PNAS RESEARCH ARTICLE | SYSTEMS BIOLOGY OPEN ACCESS

The proteome is a terminal electron acceptor

Avi I. Flamholz[a,1], Akshit Goyal[b,c], Woodward W. Fischer[d], Dianne K. Newman[a,d,1], and Rob Phillips[a,e,1]

Affiliations are included on p. 11.

Edited by Jens Nielsen, BioInnovation Institute, Hellerup, Denmark; received March 8, 2024; accepted October 30, 2024

Microbial metabolism is impressively flexible, enabling growth even when available nutrients differ greatly from biomass in redox state. *Escherichia coli*, for example, rearranges its physiology to grow on reduced and oxidized carbon sources through several forms of fermentation and respiration. To understand the limits on and evolutionary consequences of this metabolic flexibility, we developed a coarse-grained mathematical framework coupling redox chemistry with principles of cellular resource allocation. Our models inherit key qualities from both of their antecedents: i) describing diverse metabolic chemistries and ii) enforcing the simultaneous balancing of atom (e.g., carbon), electron, and energy (adenosine triphosphate) flows, as in redox models, while iii) treating biomass as both the product and catalyst of the growth process, as in resource allocation models. Assembling integrated models of respiration, fermentation, and photosynthesis clarified key microbiological phenomena, including demonstrating that autotrophs grow more slowly than heterotrophs because of constraints imposed by the intracellular production of reduced carbon. Our model further predicted that heterotrophic growth is improved by matching the redox state of biomass to the nutrient environment. Through analysis of ≈60,000 genomes and diverse proteomic datasets, we found evidence that proteins indeed accumulate amino acid substitutions promoting redox matching. We therefore propose an unexpected mode of genome evolution where substitutions neutral or even deleterious to the individual biochemical or structural functions of proteins can nonetheless be selected due to a redox-chemical benefit to the population.

microbial physiology | redox chemistry | metabolism | environmental science | protein evolution

Significance

Today, two distinct frameworks are used to study microbial growth. One approach treats microbes as coupling redox reactions to conserve energy in order to form biomass from available nutrients. Another describes the allocation of cells' finite biosynthetic capacity to different catalytic roles—e.g., transport, protein synthesis—to achieve particular metabolic rates. These frameworks have distinct strengths: Redox naturally describes a diversity of metabolic chemistries, while resource allocation explains the existence of intrinsic maximum growth rates, for example. Here, we introduce a unified redox-based resource allocation model that links physiology to environmental chemistry. This integration is conceptually useful—offering concise explanations for key microbiological phenomena—and also predictive—uncovering an unexpected mode of protein evolution.

Many microbes have a remarkable capacity for metabolic chemistry, growing on a diversity of environmentally supplied nutrients (e.g., various forms of carbon, nitrogen, and sulfur) in both oxic and anoxic settings. *Escherichia coli* strains, for example, can grow on relatively reduced (e.g., lipids) and oxidized carbon (C) sources (e.g., small organic acids) through several forms of fermentation and respiration (Fig. 1*A*). Some microbes can even switch their fundamental metabolic mode (1), eating organic molecules in some conditions (i.e., performing heterotrophy) while fixing CO_2 into organics to build biomass in others (i.e., performing autotrophy). Not all microbes are so flexible, yet this widespread metabolic flexibility is noteworthy because it is essentially absent from animals.

The language of redox chemistry—intermolecular electron (e^-) transfers—is often used to describe metabolic flexibility. Redox logic is useful because all organisms extract energy from their environments by coupling thermodynamically favorable e^- transfers to the synthesis of energy carriers like adenosine triphosphate (ATP). Using redox chemistry to "coarse-grain" metabolism is advantageous because it can describe any metabolism, enforces the simultaneous conservation of atoms and e^-, and provides an opportunity to incorporate thermodynamic limits via redox potentials (2, 3), as detailed in Fig. 2*A*, Box 1 and *SI Appendix*, section S2. Yet redox chemistry is not a sufficient description of microbial physiology. For example, microbes have intrinsic maximum growth rates that arise in nutrient surplus (λ_{max}, Fig. 1*B*), but conservation of atoms and e^- only requires that fluxes of production and consumption balance—the fluxes, and, therefore, the growth rate, can be arbitrarily large.

