POLITECNICO DI TORINO
Repository ISTITUZIONALE

A yield-cost tradeoff governs Escherichia coli's decision between fermentation and respiration in carbon-limited growth

Original
A yield-cost tradeoff governs Escherichia coli's decision between fermentation and respiration in carbon-limited growth / Mori, M.; Marinari, E.; De Martino, A.. - In: NPJ SYSTEMS BIOLOGY AND APPLICATIONS. - ISSN 2056-7189. - 5:1(2019), p. 16. [10.1038/s41540-019-0093-4]

Availability:
This version is available at: 11583/2976740 since: 2023-03-10T09:00:22Z

Publisher:
Nature Publishing Group

Published
DOI:10.1038/s41540-019-0093-4

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

12 October 2024

npj | Systems Biology and Applications

www.nature.com/npjsba

ARTICLE OPEN

A yield-cost tradeoff governs *Escherichia coli*'s decision between fermentation and respiration in carbon-limited growth

Matteo Mori[1], Enzo Marinari[2,3] and Andrea De Martino[4,5]

Living cells react to changes in growth conditions by re-shaping their proteome. This accounts for different stress-response strategies, both specific (i.e., aimed at increasing the availability of stress-mitigating proteins) and systemic (such as large-scale changes in the use of metabolic pathways aimed at a more efficient exploitation of resources). Proteome re-allocation can, however, imply significant biosynthetic costs. Whether and how such costs impact the growth performance are largely open problems. Focusing on carbon-limited *E. coli* growth, we integrate genome-scale modeling and proteomic data to address these questions at quantitative level. After deriving a simple formula linking growth rate, carbon intake, and biosynthetic costs, we show that optimal growth results from the tradeoff between yield maximization and protein burden minimization. Empirical data confirm that *E. coli* growth is indeed close to Pareto-optimal over a broad range of growth rates. Moreover, we establish that, while most of the intaken carbon is diverted into biomass precursors, the efficiency of ATP synthesis is the key driver of the yield-cost tradeoff. These findings provide a quantitative perspective on carbon overflow, the origin of growth laws and the multidimensional optimality of *E. coli* metabolism.

npj Systems Biology and Applications (2019)5:16 ; https://doi.org/10.1038/s41540-019-0093-4

INTRODUCTION

The physiology of cell growth and its relation with gene expression and metabolism have been studied quantitatively for decades. Refining the picture developed since the 1950s,[1,2] changes in growth conditions have been found to be accompanied by a re-organization of the cellular proteome, with resources being redistributed among different tasks (nutrient import, biosynthesis, etc.).[3–8] For *E. coli*, this adaptation underlies significant adjustments in cellular energetics and metabolism, with the 'acetate switch'[9–11] being the most apparent manifestation of the interplay between metabolism and gene expression. Slowly growing *E. coli* cells tend to operate close to the theoretical limit of maximum growth yield (growth rate per unit of intaken carbon);[12–15] at fast growth, instead, they display lower yields with the excretion of carbon equivalents (acetate). The transition between the two conditions is accompanied by large changes in gene expression.[6,11] A similar scenario also occurs in other proliferating cell types, like yeast[16,17] or cancer cells.[18–21] Several phenomenological theories address the question of how metabolism and gene expression coordinate to optimize growth in bacteria,[4,11,22–27] while mechanistic models of metabolism can describe changes in the usage of pathways at the crossover.[28–30] Here we combine genome-scale modeling with experimental data analysis to obtain a direct quantitative characterization of the yield-cost tradeoff in *E. coli*.

We first show that, under carbon limitation, the growth rate can be explicitly decomposed into inter-dependent contributions from the growth yield and from growth-associated protein costs. Maximum growth then implies a tradeoff between the growth yield and the proteome fraction allocated to metabolic enzymes. At fast (slow) growth, the latter (former) is optimized, and the objective function changes smoothly as carbon availability is limited and one passes from one regime to the other. We quantitatively evaluated this scenario in a genome-scale model of *E. coli*'s metabolism. In this framework, maximum growth leads to a Pareto front in the yield-cost landscape. Remarkably, an analysis of proteomic data for carbon-limited cells[6] validates this picture. To get insight into which metabolic variable drives the crossover, we further quantified the partitioning of carbon into biomass- and energy-producing pathways, both of which contribute to growth. While the relative amount of carbon converted into biomass precursors significantly exceeds that going into energy (i.e., ATP), the latter accounts for almost all of the observed change in growth yield at the crossover, implicating the ATP yield as the key biological driver of the acetate switch. The behavior of the corresponding proteome mass fractions extrapolated from data confirms this conclusion.

These results provide a quantitative characterization of the multidimensional optimality of *E. coli*[13,31] that directly addresses the crosstalk between growth and gene expression by connecting metabolic flux analysis to proteomic data.

RESULTS

Increasing the growth yield as a basic response to carbon shortage

Bacterial growth is known to be hindered by the synthesis of inefficient proteins.[3] Because of this, modulating the proteome composition is a major fitness strategy for exponentially growing

[1]Department of Physics, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA; [2]Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 2, Rome 00185, Italy; [3]INFN, Sezione di Roma 1, Piazzale Aldo Moro 2, Rome 00185, Italy; [4]Soft & Living Matter Lab, Institute of Nanotechnology (CNR-NANOTEC), c/o Dipartimento di Fisica, Sapienza Università di Roma, Piazzale Aldo Moro 2, Rome 00185, Italy and [5]Italian Institute for Genomic Medicine, via Nizza 52, Turin 10126, Italy
Correspondence: Andrea De Martino (andrea.demartino@roma1.infn.it)
These authors contributed equally: Enzo Marinari and Andrea De Martino

Received: 24 November 2018 Accepted: 12 April 2019
Published online: 01 May 2019

Published in partnership with the Systems Biology Institute

npj nature partner journals

bacteria. In *E. coli*, for instance, a substantial reshaping of the protein repertoire takes place upon carbon limitation, with ribosome-affiliated proteins and biosynthetic enzymes taking up a smaller fraction of the proteome as the growth rate μ decreases, while carbon-import and motility proteins are upregulated.[3,5,6,8] Similar changes are observed in cells subject to other modes of limitation.[6] Such modifications indicate major shifts in growth strategy in response to stress.

For a minimal phenomenological description of this fact, let J_C denote the cellular carbon intake flux (in mmol/g_{DW}h), which we assume to be subject to limitation (see Supplementary Text, Table S1 for a list of symbols used in this work). The carbon influx J_C is related to the growth rate μ via the specific intake flux q (i.e., $J_C = q \cdot \mu$), representing the amount of in-taken carbon needed to generate a unit of biomass (in units of mmol per gram of dry weight). q depends not only on the specific nutrient source, but more in general on the growth conditions (e.g., oxygen availability, osmolarity, pH). The inverse of the specific intake flux is proportional to the growth yield, which is usually expressed in terms of grams of dry weight per gram of intaken carbon; the specific carbon intake q is, however, more convenient for our purposes. For fixed specific intake q, as carbon is limited and the stress level increases, J_C decreases and μ is proportionally reduced. In turn, the proteome share ϕ_C allocated to the carbon import system will also be affected, as more proteins devoted to importing carbon will generically be needed to sustain a prescribed rate J_C under limitation. The cost associated to this macromolecular burden can be modeled via a coefficient $w_C > 0$ constraining the catabolic proteins and the carbon intake, $\phi_C = w_C \cdot J_C$, and serving as a proxy for the level of stress to which the cell is subject. Putting things together, the growth rate μ can be expressed as

$$\mu = \frac{\phi_C}{q \cdot w_C}. \tag{1}$$

Eq. (1) suggests that cells can counteract an increase of w_C (i.e., higher stress associated to carbon limitation) in two ways (Fig. 1a). The first is by increasing the protein fraction associated to carbon transport, ϕ_C: indeed, *E. coli* expresses a whole array of catabolic systems in response to carbon shortage.[5,6] The second is by decreasing the specific intake rate q, i.e. increasing the growth yield. The switch from fermentative to respiratory metabolism upon carbon limitation is an example of this type of response,[9–11] with the specific intake q reaching a minimum when acetate excretion flux is negligible. *E. coli* adopts both strategies under carbon limitation;[6] the key difference between them is that, while the former only involves a specific set of processes (i.e., those involved in carbon uptake), the metabolic rewiring needed to decrease the specific intake q implies the coordination of expression levels across multiple metabolic pathways, even not directly subject to limitation.

