case, $F_T \approx -\Delta_{\mathrm{r}} G$ and the flux becomes proportional to the driving force. This is often observed in anaerobic environments, where the thermodynamic component plays a much larger role in constraining microbial growth [36].

Beyond describing the favourability of a single reaction, thermodynamic considerations can also be used to interrogate the interplay between rate and yield in microbial metabolism at the pathway level [37, 38]. Here, yield describes the fraction of the available free energy gradient that is used for useful chemical work, such as the production of ATP. As a consequence, a high yield means a small overall $-\Delta G'$, since the energy gradient is partially conserved in ATP. In light of the foregoing, rate and yield represent a trade-off: the more efficient a pathway is in conserving energy, the lower the thermodynamic driving force becomes, resulting in a lower flux. This trade-off is the essential idea behind max-min driving force (MDF) analysis, which can be used to identify enzymes that act as thermodynamic bottlenecks in pathways. An enzyme operating close to equilibrium catalyses almost as many conversions in the reverse as in the forward direction. To overcome this inefficiency, such reactions require a high investment of cellular resources [39]. It is important to note that the concept of thermodynamic bottlenecks is quite different from that of kinetic bottlenecks (or rate-limiting steps), which exert a large control on the overall flux of the pathway [40]. MDF analysis was used to show that the Entner-Doudoroff (ED) pathway is thermodynamically more favourable than the Embden-Meyerhof-Parnas (EMP) pathway. Consequently, while the latter yields more ATP per glucose it is also more costly because it requires more enzyme. Thus, it was found that prokaryotes relying primarily on glycolysis for energy (e.g. anaerobes) tend to invest in the higher yielding, but enzymatically costlier, EMP strategy, while those that have alternative, non-glycolytic energy generating pathways (e.g. oxidative phosphorlyation in the case of aerobes) would prefer the enzymatically cheaper ED pathway [41]. Enzyme cost minimization is an extension of MDF that incorporates kinetic information to calculate the protein cost associated with a metabolic pathway [42]. This hybrid kinetic-thermodynamic approach was used to show that the ED and EMP pathways both lie close to the Pareto front of ATP yield vs. enzyme cost [43], confirming earlier studies, which rationalised the design of ATP-producing pathways by assuming that the naturally evolved pathways are close to optimal [44, 45, 6].

4

3 Thermodynamic black box approaches can be used to understand microbial metabolism

Before low cost genomic sequencing was available to interrogate the full structure of microbial metabolism, the exact metabolic transformations occurring inside a cell were relatively opaque. However, precise measurements of the heat and material (substrate consumption, product generation, and biomass production) exchanged with the environment could be made. Thus, the so-called black box techniques were developed, which depended on the measurement of macrochemical equations to summarise the conversion of nutrients into metabolic products and biomass in a single chemical formula. These techniques were used to predict cellular properties important for biotechnology without needing detailed metabolic information [13]. The entropy balance of a black box model of a cell, as depicted in Fig. 1, is given by

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \overbrace{\underbrace{\frac{\dot{Q}}{T}}_{\text{enthalpic term}} + \underbrace{\sum_i \bar{s}_i \dot{n}_i}_{\text{entropic term}} - \bar{s}_X \dot{n}_X}^{\text{compensating terms}} + \overbrace{\dot{S}_{\text{prod}}}^{\text{always positive}} \tag{4}$$

Here, $\dot{Q}$ is the rate of heat exchanged between cell and environment, $\bar{s}_i$ is the partial molar entropy of species i, $\dot{n}_i$ is the rate of import/export of metabolite i, and the subscript X denotes biomass. Importantly, due to the irreversibility of growth, entropy is produced by the cell ($\dot{S}_{\text{prod}} > 0$), leading to unavoidable energetic losses according to the second law of thermodynamics. To avoid thermal cell death, or structural disorganization
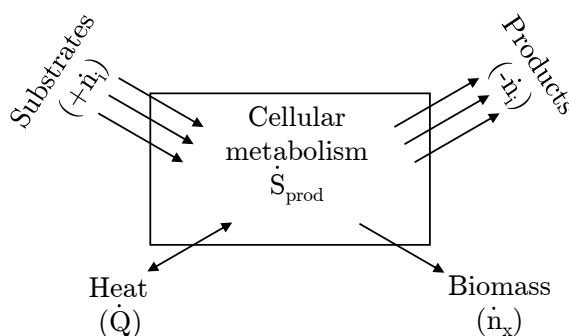

Figure 1: Black box approaches assume that only extracellular metabolites are measured and do not require detailed metabolic mechanisms to yield predictions. Due to the irreversible nature of cellular growth, entropy must be produced and leads to unavoidable energetic losses. The import rate of a variable is taken as positive. Based on Figure 19.10 from [46].

leading to cell death, a cell cannot accumulate entropy (implying $\frac{\mathrm{d}S}{\mathrm{d}t} = 0$). Thus, entropy production must be balanced by either heat or metabolite exchange with the environment, revealing one of the thermodynamic roles of catabolism: to export entropy in the form of heat or high entropy waste products to compensate for metabolic entropy production [46]. Interestingly, it was shown that microorganisms adopt different catabolic strategies to achieve this. Respiration is associated with enthalpic compensation, which means that entropy is exported predominantly as heat (large negative $\dot{Q}$), whereas fermentation is associated with entropic compensation through a high export rate (large negative $\dot{n}_i$) of high entropy compounds (large $\bar{s}_i$). In acetotrophic methanogenesis entropy export by product formation is even so large that the overall growth is endothermic (positive $\dot{Q}$). Autotrophic methanogensis, in contrast, decreases the chemical entropy in the environment, and achieves this by a very high heat production [47].

