

Article

Microbial Pathway Thermodynamics: Stoichiometric Models Unveil Anabolic and Catabolic Processes

Oliver Ebenhööh ^{1,2,*} , Josha Ebeling ¹ , Ronja Meyer ¹ , Fabian Pohlkotte ¹  and Tim Nies ^{1,*} 

¹ Institute of Quantitative and Theoretical Biology, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

² Cluster of Excellence on Plant Sciences, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

* Correspondence: oliver.ebenhoeh@hhu.de (O.E.); tim.nies@hhu.de (T.N.)

Abstract: The biotechnological exploitation of microorganisms enables the use of metabolism for the production of economically valuable substances, such as drugs or food. It is, thus, unsurprising that the investigation of microbial metabolism and its regulation has been an active research field for many decades. As a result, several theories and techniques were developed that allow for the prediction of metabolic fluxes and yields as biotechnologically relevant output parameters. One important approach is to derive macrochemical equations that describe the overall metabolic conversion of an organism and basically treat microbial metabolism as a black box. The opposite approach is to include all known metabolic reactions of an organism to assemble a genome-scale metabolic model. Interestingly, both approaches are rather successful at characterizing and predicting the expected product yield. Over the years, macrochemical equations especially have been extensively characterized in terms of their thermodynamic properties. However, a common challenge when characterizing microbial metabolism by a single equation is to split this equation into two, describing the two modes of metabolism, anabolism and catabolism. Here, we present strategies to systematically identify separate equations for anabolism and catabolism. Based on metabolic models, we systematically identify all theoretically possible catabolic routes and determine their thermodynamic efficiency. We then show how anabolic routes can be derived, and we use these to approximate biomass yield. Finally, we challenge the view of metabolism as a linear energy converter, in which the free energy gradient of catabolism drives the anabolic reactions.

Keywords: energy metabolism; elementary conversion modes; metabolic networks; energy converter



Citation: Ebenhööh, O.; Ebeling, J.; Meyer, R.; Pohlkotte, F.; Nies, T. Microbial Pathway Thermodynamics: Stoichiometric Models Unveil Anabolic and Catabolic Processes. *Life* **2024**, *14*, 247. <https://doi.org/10.3390/life14020247>

Academic Editor: Tatyana Sokolova

Received: 4 December 2023

Revised: 29 January 2024

Accepted: 5 February 2024

Published: 9 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Microbial organisms are an essential part of most ecosystems. They function as vital members of natural production chains, leading to the formation of chemical compounds that have complexity which is unreachable by current technological standards [1]. Therefore, much scientific effort has been spent to understand the metabolism of microbes [2]. Today, the human exploitation of microbial metabolism has long left the stage of merely producing fermented food, such as cheese or alcoholic beverages [3]. Microbes, often viewed as natural factories, are used in biotechnological applications as an integrated component of drug and food fabrication or bioremediation projects [4–6]. In order to assess the potential of a microbial organism for bioeconomical applications, it is important to fully understand its metabolic capabilities. This includes a thermodynamic characterization of metabolic pathways, which allows us, for example, to calculate maximally possible yields.

Gaining full knowledge about microbial metabolism was and is a complex scientific problem. This complexity derives from the multitude of microbial organisms, the diversity of metabolic pathways, and the different strategies to regulate metabolism [7–9]. In the pre-genomic era, researchers used precise measurements of input (substrate) and output (product) relationships for microbial cultures and a deep understanding of thermodynamics to design new biotechnological strategies, based on so-called macrochemical

equations [2,10–14]. A macrochemical equation summarizes the conversion of substrates into metabolic products and biomass. Describing microbial metabolism by a single macrochemical equation essentially treats microbial metabolism as a black box, ignoring all intracellular metabolic details. Still, this single chemical equation can accurately describe the overall metabolic activity and thus can serve as a means to understand and predict metabolic properties (see, e.g., [11]).

The macrochemical equation can be understood as a sum of two separate reactions which describe catabolism (the breakdown of nutrients to gain free energy in the form of ATP) and anabolism (the formation of new biomass from the nutrients using the energy in the form of ATP provided by catabolism) [15]. In this picture, microbial growth is described as a thermodynamic energy converter, where the catabolic reactions provide the required free energy to drive anabolism (see Figure 1). Here, the negative free energies of reactions of the catabolic and anabolic half-reactions, denoted by $-\Delta_{\text{cat}}G$ and $-\Delta_{\text{ana}}G$, respectively, are generalized thermodynamic forces, and the respective reaction rates, J_{cat} and J_{ana} , are generalized thermodynamic fluxes. In the theory of non-equilibrium thermodynamics, it has been shown that in general, for systems close to equilibrium, the generalized fluxes J depend linearly on generalized forces X (X is a vector combining $-\Delta_{\text{cat}}G$ and $-\Delta_{\text{ana}}G$), i.e., $J = \mathbf{L}X$, where the entries L_{ij} of the matrix \mathbf{L} are the so-called phenomenological coefficients [16]. The name phenomenological coefficient refers to the fact that the quantities cannot be derived from first principles but they have to be experimentally measured. It is often assumed that the linear relationship is also valid for catabolic and anabolic forces and fluxes [14,17–19], which would entail that the following relation holds:

$$\begin{pmatrix} J_{\text{cat}} \\ J_{\text{ana}} \end{pmatrix} = \mathbf{L} \cdot \begin{pmatrix} -\Delta_{\text{cat}}G \\ -\Delta_{\text{ana}}G \end{pmatrix}. \quad (1)$$

However, it is thus far not clear why this actually should be the case, considering that metabolically active and dividing cells operate far from thermodynamic equilibrium. We test the assumption of a linear relationship with published experimental data for *S. cerevisiae* and *E. coli* grown in chemostat cultures and observe that the measured data clearly contradict this assumption.

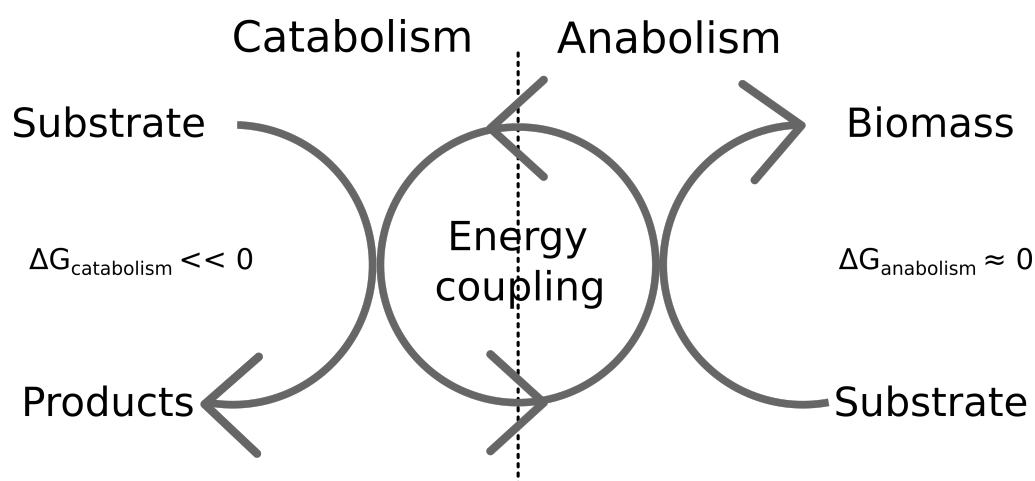


Figure 1. The view of microbial metabolism as a thermodynamic energy converter. Catabolic reactions have a large negative free energy gradient, driving anabolic reactions [14].

The experimental determination of a macrochemical growth equation requires precise measurements of all chemical substances that are consumed and produced by the growing microbes. These measurements are possible in controlled chemostat cultures, in which microbes grow on a defined growth medium. However, many microbes cannot easily be cultured in chemically defined media, which prevents an experimental determination of macrochemical growth equations. With the recent scientific advancement in genome

sequencing, genomic data became available for a huge number of organisms, including those which are difficult to culture [20]. This information allows for building stoichiometric (genome-scale) metabolic models, manually or semi-automatically [21–24]. Genome-scale metabolic models are a formalization of all known biochemical reactions of an organism. As such, they combine genomic, proteomic, and metabolic information to build an *in silico* representation that can be used to derive steady-state flux distributions [25]. The latter is achieved by the optimization of an objective function (such as biomass production) using linear programming. The inspection and analysis of genome-scale metabolic models benefit from a rich theory for metabolic networks (see, e.g., [25–29]). Thus, such models can support strategies to improve the product yield in biotechnological applications. Elementary flux modes (EFMs) are a systematic way to quantify the metabolic capabilities of an organism [26]. EFMs describe all possible pathways between substrate and product. Due to combinatorial explosion [30], it is still challenging to calculate all EFMs for larger structural metabolic models, also with modern computational facilities. To overcome this, and because for many investigations only the conversion between substrate and product is relevant, elementary conversion modes (ECMs) were introduced by Urbanczik and Wagner [31]. According to [32], elementary conversion modes are the minimal building blocks of all net conversions (input–output relationships) in a metabolic network. Plainly speaking, ECMs ignore all intracellular processes and only focus on the effects of metabolic pathways on the external metabolites. Using ECMs instead of EFMs reduces the necessary computational power drastically. Additionally, modern software, such as *ecmtool* (<https://github.com/SystemsBioinformatics/ecmtool>, last accessed 29 November 2023) which allows for the parallelization of the computation, helps us to obtain an exhaustive list of all metabolic capabilities of an organism in the form of ECMs [33].