The role of protein costs

Any increase in ϕ_C under carbon limitation has to be compensated by a reduction of the share of the rest of the proteome (Fig. 1a). In *E. coli*, the mass fraction of many biosynthetic proteins, from anabolism to protein synthesis, indeed decreases with the growth rate in an approximately linear fashion.[3,5,6] For simplicity, let us assume that the proteome fraction devoted activities other than carbon-import, which we denote as $\phi_{NC} \equiv 1 - \phi_C$, adjusts with the growth rate so that $\phi_{NC} = \phi_{NC,0} + \varepsilon_{NC}\mu$, with an offset term $\phi_{NC,0}$ and a slope ε_{NC} (see Supplementary Text, Note 1 for a more general formulation). When plugged in Eq. (1), this results in

$$\mu = \frac{1 - \phi_{NC,0}}{\varepsilon_{NC} + q \cdot w_C}. \tag{2}$$

The term $\phi_{NC,0}$ includes both constitutively expressed, "unnecessary" proteins (not required to grow in the given medium but important in fluctuating conditions[32]) and the basal abundance of metabolic enzymes. The coefficient ε_{NC}, governing the growth-dependent part of the proteome, depends instead on the expression level of metabolic enzymes as well as on the efficiency of the metabolic pathways, including energy generation and anabolic processes.[5,6,30] Because changes in proteome allocation impact the efficiency of carbon utilization, q and ε_{NC} cannot be varied independently. This will lead to tradeoffs in Eq. (2). Under sever carbon shortage (high stress, large w_C), μ is dominated by the carbon efficiency term, i.e., $\mu \propto (q \cdot w_C)^{-1}$ as in (1), and maximum growth coincides with minimum q. In carbon-rich conditions (low stress, small w_C), instead, $\mu \propto 1/\varepsilon_{NC}$ and maximum growth is achieved by minimizing the proteome cost of the other metabolic activities.

Eq. (2) is purely phenomenological. We will now re-derive it within a genome-scale model of *E. coli* metabolism based on empirical growth laws.[4] This will allow us to quantitatively describe the changes in cellular metabolism and proteome organization that occur under carbon limitation.

Empirical growth laws quantitatively link growth rate, specific carbon intake, and enzyme costs

Multiple studies have shown that, in carbon-limited growth, *E. coli*'s proteome is partitioned into sectors whose relative sizes adjust with the growth conditions.[3,5,6] For our purposes, the simplest partition involves four sectors, three responding to the growth condition [ribosome-affiliated proteins (R), metabolic enzymes (E) and proteins involved in the carbon-uptake system (C)], and the fourth with a fixed proteome share ('housekeeping' sector Q). Assuming proportionality to enzyme expression, metabolic fluxes effectively mediate proteome re-shaping.[3,5,30] In this scenario the normalization of proteome mass fractions takes the form (Fig. 1b and Methods)

$$w_C J_C + \sum_{i \in E} w_i |v_i| + w_R \mu = \phi_{max}, \tag{3}$$

where J_C is the carbon intake flux, v_i is the flux of reaction i, and ϕ_{max} is a constant set by the baseline expression levels of all proteins. The terms on the left-hand side of (3) correspond to the condition-dependent C-, E- and R-sectors. In particular (see previous sections), $w_C J_C$ describes the proteome share devoted to the limited activity, with w_C quantifying the level of nutritional stress. The R- and E-sectors, instead, form the non-limited proteome. The coefficients w_i hence represent the 'cost' of each reaction $i \in E$ in terms of the proteome share to be allocated to its enzyme per unit of net flux, while w_R quantifies the proteome share of ribosome-affiliated proteins per unit of growth rate.

Eq. (3) implies that increased nutrient stress (i.e. larger w_C) causes a flux-mediated re-shaping of the cellular proteome that in turn affects the growth rate (Fig. 1c). J_C is related to μ by the specific intake rate $q = J_C/\mu$; likewise, we define the specific protein cost $\varepsilon = \sum_{i \in E} w_i |v_i|/\mu$. Using these definitions one can isolate μ from (3), obtaining

$$\mu = \frac{\phi_{max}}{\varepsilon + w_C q + w_R}. \tag{4}$$

Eq. (4) is the quantitative counterpart of Eq. (2). As in the latter, the interplay between carbon and enzyme efficiency is controlled by the stress level w_C, while w_R and ϕ_{max} are parameters that can be derived from experiments. Note that, by definition, q and ε do not depend separately on the flux vector $\mathbf{v} = (J_C, \{v_i\})$ and on the growth rate μ but, rather, are functions of the ratio $\mathbf{v}/\mu \equiv \boldsymbol{\xi}$, a vector whose entries ξ_i give the contribution of reaction i to the cellular biomass in terms of mmol of metabolites transformed per gram of dry cell. Hence $\boldsymbol{\xi}$ characterizes the "metabolic state" of the

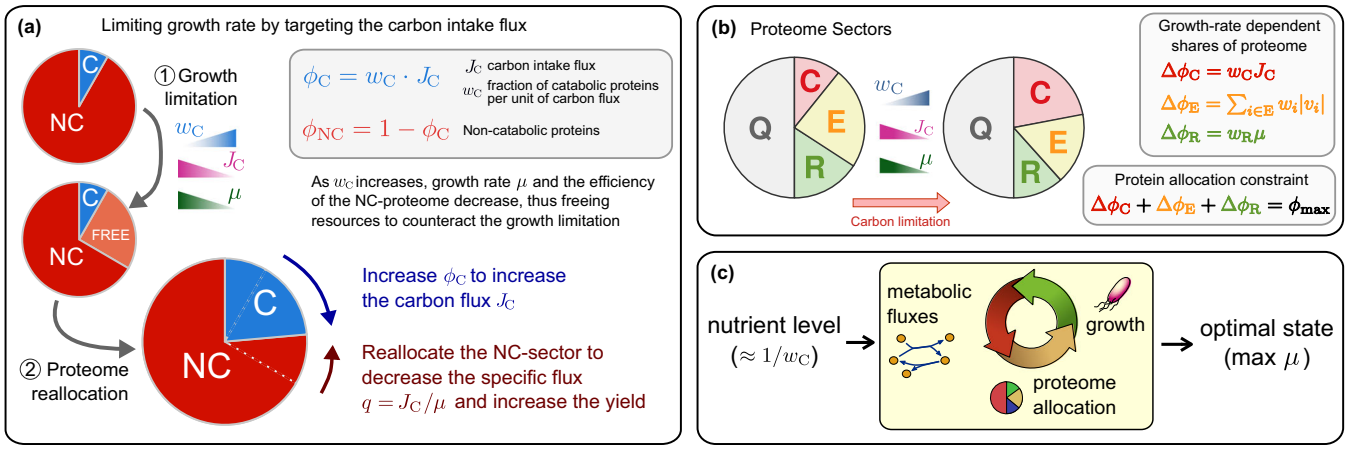

Fig. 1 Schematic view of the proteome allocation problem. **a** Effect of growth limitation via carbon starvation on proteome composition. An increase in the cost of the catabolic (C) sector reduces the growth rate. Part of the non-catabolic (NC) sector, no longer required at slower growth, is freed and can be reallocated. The decrease in growth can be counteracted by expanding the C-sector (i.e., by increasing ϕ_C) and/or by allocating part of the NC-sector to re-organize metabolism so as to decrease the specific carbon intake $q = J_C/\mu$. **b** Proteome sectors considered for *E. coli*, namely ribosomal (R), enzymatic (E), carbon-import (C), and housekeeping core (Q). Following ref. [30] we assume that all sectors but the core have condition-dependent parts denoted by $\Delta\phi_j$. By normalization, their sums are constrained as in Eq. (3). **c** A change in the nutrient level affects the cellular metabolism and gene expression. In Constrained Allocation FBA,[30] the interplay between metabolic fluxes, protein levels, and growth rate leads to optimal phenotypes with maximum growth rate

cell, and the vector $\mathbf{v} = \mu\boldsymbol{\xi}$ can in turn be seen as a parametrization of metabolic fluxes in terms of the state $\boldsymbol{\xi}$ and of the growth rate μ. From the point of view of regulation, instead, the vector $\boldsymbol{\xi}$ represents a regulation program, prescribing the growth-dependence of enzyme mass fractions.