In recent years, several groups have developed a theory of cellular resource allocation to explain how microbes rearrange their internal machinery to achieve different growth rates (4–9). These models treat microbes as self-replicating catalysts whose growth rate, λ, is determined by the manner in which they apportion nutrients (e.g., carbon) and catalytic activity (e.g., protein synthesis by ribosomes) to produce the enzymes that perform core cellular tasks like catabolism and biosynthesis (Box 1 and Fig. 2*B*).

Considering cells as resource-allocating self-replicators can explain the existence of an intrinsic maximum growth rate, λ_{max}, which arises when resources are optimally

Author contributions: A.I.F., A.G., W.W.F., D.K.N., and R.P. designed research; A.I.F. and A.G. performed research; A.I.F. and A.G. contributed new reagents/analytic tools; A.I.F. and A.G. analyzed data; and A.I.F., A.G., W.W.F., D.K.N., and R.P. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: aflamhol@caltech.edu, dkn@caltech.edu, or phillips@pboc.caltech.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2404048121/-/DCSupplemental.

Published January 3, 2025.

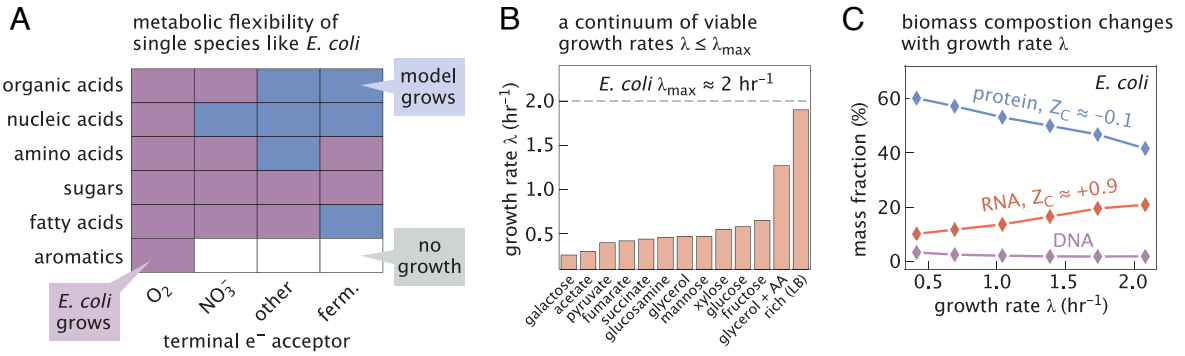

Fig. 1. *E. coli* exemplifies aspects of microbial growth that motivate integration of redox and resource allocation perspectives on metabolism. (*A*) Many heterotrophic microbes are metabolically flexible, able to grow on a diversity of carbon sources (Y axis) and terminal e^- acceptors (X axis); "other" denotes trimethylamine N-oxide or dimethyl sulfoxide and "ferm." fermentation. Purple rectangles indicate growth of *E. coli* strains, blue indicates growth predictions from a metabolic model, and white denotes no growth (*SI Appendix*, section S1.3). Such diversity of donor-acceptor pairs motivates a redox description of metabolism. (*B*) Microbes have intrinsic maximum growth rates, λ_{max}, that arises in nutrient surplus. Further, cells can grow at any growth rate $\lambda < \lambda_{max}$. This capacity is illustrated here with *E. coli* grown on various C sources (21) but also observed with variable nutrient concentrations in continuous culture. Cell-intrinsic limits on growth and metabolic rates emerge naturally from resource allocation models (Box 1). (*C*) The chemical composition of cells is coupled to the growth rate, λ. As ribosomes are ≈2/3 RNA, faster-growing cells typically contain more RNA and less protein, as shown here for *E. coli* with data from (14). While compositional shift is due to the role of ribosomes in translation, it has chemical consequences—RNA (C redox state $Z_C \approx +0.9$) is more oxidized than protein ($Z_C \approx -0.1$) and so faster growth entails the relative oxidation of biomass (15).

divided between catabolism and biosynthesis so that influxes of precursors are matched by the anabolic processes consuming them (4–9). Furthermore, because biological macromolecules play different cellular roles—metabolic reactions are catalyzed by proteins, but ribosomes are 60 to 70% RNA by mass (10, 11)—resource allocation can also explain observed changes in biomass composition that accompany increases in λ (4–9), including increases in the RNA:protein ratio (5, 12–14).