As suggested in [32], we view ECMs as building blocks of macrochemical equations. We show how genome-scale metabolic models can be used to systematically enumerate all possible catabolic pathways. With thermodynamic data, in particular the energies of formation of substrates and products obtained from the *eQuilibrator* tool [34,35], we characterize the catabolic pathways by their energy gradient, by which we understand the free energy that becomes available from converting nutrients with a high energy of formation into catabolic products with a lower energy of formation. Using the network models, we further estimate the maximal ATP production capacity for each catabolic pathway, and we thus determine their thermodynamic efficiencies. We then analyze experimental data for *Saccharomyces cerevisiae* [36] and *Escherichia coli* [37], grown under controlled chemostat conditions in defined media, to separate the macrochemical equation into its catabolic and anabolic parts and characterize their thermodynamic properties. Interpreting our findings in the context of the energy converter model, we identify the limitations of the applicability of the linear converter model, but we observe an interesting linear scaling law between growth rate and metabolic power.

2. Theory and Methods

2.1. Calculating Elementary Conversion Modes

Elementary conversion modes (ECMs, [31]) are a concept related to elementary flux modes (EFMs, [26]). However, whereas an EFM describes a minimal set of reactions in a network which can carry a stationary flux through the network, ECMs are not defined in terms of the metabolic routes through the network but by their end results, i.e., by the stoichiometry in which they convert external metabolites. As such, enumerating all ECMs of a network provides a full description of all possible metabolic conversions by the network. Here, we employ ECMs to thermodynamically characterize catabolic routes. For this, we calculate ECMs which connect external metabolites but which do not lead to the production of biomass. We perform calculations for three metabolic models, the *E. coli* core model [38], the genome-scale model iJR904 for *E. coli* str. K-12 substr. MG1655 [39], and the genome-scale model iND750 for the yeast *S. cerevisiae* [40], using *ecmtool* [32,33]. Due to their size, for the latter two (genome-scale) models, a complete enumeration of all ECMs is not easily

possible with standard computing equipment. We therefore hid all external metabolites that contain phosphate, sulfur, or nitrogen and dismissed all compounds with more than six carbon atoms. These steps reduced the number of catabolic routes considerably and allowed the calculations to be performed in a reasonable time. As input (carbon source) for *ecmtool*, we focused on simple sugars and carboxylic acids (glucose, xylose, pyruvate, 2-oxoglutarate). Moreover, we allowed oxygen to be present. The output of *ecmtool* is a matrix, in which the rows are the respective elementary conversion modes and the columns are all external metabolites that were not hidden. Excluding larger molecules and those containing sulfur, phosphorus, and nitrogen reduces the number of conversion modes and thus the number of possible catabolic routes. For example, ecologically important catabolic pathways obtaining free energy from oxidizing ammonia [41] are excluded by our simplification. For the *E. coli* core network [38], no metabolite had to be hidden, and thus the full catabolic potential of the core network could be described.

2.2. Estimating Gibbs Free Energy of Reaction and Thermodynamic Efficiency

To approximate the standard Gibbs free energy of reaction ($\Delta_{\text{cat}}G^\circ$) for each obtained ECM, we used the Python API of the *eQuilibrator* tool [35]. We extracted the Gibbs free energies of formation ($\Delta_f G^\circ$) for all external metabolites involved in an ECM. Next, we normalized the ECMs with respect to the carbon atoms of the carbon source (C-mol) and applied a Laplace transformation, adapting for temperature (298.15 K), pH (7.4), pMg (3.0), and ionic strength (0.25 M). We used Hess's law to calculate the standard Gibbs free energy of reaction for each ECM,

$$\Delta_{\text{cat}}G^\circ = \sum_{i=1}^m v_i \Delta_f G_i^\circ, \quad (2)$$

where v_i and $\Delta_f G_i^\circ$ are the stoichiometric coefficient and the Gibbs free energy of formation of the i^{th} external compound in the ECM, respectively.

In the following, we will use the standard Gibbs free energies as an objective quantity to characterize a large number of catabolic routes. It has to be noted, though, that the actual free energy gradients depend on the concentrations of substrates and products in the solution and that in physiological conditions cells never experience standard conditions. If all concentrations c_i of the external compounds are known, then the actual energy of reaction can be determined by

$$\Delta_{\text{cat}}G = \Delta_{\text{cat}}G^\circ + RT \ln \prod_{i=1}^m c_i^{v_i}. \quad (3)$$

Whereas for single biochemical reactions the concentration-dependent term in Equation (3) is often of the same order of magnitude as the standard energy of reaction and can therefore critically affect the thermodynamics and directionality of a reaction, its relative importance is much lower for overall pathways, which consist of several reactions and thus exhibit a larger total energy of reaction. The pathways considered here display standard energies of reaction between 30 and 500 kJ C-mol⁻¹, which corresponds to between 180 and 3000 kJ mol⁻¹ substrate. On the other hand, for each order-of-magnitude deviation from standard conditions, the second term results in a correction of only $RT \ln 10 = 5.7 \text{ kJ mol}^{-1}$. Moreover, the estimations of the standard energies of reaction, which are based on the group contribution method [42], come with a considerable margin of error. For example, assuming typical glucose concentrations for substrate unlimited growth of $c_{\text{Glc}} = 0.16 \text{ mol L}^{-1}$ ($\approx 30 \text{ g L}^{-1}$), dissolved oxygen of $c_{\text{O}_2} = 250 \mu\text{mol L}^{-1}$, and carbon dioxide of $c_{\text{CO}_2} = 13 \mu\text{mol L}^{-1}$ (corresponding to ambient air with 20% oxygen and 420 ppm carbon dioxide), the second term computes to $-39.4 \text{ kJ mol}^{-1}$ (at 25 °C), in which (i) still lies well within the error range of the computational prediction of the standard energy of reaction, $\Delta_{\text{cat}}G^\circ = -2927.8 \pm 49.5 \text{ kJ mol}^{-1}$ [34], and (ii) is two orders of magnitude smaller than the standard energy of reaction itself. Taking into account the total energy gradients and the errors in the prediction method of reaction energies, we can

expect that the standard energies of reaction provide a sufficiently good thermodynamic characterization of catabolic pathways.

For the calculation of the maximal ATP production for an ECM, we constrained all external fluxes to the values of the respective ECMs while maximizing ATP hydrolysis (excluding ATP maintenance):

$$\begin{aligned} &\text{maximize} && v_{\text{ATPM}}, \\ &\text{such that} && \mathbf{N} \cdot \mathbf{v} = 0 \\ & && v_{i,\text{ex}} = v_i \text{ for } i \in \text{ECM}, \\ & && v_{j,\text{ex}} = 0 \text{ for } j \notin \text{ECM}, \end{aligned} \quad (4)$$

where \mathbf{N} is the stoichiometric matrix of the metabolic model, v_{ATPM} is the flux through the reaction



and $v_{i,\text{ex}}$ are the fluxes through the reaction exchanging metabolite i , which is constrained to the stoichiometric coefficient v_i obtained by the respective ECM. The stoichiometric coefficients are normalized to one carbon mole substrate.

The thermodynamic efficiency of ATP production is calculated as

$$\eta = \frac{c_{\text{ATP}} \cdot \Delta_r G_{\text{ATPase}}}{|\Delta_{\text{cat}} G|}, \quad (6)$$

where c_{ATP} is the maximal ATP yield per carbon mole, $\Delta_r G_{\text{ATPase}}$ is the energy of reaction for ATP synthesis from ADP and inorganic phosphate, and $\Delta_{\text{cat}} G$ is the energy of reaction of the respective catabolic pathway.

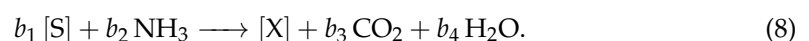
For the Gibbs free energy of ATP synthesis, we used a typical value for *E. coli* of $\Delta_r G_{\text{ATPase}} = 46.5 \text{ kJ mol}^{-1}$ [43].

2.3. Calculating the Stoichiometry of Anabolism

We assume that the substrate [S] has the normalized sum formula CH_xO_y , the biomass [X] has $\text{CH}_a\text{O}_b\text{N}_c$, and that the biomass is more reduced than the substrate, i.e.,

$$\gamma_S = 4 + x - 2y \leq \gamma_X = 4 + a - 2b - 3c. \quad (7)$$

We assume an overall stoichiometry of



Every carbon that is converted to biomass will have to be reduced by $\gamma_X - \gamma_S$. From the overall redox balance, it follows that for each carbon that is converted into biomass,

$$b_3 = \frac{\gamma_X - \gamma_S}{\gamma_S} \quad (9)$$

carbons have to be oxidized to CO_2 . From the carbon balance of (8), it follows that

$$b_1 = 1 + b_3 = \frac{1}{Y_{\text{X/S}}^{\text{max}}}. \quad (10)$$

The nitrogen and hydrogen balances entail that

$$b_2 = b \quad \text{and} \quad b_4 = \frac{b_1 x + 3b_2 - a}{2}. \quad (11)$$

It is straight-forward to generalize these calculations to include sulfur and phosphorus into the biomass.

2.4. Calculating the Optimal Anabolic Reaction from Genome-Scale Networks

To determine the maximal yield and the minimal ATP requirement for biomass formation, we perform two subsequent linear programs. First, the exchange reactions are constrained, such that only the carbon source (substrate) and oxygen can be imported (negative flux), but other metabolites can be released (positive flux). The biomass reaction is constrained to one carbon mole per unit of time. The ATP hydrolysis reaction is not constrained, which means it can run in reverse and provide ATP. Subsequently, substrate import (negative) is maximized:

$$\begin{aligned}
 &\text{maximize} && v_{\text{substrate,ex}}, \\
 &\text{such that} && \mathbf{N} \cdot \mathbf{v} = 0 \\
 &&& v_{\text{biomass}} = 1, \\
 &&& -\infty < v_{\text{substrate,ex}} \leq 0 \\
 &&& -\infty < v_{\text{O}_2,\text{ex}} \leq 0 \\
 &&& 0 \leq v_{j,\text{ex}} < \infty \text{ for other metabolites}
 \end{aligned} \tag{12}$$

The resulting optimal flux is negative, and the absolute value denotes the minimal substrate requirement to produce one carbon mole of biomass, if ATP is provided in abundance.