Note that this linear parametrization of the fluxes is correct at sufficiently fast growth, i.e., for $\mu \gtrsim 0.5$/h (Methods). This range includes the acetate switch and is the condition on which we focus here. At slower rates, the approximation will deteriorate due to the presence of non-negligible maintenance fluxes (Methods).

In silico states of maximum growth display a Pareto optimal tradeoff between specific intake and protein costs

Based on (4), maximizing μ is equivalent to minimizing the total cost

$$\mathcal{C}(\boldsymbol{\xi}) = \varepsilon(\boldsymbol{\xi}) + w_C\, q(\boldsymbol{\xi}). \tag{5}$$

More precisely, for each value of w_C, finding the flux vector $\mathbf{v}$ that maximizes μ subject to (3) is equivalent to finding the metabolic state $\boldsymbol{\xi}$ minimizing (5) (Supplementary Text Fig. S1). In other terms, the constraint (3) associates to each state $\boldsymbol{\xi}$ a growth rate $\mu(\boldsymbol{\xi})$ via (4). Such a $\mu(\boldsymbol{\xi})$ will be highest when $\boldsymbol{\xi}$ minimizes (5). Fluxes are then given by $\mathbf{v}(\boldsymbol{\xi}) = \boldsymbol{\xi}\mu(\boldsymbol{\xi})$ and satisfy the constraint (3) by construction. As w_C varies, cells must strike a compromise between the specific intake q and biosynthetic costs ε in order to maximize μ. The ensuing tradeoff between q and ε can be described by a Pareto front that separates an accessible region of the (q, ε) plane, such that each point lying therein represents a viable state $\boldsymbol{\xi}$, from an inaccessible one, with states of maximum growth lying on the front (Fig. 2a). μ increases as one moves along the Pareto front towards larger values of specific intake q, corresponding to lower growth yields.

Pareto-optimal and sub-optimal states for *E. coli* can be generated using a genome-scale model of metabolism (Fig. 2b, Methods and Supplementary Text, Note 2). Sub-optimal values of q and ε expectedly lie in the feasible region to the right of the Pareto front (black line). Markers in Fig. 2b, varying from blue (mostly respirative, lower q) to red (mostly fermentative, lower ε) show that *E. coli*'s metabolism is robustly characterized by respiration at slow growth and upregulated fermentation at fast growth, in agreement with experiments.

Empirical data support Pareto-optimality

The above picture can be validated by estimating ε from mass spectrometry data[6] and q from genome scale-modeling constrained by the empirical acetate overflow (Methods). The resulting curve in the (q, ε) plane (Fig. 2c) qualitatively agrees with the theory. At a quantitative level, the normalized protein cost ε predicted in silico for low-q states matches the observed enzyme cost at $\mu \simeq 0.7$/h. For faster growth, where acetate excretion sets in, our model underestimates the decrease in ε by only about 10%, most likely due to the reduced efficiency of respiration enzymes.[11] Likewise, at slow growth (below about 0.6/h) our prediction appears to underestimate ε, possibly due to the decrease in enzyme efficiencies that is known to set in at low μ[29,33–36] and which is not accounted for in our constraint-based framework.

The ATP yield is the key driver of the tradeoff

The structure of metabolic states shown in Fig. 2b suggests a tight link between growth and energy metabolism, as fast (slow) growth favors the use of low (high) ATP yield pathways.[11] Whether the efficiency of energy biosynthesis is the driver of carbon overflow is however not obvious from Eq. (5). In schematic terms (Fig. 3a), one can think that cells channel the in-taken carbon through pathways contributing respectively to the overall production of energy and to the synthesis of biomass precursors. However, about 1/3 of the energy production sustains anabolic processes underlying the synthesis of biomass building blocks (Supplementary Text Fig. S2), and therefore the corresponding carbon flux can also be associated to the biosynthesis of biomass precursors. Hence, we focus on the fluxes denoted respectively as $J_{C\to B}$ and $J_{C\to E}$ in Fig. 3a, corresponding to the carbon flux associated to biomass biosynthesis and the carbon flux associated to the production of energy feeding the cellular processes downstream of metabolism, e.g., protein synthesis, respectively. The two fluxes satisfy

$$J_C = J_{C\to B} + J_{C\to E}, \tag{6}$$

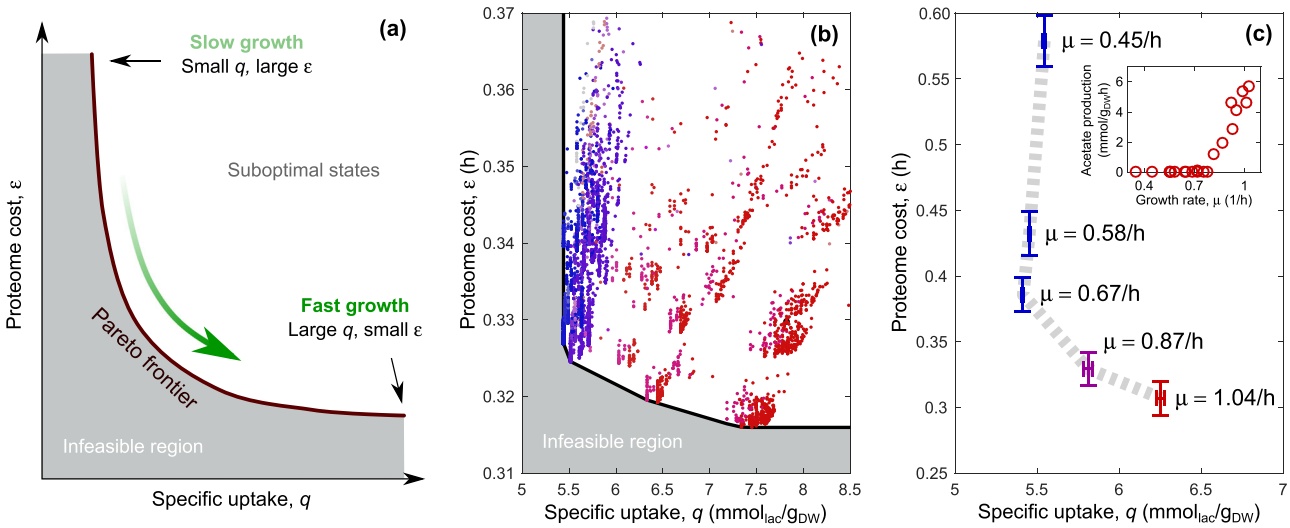

Fig. 2 Tradeoff between minimum specific carbon intake and minimum enzyme cost in *E. coli*. **a** Multi-objective optimality and Pareto front. Two cost functions (specific carbon uptake q and specific proteome cost ε) are shown, together with the feasible (in white) and infeasible (in gray) regions, separated by the Pareto front. Optimal (i.e., growth maximizing) solutions lie on the latter. At slow growth, the optimal metabolic state corresponds to small q and large ε. As the growth rate increases (green arrow), optimal states move along the frontier, increasing q and reducing ε. **b** In silico prediction for optimal *E. coli* growth on lactose-limited minimal medium. The black line corresponds to the computed Pareto front, while individual points in the feasible region describe sub-optimal solutions. Blue (red) markers represent solutions dominated by respiration (fermentation), while purple markers denote mixtures. **c** *E. coli* states obtained by integrating mass spectroscopy data for lactose-limited growth from ref. [6] with in silico yields qualitatively reproduce (with quantitative accuracy for the yield) the predicted Pareto front. The values of μ reported next to the experimental points represent the experimental growth rates; colors represent the switch between respiration (in blue) and fermentation (in red), as determined experimentally (inset, data from Basan et al.[11]). Error bars represent the standard deviation