In addition to its central role in ribosomes, RNA is also fairly oxidized (e^- poor), with a formal carbon redox state of $Z_C \approx +0.9$ charge units per carbon. In contrast, values for proteins are about 1 e^-/C more reduced (e^- rich), ranging from −0.1 to −0.3 (3, 15). As such, increasing RNA:protein ratios entail the oxidation of biomass carbon during faster growth (Fig. 1*C*). To understand how concerted changes in biomass redox state should affect growth, we developed an integrated redox chemical model of microbial resource allocation to study intrinsic limits on growth and metabolic rates in coarse-grained, yet chemically specific language (Box 1 and Fig. 2*C*). As microbial communities are pivotal contributors to global recycling of carbon, nitrogen, phosphorus, and sulfur (16), theories describing limits on the chemical activities of microbes offer a conceptual framework for reasoning about the history and future of biogeochemical cycles (17–20). Here, we describe such a theory, draw lessons and predictions from it, and offer empirical support for those predictions from the genomes and proteomes of diverse microbes.

Results

A Redox Chemical Model of Microbial Self-Replication. To merge redox and resource-economic descriptions, we track the redox state of carbon (Z_C) in growth substrates (e.g., an organic C source like glucose) and products (e.g., CO_2, biomass). Z_C counts the average number of valence e^- associated with C atoms in a molecule (3). Electrons are counted by their charge, with negative Z_C denoting excess e^- compared to neutrally charged C atoms (4 valence e^-) and positive Z_C denoting paucity. See the Box 1 and *SI Appendix*, section S2 for detail on redox calculations.

Consider heterotrophic respiration—the net transfer of e^- from organic molecules to a terminal acceptor like O_2 as depicted in Fig. 2. This metabolism couples three redox reactions i) oxidation, ii) reduction, and iii) anabolism. Oxidation extracts e^- from the reduced C source (C_{red}) while reduction drives respiratory ATP production by harnessing the flow of e^- to O_2 to generate a proton-motive force driving an ATP synthase (2, 23). Not all C atoms can be oxidized to make ATP, however, or there will be no growth or maintenance of biomass. Rather, some organic C is used in anabolism to make new biomass, with a mass-specific flux of

$$\frac{1}{M}\frac{dM}{dt} = \lambda = m_C \cdot J_{\text{ana}}. \quad [1]$$

Here, M is the carbon mass of cells, λ is the growth rate (C mass doublings per hour), J_{ana} the anabolic flux (mol C/g C/h), and m_C is the molar mass of C. In cells, e^- are carried between these processes by soluble molecules like nicotinamides, flavins, and quinones which we represent as a single generic redox couple ECH/EC^+ for "electron carrier," ECH being the reduced form carrying two e^-. Reduction of the terminal acceptor exchanges $-\nu^e_{\text{red}}$ ECH for ν^a_{red} ATP, where stoichiometric coefficients ν_α are marked in Fig. 2*C*. Coefficients ν_α adopt negative values for reactions consuming a molecule of interest, e.g., anabolism typically consumes ATP and so $\nu^a_{\text{ana}} < 0$. As C sources can differ from biomass in Z_C, anabolism is not necessarily redox neutral, producing ν^e_{ana} ECH per C atom.

Conservation of mass requires conservation of electrons, which occurs when

$$\frac{d[\text{ECH}]}{dt} = \nu^e_{\text{ox}} J_{\text{ox}} + \nu^e_{\text{red}} J_{\text{red}} + \nu^e_{\text{ana}} J_{\text{ana}} - \lambda[\text{ECH}] = 0, \quad [2]$$

where λ[ECH] accounts for dilution of the e^- carrier due to growth. $\nu^e_{\text{ox}} = \frac{1}{2}(Z_{\text{C,ox}} - Z_{\text{C,red}})$ and $\nu^e_{\text{ana}} = \frac{1}{2}(Z_{\text{C,B}} - Z_{\text{C,red}})$ are calculated from Z_C values for biomass ($Z_{\text{C,B}}$), the C source (C_{red} with $Z_{\text{C,red}}$), and the oxidized product (e.g., $Z_{\text{C,ox}} = +4$ for CO_2). So a model tracking Z_C conserves e^- in addition to C atoms and biosynthetic activity.