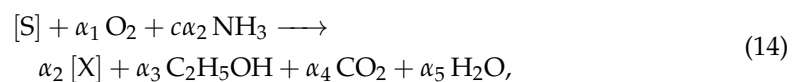
In the second step, the determined minimal substrate requirement $v_{\text{substrate}}^{\text{opt}}$ is fixed, and the ATP requirement is minimized by maximizing the (negative) flux through reaction (5):

$$\begin{aligned}
 &\text{maximize} && v_{\text{ATPM}}, \\
 &\text{such that} && \mathbf{N} \cdot \mathbf{v} = 0 \\
 &&& v_{\text{biomass}} = 1, \\
 &&& v_{\text{substrate,ex}} = v_{\text{substrate}}^{\text{opt}} \\
 &&& -\infty < v_{\text{O}_2,\text{ex}} \leq 0 \\
 &&& 0 \leq v_{j,\text{ex}} < \infty \text{ for other metabolites}
 \end{aligned} \tag{13}$$

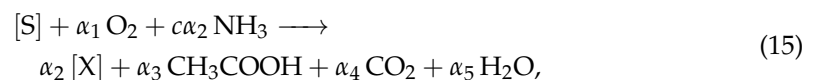
The absolute value of the optimal flux $|v_{\text{ATPM}}|$ gives a minimal ATP requirement for the production of one carbon mole biomass.

2.5. Calculating the Stoichiometry of Catabolism

Macrochemical equations of the form



for *S. cerevisiae* (see [36]) and



for *E. coli* were obtained from the original publications. Here, we use the notation [S] for one carbon mole of substrate and [X] for one carbon mole of biomass. The sum formula of biomass is assumed to be given by $\text{CH}_a\text{O}_b\text{N}_c$ (hence the factor c in the stoichiometry of NH_3), and it is given in both cases in the original publication. The stoichiometric coefficients were obtained as follows. For *S. cerevisiae*, Table 1 in [36] already provides the stoichiometric coefficients for Equation (22), which were, for our calculations, converted into carbon moles. For *E. coli*, we converted data from Table 2 in [37] which are given in $\text{g g}^{-1} \text{h}^{-1}$ to $\text{C-mol C-mol}^{-1} \text{h}^{-1}$, using the molecular weights of the chemical compounds as well as the biomass, normalized to one carbon mole.

The coefficients for the catabolic reaction



are simply determined by calculating (22) – $\alpha_2 \cdot (8)$, resulting in the coefficients

$$c_1 = 1 - b_1\alpha_2 \quad (17)$$

$$c_2 = \alpha_1 \quad (18)$$

$$c_3 = \alpha_4 - b_3\alpha_2 \quad (19)$$

$$c_4 = \alpha_3 \quad (20)$$

$$c_5 = \alpha_5 - b_4\alpha_2. \quad (21)$$

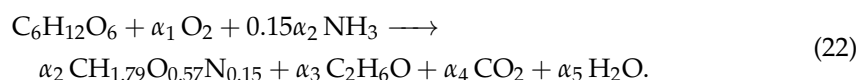
Subsequently, it is convenient to normalize this equation to the consumption of one carbon mole of substrate, i.e., dividing all coefficients by c_1 .

In the case of *E. coli*, acetate was excreted instead of ethanol at the onset of overflow metabolism [37]. In the calculation, ethanol can simply be replaced by acetic acid and the calculation remains identical.

2.6. Extraction and Processing of Experimental Data

For the experimental determination of anabolic and catabolic fluxes, it is necessary to precisely measure the macrochemical growth equation. Chemostat cultures are best suited for such measurements, because growth rates can be precisely controlled by setting the dilution rate, and cultures will approach a stationary growth. Basically, all input and output fluxes need to be quantified. This entails that nutrient concentrations in the feed and in the culture vessel need to be determined. Moreover, gas exchange rates for oxygen and CO₂ need to be measured. Further, ideally, all metabolic products should be quantified. A commonly applied criterion to check whether all products have been measured is the so-called carbon recovery coefficient, which quantifies the fraction of carbons that were provided with the nutrients, which are recovered in the biomass and the metabolic products.

Specifically, we obtain data from two sources. For *S. cerevisiae*, the data source is Table 1 in [36], which provides the stoichiometric coefficients of the macrochemical equation in the form



Here, we convert these coefficients to match the macrochemical equation in the form of Equation (22), where all carbon-containing chemical species are given in carbon moles. This entails the division of all coefficients by 6 (glucose has 6 carbons), except α_3 , which is divided by 3, because ethanol has 2 carbons. With these converted stoichiometric coefficients, we proceed as described above.

For *E. coli*, the data are extracted from Table 2 in [37]. There, biomass formation rate is given in gram biomass per liter and hour. The exchange rates of substrates and products are given in gram substrate per gram biomass per hour. To convert these numbers into stoichiometric coefficients corresponding to Equation (15), we first convert these numbers from gram to carbon mole. For biomass formation, the values are divided by the mass of one carbon mole biomass (26.11 g C-mol^{−1}). Likewise, the exchange rates are multiplied by the mass of one carbon mole biomass and divided by the molecular mass of the respective substrate. Then, the resulting values are converted to stoichiometric coefficients as follows: α_2 (C-mol biomass per C-mol glucose) results simply from dividing the growth rate by the specific glucose uptake rate (in C-mol C-mol^{−1} h^{−1}). Likewise, α_3 (for acetate) and α_4 (for CO₂) are calculated. Finally, the stoichiometric coefficient for O₂ (α_1) is obtained by the oxygen balance of Equation (15), and the coefficient for H₂O by the hydrogen balance.

3. Results

3.1. Calculating Elementary Conversion Modes to Characterize Catabolic Pathways

We use elementary conversion modes (ECMs, see [32,33]) to assess the metabolic capabilities of several genome-scale metabolic networks. To illustrate our approach, we begin our analysis with the *E. coli* core network [38], an *E. coli* core metabolism model of reduced complexity, with only 72 metabolites connected by 95 reactions, of which 20 are exchange reactions. To systematically describe all theoretically possible catabolic routes, we use *ecmtool* [33] to calculate all ECMs which do not produce biomass. The resulting ECMs describe all possible routes and their stoichiometries by which external substances can be interconverted.

We then calculate for each individual ECM the standard Gibbs free energy of reaction, based on the standard energies of formation estimated by *eQuilibrator* [34]. The standard energies of catabolism $\Delta_{\text{cat}}G^\circ$, normalized to one carbon mole of consumed substrate, are displayed in Figure 2. Most of the catabolic pathways have relatively low energy gradients between approximately 20 and 50 kJ C-mol⁻¹. To this large group of pathways belong key catabolic routes, such as the fermentation of glucose to lactate or ethanol. Few catabolic routes exhibit Gibbs free energies of reactions with a higher energy gradient than 50 kJ C-mol⁻¹ (17 ECMs). The pathways with the largest absolute Gibbs free energy of reaction are the combustion of glucose ($\Delta_{\text{cat}}G^\circ \approx -488$ kJ C-mol⁻¹) and the production of formate from glucose ($\Delta_{\text{cat}}G^\circ \approx -325$ kJ C-mol⁻¹). In particular, oxygen-using ECMs belong to the group with the largest absolute $\Delta_{\text{cat}}G^\circ$ (compare red bars in Figure 2). As shown in Figure 2, the usage of nitrogen does not appear to be an indicator of whether the respective ECM has a high or low energy gradient.

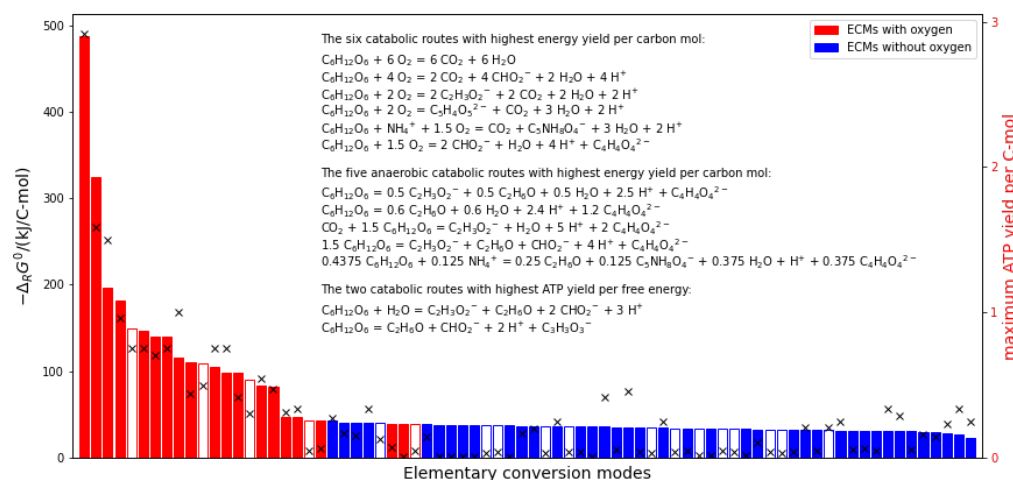


Figure 2. Standard Gibbs free energies of all catabolic pathways, normalized to carbon mole. The catabolic pathways were derived using elementary conversion modes (ECMs) calculated from the *E. coli* core network. Red symbolizes ECMs that use oxygen, while blue denotes ECMs not using oxygen. Filled bars belong to ECMs that include no compounds with the element nitrogen, while empty ones include nitrogen-containing metabolites. The black crosses indicate the maximal yield of ATP per carbon mole nutrient for each ECM (right axis).