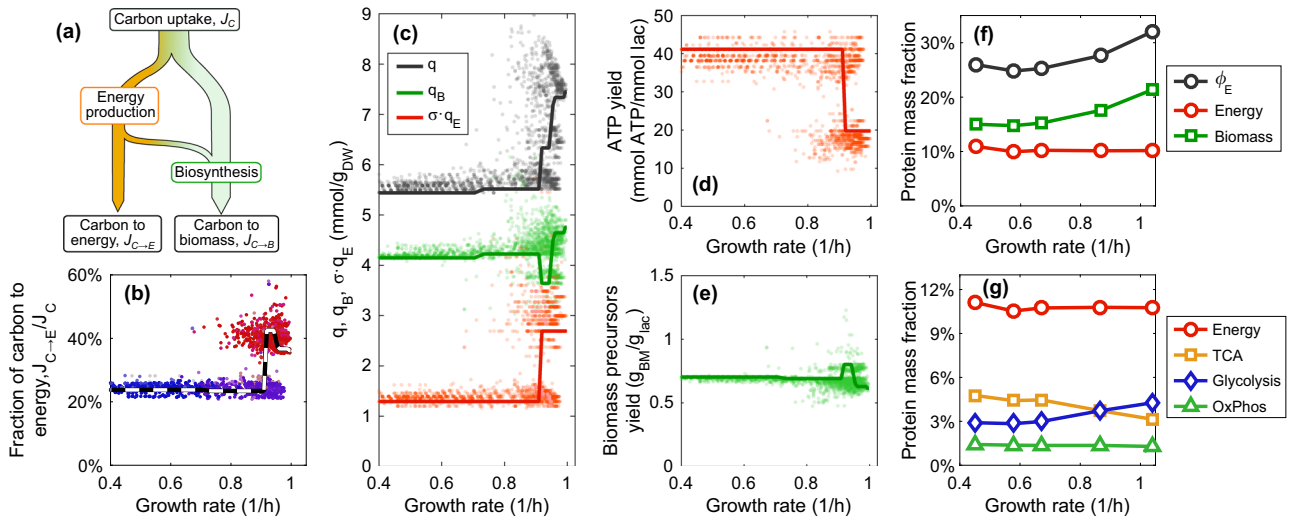

Fig. 3 Carbon partitioning and the role of the ATP yield. **a** Scheme of the partitioning of the carbon intake flux into energy production and biomass precursors. Some of the energy produced by the cell is used for the synthesis of biomass precursors, and is hence accounted for in the overall precursors biosynthesis flux (carbon flux $J_{C\to B}$); the rest of the carbon ($J_{C\to E}$) is spent to generate the energy necessary for biosynthetic activities downstream of metabolism, e.g., translation. **b** The relative carbon-to-energy flux estimated by sampling a genome-scale metabolic model roughly varies between 25 and 40% of the overall carbon influx. Similarly to Fig. 2b, the color indicates the type of energetic metabolism (blue, respiration; red, fermentation), while the dashed line indicates the optimal solution corresponding to the Pareto front. **c** Behavior of the specific uptake q and of its biomass- (q_B) and energy-related components (σq_E), see Eq. (7). The change in q is mostly accounted for by the change in σq_E. Solid lines represent the Pareto-optimal solution. **d** ATP yield, $1/q_E$, and **e** biomass yield as a function of μ. The sharp decrease in the former corresponds to the respiration-to-fermentation crossover leading to carbon overflow. Solid lines represent the Pareto-optimal solution. The biomass yield is given by the average molar mass of the biomass components, over q_B. **f** Mass fractions of proteins associated to energy and biomass synthesis within the enzymatic sector (fraction ϕ_E). **g** The respiration-to-fermentation crossover occurs within the energy-related mass fraction, which remains roughly constant as μ is changed. Proteomics data from ref. [6]

with J_C being the total the carbon import rate. Based on (6), we can write the specific intake $q = J_C/\mu$ as

$$q = q_B + \sigma q_E. \tag{7}$$

The quantity $q_B = J_{C\to B}/\mu$ represents the amount of carbon used to synthesize biomass precursors (in units of mmol of carbon per gram of dry weight). The term $q_E = J_{C\to E}/(\sigma\mu)$ instead stands for the carbon intake associated to ATP production (i.e., the inverse number of ATP molecules regenerated per lactose molecule), where we accounted for the fact that at fast growth the overall energy requirement J_E of processes downstream of metabolism (e.g. translation) is proportional to the growth rate μ via a fixed growth-associated maintenance rate σ, i.e. $J_E = \sigma\mu$ (Methods). To assess the partitioning of carbon into energy production and biomass precursors one has to estimate q_E and q_B. This can be done using a genome-scale model of *E. coli*'s metabolic network upon generalizing the parametrization $\mathbf{v} = \mu\boldsymbol{\xi}$ to the linear combination $\mathbf{v} = J_E\boldsymbol{\eta} + \mu\boldsymbol{\beta}$, where the vectors $\boldsymbol{\eta}$ and $\boldsymbol{\beta}$ can be univocally determined by perturbing the energetic demand of the cell (Supplementary Text, Note 3). In this setting, the metabolic state of the cell is described by the pair of vectors ($\boldsymbol{\eta}$, $\boldsymbol{\beta}$) rather than by the single vector $\boldsymbol{\xi}$, but it is still possible to compute optimal and sub-optimal states together with the corresponding values of q_B and q_E (Supplementary Text, Note 3C).

Our results indicate that the carbon flux J_C is diverted preferentially into the synthesis of biomass precursors, with $J_{C\to B}$ amounting to 60–75% of J_C (Fig. 3b). Still, most of the variation in the specific intake q with μ is accounted for by a significant increase in q_E (Fig. 3c), corresponding to a dramatic decrease in ATP yield ($1/q_E$, Fig. 3d) from 40 ATP per lactose for respiration-based states to about 20 for fermentation-based ones, in agreement with estimates from ref. [11] On the other hand, the biomass yield ($1/q_B$, Fig. 3e) does not display any significant change. In other terms, the energy yield appears to be the main driver of the acetate switch despite the significantly uneven partitioning of carbon between ATP and biomass precursor synthesis pathways.

By exploiting the same decomposition of $\mathbf{v}$ we are also able to infer whether individual reactions are mostly associated to the synthesis of energy or biomass precursors. This information is encoded in the entries of $\boldsymbol{\eta}$ and $\boldsymbol{\beta}$, which relate fluxes respectively to the energy and the biomass class. Non-zero entries of $\boldsymbol{\eta}$ occur mostly for reactions belonging to glycolysis, TCA cycle, and oxidative phosphorylation pathways. The corresponding proteome mass fractions derived from proteomics sum up to about 11% of the total in all conditions (Fig. 3f), while the crossover between TCA and glycolytic proteins starts at growth rates faster than 0.7/h (Figs 3g and 2c, inset) in agreement with previous determinations.[11] In turn, the mass fraction of proteins associated to the production of biomass precursors increases almost proportionally to the growth rate for $\mu \gtrsim 0.7$/h (Fig. 3f), implying that the decrease in $\varepsilon = \phi_E/\mu$ observed at fast growth (Fig. 2c) is mostly due to the fact that the proteome share allocated to energy metabolism remains roughly constant as μ changes.

A two-state model of *E. coli*'s acetate switch

These results strongly implicate energy metabolism in the increase in the specific carbon uptake q and the corresponding drop in the protein cost ε observed at fast growth. The presence of two well-defined types of states (associated respectively to respiration and fermentation), the shift in the efficiency of the optimal energetic metabolism, and the lack of significant changes in the efficiency of anabolic processes point to a minimal two-state model of *E. coli*'s acetate switch. We classify states on the Pareto front in two broad sets (Fig. 4a). The first corresponds to a 'fermentation' phenotype with large specific carbon intake ($q_{\rm fer} \gtrsim 6\ {\rm g_{lac}/g_{DW}}$, Fig. 2c) but low specific protein cost ($\varepsilon_{\rm fer} \simeq 0.3$ h), showing carbon overflow and robust fermentative flux. The second one has lower carbon intake ($q_{\rm res} \gtrsim 5.5\ {\rm g_{lac}/g_{DW}}$) but higher protein costs ($\varepsilon_{\rm res} \simeq 0.35$ h) and relies on respiration for ATP production. We assume that generic flux patterns are described by linear combinations of these phenotypes with parameter α ($0 \le \alpha \le 1$), thereby giving a combined specific carbon intake $q(\alpha) = \alpha q_{\rm res} + (1-\alpha) q_{\rm fer}$ and protein cost $\varepsilon(\alpha) = \alpha\varepsilon_{\rm res} + (1-\alpha)\varepsilon_{\rm fer}$. The growth rate μ is then

$$\mu(\alpha) = \frac{\phi_{max}}{w_R + \varepsilon(\alpha) + w_C\, q(\alpha)} \tag{8}$$