In addition to balancing redox reactions, cells must also extract sufficient ATP for maintenance (b) and biosynthesis ($\nu^a_{\text{ana}} J_{\text{ana}}$), implying that

$$\frac{d[\text{ATP}]}{dt} = \nu^a_{\text{ox}} J_{\text{ox}} + \nu^a_{\text{red}} J_{\text{red}} + \nu^a_{\text{ana}} J_{\text{ana}} - \lambda[\text{ATP}] - b = 0. \quad [3]$$

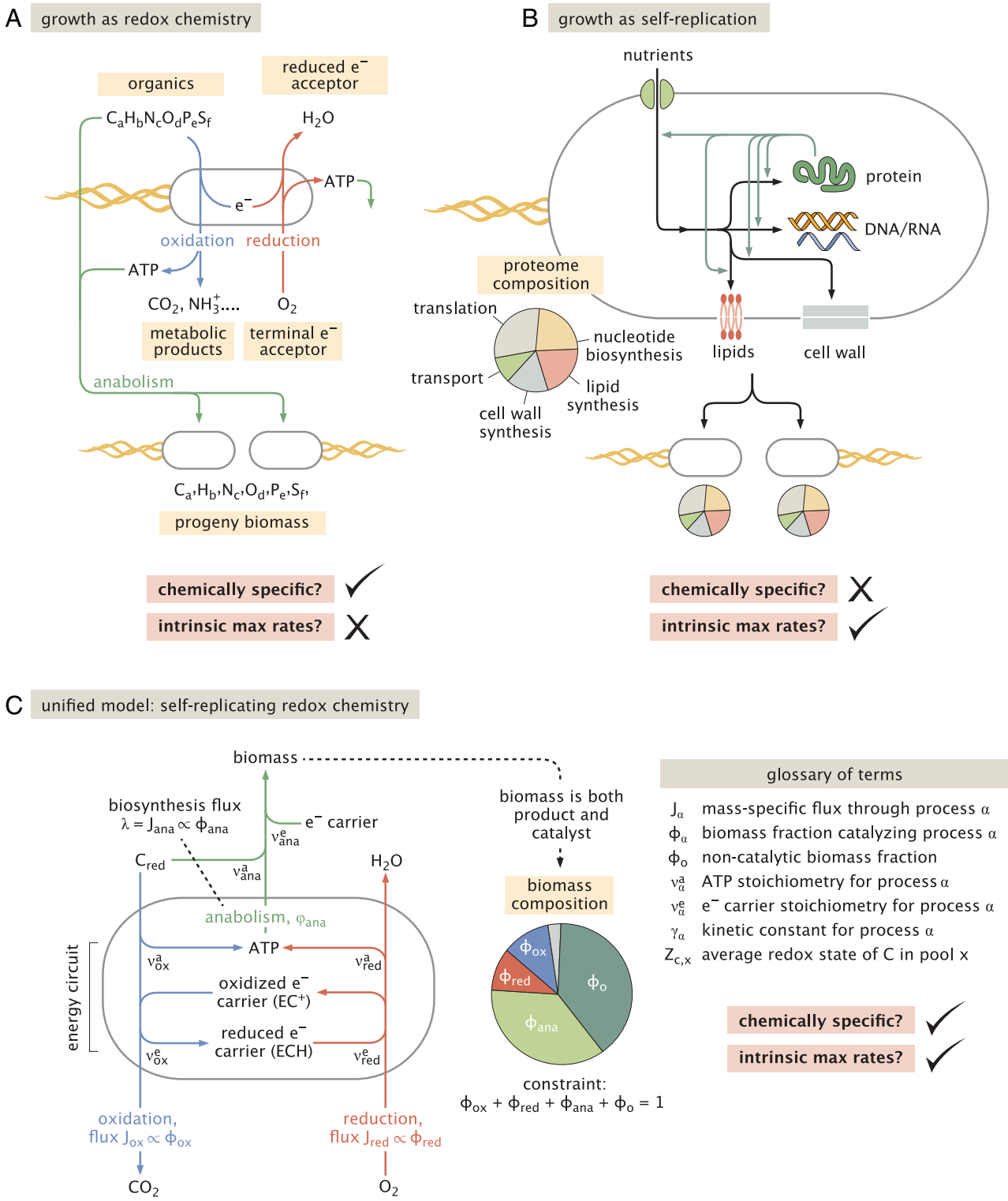

Fig. 2. Redox chemistry and cellular resource allocation are complementary frameworks describing the coupling of metabolism and growth. (*A*) A redox framing of metabolism treats growth as a electrochemical process converting nutrients into energy (ATP) and biomass. The chemical composition of nutrients, biomass, and secretions are enumerated so that conservation of atoms and electrons (e^-) can be verified. (*B*) A resource allocation perspective, in contrast, considers the subdivision of cellular biosynthetic activity to a variety of tasks like translation, precursor biosynthesis, and nutrient transport. Proteins are the dominant catalysts of cellular activities, so this allocation is typically represented as a pie chart subdividing total ribosome activity. These models have complementary strengths. Redox models like (*A*) are chemically specific, expressing differences between sugar and fatty acid nutrients, O_2 and sulfate as e^- acceptors. Allocation frameworks like (*B*), in contrast, lack chemical specificity, but recapitulate cell-intrinsic limits on growth and metabolic rates, as explained in Box 1. (*C*) We integrated these two frameworks to construct a unified model of microbial metabolism. Reactions are written per C atom with electron (ν^e_α) and ATP (ν^a_α) stoichiometries as labeled. Notice that the pie chart here subdivides total biosynthetic capacity, not just protein translation. The unlabeled gray sector indicates that more processes can be modeled, as discussed in the text. The anabolic flux J_{ana} equals the growth rate λ in appropriate units (Eq. **1**). Key terms are defined in the glossary; model equations are described in the text and *SI Appendix*, sections S3–S6.

Box 1.

Redox chemical and resource allocation approaches to microbial physiology.