For each catabolic route, we use the metabolic model to calculate the maximal ATP yield. For this, the exchange reactions are constrained to the stoichiometries of the respective ECM, and subsequently the flux through ATP hydrolysis is maximized (see Theory and Methods; Section 2). The resulting maximal ATP yields per carbon mole substrate are indicated by black crosses in Figure 2. While as a tendency high energy gradient pathways also allow for a higher ATP yield, there are a considerable number of ECMs with very low ATP yield (34 ECMs exhibit a maximal ATP yield of less than 0.1 mol ATP per C-mol).

3.2. Thermodynamic Efficiency of Catabolic Routes

To investigate whether these general patterns are also conserved in more complete and therefore realistic genome-scale models, we repeat our analysis for the iJR904 metabolic network model [39] of *E. coli* as well as for the iND750 metabolic network model [40] of the yeast *S. cerevisiae*. Figure 3 illustrates the results for *S. cerevisiae*, obtained with the iND750 model. We identify all ECMs using one of the four carbon sources glucose, xylose, α -ketoglutarate, and pyruvate. Also, here, most ECMs yield a low energy gradient, while those with the highest gradient correspond to the full oxidation of glucose and xylose. Full respiration for both sugars releases around $488 \text{ kJ C-mol}^{-1}$, and they are, thus, the ECMs with the highest free energy gradient. These two pathways also display the highest ATP yield per carbon mole (approx. 2.92 mol/C-mol), as indicated in Figure 3, left panel. In contrast to this, most other catabolic routes release energy in a range between 40 and $200 \text{ kJ C-mol}^{-1}$ with low ATP production. Among these routes are, besides other metabolic modes, fermentation reactions, such as the metabolization of glucose to lactate or ethanol.

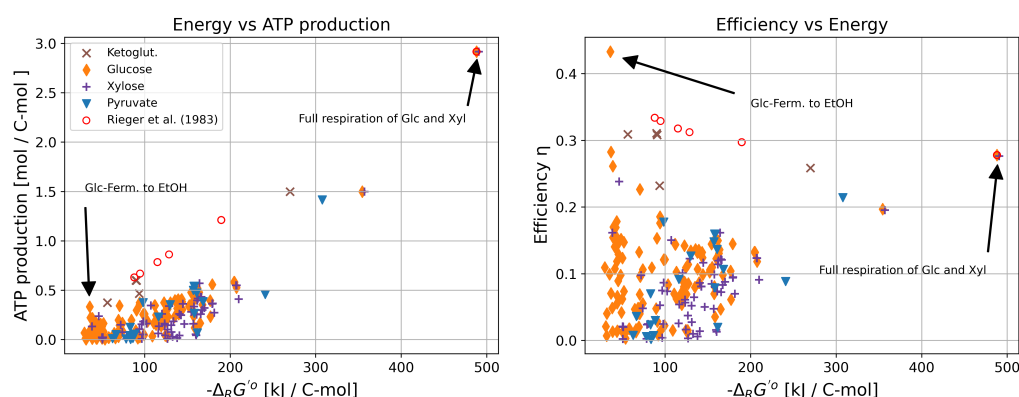


Figure 3. Thermodynamic characterization of catabolic routes in the *S. cerevisiae* genome-scale model (iND750) for α -ketoglutarate, glucose, xylose, and pyruvate as carbon sources. Additionally, oxygen is allowed to be a substrate in the calculation of the elementary conversion modes. The efficiency is based on a typical value of 46.5 kJ/mol for the production of ATP in *E. coli* [43]. For comparison we also calculated ATP production and efficiency based on data by [36]. Only ECMs are plotted with an ATP yield higher than 0.001 mol/C-mol .

The thermodynamic efficiency η of a thermodynamic engine is defined as the fraction of work generated per input energy provided to power the system. Defining the free energy used to drive ATP synthesis as the useful chemical work, we calculate the thermodynamic efficiency (η , Equation (6)). We assume the typical value of $\Delta_r G_{\text{ATPase}} = 46.5 \text{ kJ mol}^{-1}$ for ATP synthesis [43]. Further, we approximate the Gibbs free energy of catabolism by the corresponding standard Gibbs free energy of reaction, because changes in substrate and product concentrations in the medium are likely to have only a minor effect on the quantity (see Methods for discussion and examples). The determined efficiencies η are depicted in the right panel of Figure 3. Interestingly, the pathways with the highest energy gradient are not the most efficient. For instance, under full respiration of glucose, only 28% of the released free energy is converted to chemical work producing ATP. In contrast, the fermentation of glucose to lactate exhibits an efficiency of 43%. The fermentation of glucose to lactate is one of the most efficient reactions, with higher efficiencies only found in fermentation processes involving ethanol production.

3.3. Deriving Anabolic Information from a Metabolic Network

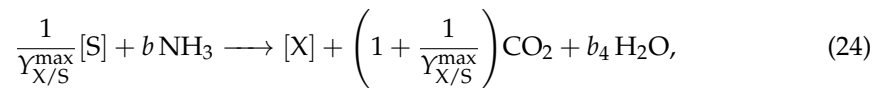
Likewise, anabolic pathways can also be investigated in separation. In the following, we compare predictions from a macrochemical description of anabolism, which ignores details of the intracellular reaction network, with those resulting from employing genome-scale metabolic models. Specifically, we calculate maximal theoretical yields based only on

the chemical properties of the nutrients and the biomass, and then we analyze how closely they can be realized by the actual metabolism of *E. coli* and *S. cerevisiae*.

The theoretical limit to how much carbon in the nutrient can be converted to biomass carbon is given by the degree of reduction. If biomass is more reduced than substrate, a fraction of the substrate carbons needs to be oxidized in order to ensure the overall redox balance. Specifically, if γ_S and γ_X are the degrees of reduction of substrate and biomass, respectively, then the theoretical maximal yield is given by [10]

$$\gamma_{X/S}^{\max} = \begin{cases} \frac{\gamma_S}{\gamma_X} & \text{for } \gamma_S \leq \gamma_X \\ 1 & \text{else} \end{cases} \quad (23)$$

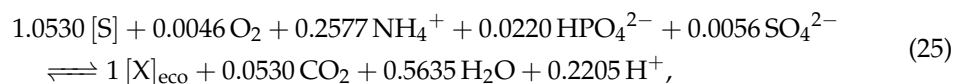
Based on the elemental composition of the biomass and the substrate, an ideal anabolic reaction stoichiometry can be determined. We assume we have a substrate with a sum formula $[S] = \text{CH}_x\text{O}_y$ (normalized to C-mol) and biomass with $[X] = \text{CH}_a\text{O}_b\text{N}_c$, $\gamma_S = 4 + x - 2y$ and $\gamma_X = 4 + a - 2b - 3c$ (assuming ammonia as nitrogen source, see [10]). If $\gamma_S \leq \gamma_X$, the stoichiometry reads



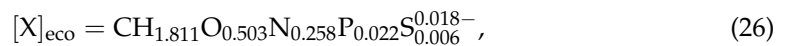
with $b_4 = (x/\gamma_{X/S}^{\max} + 3b - a)/2$ (see Theory and Methods; Section 2).

This equation allows for the calculation of the standard energy of reaction of anabolism, $\Delta_{\text{ana}}G^\circ$. To estimate the Gibbs free energy of formation of biomass, which is required to determine the energy of reaction, we employ the empirical method proposed by Battley [44].

Constraint-based models can be employed to investigate to what extent such an ideal anabolic reaction can be realized by a microorganism's metabolism. We employ the genome-scale networks for *S. cerevisiae* and *E. coli* and minimize the nutrient uptake for fixed biomass production, while allowing ATP to be provided externally (see Theory and Methods; Section 2). A subsequent optimization, in which the minimal nutrient uptake is fixed and the required reverse ATP hydrolysis is minimized, allows for determining the minimal ATP requirement per carbon mole biomass formed. For the iJR904 model of *E. coli* metabolism, we obtain the following optimal anabolic stoichiometry for growth on glucose:



where $[S]$ denotes 1 C-mol of substrate ($\frac{1}{6}\text{C}_6\text{H}_{12}\text{O}_6$) and $[X]_{\text{eco}}$ 1 C-mol of *E. coli* biomass, with the sum formula determined by the biomass reaction of the iJR904 model



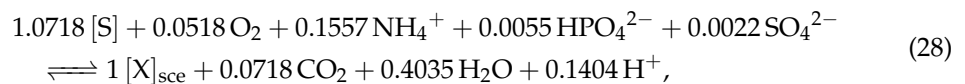
and a degree of reduction and energy of formation of

$$\gamma_{X,\text{eco}} = 4.193, \quad \Delta_f G_{X,\text{eco}}^\circ = -101.10 \text{ kJ C-mol}^{-1}. \quad (27)$$

The resulting simulated maximal yield of $\gamma_{X/S}^{\text{sim}} = \frac{1}{1.053} = 95.0\%$ is slightly lower than that expected by Equation (23), which results in $\gamma_{X/S}^{\max} = \frac{4}{4.193} = 95.4\%$. This is explained by the fact that small amounts of oxygen are also required for pure anabolic biomass formation. In iJR904, this is caused by a minimal required flux through a cytochrome oxidase, which requires molecular oxygen as substrate.

A subsequent optimization reveals a minimal requirement of 1.766 mol ATP per carbon biomass produced.