A highly simplified, but very useful and illustrative black box model, is the linear Gibbs energy converter, as

illustrated in Fig. 2. In this model, catabolism drives anabolism, and these two processes are coupled through
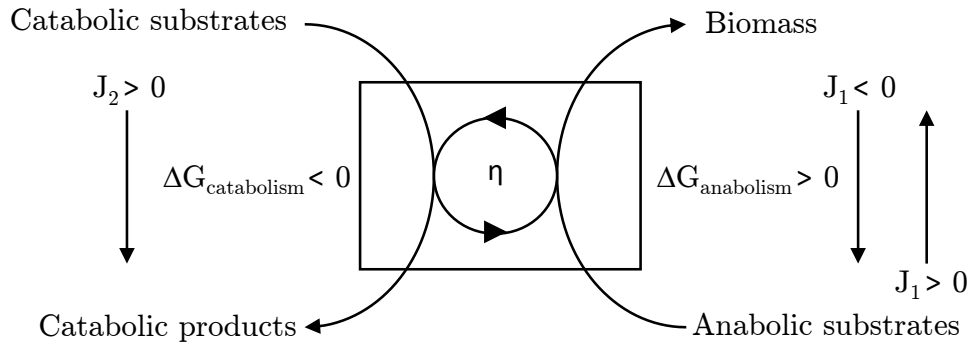

Figure 2: Gibbs energy converter model. Catabolism and anabolism are represented as coupled processes, where the Gibbs energy liberated by exergonic catabolic reactions drive the endergonic anabolic reactions.

the so-called Onsager coefficients (L_{ij}), as shown in Eq. (5). Here J_1 and J_2 are the fluxes of anabolism and catabolism, respectively; and $X_1 = -\Delta G_{\mathrm{anabolism}}$ and $X_2 = -\Delta G_{\mathrm{catabolism}}$ are the respective driving forces of these fluxes [48, 49, 50]. Note that catabolism would spontaneously proceed in the forward direction ($J_2 > 0$), while anabolism, since it is endergonic, would spontaneously proceed in the negative direction ($J_1 < 0$), i.e. break down biomass into its constitutive components.

$$J_1 = L_{11}X_1 + L_{12}X_2$$
$$J_2 = L_{21}X_1 + L_{22}X_2$$

(5)

The coupling between anabolism and catabolism is effected by L_{12} and L_{21}. For the purpose of studying the coupling between catabolism and anabolism, we can assume that Onsager's reciprocal relations hold, and $L_{12} = L_{21}$ [51]. The degree of coupling is defined as $q = \frac{L_{12}}{\sqrt{L_{11}L_{22}}}$. Due to the second law of thermodynamics, L_{11} and L_{22} must both be positive, but the coupling terms may be of any sign. The biologically important case occurs when they are positive, hence it can be shown that $0 \leq q \leq 1$. In this regime catabolic free energy is channeled towards the biosynthesis of new biomass ($J_1 > 0$), at the cost of catabolic flux (J_2 decreases).

This simple model can be used to explain why maintenance energy is necessary in systems that are not perfectly coupled ($q < 1$). In Fig. 3a, the normalised flow ratio j is plotted over the normalised force ratio x (definitions are given in the Figure caption). Quantitatively, it can be shown that when $J_1 = 0$ (zero net growth rate), $x = -q$ and the catabolic flux J_2 at this point is nonzero and proportional to $(1 - q^2)$ [48]. This catabolic flux represents the entropic resistance that needs to be overcome before new biomass can be produced. Maintenance energy is experimentally observed in chemostat cultures at the limit of zero growth, and, in the context of the energy converter model, corresponds to this imperfectly coupled regime.

The efficiency, defined as $\eta = -\frac{J_1 X_1}{J_2 X_2}$, can be interpreted as the yield of the overall growth process, while the anabolic flux power ($P = J_1 X_1$) reflects growth rate, as discussed in Section 2 for single reactions and pathways. In Figures 3b – 3c it is demonstrated that the optimum efficiency (yield) and biomass flux (rate) do not coincide, clearly indicating the existence of a trade-off between microbial growth and yield, as is often observed in nature [52].