Cells are made of a collection of small and macro-molecules called biomass. The anabolic reactions forming biomass require energy because the product is an ordered assemblage that maintains structural integrity, senses external changes, and catalyzes reactions. To repair biomass or grow, cells must also draw atoms and bonding e^- from the outside world. These form the covalent bonds—e.g., C-C, C-N, and Fe-S—found in all biological molecules.

To sustain life, therefore, metabolism must extract energy, atoms, and e^- in the correct proportions. Coordinating, or "balancing" these flows in a changing environment is a nontrivial challenge—one met by regulatory control systems collectively termed "physiology."

The challenge of balancing energy, atoms, and e^- is emphasized by writing reduction-oxidation (redox) reactions describing a metabolism, in this case, aerobic respiration of reduced organic matter (C_{red}) as in Fig. 2*A*. Respiring cells withdraw e^- from C_{red}, i.e., oxidize it. Those e^- are either donated to O_2 ("reduction," fraction f) forming an "energy circuit" that generates ATP, or reacted with ATP and C_{red} to form biomass (fraction $\bar{f} = 1 - f$). These reactions are summarized as

$$\begin{array}{lll|l}
C_{red} \xrightarrow{\text{oxidation}} CO_2 + \nu^e_{ox}\,e^- + \nu^a_{ox}\,ATP & \times 1 & & E'^{\circ}_{ox} < 0\ V \\
O_2 + \nu^e_{red}\,e^- \xrightarrow{\text{reduction}} 2\,H_2O + \nu^a_{red}\,ATP & \times f & & E'^{\circ}_{red} \approx 0.8\ V \\
C_{red} + \nu^e_{ana}\,e^- + \nu^a_{ana}\,ATP \xrightarrow{\text{anabolism}} biomass & \times(1-f) & & E'^{\circ}_{ana} < 0\ V \\
\hline
(2-f)\,C_{red} + f\,O_2 + \underline{0\,e^-} + \underline{0\,ATP} \xrightarrow{\text{net}} (1-f)\,biomass + CO_2 + 2f\,H_2O & & &
\end{array}$$

Here, ν^e_α and ν^a_α are stoichiometric coefficients for e^- and ATP, respectively. These redox reactions—oxidation, reduction, and anabolism—sum up to a full metabolism ("net"). Conservation of mass demands zero net production of e^-, meaning $0 = \nu^e_{ox} - f\nu^e_{red} - \bar{f}\nu^e_{ana}$ as underlined above. Similarly, ATP mass-balance requires $0 = \nu^a_{ox} + f\nu^a_{red} - \bar{f}\nu^a_{ana}$.