For the iND750 metabolic model of *S. cerevisiae*, we obtain



with

$$[\text{X}]_{\text{sce}} = \text{CH}_{1.8243}\text{O}_{0.6589}\text{N}_{0.1557}\text{P}_{0.0055}\text{S}_{0.0022} \quad (29)$$

and a degree of reduction and energy of formation of

$$\gamma_{\text{X},\text{sce}} = 4.080, \quad \Delta_f G_{\text{X},\text{sce}}^\circ = -128.76 \text{ kJ C-mol}^{-1}. \quad (30)$$

Here, the discrepancy between the computationally determined maximal yield of $\gamma_{\text{X/S}}^{\text{sim}} = \frac{1}{1.0718} = 93.3\%$ and the $\gamma_{\text{X/S}}^{\text{max}} = 98.0\%$ expected from Equation (23) is even larger. The minimal requirement of ATP to produce biomass is predicted to be slightly larger than for *E. coli* with 2.031 mol ATP per C-mol biomass.

We repeated the calculations for different carbon sources. The results are summarized in Table 1. In general, the expected trend can be observed that more oxidized carbon sources result in lower maximal yields (see also Equation (23)). Moreover, the maximal simulated yields predicted by the model are usually very close to the maximal yield predicted by the degree of reduction alone. This demonstrates that the metabolic networks of *S. cerevisiae* and *E. coli* are structured in such a way that the theoretical yield limits, resulting from the degrees of reduction of substrate and biomass alone, can be closely approximated. Only for *S. cerevisiae* growing on oxoglutarate, yields predicted by the model are considerably lower. The reason for this is that the network defined by the iND750 model is not capable of producing biomass from oxoglutarate without metabolic side products. The optimal solution produces 0.073 mol xanthine ($\text{C}_5\text{H}_4\text{N}_4\text{O}_2$) per C-mol biomass. This leads to a reduced carbon (and in fact, nitrogen) yield, but a larger free energy gradient of anabolism compared to other carbon sources, which can be readily explained by the highly oxidized side product.

Table 1. Thermodynamic properties of anabolic pathways. The theoretical maximal yield $\gamma_{\text{X/S}}^{\text{max}}$ is calculated according to Equation (23). The maximal yield $\gamma_{\text{X/S}}^{\text{sim}}$ predicted by the metabolic model is determined using the linear program (12). The minimal anabolic ATP requirement per carbon mole biomass $a_{\text{ATP},\text{min}}$ is determined using the linear program (13). The standard energies of reaction for anabolism, $\Delta_{\text{ana}} G^\circ$, are determined from the overall anabolic stoichiometries like those given for glucose in Equations (25) and (28).

Organism	Carbon Source	$\gamma_{\text{X/S}}^{\text{max}}$	$\gamma_{\text{X/S}}^{\text{sim}}$	$a_{\text{ATP},\text{min}}$ (mol C-mol ⁻¹)	$\Delta_{\text{ana}} G^\circ$ (kJ C-mol ⁻¹)
<i>E. coli</i>	glucose	0.954	0.950	1.766	−54.81
	xylose	0.954	0.950	1.813	−56.74
	oxoglutarate	0.763	0.730	2.051	−44.28
	pyruvate	0.795	0.791	2.268	−32.44
<i>S. cerevisiae</i>	glucose	0.980	0.933	2.031	−74.51
	xylose	0.980	0.933	2.245	−77.85
	oxoglutarate	0.784	0.457	1.453	−279.33
	pyruvate	0.817	0.815	2.684	−29.62

3.4. Separating Catabolism from Anabolism Based on Chemostat Data

In a controlled continuous microbial cultivation system, such as a chemostat [45,46], it is possible to grow microbial cultures at a steady state with pre-defined growth rates. In chemostat cultures, the dilution rate D is an experimental control parameter and, when the culture has reached a stationary state, the specific growth rate is forced to be equal to the dilution rate. Measuring nutrient and gas exchange rates as well as nutrient and product concentrations in the reactor allows for the experimental determination of the overall growth stoichiometries [11–13,19].

In the following, we employ experimentally determined macrochemical equations for the growth of *S. cerevisiae* [36] and *E. coli* [37] to calculate catabolic stoichiometries, ATP production potential, and thermodynamic efficiencies for each condition. Both publications contain a full characterization of macrochemical growth parameters for chemostat cultures at different dilution rates. Specifically, gas exchange rates were quantified to determine oxygen uptake and CO₂ evolution rates, and substrate, product, and biomass concentrations in the medium were measured, allowing for the calculation of the respective exchange rates. Together with the elemental composition of the biomass, which is also reported in both sources, these data provide a precise quantification of the macrochemical growth equation for different dilution rates. We then separated the macrochemical equations by first identifying the ideal anabolic stoichiometry based on the degrees of reduction of substrate and biomass (see Methods, Section 2.3) and then subtracting this anabolic stoichiometry from the macrochemical equation (see Methods, Section 2.5).

The determined catabolic coefficients are summarized in Figure 4. It can clearly be seen that the onset of overflow metabolism, when glucose is partly fermented even in the presence of sufficient oxygen, occurs at growth rates of around 0.3 h^{−1} for *S. cerevisiae* and around 0.4 h^{−1} for *E. coli*. With the catabolic coefficients, we calculate the standard Gibbs free energy of reaction of the overall catabolic conversion, where we obtained the standard Gibbs free energies of formation, required for this calculation, from the equilibrator tool [34]. It can be observed (blue lines in Figure 4) that with the onset of overflow metabolism, the Gibbs free energy gradients are also reduced significantly.

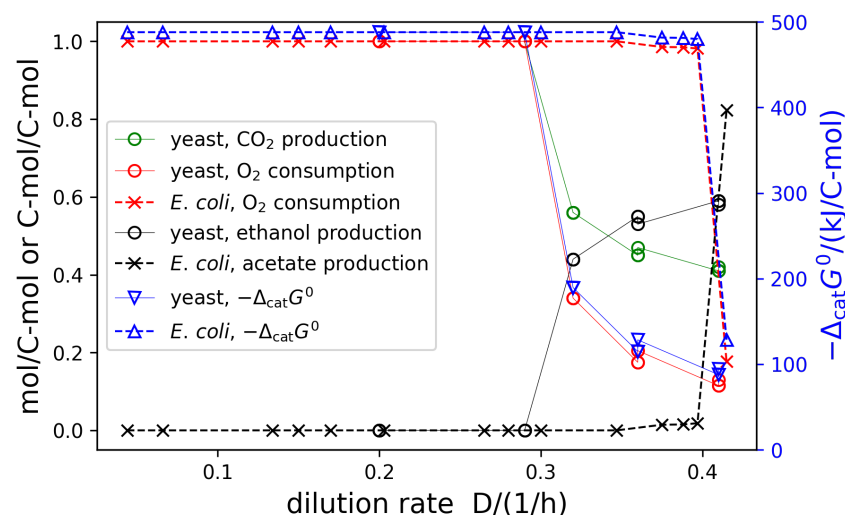


Figure 4. The catabolic stoichiometric coefficients and thermodynamic driving forces determined for the chemostat growth of *E. coli* [36] and *S. cerevisiae* [37] at different dilution rates. All coefficients are given in mol/C-mol substrate, except for ethanol and acetate, which are given in C-mol/C-mol substrate. For *E. coli*, CO₂ production is identical to O₂ consumption. The thermodynamic driving force is given as the standard energy of reaction of the overall catabolic conversion, normalized to one carbon mole of substrate.

3.5. Is the Linear Energy Converter a Good Model for Microbial Growth?

With the theory developed on how to separate catabolism and anabolism from a macrochemical equation, we are now in the position to systematically test whether experimental data of chemostat growth are actually in agreement with the assumption of a linear flux–force relationship (see Equation (1)) between the anabolic and catabolic driving forces and the corresponding fluxes.

For this, we will focus on the experimental data for *S. cerevisiae* [36]. The reason for this is that in the data for *E. coli* [37] the carbon recovery coefficient, which quantifies the fraction of carbons ‘recovered’ in the biomass and metabolic byproducts, compared to the carbons provided by the nutrients, was lower than 90% for large growth rates (above

$\approx 0.4 \text{ h}^{-1}$). This indicates that not all metabolic products were measured, thus preventing a reliable calculation of catabolic stoichiometries.

In the thermodynamic interpretation of microbial growth as a linear energy converter [18,19], catabolic and anabolic fluxes are assumed to linearly depend on the catabolic and anabolic forces, as illustrated by Equation (1). Here, the anabolic flux equals the growth rate, which in turn is set by the dilution rate of the chemostat, $J_{\text{ana}} = D$. The catabolic flux is calculated by separating the catabolic and anabolic equations (see above). As discussed in Section 2 (Theory and Methods), we approximate the actual Gibbs free energies by the standard energies, and we motivate this approximation by the large overall energy gradients. Moreover, we assume that the anabolic Gibbs free energy $\Delta_{\text{ana}}G$ is approximately constant over different dilution rates, because the stoichiometry of anabolism remains constant, according to Equation (8). Based on the experimentally determined biomass composition of *S. cerevisiae* [36] ($\text{CH}_{1.79}\text{O}_{0.57}\text{N}_{0.15}$), we calculate the standard energy of formation of biomass with the empirical method of Battley [44] as

$$\Delta_f G_X^\circ = -104.9 \text{ kJ C-mol}^{-1}, \quad (31)$$

and with that the anabolic energy of reaction as

$$\Delta_{\text{ana}} G^\circ = -42.2 \text{ kJ C-mol}^{-1}. \quad (32)$$

With this full knowledge of catabolic and anabolic fluxes and forces, we can challenge the linear converter model. In Figure 5 (Figure S2 for *E. coli*), various fluxes (catabolic glucose consumption, anabolic (growth) rate, total glucose consumption) are displayed as functions over the catabolic driving force or, alternatively (x -axis on top), the force ratio $x = \Delta_{\text{cat}} G^\circ / \Delta_{\text{ana}} G^\circ$. It can clearly be observed that the fluxes do not linearly depend on the forces, as would be predicted by the linear converter model. On the contrary, fluxes are larger for smaller forces. This indicates that the lower the energy gradient driving the growth, the faster microbes grow under the condition of overflow metabolism. This stark discrepancy from a linear converter model can readily be explained by considering that overflow metabolism results from active regulatory processes inside the cells. Often, overflow metabolism is explained by capacity constraints within the cell: whereas respiration results in a considerably higher yield, it also requires higher protein investment than fermentation, and therefore, at very high growth rates, it is more efficient to ‘waste’ substrate and operate a lower yield pathway (see, e.g., [2,47,48]). This entails that there exist feedback regulation mechanisms which result in a highly non-linear relation between the observed fluxes and the driving forces. It can be concluded that a linear energy converter model is too simplistic and is not in agreement with experimental data, at least over conditions in which the catabolic pathways change. The reason for this is to be sought in feedback mechanisms by which cells adapt their metabolism to external conditions.