While the linear Gibbs energy converter model is intuitively appealing, the model depends on how catabolism and anabolism are split with respect to the black box macrochemical equation, which is a non-trivial problem of definition since many catabolic reactions are actually amphibolic [13]. The problem becomes even greater if
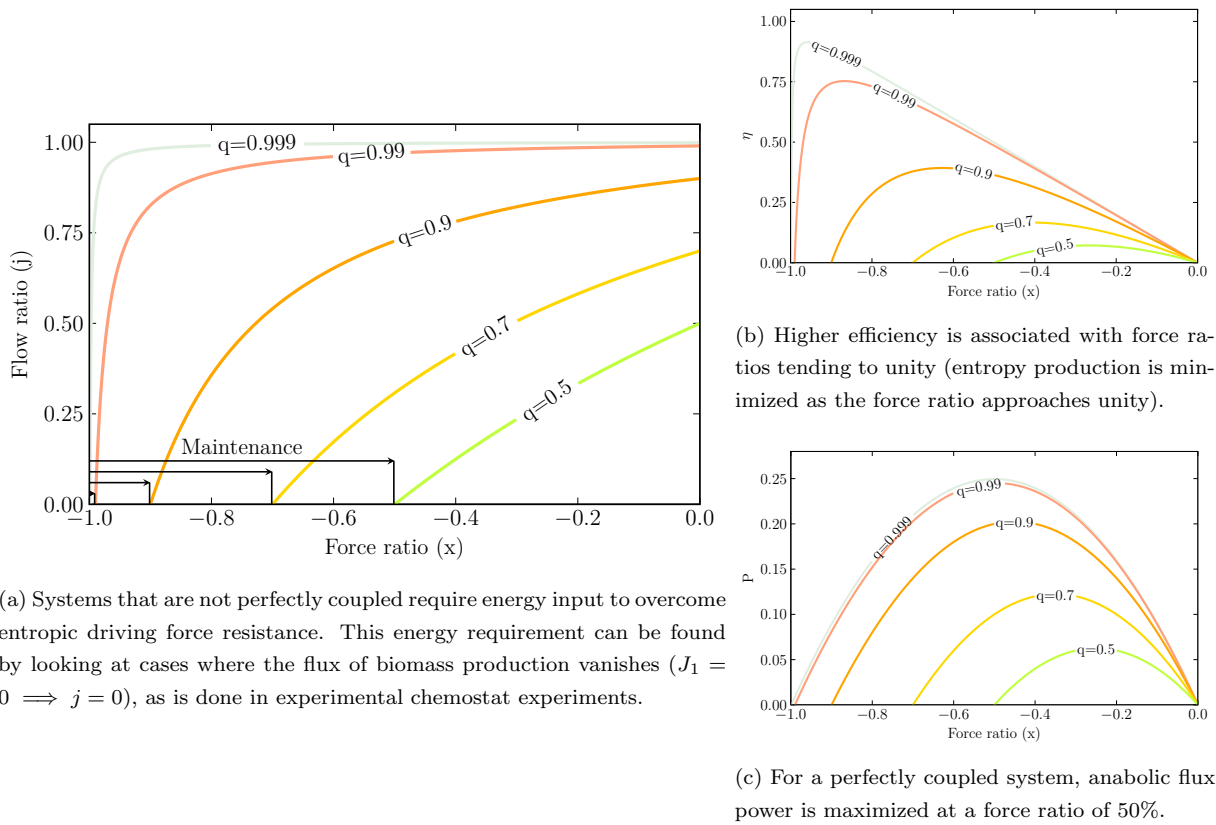
6

(a) Systems that are not perfectly coupled require energy input to overcome entropic driving force resistance. This energy requirement can be found by looking at cases where the flux of biomass production vanishes ($J_1 = 0 \implies j = 0$), as is done in experimental chemostat experiments.

(b) Higher efficiency is associated with force ratios tending to unity (entropy production is minimized as the force ratio approaches unity).

(c) For a perfectly coupled system, anabolic flux power is maximized at a force ratio of 50%.

Figure 3: The normalized force ratio $\left(x = \frac{X_1}{X_2}\sqrt{\frac{L_{11}}{L_{22}}}\right)$ and normalized flux ratio $\left(j = \frac{J_1}{J_2}\sqrt{\frac{L_{22}}{L_{11}}}\right)$ as a function of coupling strength, q, can be used to interrogate the predictions of the Gibbs linear energy transducer model. Positive flow ratio, j, indicates that biomass is being produced. Figure based on Figures 1, 2, and 5 from [48]. Code used to generate the plots are available at `https://gitlab.com/qtb-hhu/thermodynamic-linear-converter-review-2021`.

several carbon or nitrogen sources are available simultaneously. Furthermore, measuring the coupling coefficient is not straightforward. Assuming perfect coupling ($q = 1$) and maximizing anabolic flux power (P), the resultant predicted thermodynamic efficiency is 50%. This efficiency seems to match measured values for aerobes ($\eta_{\text{observed}} \approx 30 - 60\%$), but grossly overestimates anaerobic efficiencies, suggesting that the linear converter theory might be too simplistic [53].

Other approaches make direct use of the Gibbs energy balance, as shown in Eq. (6). Here μ_i is the chemical potential of species i, and $\dot{W}$ the rate of work done on the system.

$$\frac{\mathrm{d}G}{\mathrm{d}t} = \dot{W} + \sum_i \mu_i \dot{n}_i - \mu_X \dot{n}_X - T\dot{S}_{\text{prod}} \tag{6}$$

After some simplifications, it can be shown that $\Delta_{\text{r}}G_X \propto -T\dot{S}_{\text{prod}}$, suggesting that the total Gibbs energy dissipated ($\Delta_{\text{r}}G_X$) by a microbe is the driving force of metabolism [54, 55]. Based on this, some black box approaches have been developed that make use of Gibbs energy dissipation as a predictive tool [15]. Despite the complexity of metabolism eschewed, these techniques have been used to model biomass yield ($Y_{X/S}$) on a variety of substrates, with an average error of 13 – 19 % [56] and 13 – 23% [57], as well as cellular maintenance coefficients (average error 32 – 41 %) [58], under aerobic and anaerobic conditions. It is noteworthy to point out that these approaches were all developed before inexpensive sequencing technology was readily available to interrogate the details of cellular metabolism, yet they are remarkably predictive [59].

4 Genome-scale models of metabolism: peeking in the black box

Due to the advent of low-cost genomic sequencing, and the existence of large databases mapping genes to functions, it has become possible to rapidly characterize the metabolic capability of an organism *in silico*. Consequently, an increasing number of highly detailed genome-scale metabolic models (GEMs) have been developed to systematically understand and engineer microorganisms [60]. If reaction rates v are known, the temporal change of metabolite concentrations x, is given by $\frac{dx}{dt} = Sv(x, t)$, where S is the stoichiometric matrix describing the stoichiometry of the metabolic network encoded by the GEM. However, kinetically characterizing each metabolic reaction in an organism is infeasible due to the scale, complexity, and interactions endemic to microbial metabolism, although attempts have been made for two organisms [61, 62]. Consequently, simplifications are typically made to assist analysis. A fundamental assumption made in most GEM studies is that cellular metabolism is approximately in a stationary state. This simplifies the dynamic equation to $Sv = 0$, which relates the stationary reaction fluxes v to the system's stoichiometry alone. Because this central equation is under-determined, further assumptions are necessary. Flux balance analysis (FBA) casts a GEM into a linear program, as shown in Eq. (7).