Considering glucose oxidation to CO_2, $\nu^e_{ox} = 4\ e^-/C$ can be calculated from formal C oxidation states (Z_C) of substrates and products, as elaborated in *SI Appendix*, section S2. Movement of e^- toward more positive reduction potentials, E'°, is favorable, and so cells can extract energy as ATP by transferring e^- from C_{red} to O_2—here, O_2 is the "terminal e^- acceptor." The standard free energy of this reaction—the sum of two half-reactions—is $\Delta_r G'^{\circ} = -nF\Delta E'^{\circ} = -nF(E'^{\circ}_{red} - E'^{\circ}_{ox})$, where n is the number of e^- transferred and F the Faraday constant. After accounting for reactant concentrations (*SI Appendix*, section S2) we arrive at a reaction energy $\Delta_r G'$ that is constrained by the second law as follows

$$\underbrace{\Delta_r G'}_{\text{oxidation + reduction energy}} + \underbrace{(\nu^a_{ox} + \nu^a_{red})}_{\text{ATP yield}}\ \underbrace{\Delta_r G'_{ATP}}_{\text{ATP energy}} \le 0.$$

This example highlights two merits of a redox framing of metabolism. First, by enumerating the chemistry of e^- donors (e.g., C_{red}), acceptors (e.g., O_2), and biomass, redox models enforce conservation of elements (here C) and e^-. Formal oxidation states (e.g., Z_C) aid in verifying conservation of e^-—redox formalism used for counting. Second, tables of E'° values are used to ensure that ATP yields are consistent with the second law—redox formalism for energetics.

Redox descriptions of metabolism are also chemically specific—preserving distinctions between glucose, and pyruvate, for instance, and distinguishing respiration from fermentation or photosynthesis (*SI Appendix*, Figs. S1, S3, and S4). However, redox principles fail to explain key facets of growth, especially intrinsic limits on growth and metabolic rates. In Fig. 2*A*, rates are limited by the supply of key nutrients (e.g., O_2), but have no biological limits.

Maximum rates arise because of the limited activities of biological catalysts. Proteins are made by ribosomes, for example, at maximum elongation rates of ≈10 amino acids/s (AA/s). Growth can proceed no faster than ribosomes reproduce the proteome, so a simple estimate of the ribosome-limited growth rate is given by

$$\lambda_{rib} \approx \frac{N_{rib}\,\frac{ribosomes}{cell} \cdot 4 \times 10^4\,\frac{AA}{h \cdot rib.}}{10^9\,\frac{AA}{cell}} = 4 \times 10^{-5} \cdot N_{rib}\ h^{-1}.$$

Here, $\lambda_{rib} \approx 1\ h^{-1}$ when cells contain $\approx 2 \times 10^4$ active ribosomes, in rough agreement with measurements (7, 22).

So the fact of autocatalysis—that biomass is both the product and catalyst of anabolism—sets finite limits on growth. Ribosomes are of special interest because proteins perform the catalytic work in cells. Models can encode these constraints by expressing the allocation of a finite ribosome activity to the production of proteins with different activities (4–9). This division of labor is often presented as a pie chart, as in Fig. 2*B*.

(see Box 1, continued)

Box 1, (continued).

Resource allocation models recapitulate cell-intrinsic limits on growth, yet lack chemical specificity. This motivated our integration of redox principles. As redox framings treat the whole of biomass as a product, a key step was drawing pie charts subdividing all of biomass rather than protein, as in Fig. 2*C*. To sustain growth in this unified model, cells must allocate resources to synthesis of oxidative (C mass fraction ϕ_{ox}), reductive (ϕ_{red}), and anabolic (ϕ_{ana}) enzymes so that fluxes J_{ox}, J_{red}, and J_{ana} balance, extracting sufficient ATP while conserving atoms, cofactors, e^-, and total biosynthetic activity. ϕ_O denotes the remaining noncatalytic mass, encompassing structural components like lipids. Equations are described in the text and derived in *SI Appendix*, sections S3–S6.