Similar to the calculation of the maximal ATP yields of the different catabolic pathways, we determine the maximal ATP yield for the observed catabolic stoichiometries by constraining the genome-scale networks to the observed catabolic stoichiometries. The red circles in Figure 3 present the result of this calculation (left panel) and the respective thermodynamic efficiencies (right panel). Interestingly, the ATP yields as well as the efficiencies are higher than those for elementary conversion modes with similar energy gradients. This can be explained by considering that the experimentally observed catabolic stoichiometries are a linear combination of two conversion modes (full respiration and fermentation) only and that the fermentation pathways are especially identified to have the highest thermodynamic efficiency (see Figures 2 and 3).

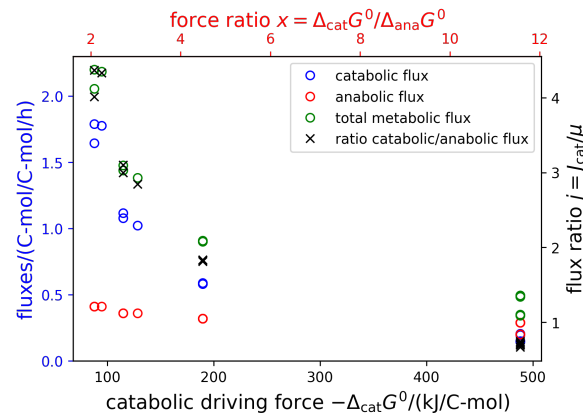


Figure 5. Metabolic fluxes as functions of the catabolic driving force. Shown are the catabolic (blue), anabolic (red), and total (green) glucose consumption rates in dependence of the catabolic driving force, $-\Delta_{\text{cat}} G^\circ$. On the x-axis on the top, the force ratio $x = \Delta_{\text{cat}} G^\circ / \Delta_{\text{ana}} G^\circ$ is given.

With the usual definition of the power as the product of flux and force, we can quantify the catabolic and anabolic powers $P_{\text{cat}} = -J_{\text{cat}} \cdot \Delta_{\text{cat}} G$ and $P_{\text{ana}} = -D \cdot \Delta_{\text{ana}} G$, as well as the power of ATP production $P_{\text{ATP}} = c_{\text{ATP}} \cdot J_{\text{cat}} \cdot \Delta_r G_{\text{ATPase}}$ in kJ C-mol^{-1} . Figure 6 shows that the powers increase approximately linearly with growth rate, despite the fact that metabolism changes considerably for fast growth rates. Moreover, *E. coli* and *S. cerevisiae* behave quite similarly. For large growth rates, the powers of *E. coli* appear to be somewhat lower, but this could be a result of incomplete carbon recovery in the experiments, because besides acetate no other fermentation products were measured. This could lead to an underestimation of the Gibbs free energy gradients and consequently of the ATP yields.

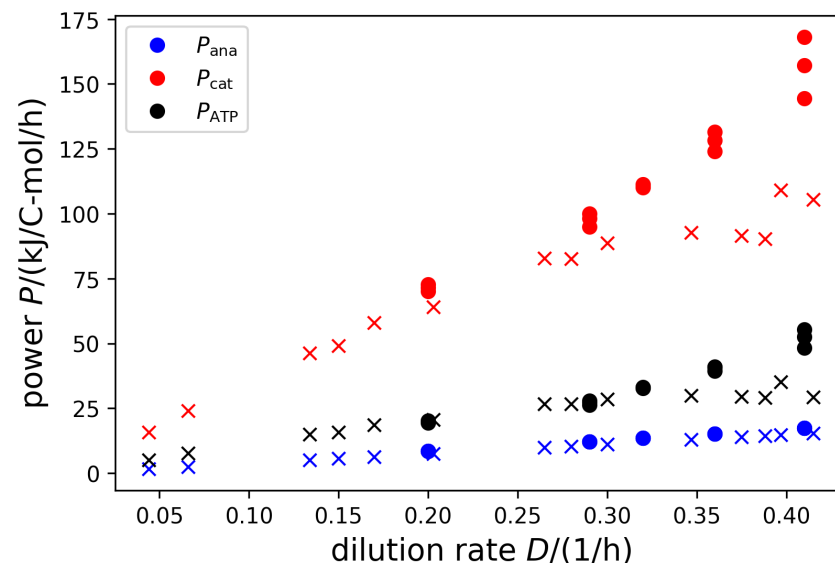


Figure 6. Catabolic and anabolic powers, as well as the power of ATP synthesis as a function of growth rate. Catabolic power is depicted in red, anabolic power in blue, and ATP synthase power in black. Circles present results for *S. cerevisiae*, crosses for *E. coli*.

4. Discussion and Conclusions

Microbial organisms are a cornerstone of the modern biotechnological industry. They are invaluable for producing pharmaceuticals, food, and construction materials [3]. Today, by using sophisticated genetic techniques and engineering, microorganisms can be used to tackle modern problems of society, such as the remediation of waste land, the production of drugs, or finding environmentally sustainable building materials [49–51]. However, such advancements in exploiting bacteria and unicellular eukaryotes were only possible with

a thorough understanding of their metabolism. Multiple theories and techniques have been developed to gain knowledge about metabolic pathways. One of the most promising strategies currently is to develop genome-scale metabolic networks encoding almost all metabolic information available for an organism [52].

Genome-scale metabolic models are used to understand and probe the metabolic capabilities of an organism and allow for the calculation of maximal yields [53]. The construction of such models only became possible with advances in sequencing technologies and protein function prediction, through which more and more fully sequenced genomes and better annotated proteomes become available [20,54]. Before genome-scale metabolic networks existed, many scientists and engineers relied on macrochemical equations. These equations describe the metabolism of organisms as a black box by one overall chemical equation. Over the decades, this approach has been characterized extensively for its applicability and in the context of thermodynamics. However, a challenge when applying macrochemical equations is separating anabolism from catabolism so that both metabolic modes can be studied individually. Here, with the help of genome-scale models and thermodynamic calculations, we showed how we can extract both anabolic and catabolic conversions. We characterized both metabolic modes and challenged commonly used viewpoints on microbial metabolism, such as its representation as a linear energy converter.

By using elementary conversion modes (ECMs), which are an alternative to elementary flux modes, we could systematically enumerate catabolic routes from genome-scale networks of *E. coli* and *S. cerevisiae* (see Figures S1 and 3). Combined with thermodynamic data of the Gibbs free energy of formation for all metabolites, as provided by the eQuilibrator tool, it is simple to derive the standard Gibbs free energy of reactions for all input–output relationships (ECMs). By doing so, we calculated the catabolic driving force of microbial growth for all theoretically possible routes.

However, the second law of thermodynamics implies that not all of the available free energy can be used to perform useful chemical work. Combining ECMs with constraint-based modeling of genome-scale networks, we calculated the thermodynamic efficiency of the ATP production of each catabolic route. Interestingly, we found the most efficient pathways exhibited a thermodynamic efficiency between approximately 30 and 45%. Interestingly, most catabolic routes show a considerably lower efficiency below 20%. It should be noted, though, that the values for the efficiency have to be interpreted with care. Firstly, we assumed standard energies of reaction for the catabolic routes. The actual concentrations of nutrients and catabolic products in the medium may slightly affect these energy values. Secondly, we assumed a fixed value for the energy of reaction for ATP synthesis, which of course may change for different physiological conditions and depends primarily on the ATP–ADP ratio and the concentration of inorganic phosphate. Taking this into account, the highest efficiency, which is observed for the fermentation pathways, is very close to the 50% that is predicted to yield the highest ATP production rates by simple linear thermodynamic energy converter models [55]. It is remarkable that the actually realized catabolic pathways in chemostat cultures (see red circles in Figure 3) provide a higher efficiency than elementary pathways with a similar free energy gradient. This observation stresses the important role of the pure respiration and fermentation pathways of catabolism. Because of their high efficiencies, operating them in combination always provides a higher thermodynamic efficiency than any single elementary conversion mode.

While a linear energy converter model seems adequate to predict optimal thermodynamic efficiencies of ATP-producing pathways with reasonable accuracy [55], our interpretation of experimental data shows that this is not the case when microbial growth is considered. Our results clearly demonstrate that the flux–force relationship is not linear and that, in fact, anabolic and catabolic fluxes *decrease* with increasing catabolic driving force. In other words, the faster microbes grow, the lower the energy gradient that drives this growth. This observation, however, holds for conditions during which catabolism exhibits rather drastic changes, from pure respiration at low growth rates to predominant fermentation at high growth rates. For growth in batch cultures on different substrate

concentrations, however, it was shown that the linear converter model yielded very good results, which indeed fit the data better than a simple Monod equation [56]. It can therefore be hypothesized that the linear energy converter model is adequate as long as the catabolic mode does not change, and thus the driving force is mainly influenced by substrate concentration, but fails as too simplistic if experimental conditions encompass a change in catabolic pathways. In order to experimentally test this hypothesis, systematic growth data for different microorganisms grown under different nutrient concentrations and dilution rates would be necessary. As discussed in detail in the Results section, such experiments must be designed to yield exact information about the media composition, growth rate, biomass composition, and formation rates of metabolic product. In short, all required information to obtain a precise macrochemical equation of metabolism is required to calculate the anabolic and catabolic fluxes separately, which allows for testing the validity of the linear converter model.