(see Eq. (4)). This model predicts (Supplementary Text, Note 4) that, at optimal growth, a transition between the fermentation phenotype (fast growth) and the respiration phenotype (slow growth) occurs when the extra cost of importing carbon in the former matches the extra protein cost of respiration in the latter (Fig. 4b). The transition occurs when w_C takes the value

$$w_C^{\rm ac} \equiv \frac{\varepsilon_{\rm res} - \varepsilon_{\rm fer}}{q_{\rm fer} - q_{\rm res}} \simeq 0.1\,\frac{\rm g_{DW}\cdot h}{\rm mmol\ lac}, \tag{9}$$

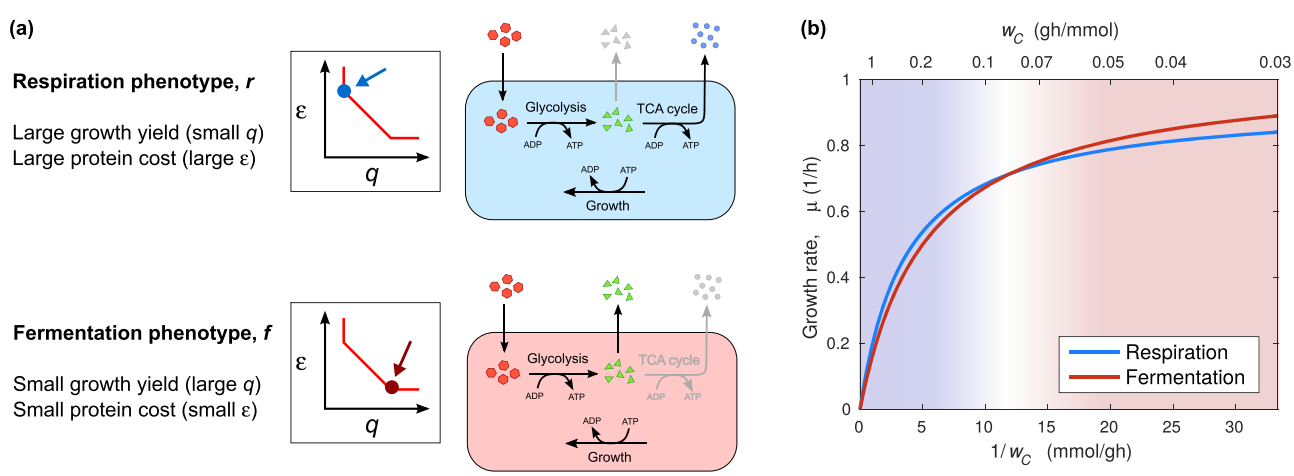

Fig. 4 Phenomenological two-state view of *E. coli* carbon-limited growth. **a** Respiration and fermentation phenotypes as characterized by the multi-objective optimal states on the Pareto front of *E. coli*'s metabolism. The respiration phenotype has a larger yield (smaller q) and larger specific protein costs, while the fermentation phenotype carries lower yields (higher q) and smaller cost. **b** Growth rate (μ) versus carbon-intake cost w_C as obtained from the phenomenological two-state model discussed in the text (see Supplementary Text, Note 4). For each w_C, the optimal phenotype is the one for which μ is largest. The switch from the fermentation to the respiration phenotype occurs when w_C matches the extra protein cost required by respiration

corresponding to $\mu_{ac} \simeq 0.7/h$, close to the experimentally determined onset of the acetate switch (Fig. 2c). While a discontinuous respiration-fermentation transition is unrealistic, this model emphasizes the role of the ATP yield and the interplay between specific carbon intake and protein costs in determining carbon overflow in *E. coli*. The value w_C^{ac} above which the fermentation phenotype becomes dominant and acetate excretion sets in can be interpreted as the external carbon level for which the two strategies provide the same cost-benefit balance, therefore giving the same growth rate. A comparison between this toy model and other phenomenological approaches is given in Supplementary Text, Note 4.

DISCUSSION

We have provided a description of *E. coli*'s carbon-limited growth in terms of a multi-objective optimization problem in which the cell optimizes for both the efficiency of carbon utilization and the protein cost of metabolism. As the external carbon level is varied, the optimal solution interpolates between the two functions, focusing the optimization on protein costs (carbon efficiency) in carbon-rich (carbon-poor) media. The emergent scenario depicts, with quantitative accuracy, the economics underlying the yield-cost tradeoff in *E. coli*. States associated to respiration or fermentation lie in different regions of the plane spanned by the specific carbon uptake q and the specific protein cost ε. Because of the proteome allocation constraint Eq. (3), a predominantly respirative, yield-maximizing strategy is optimal at slow growth, while a cost-minimizing strategy relying on fermentation sets in at fast growth. The Pareto front bends between $\mu \simeq 0.6/h$ and $\mu \simeq 0.9/h$, signaling the gradual shift from one strategy to the other. Remarkably, this range contains the empirical acetate onset point, which lies at $\mu \simeq 0.7/h$.[11] To clarify which variable drives the change in growth yield observed along the crossover, we developed a genome-scale in silico framework to decompose the carbon influx into terms contributing respectively to biomass and energy synthesis. We found that the partitioning is significantly uneven, with most carbon going into biomass precursors. The energy-biosynthetic component, however, accounts for most of the change seen in the growth yield. The efficiency of ATP production hence appears as the key driver of the yield-cost tradeoff, and a minimal two-state model based on this result can indeed recapitulate the overflow scenario.

The analysis of optimal growth in terms of a yield-cost tradeoff and the decomposition of the carbon uptake into energy and biomass contributions are the main results of this work. When integrated with proteomic data, such a decomposition allows to comprehend the behavior of metabolic fluxes and proteome mass fractions in a unified framework. It is noteworthy that a detailed study of carbon partitioning had so far been carried out only within streamlined models.[37–40]

The two-state scheme presented here differs from other coarse-grained models of carbon overflow in many respects. For instance, the yields and cost parameters for respiration and fermentation used in ref. [11] refer to the ATP yield (as opposed to the growth yield) and to specific "respiration" and "fermentation" proteomes. Here, both pathways belong to the same *E*-sector and the focus is on the global re-allocation of the proteome rather than the up- or down-regulation of specific pathways. Secondly, the cost of carbon uptake (i.e., w_C) is implicitly assumed to be nil in ref.,[11] implying that metabolism is completely determined by the normalization of proteome fractions and by the carbon and energy flux balance. Under Pareto optimality, instead, it derives from the tradeoff between the *C*- and *E*-sectors. This makes it in principle possible to describe strains with different "acetate overflow lines" (e.g., mutants[41,42] or "acetate feeding" strains obtained in evolution experiments[43,44]), which correspond to feasible (albeit sub-optimal) cellular states that would be harder to describe by the model of ref., [11] which characterizes an "optimal" strain. Experiments probing the fitness advantages of different metabolic strategies in different environmental and ecological contexts might shed light on the evolution of the acetate switch.

By extending the model of Basan et al.,[11] Vazquez and Oltvai[45] have linked overflow metabolism to a macromolecular crowding constraint along the lines of refs. [46,47] (A similar constraint plays an important role in the self-replicating cell model discussed in ref. [22]) For *E. coli*, this scenario appears to be at odds with the empirical fact that the cell volume adjusts in response to changes in the macromolecular composition of the cell so as to maintain the density roughly constant across distinct conditions.[48,49] While this reinforces the idea that cytoplasmic crowding is finely regulated,[50] a constant intracellular density is unlikely to be the main driver of the acetate switch. In addition, however, ref. [45] points out that, at slow growth, an increase of the proteome share of proteins other than those associated to respiration and fermentation has to take place. Our results are in line with this observation. In fact, proteins included in the *C*-sector are upregulated at low μ, in agreement with quantitative measurements.[6]

It is known that many different organisms share *E. coli*'s behavior in terms of growth laws[4] and carbon overflow. The detailed picture may, however, differ significantly from *E. coli*'s. For instance, recent work on *L. lactis*, an industrial bacterium that displays a transition between different types of fermentations rather than between fermentation and respiration, indicates that protein costs are not a determinant factor in its growth strategies.[51] On the other hand, many features observed in *E. coli* are shared by *S. cerevisiæ*,[52] despite the fact that carbon overflow in yeast appears to respond to the carbon intake flux rather than to the macroscopic growth rate.[53] Likewise, relationships like Eqs. (1) and (2) hold under more general types of limitations than carbon shortage, provided J_C is replaced by the flux of the limited activity and w_C by a proxy for the level of stress to which the cell is subject. The tradeoff theory presented here is therefore likely to be more broadly applicable than the case study on which we have focused.