One fundamental aspect of growth absent from these equations is autocatalysis: Biomass is both the product and the catalyst of growth (4–9). As such, each of the mass-specific fluxes J_{ox}, J_{red}, and J_{ana} (generically J_α), depend on the amount of cellular catalyst (C mass $\phi_\alpha M$) and its mass-specific activity (γ_α, mol/g C/h)

$$J_\alpha = \gamma_\alpha \phi_\alpha f_\alpha(\mathbf{c}). \quad [4]$$

Cells manipulate ϕ_α by regulating the synthesis and activity of enzymes. Here, $f_\alpha(\mathbf{c})$ is relative catalytic activity as a function of reactant concentrations (via the concentration vector $\mathbf{c}$), accounting for substrate saturation, product inhibition, and concentration-dependent regulation, as described in *SI Appendix*, section S4.

Recognizing the autocatalytic constraint on growth, we subdivide biomass into its catalytic (ϕ_{ox}, ϕ_{red}, and ϕ_{ana}) and noncatalytic (ϕ_O) C mass fractions so that a proportionality between, for example, J_{red} and ϕ_{red} can be enforced. An allocation constraint then requires that all of biomass is accounted for, as given by

$$\underbrace{\phi_{ox} + \phi_{red} + \phi_{ana}}_{\text{catalytic C mass fractions}} = 1 - \phi_O. \quad [5]$$

As a result of this constraint, and in contrast with purely redox descriptions of metabolism, our unified model displays a maximum growth rate λ_{max} because allocation of more biomass to one process (e.g., increasing ϕ_{red}) displaces another (e.g., ϕ_{ana}, Fig. 3*A*).

Chemical Limits on Growth and Metabolic Rates. While fluxes J_α are potentially nonlinear functions of substrate and product concentrations, the above model can be simplified by assuming reactions are irreversible and substrate-saturated, i.e., $f_\alpha(\mathbf{c}) = 1$. We study this simplified model to gain intuition about the chemical requirements of growth, exploring nonlinear models in supplement (*SI Appendix*, section S7).

Since cells must apportion finite catalytic capacity (i.e., enzyme mass) between C source oxidation, acceptor reduction, and biosynthesis, a maximum growth rate λ_{max} arises when increasing the anabolic allocation ϕ_{ana} no longer increases the biosynthetic flux $m_C \cdot J_{ana} = \lambda$. This occurs because increasing ϕ_{ana} further displaces oxidation or reduction, lowering production of ATP or ECH beneath the level needed to support J_{ana} and, therefore, λ.

Given a set of parameter values and limits on ϕ_O, one can calculate λ_{max} and J^*_{red}, a flux beyond which additional respiratory CO_2 production no longer increases λ (Fig. 3*A* and *SI Appendix*, Fig. S6). These values depend on biochemical kinetics (γ_α), stoichiometries (e.g., ν^a_{red} ATP produced per O_2 reduced), and the redox states of nutrients and biomass ($Z_{C,red}$ and $Z_{C,B}$). As such, our model permits principled calculation of carbon use efficiencies—values representing the fraction of C taken up from the environment that is incorporated into biomass—which are widely used to understand and predict microbial carbon cycling in natural settings (24, 25).

Understanding the Role of Noncatalytic Compounds in Heterotrophic Growth. Presuming ATP and e^- carrier concentrations are constrained to typical physiological values [≈1 to 10 mM, (22)], our model involves three equations and four free variables ϕ_{ox}, ϕ_{red}, ϕ_{ana}, ϕ_O (Fig. 2*C*). When the growth rate is maximized, i.e., when $\lambda = \lambda_{max}$, the noncatalytic biomass fraction, ϕ_O, must reach the minimum allowed value $\phi_{O,min}$ (Fig. 3*B*). Otherwise cells could grow faster by decreasing ϕ_O and redistributing that mass to catalytic tasks. It follows that $\phi_O > \phi_{O,min}$ when $\lambda < \lambda_{max}$, i.e., that cells make more noncatalytic biomass when forced to grow at submaximal rates due to nutrient limitation, for example. In the linearized model, therefore, accepting variable $\phi_O > \phi_{O,min}$ permits a continuum of submaximal growth rates $\lambda < \lambda_{max}$, with ϕ_O increasing as λ decreases (5). This observation reveals how noncatalytic C storage molecules like glycogen (26) and lipid bodies (27, 28) can serve as a means of balancing fluxes of carbon, energy, and reductant to