An interesting observation is that the output powers scale approximately linear with growth rate and that the proportionality is very similar for two organisms as different as the bacterium *E. coli* and the eukaryote *S. cerevisiae*. For the anabolic power (growth rate times anabolic driving force), this result is trivial because we assumed the anabolic force, $-\Delta_{\text{ana}}G$, to be constant. However, for the catabolic power (nutrient consumption rate for catabolism times catabolic driving force), this result is far from obvious. For technical systems, such as ships [57], bikes, cars, or trains (see [58], Chapter IIIA), the power increases over-proportionally with speed, approaching an approximately quadratic relationship. The linear power–growth rate relationship entails that, employing engineering terms, a ‘resistance’ that needs to be overcome by the thermodynamic driving forces when producing new biomass is a constant rather than dependent on the biomass production rate. Moreover, the force, corresponding to the slope of the power–growth rate curves, appears to be the same for *E. coli* and *S. cerevisiae*. Whether these laws are of a universal nature remains to be tested with systematic experiments of more microbial species grown on different nutrient sources.

The present study is largely theoretical and it provides advances in our methodology of how to study microbial metabolism, both through the more classical black box approach based on macrochemical equations as well as with constraint-based models. A direct application of our theories helps in predicting maximal yields for different nutrients and thus may also support the design of cost- and resource-efficient growth media. A recent approach to avoid the emergence of non-producing mutants in biotechnological applications is to genetically engineer microbes such that product formation is directly coupled with growth [59–61]. Our theories are directly applicable to such scenarios and allow for estimating the expected effects on yield and thermodynamic driving forces.

From a theoretical perspective, a main challenge in the future will be to understand the underlying mechanisms that cause microbial growth to deviate so strongly from the behavior expected for a linear energy converter. The next challenge will then be to employ such understanding to find equations that generalize the simplistic linear converter model to also include scenarios in which catabolic pathways change. Such equations would be highly useful, because they will allow for predicting not only the yield but also the growth rates of microorganisms.

In summary, we could show how combining black box macrochemical approaches and genome-scale metabolic models can help to systematically characterize catabolic routes and find separate chemical equations for anabolism and catabolism. Interpreting experimental data from chemostats with our theoretical models reveals that the efficiency of catabolism appears optimal, both for *E. coli* and the yeast *S. cerevisiae*, over a wide range of growth rates. Moreover, our analyses allow us to speculate that the linear power–growth rate relationship is a universal property of microbial growth.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/life14020247/s1>, Figure S1: Thermodynamic characterisation of catabolic routes in *E. coli* genome-scale model (iJR904) for α -ketoglutarate, glucose, xylose, and

pyruvate as carbon source. Additionally, oxygen is allowed to be a substrate in the calculation of the elementary conversion modes. The efficiency is based on a typical value of 46.5 kJ/mol for production of ATP in *E. coli* [43]. Only ECMs are plotted with a ATP yield higher than 0.001 mol/C-mol; Figure S2: Metabolic fluxes as function of the catabolic driving force for data for *Escherichia coli* [37]. Shown are the catabolic (blue), anabolic (red) and total (green) glucose consumption rates in dependence of the catabolic driving force, $-\Delta_{cat}G^0$. On the x -axis on the top, the force ratio $x = \Delta_{cat}G^0 / \Delta_{ana}G^0$ is given.

Author Contributions: Conceptualization, O.E.; methodology, O.E.; formal analysis, O.E., T.N., F.P., R.M. and J.E.; writing—original draft preparation, O.E., T.N. and F.P.; writing—review and editing, O.E., T.N., F.P., R.M. and J.E.; visualization, O.E., T.N., F.P. and R.M.; supervision, O.E.; funding acquisition, O.E. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Deutsche Forschungsgemeinschaft (DFG), project ID 391465903/GRK 2466 (T.N.), and the DFG by the Collaborative Research Center SFB1535, project ID 458090666/CRC1535/1 (F.P.). OE was supported by Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy—EXC-2048/1—project ID 390686111.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: The original contributions presented in the study are included in the article/Supplementary Materials; further inquiries can be directed to the corresponding authors. The code can be found at <https://gitlab.com/qtb-hhu/thermodynamics-task-force/2023-energy-metabolism-of-microorganisms>, last accessed on 4 January 2024.

Acknowledgments: The authors would like to thank St. Elmo Wilken for fruitful discussions and valuable input.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

EFMs elementary flux modes
ECMs elementary conversion modes

References

1. Ramachandra Rao, S.; Ravishankar, G. Plant cell cultures: Chemical factories of secondary metabolites. *Biotechnol. Adv.* **2002**, *20*, 101–153. [\[CrossRef\]](#)
2. Saadat, N.P.; Nies, T.; Rousset, Y.; Ebenhöf, O. Thermodynamic limits and optimality of microbial growth. *Entropy* **2020**, *22*, 277. [\[CrossRef\]](#)
3. Buchholz, K.; Collins, J. The roots—A short history of industrial microbiology and biotechnology. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 3747–3762. [\[CrossRef\]](#)
4. Waters, A.L.; Hill, R.T.; Place, A.R.; Hamann, M.T. The expanding role of marine microbes in pharmaceutical development. *Curr. Opin. Biotechnol.* **2010**, *21*, 780–786. [\[CrossRef\]](#)
5. Thodey, K.; Galanie, S.; Smolke, C.D. A microbial biomanufacturing platform for natural and semisynthetic opioids. *Nat. Chem. Biol.* **2014**, *10*, 837–844. [\[CrossRef\]](#)
6. Das, S.; Dash, H.R. Microbial bioremediation: A potential tool for restoration of contaminated areas. In *Microbial Biodegradation and Bioremediation*; Elsevier: Amsterdam, The Netherlands, 2014; pp. 1–21. [\[CrossRef\]](#)
7. Downs, D.M. Understanding microbial metabolism. *Annu. Rev. Microbiol.* **2006**, *60*, 533–559. [\[CrossRef\]](#)
8. Heinemann, M.; Sauer, U. Systems biology of microbial metabolism. *Curr. Opin. Microbiol.* **2010**, *13*, 337–343. [\[CrossRef\]](#)
9. Chubukov, V.; Gerosa, L.; Kochanowski, K.; Sauer, U. Coordination of microbial metabolism. *Nat. Rev. Microbiol.* **2014**, *12*, 327–340. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Roels, J. Application of macroscopic principles to microbial metabolism. *Biotechnol. Bioeng.* **1980**, *22*, 2457–2514. [\[CrossRef\]](#)
11. Heijnen, J.; Van Dijken, J. In search of a thermodynamic description of biomass yields for the chemotrophic growth of microorganisms. *Biotechnol. Bioeng.* **1992**, *39*, 833–858. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Heijnen, S.J. Thermodynamics of microbial growth and its implications for process design. *Trends Biotechnol.* **1994**, *12*, 483–492. [\[CrossRef\]](#)