$$
\begin{aligned}
&\max_{v} && \mu(v) \\
&\text{s. t.} && Sv = 0 \\
& && v_{\text{LB}} \leq v \leq v_{\text{UB}}
\end{aligned}
\tag{7}
$$

Often, a so-called biomass objective function is used as μ; this ad hoc linear sum of weighted fluxes describes the proportion of macromolecular constituents needed by a microorganism to grow [63], and is related to the macrochemical equation used in black box models. Flux measurements and thermodynamic considerations are used to set the flux bounds, v_{LB} and v_{UB} [64]. It is particularly important to constrain directions of reactions to reduce the flux variability in a model, and to remove thermodynamically infeasible cycles. Typically, physiologically observed metabolite concentration bounds are used in conjunction with Eq. (2) to bound $\Delta_{\text{r}}G'$, and subsequently assign direction to reactions [65].

While the aforementioned approach ensures that individual reactions obey the laws of thermodynamics, it does not ensure that the entire reaction network is thermodynamically consistent. This can be achieved by imposing additional thermodynamic constraints on the classic FBA algorithm. To this end, the two most robust methods are thermodynamic-FBA [66, 67] and loopless-FBA [68]. The former method directly incorporates thermodynamic parameters, i.e. $\Delta_{\text{r}}G'^{\circ}$, for each reaction, as well as bounds on metabolite concentrations achievable by the cell. By incorporating Eqs. (2) and (3) into (7), the optimization problem is converted into a mixed integer linear program (MILP). The latter method does not require thermodynamic information, however it also reformulates (7) as a MILP. In both cases it is guaranteed that solutions represent thermodynamically consistent flux distributions, albeit at the cost of solving a more computationally intensive optimization problem. Interestingly, while loopless-FBA does not directly use thermodynamic parameters, it can be restated as a simplification of thermodynamic-FBA, in which internal metabolite concentrations are unbounded [69].

Thermodynamic-FBA has been used to investigate the metabolism of *Escherichia coli*. Interestingly, by imposing thermodynamic constraints on its GEM, fewer reactions were found to be directionally constrained than were manually constrained in the original FBA formulation, suggesting that the individual reaction based approach may be too conservative [70]. However, this analysis also highlighted the sensitivity of thermodynamic-FBA to uncertainties in $\Delta_{\text{r}}G'^{\circ}$, as well as metabolite concentration bounds, suggesting that the loopless approach may be more robust, because it does not depend on specific values of the reaction energies. However, computational complexity issues plague both of these MILP-based methods, although faster variants of loopless-FBA have been developed [71]. For this reason, alternative methods that remove thermodynamically infeasible cycles, but do not interrogate reaction directions, are more often used in practice [72, 73].

The rate/yield trade-off can also be quantitatively analyzed with GEMs, if kinetic data is also available. Enzyme flux cost minimization is a technique that combines GEM based analyses with enzyme cost minimization analysis. By using this technique, it was found that environmental, as well as kinetic, parameters affect the trade-off between rate and yield in *E. coli* [74]. The flexibility of constraint based analysis also facilitates its extension, e.g. by incorporating phenomenological observations as constraints on fluxes. Recently, an upper bound on the rate of Gibbs free energy dissipation of cellular metabolism was introduced to further constrain the solution space of FBA [75]. The additional constraints allowed FBA to accurately predict overflow metabolism in *Saccharomyces cerevisiae* and *E. coli*. As argued in [76], a common feature of all models explaining overflow metabolism is the existence of two constraints, which restrict metabolism under different conditions. In [75], these constraints were based on thermodynamic considerations. Interestingly, the same results were obtained by using a far simpler black box approach [77], which suggests that the two approaches could be synergistically combined to yield novel insights.

5   Connecting the black box and genome-scale metabolic modeling approaches

Given the increasing ubiquity of GEMs, a rich tool set has been developed to quantitatively analyze metabolism at the genome-scale [78]. However, due to the computational complexity associated with thermodynamic-FBA approaches, it seems that most thermodynamic analyses of GEMs are restricted to reaction direction assignments. In contrast, the historic black box approaches made prodigious use of thermodynamics to overcome their inability to inspect the metabolic features of an organism in detail. Given the accuracy of yield predictions made by using only black box techniques, one can only wonder to what degree deeper thermodynamic integration of GEMs can increase their predictive abilities.

Promisingly, the development of sophisticated thermodynamic databases has already enabled detailed analysis of pathways [41, 43], unavailable a mere decade before their release [79]. These databases can be easily integrated with FBA based approaches. In Fig. 4 we illustrate this integration by calculating the total Gibbs energy dissipated for a variety of organisms and carbon sources, as well as the associated biomass yields.