METHODS

Derivation of the proteome constraint, Eq. (3)

The proteome normalization condition using the *C*-, *R*-, *E*- and *Q*-sectors reads

$$\phi_C + \phi_R + \phi_E = \underbrace{1 - \phi_Q}_{\text{fixed}}. \tag{10}$$

Based on the bacterial 'growth laws' characterized in refs., [3,6] each of the three condition-dependent terms in Eq. (10) is of the form $\phi_X = \phi_{X,0} + \Delta\phi_X$, where $\phi_{X,0}$ is a constant offset value, the minimal value of $\Delta\phi_X$ is zero, and $X \in \{C, R, E\}$. In turn, if metabolism mediates the adjustment of each of the three $\Delta\phi_X$ terms, these can be expressed in terms of their driving fluxes.[3,5,6,30] In particular, using a linear enzyme-flux relation $\phi_i = \phi_{i,0} + w_i|v_i|$ one gets to the following expressions for the condition-dependent parts of the proteome:

$$\begin{aligned} \Delta\phi_C(J_C) &= w_C J_C, \\ \Delta\phi_R(\mu) &= w_R \mu, \\ \Delta\phi_E(\mathrm{v}) &= \sum_{i\in E} w_i |v_i|. \end{aligned} \tag{11}$$

Substituting these into Eq. (10) immediately leads to Eq. (3) upon including all condition-independent terms in the constant $\phi_{\max} = 1 - \phi_{C,0} - \phi_{R,0} - \phi_{E,0}$, where $\phi_{E,0}$ corresponds to the sum of the offsets of the individual metabolic enzymes: $\phi_{E,0} = \sum_{i\in E} \phi_{i,0}$.

Metabolic network reconstruction

All computations were carried out on the iJR904 GSM/GPR genome-scale metabolic model[54] with a lactose-limited minimal medium.

Constraint-based model to compute optimal metabolic fluxes

In order to compute reaction fluxes corresponding to maximum growth within a genome-scale model, we have resorted to Constrained Allocation Flux Balance Analysis (CAFBA). Flux Balance Analysis (FBA[55]) approaches to metabolic network modeling search for optimal flux vectors $\mathbf{v} = \{v_i\}$ within the space $\mathcal{F}$ defined by the mass balance conditions $\mathbf{Sv} = \mathbf{0}$, $\mathbf{S}$ denoting the stoichiometric matrix, and by lower and upper bounds for the fluxes:

$$\mathcal{F} = \{\mathrm{v} : \mathrm{Sv} = 0 \text{ and } \mathrm{v}_{\min} \le \mathrm{v} \le \mathrm{v}_{\max}\}. \tag{12}$$

The most common bounds encode for thermodynamic constraints imposing that $v_i \ge 0$ for irreversible reactions. Other fluxes might instead be set to prescribed values, as the ATP hydrolysis flux related to maintenance (ATPm), which is fixed at 7.6 mmol/g_{DW}h in the iJR904 model. FBA usually looks for the flux vector $\mathrm{v} \in \mathcal{F}$ that maximizes the growth rate μ subject to a cap on the nutrient intake, and is solved by Linear Programming (LP).[55]

CAFBA[30] is a modification of FBA where $\mathcal{F}$ is further constrained by Eq. (3). CAFBA is explicitly stated as

$$\max_{\mathrm{v}\in\mathcal{F}} \mu \quad \text{subj. to} \quad w_C J_C + w_R \mu + \sum_{i\in E} w_i |v_i| = \phi_{\max}. \tag{13}$$

As shown in ref., [30] CAFBA can be solved by LP as long as the protein constraint is growth-limiting. Importantly, no bounds on the influx of nutrients present in the growth medium (e.g., the carbon source) are required. In CAFBA, different degrees of limitation of the carbon source are modeled by increasing the value of w_C, which is thus the main control parameter. We set $\phi_{\max}$ and w_R to their *E. coli*-specific empirical values of 0.48 and 0.169/h, respectively.[3] In addition, the costs w_i of reactions in the *E*-sector were assumed to take on the same value for each i, namely $w_E = 8.3 \times 10^{-4}$ g_{DW}h/mmol.[30] This value imposes a maximum achievable growth rate, obtained in the limit $w_C \to 0$, of 1/h, comparable to that seen in.[6] (See ref. [30] for a discussion of alternative choices.) For sakes of simplicity, we have also assumed that the stoichiometry of the biomass synthesis reaction whose rate quantifies μ is growth-rate independent.[55] The case of growth-rate dependent biomass composition does not yield significant differences with the scenario we present.[30]

Growth-associated and maintenance energy fluxes

In genome-scale models of metabolism, the energy requirements of the cell are described by an ATP hydrolysis flux $J_E = \sigma_0 + \sigma\mu$ formed by a growth-independent "maintenance" flux σ_0 (the ATPm) and a growth-associated term $\sigma\mu$. These terms account for process not included in the metabolic network, like protein synthesis. In *E. coli*, the ratio σ_0/σ is estimated to lie between 0.05 and 0.15/h.[30,56,57] As a consequence, σ_0 can be neglected at growth rates above roughly 0.5/h.

Computation of the Pareto front

The Pareto front shown in Fig. 2b was computed by solving CAFBA for different values of w_C after silencing ATPm. To compensate for the lack of maintenance-associated energy costs, we increased the growth-associated ATP hydrolysis rate by an amount equal to the ATPm flux (i.e., 7.6 $\mathrm{mmol}_{\mathrm{ATP}}/g_{DW}$), so that the total ATP hydrolysis flux at the maximum achievable growth rate of 1/h is the same as in the default model. The difference in the overall ATP hydrolysis flux (including maintenance and growth-rate dependent components) between this implementation of CAFBA and the standard one[30] is within 15% for growth rates above 0.5/h. For each CAFBA-derived optimal state $\boldsymbol{\xi} = \mathbf{v}/\mu$, the specific intake $q = J_C/\mu$ and the specific protein cost $\varepsilon = w_E \sum_i |v_i|/\mu$ were computed, returning a set of points (one per state) in the (q, ε) plane. The Pareto front is obtained by joining all optimal states via straight lines; as w_C is varied, the optimal solution jumps between the states defining the front, leading to the discontinuities seen in Fig. 2b (see Supplementary Note 2 for mathematical details and an illustration of the optimal states in CAFBA). Such abrupt shifts are at odds with the smooth behavior observed in experiments[11] and are due to our choice to work with at fixed weights w_i.[30] The version of CAFBA with randomized weights described in ref. [30] makes use of an averaging procedure that yields the same qualitative picture but quantitative agreement with experiments. Here we have chosen to stick to the non-random case for sakes of simplicity.

Generation of sub-optimal CAFBA solutions

In order to generate the sub-optimal CAFBA solutions shown in Fig. 2b, we computed the values of q and ε for flux vectors $\mathbf{v} = (J_C, \{v_i\})$ different from the optimal ones. To ensure that such states lie close to the Pareto front, we used flux vectors that are optimal for a version of CAFBA in which homogeneous costs ($w_i = w_E$ for each i) are replaced by independent identically-distributed random costs with mean w_E and dispersion δ, as discussed in ref. [30] After generating a large number of such vectors for different values of w_C, we computed the corresponding metabolic states $\boldsymbol{\xi}$ by normalizing each of them by its growth rate. For each such $\boldsymbol{\xi}$, we then computed the specific uptake $q = \xi_C$ and protein cost $\varepsilon = w_E \sum_i |\xi_i|$, using homogeneous weights for the latter. This method yields viable but sub-optimal solutions to CAFBA with homogeneous weights. The typical distance of sub-optimal solutions from the Pareto front is controlled by the dispersion δ of the individual costs w_i.[30] Sampled solutions approach the Pareto front as $\delta \to 0$. Upon increasing δ, protein costs get more variable and metabolic states far from the Pareto front become increasingly likely.