13. Von Stockar, U. *Biothermodynamics: The Role of Thermodynamics in Biochemical Engineering*; PPUR Presses Polytechniques: Lausanne, Switzerland, 2013.
14. Wilken, S.E.; Frazão, V.V.; Saadat, N.P.; Ebenhöf, O. The view of microbes as energy converters illustrates the trade-off between growth rate and yield. *Biochem. Soc. Trans.* **2021**, *49*, 1663–1674. [[CrossRef](#)] [[PubMed](#)]
15. Von Stockar, U.; Liu, J. Does microbial life always feed on negative entropy? Thermodynamic analysis of microbial growth. *Biochim. Biophys. Acta* **1999**, *1412*, 191–211. [[CrossRef](#)] [[PubMed](#)]
16. Onsager, L. Reciprocal Relations in Irreversible Processes. I. *Phys. Rev.* **1931**, *37*, 405–426. [[CrossRef](#)]
17. Stucki, J.W. The optimal efficiency and the economic degrees of coupling of oxidative phosphorylation. *Eur. J. Biochem.* **1980**, *109*, 269–283. [[CrossRef](#)]
18. Von Stockar, U.; Vojinović, V.; Maskow, T.; Liu, J. Can microbial growth yield be estimated using simple thermodynamic analogies to technical processes? *Chem. Eng. Process. Process. Intensif.* **2008**, *47*, 980–990. [[CrossRef](#)]
19. Von Stockar, U. Biothermodynamics of live cells: A tool for biotechnology and biochemical engineering. *J. Non-Equilib. Thermodyn.* **2010**, *35*, 415–475. [[CrossRef](#)]
20. Mukherjee, S.; Stamatis, D.; Li, C.; Ovchinnikova, G.; Bertsch, J.; Sundaramurthi, J.; Kandimalla, M.; Nicolopoulos, P.; Favognano, A.; Chen, I.M.; et al. Twenty-five years of Genomes OnLine Database (GOLD): Data updates and new features in v.9. *Nucleic Acids Res.* **2022**, *51*, D957–D963. [[CrossRef](#)] [[PubMed](#)]
21. Francke, C.; Siezen, R.J.; Teusink, B. Reconstructing the metabolic network of a bacterium from its genome. *Trends Microbiol.* **2005**, *13*, 550–558. [[CrossRef](#)]
22. Machado, D.; Andrejev, S.; Tramontano, M.; Patil, K.R. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Res.* **2018**, *46*, 7542–7553. [[CrossRef](#)]
23. Mendoza, S.N.; Olivier, B.G.; Molenaar, D.; Teusink, B. A systematic assessment of current genome-scale metabolic reconstruction tools. *Genome Biol.* **2019**, *20*, 1–20. [[CrossRef](#)]
24. Saadat, N.P.; van Aalst, M.; Ebenhöf, O. Network reconstruction and modelling made reproducible with moped. *Metabolites* **2022**, *12*, 275. [[CrossRef](#)] [[PubMed](#)]
25. Orth, J.D.; Thiele, I.; Palsson, B.Ø. What is flux balance analysis? *Nat. Biotechnol.* **2010**, *28*, 245–248. [[CrossRef](#)] [[PubMed](#)]
26. Schuster, S.; Hilgetag, C. On elementary flux modes in biochemical reaction systems at steady state. *J. Biol. Syst.* **1994**, *2*, 165–182. [[CrossRef](#)]
27. Poolman, M.G.; Sebu, C.; Pidcock, M.K.; Fell, D.A. Modular decomposition of metabolic systems via null-space analysis. *J. Theor. Biol.* **2007**, *249*, 691–705. [[CrossRef](#)] [[PubMed](#)]
28. Trinh, C.T.; Wlaschin, A.; Sreenc, F. Elementary mode analysis: A useful metabolic pathway analysis tool for characterizing cellular metabolism. *Appl. Microbiol. Biotechnol.* **2008**, *81*, 813–826. [[CrossRef](#)] [[PubMed](#)]
29. Klamt, S.; Regensburger, G.; Gerstl, M.P.; Jungreuthmayer, C.; Schuster, S.; Mahadevan, R.; Zanghellini, J.; Müller, S. From elementary flux modes to elementary flux vectors: Metabolic pathway analysis with arbitrary linear flux constraints. *PLoS Comput. Biol.* **2017**, *13*, e1005409. [[CrossRef](#)] [[PubMed](#)]
30. Klamt, S.; Stelling, J. Combinatorial complexity of pathway analysis in metabolic networks. *Mol. Biol. Rep.* **2002**, *29*, 233–236. [[CrossRef](#)] [[PubMed](#)]
31. Urbanczik, R.; Wagner, C. Functional stoichiometric analysis of metabolic networks. *Bioinformatics* **2005**, *21*, 4176–4180. [[CrossRef](#)]
32. Clement, T.J.; Baalhuis, E.B.; Teusink, B.; Bruggeman, F.J.; Planqué, R.; de Groot, D.H. Unlocking elementary conversion modes: Ecmtool unveils all capabilities of metabolic networks. *Patterns* **2021**, *2*, 100177. [[CrossRef](#)]
33. Buchner, B.; Clement, T.J.; de Groot, D.H.; Zanghellini, J. ecmtool: Fast and memory-efficient enumeration of elementary conversion modes. *Bioinformatics* **2023**, *39*, btad095. [[CrossRef](#)] [[PubMed](#)]
34. Beber, M.E.; Gollub, M.G.; Mozaffari, D.; Shebek, K.M.; Flamholz, A.I.; Milo, R.; Noor, E. eQuilibrator 3.0: A database solution for thermodynamic constant estimation. *Nucleic Acids Res.* **2022**, *50*, D603–D609. [[CrossRef](#)] [[PubMed](#)]
35. Noor, E.; Haraldsdóttir, H.S.; Milo, R.; Fleming, R.M. Consistent estimation of Gibbs energy using component contributions. *PLoS Comput. Biol.* **2013**, *9*, e1003098. [[CrossRef](#)] [[PubMed](#)]
36. Rieger, M.; Käppeli, O.; Fiechter, A. The Role of Limited Respiration in the Incomplete Oxidation of Glucose by *Saccharomyces Cerevisiae*. *Microbiology* **1983**, *129*, 653–661. [[CrossRef](#)]
37. Kayser, A.; Weber, J.; Hecht, V.; Rinas, U. Metabolic flux analysis of *Escherichia coli* in glucose-limited continuous culture. I. Growth-rate-dependent metabolic efficiency at steady state. *Microbiology* **2005**, *151*, 693–706. [[CrossRef](#)] [[PubMed](#)]
38. Orth, J.D.; Fleming, R.M.T.; Palsson, B.O. Reconstruction and Use of Microbial Metabolic Networks: The Core *Escherichia coli* Metabolic Model as an Educational Guide. *EcoSal Plus* **2010**, *4*. [[CrossRef](#)] [[PubMed](#)]
39. Reed, J.L.; Vo, T.D.; Schilling, C.H.; Palsson, B.O. An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biol.* **2003**, *4*, R54. [[CrossRef](#)] [[PubMed](#)]
40. Duarte, N.C.; Herrgård, M.J.; Palsson, B.Ø. Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model. *Genome Res.* **2004**, *14*, 1298–1309. [[CrossRef](#)]
41. Lehtovirta-Morley, L.E. Ammonia oxidation: Ecology, physiology, biochemistry and why they must all come together. *FEMS Microbiol. Lett.* **2018**, *365*, fny058. [[CrossRef](#)]
42. Mavrovouniotis, M.L. Group contributions for estimating standard gibbs energies of formation of biochemical compounds in aqueous solution. *Biotechnol. Bioeng.* **1990**, *36*, 1070–1082. [[CrossRef](#)]

43. Tran, Q.H.; Uden, G. Changes in the proton potential and the cellular energetics of *Escherichia coli* during growth by aerobic and anaerobic respiration or by fermentation. *Eur. J. Biochem.* **1998**, *251*, 538–543. [\[CrossRef\]](#)
44. Battley, E.H. Calculation of entropy change accompanying growth of *Escherichia coli* K-12 on succinic acid. *Biotechnol. Bioeng.* **1993**, *41*, 422–428. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Monod, J. La technique de culture continue: Theorie et applications. *Sel. Pap. Mol. Biol. Jacques Monod* **1978**, *79*, 390–410.
46. Hoskisson, P.A.; Hobbs, G. Continuous culture—making a comeback? *Microbiology* **2005**, *151*, 3153–3159. [\[CrossRef\]](#) [\[PubMed\]](#)
47. De Groot, D.H.; Lischke, J.; Muolo, R.; Planqué, R.; Bruggeman, F.J.; Teusink, B. The common message of constraint-based optimization approaches: Overflow metabolism is caused by two growth-limiting constraints. *Cell. Mol. Life Sci.* **2019**, *77*, 441–453. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Niebel, B.; Leupold, S.; Heinemann, M. An upper limit on Gibbs energy dissipation governs cellular metabolism. *Nat. Metab.* **2019**, *1*, 125–132. [\[CrossRef\]](#) [\[PubMed\]](#)
49. Stryjewska, A.; Kiepusa, K.; Librowski, T.; Lochyński, S. Biotechnology and genetic engineering in the new drug development. Part I. DNA technology and recombinant proteins. *Pharmacol. Rep.* **2013**, *65*, 1075–1085. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Verma, S.; Kuila, A. Bioremediation of heavy metals by microbial process. *Environ. Technol. Innov.* **2019**, *14*, 100369. [\[CrossRef\]](#)
51. Stabnikov, V.; Ivanov, V.; Chu, J. Construction Biotechnology: A new area of biotechnological research and applications. *World J. Microbiol. Biotechnol.* **2015**, *31*, 1303–1314. [\[CrossRef\]](#)
52. Fang, X.; Lloyd, C.J.; Palsson, B.O. Reconstructing organisms in silico: Genome-scale models and their emerging applications. *Nat. Rev. Microbiol.* **2020**, *18*, 731–743. [\[CrossRef\]](#)
53. Rupp, E.; Papin, J.A.; de Figueiredo, L.F.; Schuster, S. Metabolic reconstruction, constraint-based analysis and game theory to probe genome-scale metabolic networks. *Curr. Opin. Biotechnol.* **2010**, *21*, 502–510. [\[CrossRef\]](#)
54. Bonetta, R.; Valentino, G. Machine learning techniques for protein function prediction. *Proteins Struct. Funct. Bioinform.* **2020**, *88*, 397–413. [\[CrossRef\]](#)
55. Werner, S.; Diekert, G.; Schuster, S. Revisiting the thermodynamic theory of optimal ATP stoichiometries by analysis of various ATP-producing metabolic pathways. *J. Mol. Evol.* **2010**, *71*, 346–355. [\[CrossRef\]](#)
56. Westerhoff, H.V.; Lolkema, J.S.; Otto, R.; Hellingwerf, K.J. Thermodynamics of growth non-equilibrium thermodynamics of bacterial growth the phenomenological and the Mosaic approach. *Biochim. Biophys. Acta BBA-Rev. Bioenerg.* **1982**, *683*, 181–220. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Berthelsen, F.H.; Nielsen, U.D. Prediction of ships' speed-power relationship at speed intervals below the design speed. *Transp. Res. Part D Transp. Environ.* **2021**, *99*, 102996. [\[CrossRef\]](#)
58. MacKay, D.J. *Sustainable Energy-without the Hot Air*; Bloomsbury Publishing: London, UK, 2016.
59. Harder, B.J.; Bettenbrock, K.; Klamt, S. Model-based metabolic engineering enables high yield itaconic acid production by *Escherichia coli*. *Metab. Eng.* **2016**, *38*, 29–37. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Du, W.; Jongbloets, J.A.; van Boxtel, C.; Pineda Hernández, H.; Lips, D.; Oliver, B.G.; Hellingwerf, K.J.; Branco dos Santos, F. Alignment of microbial fitness with engineered product formation: Obligatory coupling between acetate production and photoautotrophic growth. *Biotechnol. Biofuels* **2018**, *11*, 38. [\[CrossRef\]](#) [\[PubMed\]](#)
61. Boecker, S.; Schulze, P.; Klamt, S. Growth-coupled anaerobic production of isobutanol from glucose in minimal medium with *Escherichia coli*. *Biotechnol. Biofuels Bioprod.* **2023**, *16*, 148. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.