While it is intuitively clear that catabolic processes drive biomass formation, as modelled in the black box approaches of Section 3, it is less clear how this idea can be applied to GEMs. While GEMs reveal exactly which biochemical transformations are taking place, they do not couple energetics to yield, e.g. there is no disadvantage in increasing the ATP yield of catabolism. However, a rate/yield trade-off must exist. By increasing the number of ATPs generated through catabolism, the Gibbs free energy change of the pathway will approach equilibrium. In light of Sections 2 and 3, this would reduce the rate, since the driving force is diminished. Consequently, metabolism would slow down, reducing the benefit of increasing the ATP yield. This effect is not captured by current constraint-based approaches [86]. It is still an open research question on how to best leverage the bounty of information contained in GEMs, while also incorporating thermodynamic driving force considerations. Addressing this shortcoming by making use of the linear energy converter ideas touched upon in Section 3 could clear the way for establishing the connection between rate and yield at the genome-scale.
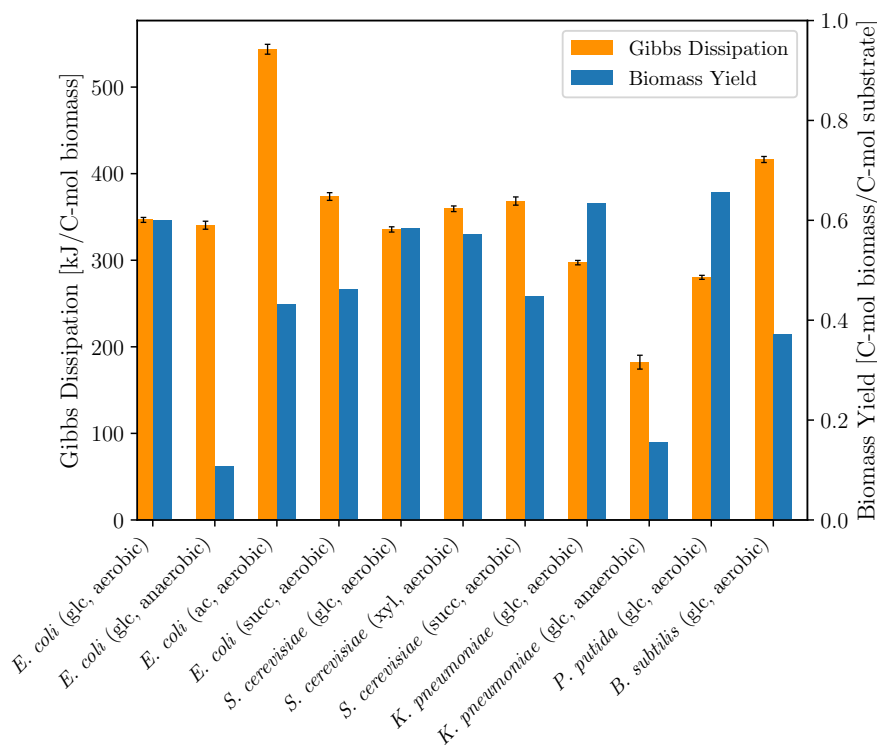
Figure 4: Gibbs dissipation and biomass yields calculated using FBA and the GEMs of *E. coli* [80], *S. cerevisiae* [81], *Klebsiella pneumoniae* [82], *Pseudomonas putida* [83], and *Bacillus subtilis* [84]. Thermodynamic data was sourced from [19] and [85]. Carbon source abbreviations: glc/glucose, ac/acetate, succ/succinate, xyl/xylose. Error bars represent the propagated standard deviation of the mean Gibbs energy of formation for extracellular metabolites and biomass. Code used to generate the plot is available at https://gitlab.com/qtb-hhu/thermodynamic-linear-converter-review-2021.

6 Conflicts of interest, acknowledgements, funding information and author contribution

Author Contributions

Funding

Conflicts of Interest

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

[1] J. M. Smith, H. C. Van Ness, and M. M. Abbott. *Introduction to chemical engineering thermodynamics*. McGraw-Hill, 2005.

[2] Georgy Lebon, David Jou, and José Casas-Vázquez. *Understanding non-equilibrium thermodynamics*, volume 295. Springer, 2008.

[3] Signe Kjelstrup, Dick Bedeaux, Eivind Johannessen, and Joachim Gross. *Non-equilibrium thermodynamics for engineers*. World Scientific, 2017.

[4] Shoichi Toyabe, Takahiro Watanabe-Nakayama, Tetsuaki Okamoto, Seishi Kudo, and Eiro Muneyuki. Thermodynamic efficiency and mechanochemical coupling of f1-atpase. *Proceedings of the National Academy of Sciences*, 108(44):17951–17956, 2011.

[5] Christopher P Kempes, David Wolpert, Zachary Cohen, and Juan Pérez-Mercader. The thermodynamic efficiency of computations made in cells across the range of life. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 375(2109):20160343, 2017.

[6] Steven J Court, Bartlomiej Waclaw, and Rosalind J Allen. Lower glycolysis carries a higher flux than any biochemically possible alternative. *Nature communications*, 6(1):1–8, 2015.

[7] Arren Bar-Even, Avi Flamholz, Elad Noor, and Ron Milo. Thermodynamic constraints shape the structure of carbon fixation pathways. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 1817(9):1646–1659, 2012.

[8] Tobias Großkopf and Orkun S Soyer. Microbial diversity arising from thermodynamic constraints. *The ISME journal*, 10(11):2725–2733, 2016.

[9] Irwa Issa, Sebastian Delbrück, and Ulrich Hamm. Bioeconomy from experts' perspectives–results of a global expert survey. *PloS one*, 14(5):e0215917, 2019.

[10] Kashif Hayat, Saiqa Menhas, Jochen Bundschuh, and Hassan Javed Chaudhary. Microbial biotechnology as an emerging industrial wastewater treatment process for arsenic mitigation: A critical review. *Journal of Cleaner Production*, 151:427–438, may 2017.

[11] Jin-Song Guo, Fang Fang, Peng Yan, and You-Peng Chen. Sludge reduction based on microbial metabolism for sustainable wastewater treatment. *Bioresource Technology*, 297:122506, feb 2020.

[12] Aditya Sarnaik, Arren Liu, David Nielsen, and Arul M Varman. High-throughput screening for efficient microbial biotechnology. *Current Opinion in Biotechnology*, 64:141–150, aug 2020.

[13] J. J. Heijnen and J. P. Van Dijken. In search of a thermodynamic description of biomass yields for the chemotrophic growth of microorganisms. *Biotechnol Bioeng*, 39(8):833–858, Apr 1992.