Computation of ATP and biomass yields

Optimal and sub-optimal CAFBA solutions shown in Fig. 3 were generated as follows. For each solution $\mathbf{v}$ to CAFBA with randomized protein costs (obtained as described above), we computed a "perturbed" solution $\mathbf{v}'$ using a small but non-zero ATPm ($\sigma_0 = 10^{-2}$ mmol ATP/g_{DW}h). The vectors $\boldsymbol{\beta}$ and $\boldsymbol{\eta}$, in terms of which flux vectors can be decomposed, can be obtained uniquely as linear combinations of $\mathbf{v}$ and $\mathbf{v}'$. The growth rate corresponding to a sub-optimal solution to the CAFBA problem with homogeneous proteins costs are instead computed from a generalization of Eq. (4) that includes an explicit dependence on σ_0. Details are found in Supplementary Text, Note 3.

Computation of protein costs from mass spectrometry data and comparison with the predicted Pareto front

Mass spectrometry data from[6] quantify protein levels for *E. coli* NQ381 (a strain with titratable LacY enzyme derived from the prototrophic K-12 strain NCM3722) grown in minimal lactose media. Five different growth rates were obtained by inducing different levels of LacY. Relative protein abundances were converted into absolute mass fractions following the method used in.[11] Briefly, we rescaled relative mass fractions with absolute abundances from ribosome profiling data obtained for *E. coli* strain MG1655[58]; strain-specific differences were accounted for by using the proteomics dataset from ref. [8] as a bridge between NCM3722 and MG1655. The specific cost ε of the *E*-sector shown in Fig. 2c was obtained as follows. First, reactions were assigned to the *E*-sector according to the partition used in ref. [30] Next, for each reaction, we obtained a corresponding list of enzymes from the Gene-Protein-Reaction matrix included in the iJR904 reconstruction. We denote by $n_{i,\mathrm{tot}}$ the number of reactions in which enzyme i participates (irrespective of whether they are assigned to the *E*-sector or not), and by $n_{i,E}$ the number of such processes included in the *E*-sector. Given the experimental protein mass fractions ϕ_i, our estimate for the mass fraction of the *E*-sector ϕ_E, and its proteome cost ε are given respectively by

$$\phi_E = \sum_{i\in E} \frac{n_{i,E}}{n_{i,\mathrm{tot}}} \phi_i \quad \text{and} \quad \varepsilon = \frac{\phi_E}{\mu}. \tag{14}$$

Unfortunately, growth yields for these dataset are not available. Instead, these were estimated with FBA by minimizing the carbon intake flux while constraining the growth rates and acetate fluxes to their experimental values. Notice that yields may vary considerably across experiments, and that their value depends on those of maintenance and growth-dependent ATP hydrolysis rates. For our purposes, though, the key feature is the decrease that occurs at fast growth rates (above approx. 0.7/h) due to acetate excretion, which is robust and independent on the carbon source used.[11,30]

Computation of energy- and biomass-associated proteome fractions

To provide a condition-independent partitioning of the protein mass fraction of the experimental *E*-sector into a biomass- and an energy-associated component we adopted a criterion that relies on the flux decomposition $\mathbf{v} = \boldsymbol{\beta}\mu + J_E\boldsymbol{\eta}$ and on the whole set of sampled metabolic states shown in Fig. 3. Because of the inherent conceptual difficulty in disentangling the contribution of energy to biomass synthesis, our choice is rather conservative with respect to the biomass-associated proteome fraction. Specifically, a reaction is considered "energy-associated" (index ε) if the following conditions are satisfied simultaneously in at least one of

the metabolic states shown in Fig. 3: (i) its flux v_i is significantly non-zero (specifically, $|v_i| > 10^{-6}$ mmol/g_{DW}h); (ii) its η_i component is sufficiently large compared to β_i (specifically, $|\eta_i| > 0.1 \cdot |\beta_i|$). Otherwise, it is "biomass-associated" (index $\mathcal{B}$). The biomass- and energy-associated proteome fractions are then defined as

$$\phi_E^{\mathcal{B}} = \sum_{i \in \mathcal{B}} \frac{n_{i,E}}{n_{i,\mathrm{tot}}} \phi_i \quad \text{and} \quad \phi_E^{\varepsilon} = \sum_{i \in \varepsilon} \frac{n_{i,E}}{n_{i,\mathrm{tot}}} \phi_i, \tag{15}$$

respectively. These quantities are shown in Fig. 3f along with $\phi_E = \phi_E^{\mathcal{B}} + \phi_E^{\varepsilon}$. The majority of proteins in ε belongs to basic energy-producing pathways (Fig. 3g). The pathway membership of each reaction was based on the "subsystems" description in the iJR904 model ('Citric Acid Cycle' for TCA cycle, 'Glycolysis/Gluconeogenesis' and 'Pyruvate Metabolism' for Glycolysis, and 'Oxidative Phosphorylation' for OxPhos). Notice that the size and composition of the energy sector ε resulting from our conservative classification is in agreement with that found in ref.,[11] where reactions were subdivided manually.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

All data generated or analysed during this study are included in this article (and its Supplementary information files).

ACKNOWLEDGEMENTS

M.M. thanks Terence Hwa for discussion and funding through National Institutes of Health (NIH; grant 1R01GM109069) and the Simons Foundation (grant 330378). E.M. acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No [694925]). A.D.M. acknowledges support from the European Union's Horizon 2020 Research and Innovation Staff Exchange program MSCA-RISE-2016 under Grant Agreement Nr 734439 (INFERNET).

AUTHOR CONTRIBUTIONS

Conceptualization, M.M., E.M. and A.D.M.; Software and data analysis, M.M.; Investigation, M.M., E.M. and A.D.M.; Writing, M.M. and A.D.M. All authors reviewed the manuscript and approved the final version.

ADDITIONAL INFORMATION

Supplementary Information accompanies the paper on the *npj Systems Biology and Applications* website (https://doi.org/10.1038/s41540-019-0093-4).