[14] Hugo F Cueto-Rojas, AJA Van Maris, S Aljoscha Wahl, and JJ Heijnen. Thermodynamics-based design of microbial cell factories for anaerobic product formation. *Trends in biotechnology*, 33(9):534–546, 2015.

[15] Robbert Kleerebezem and Mark CM Van Loosdrecht. A generalized method for thermodynamic state analysis of environmental systems. *Critical Reviews in Environmental Science and Technology*, 40(1):1–54, 2010.

[16] Changdai Gu, Gi Bae Kim, Won Jun Kim, Hyun Uk Kim, and Sang Yup Lee. Current status and applications of genome-scale metabolic models. *Genome biology*, 20(1):1–18, 2019.

[17] Daniele De Martino, Fabrizio Capuani, Matteo Mori, Andrea De Martino, and Enzo Marinari. Counting and correcting thermodynamically infeasible flux cycles in genome-scale metabolic networks. *Metabolites*, 3:946, October 2013.

[18] Robert N Goldberg, Yadu B Tewari, and Talapady N Bhat. Thermodynamics of enzyme-catalyzed reactions—a database for quantitative biochemistry. *Bioinformatics*, 20(16):2874–2877, 2004.

[19] Avi Flamholz, Elad Noor, Arren Bar-Even, and Ron Milo. EQuilibrator - The biochemical thermodynamics calculator. *Nucleic Acids Research*, 40(D1):770–775, 2012.

[20] Elad Noor, Arren Bar-Even, Avi Flamholz, Yaniv Lubling, Dan Davidi, and Ron Milo. An integrated open framework for thermodynamics of reactions that combines accuracy and coverage. *Bioinformatics*, 28(15):2037–2044, 2012.

[21] Elad Noor, Hulda S. Haraldsdóttir, Ron Milo, and Ronan M.T. Fleming. Consistent Estimation of Gibbs Energy Using Component Contributions. *PLoS Computational Biology*, 9(7), 2013.

[22] Robert A. Alberty, Athel Cornish-Bowden, Robert N. Goldberg, Gordon G. Hammes, Keith Tipton, and Hans V. Westerhoff. Recommendations for terminology and databases for biochemical thermodynamics. *Biophysical Chemistry*, 155(2-3):89–103, 2011.

[23] Thomas G. Waddell, Pavle Repovic, Reinhart Heinrich, Enrique Melendez-hevia, and Francisco Montero. Optimization of glycolysis: a new look at the efficiency of energy coupling. *Biochemical Education*, 25(4):204–205, 1997.

[24] Jørgen Christoffersen, Margaret R. Christoffersen, and Axel Hunding. Comments on optimization of glycolysis. *Biochemical Education*, 26(4):290–291, 1998.

[25] Thomas G. Waddell, Pavle Repovic, Enrique Meléndez-Hevia, Reinhart Heinrich, and Francisco Montero. Optimization of glycolysis: New discussions. *Biochemical Education*, 27(1):12–13, 1999.

[26] Keith L. Manchester. Optimization of energy coupling: What is all the argument about? *Biochemical Education*, 28(1):18–19, 2000.

[27] J Carlos Aledo. Metabolic pathways : does the actual Gibbs free-energy change affect the flux rate? *Biochemistry and molecular biology education*, 29:142–143, 2001.

[28] Hagai Rottenberg. The thermodynamic description of enzyme-catalyzed reactions: The linear relation between the reaction rate and the affinity. *Biophysical Journal*, 13(6):503–511, 1973.

[29] R. van der Meer, H. V. Westerhoff, and K. Van Dam. Linear relation between rate and thermodynamic force in enzyme-catalyzed reactions. *Biochimica et Biophysica Acta - Bioenergetics*, 591(2):488–493, 1980.

[30] C. Y. Hoh and R. Cord-Ruwisch. A practical kinetic model that considers endproduct inhibition in anaerobic digestion processes by including the equilibrium constant. *Biotechnology and Bioengineering*, 51(5):597–604, 1996.

[31] Jens Nielsen. Metabolic control analysis of biochemical pathways based on a thermokinetic description of reaction rates. *Biochemical Journal*, 138:133–138, 1997.

[32] Qusheng Jin and Craig M. Bethke. A new rate law describing microbial respiration. *Applied and Environmental Microbiology*, 69(4):2340–2348, 2003.

[33] Elad Noor, Avi Flamholz, Wolfram Liebermeister, Arren Bar-Even, and Ron Milo. A note on the kinetics of enzyme action: A decomposition that highlights thermodynamic effects. *FEBS Letters*, 587(17):2772–2777, 2013.

[34] Daniel A. Beard and Hong Qian. Relationship between thermodynamic driving force and one-way fluxes in reversible processes. *PLoS ONE*, 2(1):1–4, 2007.

[35] JBS Haldane. Enzymes longmans. *Green and Co, UK*, 1930.

[36] Jorge Rodríguez, Juan M Lema, and Robbert Kleerebezem. Energy-based models for environmental biotechnology. *Trends in biotechnology*, 26(7):366–374, 2008.

[37] T. Pfeiffer, S. Schuster, and S. Bonhoeffer. Cooperation and competition in the evolution of ATP-producing pathways. *Science*, 293(5534):1436, 2001.

[38] Q. Jin and C. M. Bethke. The thermodynamics and kinetics of microbial metabolism. *American Journal of Science*, 307(4):643–677, April 2007.

[39] Elad Noor, Arren Bar-Even, Avi Flamholz, Ed Reznik, Wolfram Liebermeister, and Ron Milo. Pathway thermodynamics highlights kinetic obstacles in central metabolism. *PLoS Comput Biol*, 10(2):e1003483, 2014.

[40] R. Heinrich and T. A. Rapoport. A linear steady-state treatment of enzymatic chains. general properties, control and effector strength. *Eur J Biochem*, 42(1):89–95, Feb 1974.