Competing interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

1. Schaechter, M., Maaløe, O. & Kjeldgaard, N. O. Dependency on medium and temperature of cell size and chemical composition during balanced growth of Salmonella typhimurium. *Microbiology* **19**, 592–606 (1958).
2. Kjeldgaard, N. O., Maaløe, O. & Schaechter, M. The transition between different physiological states during balanced growth of Salmonella typhimurium. *Microbiology* **19**, 607–616 (1958).
3. Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z. & Hwa, T. Interdependence of cell growth and gene expression: origins and consequences. *Science* **330**, 1099–1102 (2010).
4. Scott, M. & Hwa, T. Bacterial growth laws and their applications. *Curr. Opin. Biotechnol.* **22**, 559–565 (2011).
5. You, C. et al. Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature* **500**, 301–306 (2013).
6. Hui, S. et al. Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Mol. Sys. Biol.* **11**, 784 (2015).
7. Erickson, D. W. et al. A global resource allocation strategy governs growth transition kinetics of Escherichia coli. *Nature* **551**, 119–123 (2017).
8. Schmidt, A. et al. The quantitative and condition-dependent Escherichia coli proteome. *Nat. Biotechnol.* **34**, 104–110 (2016).
9. Wolfe, A. J. The acetate switch. *Microbiol Molec. Biol. Rev.* **69**, 12–50 (2005).
10. Vemuri, G. N., Altman, E., Sangurdekar, D. P., Khodursky, A. B. & Eiteman, M. A. Overflow metabolism in Escherichia coli during steady-state growth: transcriptional regulation and effect of the redox ratio. *Appl. Environm. Microbiol.* **72**, 3653–3661 (2006).
11. Basan, M. et al. Overflow metabolism in Escherichia coli results from efficient proteome allocation. *Nature* **528**, 99–104 (2015).
12. Schuetz, R., Kuepfer, L. & Sauer, U. Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli. *Mol. Sys. Biol.* **3**, 119 (2007).
13. Schuetz, R., Zamboni, N., Zampieri, M., Heinemann, M. & Sauer, U. Multidimensional optimality of microbial metabolism. *Science* **336**, 601–604 (2012).
14. Ibarra, R. U., Edwards, J. S. & Palsson, B. O. Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* **420**, 186–189 (2002).
15. Price, N. D., Reed, J. L. & Palsson, B. O. Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat. Rev. Microbiol.* **2**, 886–897 (2004).
16. De Deken, R. H. The Crabtree effect: a regulatory system in yeast. *Microbiology* **44**, 149–156 (1966).
17. Postma, E., Verduyn, C. O., Scheffers, W. A. & Van Dijken, J. P. Enzymic analysis of the crabtree effect in glucose-limited chemostat cultures of Saccharomyces cerevisiae. *Appl. Envir. Microbiol.* **55**, 468–477 (1989).
18. Hsu, P. P. & Sabatini, D. M. Cancer cell metabolism: Warburg and beyond. *Cell* **134**, 703–707 (2008).
19. Diaz-Ruiz, R., Uribe-Carvajal, S., Devin, A. & Rigoulet, M. Tumor cell energy metabolism and its common features with yeast metabolism. *Biochim Biophys. Acta Rev. Cancer* **1796**, 252–265 (2009).
20. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029–1033 (2009).
21. Schuster, S., Boley, D., Möller, P., Stark, H. & Kaleta, C. Mathematical models for explaining the Warburg effect: a review focussed on ATP and biomass production. *Biochem. Soc. Trans.* **43**, 1187–1194 (2015).
22. Molenaar, D., Van Berlo, R., De Ridder, D. & Teusink, B. Shifts in growth strategies reflect tradeoffs in cellular economics. *Mol. Sys. Biol.* **5**, 323 (2009).
23. Carlson, R. P. & Taffs, R. L. Molecular-level tradeoffs and metabolic adaptation to simultaneous stressors. *Curr. Opin. Biotechnol.* **21**, 670–676 (2010).
24. Flamholz, A., Noor, E., Bar-Even, A., Liebermeister, W. & Milo, R. Glycolytic strategy as a tradeoff between energy yield and protein cost. *Proc. Natl Acad. Sci. USA* **110**, 10039–10044 (2013).
25. Maitra, A. & Dill, K. A. Bacterial growth laws reflect the evolutionary importance of energy efficiency. *Proc. Natl Acad. Sci. USA* **112**, 406–411 (2015).
26. Szenk, M., Dill, K. A. & de Graff, A. M. Why do fast-growing bacteria enter overflow metabolism? Testing the membrane real estate hypothesis. *Cell Syst.* **5**, 95–104 (2017).
27. Wortel, M. T., Noor, E., Ferris, M., Bruggeman, F. J. & Liebermeister, W. Metabolic enzyme cost explains variable tradeoffs between microbial growth rate and yield. *PLoS Comp. Biol.* **14**, e1006010 (2018).
28. Goelzer, A. & Fromion, V. Bacterial growth rate reflects a bottleneck in resource allocation. *Biochim Biophys. Acta Gen. Subj.* **1810**, 978–988 (2011).
29. O'Brien, E. J., Lerman, J. A., Chang, R. L., Hyduke, D. R. & Palsson, B. O. Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol. Sys. Biol.* **9**, 693 (2013).
30. Mori, M., Hwa, T., Martin, O. C., De Martino, A. & Marinari, E. Constrained allocation flux balance analysis. *PLoS Comp. Biol.* **12**, e1004913 (2016).
31. Hart, Y. et al. Inferring biological tasks using Pareto analysis of high-dimensional data. *Nat. Methods* **12**, 233–235 (2015).
32. Mori, M., Schink, S., Erickson, D. W., Gerland, U. & Hwa, T. Quantifying the benefit of a proteome reserve in fluctuating environments. *Nat. Comm.* **8**, 1225 (2017).
33. Bennett, B. D., Yuan, J., Kimball, E. H. & Rabinowitz, J. D. Absolute quantitation of intracellular metabolite concentrations by an isotope ratio-based approach. *Nat. Prot.* **3**, 1299–1311 (2008).
34. Boer, V. M., Crutchfield, C. A., Bradley, P. H., Botstein, D. & Rabinowitz, J. D. Growth-limiting intracellular metabolites in yeast growing under diverse nutrient limitations. *Mol. Biol. Cell* **21**, 198–211 (2010).
35. Valgepea, K., Adamberg, K., Seiman, A. & Vilu, R. Escherichia coli achieves faster growth by increasing catalytic and translation rates of proteins. *Mol. BioSys.* **9**, 2344–2358 (2013).
36. O'Brien, E. J., Utrilla, J. & Palsson, B. O. Quantification and classification of E. coli proteome utilization and unused protein costs across environments. *PLoS Comp. Biol.* **12**, e1004998 (2016).
37. Holms, W. H. The central metabolic pathways of Escherichia coli: relationship between flux and control at a branch point, efficiency of conversion to biomass,

and excretion of acetate. In *Current topics in cellular regulation* Vol. 28, pp. 69–105 (Academic Press, 1986).
38. Pramanik, J. & Keasling, J. D. Stoichiometric model of Escherichia coli metabolism: incorporation of growth-rate dependent biomass composition and mechanistic energy requirements. *Biotechnol. Bioeng.* **56**, 398–421 (1997).
39. Carlson, R. & Srienc, F. Fundamental Escherichia coli biochemical pathways for biomass and energy production: identification of reactions. *Biotechnol. Bioeng.* **85**, 1–9 (2004).
40. Carlson, R. & Srienc, F. Fundamental Escherichia coli biochemical pathways for biomass and energy production: creation of overall flux states. *Biotechnol. Bioeng.* **86**, 149–162 (2004).
41. Castano-Cerezo, S. et al. An insight into the role of phosphotransacetylase (pta) and the acetate/acetyl-CoA node in Escherichia coli. *Micro. Cell Fact.* **8**, 54 (2009).
42. Valgepea, K. et al. Systems biology approach reveals that overflow metabolism of acetate in Escherichia coli is triggered by carbon catabolite repression of acetyl-CoA synthetase. *BMC Sys. Biol.* **4**, 166 (2010).
43. Helling, R. B., Vargas, C. N. & Adams, J. Evolution of Escherichia coli during growth in a constant environment. *Genetics* **116**, 349–358 (1987).
44. Treves, D. S., Manning, S. & Adams, J. Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of Escherichia coli. *Mol. Biol. Evol.* **15**, 789–797 (1998).
45. Vazquez, A. & Oltvai, Z. N. Macromolecular crowding explains overflow metabolism in cells. *Sci. Rep.* **6**, 31007 (2016).
46. Beg, Q. K. et al. Intracellular crowding defines the mode and sequence of substrate uptake by Escherichia coli and constrains its metabolic activity. *Proc. Natl Acad. Sci. USA* **104**, 12663–12668 (2007).
47. Vazquez, A. et al. Impact of the solvent capacity constraint on E. coli metabolism. *BMC Sys. Biol* **2**, 7 (2008).
48. Woldringh, C. L., Binnerts, J. S. & Mans, A. Variation in Escherichia coli buoyant density measured in Percoll gradients. *J. Bacteriol.* **148**, 58–63 (1981).
49. Basan, M. et al. Inflating bacterial cells by increased protein synthesis. *Mol Sys. Biol* **11**, 836 (2015).
50. Vazquez, A. Optimal cytoplasmatic density and flux balance model under macromolecular crowding effects. *J. Theor. Biol.* **264**, 356–359 (2010).
51. Goel, A. et al. Protein costs do not explain evolution of metabolic strategies and regulation of ribosomal content: does protein investment explain an anaerobic bacterial Crabtree effect? *Molec. Microbiol.* **97**, 77–92 (2015).
52. Metzl-Raz, E. et al. Principles of cellular resource allocation revealed by condition-dependent proteome profiling. *eLife* **6**, e28034 (2017).
53. Huberts, D. H., Niebel, B. & Heinemann, M. A flux-sensing mechanism could regulate the switch between respiration and fermentation. *FEMS Yeast Res.* **12**, 118–128 (2012).
54. Reed, J. L., Vo, T. D., Schilling, C. H. & Palsson, B. O. An expanded genome-scale model of Escherichia coli K-12 (i JR904 GSM/GPR). *Genome Biol.* **4**, R54 (2003).
55. Orth, J. D., Thiele, I. & Palsson, B. O. What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–248 (2010).
56. Neijssel, O., Teixeira De Mattos, M. J. & Tempest, D. In *Escherichia coli and Salmonella: Cellular and Molecular Biology* (ASM Press, 1996).
57. Monk, J. M. et al. iML1515, a knowledgebase that computes Escherichia coli traits. *Nat. Biotechnol.* **35**, 904–908 (2017).
58. Li, G. W., Burkhardt, D., Gross, C. & Weissman, J. S. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* **157**, 624–635 (2014).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019