[41] Avi Flamholz, Elad Noor, Arren Bar-Even, Wolfram Liebermeister, and Ron Milo. Glycolytic strategy as a tradeoff between energy yield and protein cost. *Proceedings of the National Academy of Sciences of the United States of America*, 110(24):10039–10044, 2013.

[42] Elad Noor, Avi Flamholz, Arren Bar-Even, Dan Davidi, Ron Milo, and Wolfram Liebermeister. The protein cost of metabolic fluxes: prediction from enzymatic rate laws and cost minimization. *PLoS computational biology*, 12(11):e1005167, 2016.

[43] Chiam Yu Ng, Lin Wang, Anupam Chowdhury, and Costas D Maranas. Pareto optimality explanation of the glycolytic alternatives in nature. *Scientific reports*, 9(1):1–15, 2019.

[44] A. Stephani, J. C. Nuño, and R. Heinrich. Optimal stoichiometric designs of atp-producing systems as determined by an evolutionary algorithm. *J Theor Biol*, 199(1):45–61, Jul 1999.

[45] Oliver Ebenhöh and Reinhart Heinrich. Evolutionary optimization of metabolic pathways. theoretical reconstruction of the stoichiometry of atp and nadh producing systems. *Bulletin of mathematical biology*, 63(1):21–55, 2001.

[46] Urs von Stockar. *Biothermodynamics: The role of thermodynamics in biochemical engineering.* PPUR Presses polytechniques, 2013.

[47] U Von Stockar and J-S Liu. Does microbial life always feed on negative entropy? thermodynamic analysis of microbial growth. *Biochimica et Biophysica Acta (BBA)-Bioenergetics*, 1412(3):191–211, 1999.

[48] Jörg W Stucki. The optimal efficiency and the economic degrees of coupling of oxidative phosphorylation. *European Journal of Biochemistry*, 109(1):269–283, 1980.

[49] Hans V Westerhoff, Klaas J Hellingwerf, and Karel Van Dam. Thermodynamic efficiency of microbial growth is low but optimal for maximal growth rate. *Proceedings of the National Academy of Sciences*, 80(1):305–309, 1983. maximal growth; degree of reduction; anabolism.

[50] O. Kedem and S. R. Caplan. Degree of coupling and its relation to efficiency of energy conversion. *Trans. Faraday Soc.*, 61:1897–1911, 1965.

[51] Lars Onsager. Reciprocal relations in irreversible processes. i. *Phys. Rev.*, 37:405–426, Feb 1931.

[52] David A Lipson. The complex relationship between microbial growth rate and yield and its implications for ecosystem processes. *Frontiers in microbiology*, 6:615, 2015.

[53] Urs von Stockar, Vojislav Vojinović, Thomas Maskow, and Jingsong Liu. Can microbial growth yield be estimated using simple thermodynamic analogies to technical processes? *Chemical Engineering and Processing: Process Intensification*, 47(6):980–990, 2008.

[54] Urs Von Stockar, Thomas Maskow, Jingsong Liu, Ian W Marison, and Rodrigo Patino. Thermodynamics of microbial growth and metabolism: an analysis of the current situation. *Journal of Biotechnology*, 121(4):517–533, 2006.

[55] Urs von Stockar. Biothermodynamics of live cells: a tool for biotechnology and biochemical engineering. *Journal of Non-Equilibrium Thermodynamics*, 35(4):415–475, 2010.

[56] JJ Heijnen, Mark CM van Loosdrecht, and L Tijhuis. A black box mathematical model to calculate auto- and heterotrophic biomass yields based on gibbs energy dissipation. *Biotechnology and bioengineering*, 40(10):1139–1154, 1992.

[57] Perry L McCarty. Thermodynamic electron equivalents model for bacterial yield prediction: modifications and comparative evaluations. *Biotechnology and bioengineering*, 97(2):377–388, 2007.

[58] L Tijhuis, Mark CM Van Loosdrecht, and JJ v Heijnen. A thermodynamically based correlation for maintenance gibbs energy requirements in aerobic and anaerobic chemotrophic growth. *Biotechnology and bioengineering*, 42(4):509–519, 1993.

[59] J-S Liu, V Vojinović, Rodrigo Patiño, Th Maskow, and Urs von Stockar. A comparison of various gibbs energy dissipation correlations for predicting microbial growth yields. *Thermochimica Acta*, 458(1-2):38–46, 2007.

[60] Xin Fang, Colton J. Lloyd, and Bernhard O. Palsson. Reconstructing organisms in silico: genome-scale models and their emerging applications. *Nature Reviews Microbiology*, 18(December):23–26, 2020.

[61] Jonathan R Karr, Jayodita C Sanghvi, Derek N Macklin, Miriam V Gutschow, Jared M Jacobs, Benjamin Bolival Jr, Nacyra Assad-Garcia, John I Glass, and Markus W Covert. A whole-cell computational model predicts phenotype from genotype. *Cell*, 150(2):389–401, 2012.

[62] Derek N Macklin, Travis A Ahn-Horst, Heejo Choi, Nicholas A Ruggero, Javier Carrera, John C Mason, Gwanggyu Sun, Eran Agmon, Mialy M DeFelice, Inbal Maayan, et al. Simultaneous cross-evaluation of heterogeneous e. coli datasets via mechanistic simulation. *Science*, 369(6502), 2020.

[63] Adam M Feist and Bernhard O Palsson. The biomass objective function. *Current opinion in microbiology*, 13(3):344–349, 2010.

[64] Ines Thiele and Bernhard Ø Palsson. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature protocols*, 5(1):93, 2010.

[65] Ronan MT Fleming, Ines Thiele, and Heinz-Peter Nasheuer. Quantitative assignment of reaction directionality in constraint-based models of metabolism: application to escherichia coli. *Biophysical chemistry*, 145(2-3):47–56, 2009.

[66] Christopher S Henry, Linda J Broadbelt, and Vassily Hatzimanikatis. Thermodynamics-based metabolic flux analysis. *Biophysical journal*, 92(5):1792–1805, 2007.

[67] Michael L Mavrovouniotis. Identification of localized and distributed bottlenecks in metabolic pathways. In *Ismb*, volume 93, pages 273–283, 1993.

[68] Jan Schellenberger, Nathan E Lewis, and Bernhard Ø Palsson. Elimination of thermodynamically infeasible loops in steady-state metabolic models. *Biophysical journal*, 100(3):544–553, 2011.

[69] Elad Noor. Removing both internal and unrealistic energy-generating cycles in flux balance analysis. *arXiv preprint arXiv:1803.04999*, 2018.

[70] Joshua J Hamilton, Vivek Dwivedi, and Jennifer L Reed. Quantitative assessment of thermodynamic constraints on the solution space of genome-scale metabolic models. *Biophysical journal*, 105(2):512–522, 2013.

[71] Siu HJ Chan, Lin Wang, Satyakam Dash, and Costas D Maranas. Accelerating flux balance calculations in genome-scale metabolic models by localizing the application of loopless constraints. *Bioinformatics*, 34(24):4248–4255, 2018.

[72] Nathan E Lewis, Kim K Hixson, Tom M Conrad, Joshua A Lerman, Pep Charusanti, Ashoka D Polpitiya, Joshua N Adkins, Gunnar Schramm, Samuel O Purvine, Daniel Lopez-Ferrer, et al. Omic data from evolved e. coli are consistent with computed optimal growth from genome-scale models. *Molecular systems biology*, 6(1):390, 2010.

[73] Abdelmoneim Amer Desouki, Florian Jarre, Gabriel Gelius-Dietrich, and Martin J Lercher. Cyclefreeflux: efficient removal of thermodynamically infeasible loops from flux distributions. *Bioinformatics*, 31(13):2159–2165, 2015.

[74] Meike T Wortel, Elad Noor, Michael Ferris, Frank J Bruggeman, and Wolfram Liebermeister. Metabolic enzyme cost explains variable trade-offs between microbial growth rate and yield. *PLoS computational biology*, 14(2):e1006010, 2018.

[75] Bastian Niebel, Simeon Leupold, and Matthias Heinemann. An upper limit on Gibbs energy dissipation governs cellular metabolism. *Nature Metabolism*, 1(1):125–132, 2019.

[76] Daan H De Groot, Julia Lischke, Riccardo Muolo, Robert Planqué, Frank J Bruggeman, and Bas Teusink. The common message of constraint-based optimization approaches: overflow metabolism is caused by two growth-limiting constraints. *Cellular and Molecular Life Sciences*, 77(3):441–453, 2020.

[77] Nima P Saadat, Tim Nies, Yvan Rousset, and Oliver Ebenhöh. Thermodynamic limits and optimality of microbial growth. *Entropy*, 22(3):277, 2020.

[78] Laurent Heirendt, Sylvain Arreckx, Thomas Pfau, Sebastián N Mendoza, Anne Richelle, Almut Heinken, Hulda S Haraldsdóttir, Jacek Wachowiak, Sarah M Keating, Vanja Vlasov, et al. Creation and analysis of biochemical constraint-based models using the cobra toolbox v. 3.0. *Nature protocols*, 14(3):639–702, 2019.

[79] Thomas Maskow and Urs von Stockar. How reliable are thermodynamic feasibility statements of biochemical pathways? *Biotechnology and Bioengineering*, 92(2):223–230, 2005.

[80] Jonathan M Monk, Colton J Lloyd, Elizabeth Brunk, Nathan Mih, Anand Sastry, Zachary King, Rikiya Takeuchi, Wataru Nomura, Zhen Zhang, Hirotada Mori, et al. i ml1515, a knowledgebase that computes escherichia coli traits. *Nature biotechnology*, 35(10):904–908, 2017.

[81] Monica L Mo, Bernhard Ø Palsson, and Markus J Herrgård. Connecting extracellular metabolomic measurements to intracellular flux states in yeast. *BMC systems biology*, 3(1):1–17, 2009.

[82] Yu-Chieh Liao, Tzu-Wen Huang, Feng-Chi Chen, Pep Charusanti, Jay SJ Hong, Hwan-You Chang, Shih-Feng Tsai, Bernhard O Palsson, and Chao A Hsiung. An experimentally validated genome-scale metabolic reconstruction of klebsiella pneumoniae mgh 78578, iyl1228. *Journal of bacteriology*, 193(7):1710–1717, 2011.

[83] Juan Nogales, Bernhard Ø Palsson, and Ines Thiele. A genome-scale metabolic reconstruction of pseudomonas putida kt2440: i jn746 as a cell factory. *BMC systems biology*, 2(1):1–20, 2008.

[84] You-Kwan Oh, Bernhard O Palsson, Sung M Park, Christophe H Schilling, and Radhakrishnan Mahadevan. Genome-scale reconstruction of metabolic network in bacillus subtilis based on high-throughput phenotyping and gene essentiality data. *Journal of Biological Chemistry*, 282(39):28791–28799, 2007.

[85] Marko Popovic. Thermodynamic properties of microorganisms: determination and analysis of enthalpy, entropy, and gibbs free energy of biomass, cells and colonies of 32 microorganism species. *Heliyon*, 5(6):e01950, 2019.

[86] Stefan Schuster, Thomas Pfeiffer, and David A Fell. Is maximization of molar yield in metabolic networks favoured by evolution? *Journal of theoretical biology*, 252(3):497–504, 2008.