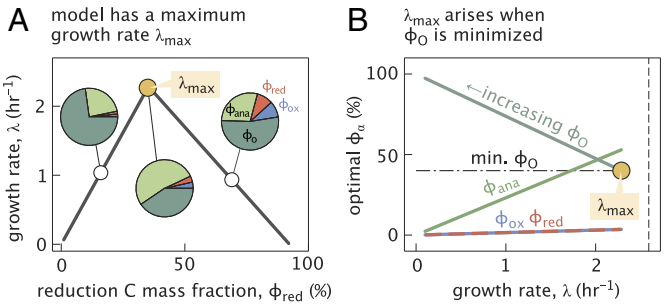

Fig. 3. The integrated model has a maximum growth rate. Panels give results from optimization of the linearized model described in the text. (*A*) A maximum growth rate, λ_{max} (yellow circle), arises because ϕ_α trade-off with each other, as shown by pie charts giving the biomass composition at marked points. When ϕ_{red} is small (red sector), for example, increasing ϕ_{red} increases λ by providing ATP for anabolism (green sector). When ϕ_{red} is large, increasing ϕ_{red} decreases λ by displacing ϕ_{ana}. The exact value of λ_{max} depends on model parameters (*SI Appendix*, Fig. S15). (*B*) Our model tracks carbon, but omits nitrogen, phosphorus, and other key nutrients. We can nonetheless simulate nutrient limitation by setting a maximum growth rate. This enables us to plot C mass fractions ϕ_α as a function of λ, analogous to chemostat experiments (14). The green line gives ϕ_{ana}; dashed red and blue give ϕ_{ox} and ϕ_{red}, respectively. λ_{max} arises when ϕ_O (dark green) achieves its minimum allowed value (dotted-dashed line). The vertical dashed line (black) gives a simple estimate of $\lambda_{max} \approx \gamma_{ana}(1 - \phi_{O,min})$ made by assuming all catalytic mass is anabolic.

enable growth at a rate set by exogenous limitations. Consistent with this view, microbes typically accumulate C storage molecules in nutrient-limited conditions, with the amount inversely related to λ (26–31).

A Need for Regulatory Mechanisms Balancing Flows of Energy, Reductant, and Carbon. Because microbes can grow on a variety of carbon sources (Fig. 1*A*), we considered respiratory growth as a function of C source redox state, $Z_{C,red}$. This value affects both oxidative e^- yield (stoichiometric coefficient ν^e_{ox}) and anabolic demand for e^- (ν^e_{ana}). Holding all other parameter values constant, we found that growth was only feasible when $Z_{C,red}$ falls in a defined range; outside this range $\lambda_{max} = 0$ (Fig. 4*A*). These limits on growth result from our assumption that anabolism consumes ATP and ECH in a defined ratio (ν^a_{ana}/ν^e_{ana} per C, Fig. 2*C*). If the C source is too reduced [e.g., $Z_{C,red} \approx -2$, as for fatty acids (3)] its oxidation yields more ECH (ν^e_{ox} increases) but anabolism consumes fewer ($-\nu^e_{ana}$ decreases). These conflicting trends make it impossible for the modeled cell to grow on very reduced substrates as there is no sink for excess reductant other than producing unneeded ATP. Likewise, problems arise if the C source is too oxidized, and these challenges persist in a nonlinear model where fluxes J_α are saturating functions of concentrations (*SI Appendix*, Figs. S12–S14).

In photosynthetic organisms, imbalances between the supply of and demand for ATP and ECH are addressed by various kinds of light-powered alternative e^- flow (32, 33). In contrast to linear e^- flow, which produces reductant and ATP in a characteristic ratio, alternative flows produce only ATP and can therefore be used to adjust the ATP:ECH production ratio to match anabolic needs, e.g., the Calvin-Benson cycle consumes 3 ATP and 4 e^- per CO_2 fixed. Based on the $Z_{C,red}$ limits of our simplified model, it is apparent that heterotrophy also requires some mechanism of ATP:ECH flux ratio balancing. Such a mechanism could include regulation of ATP production stoichiometry (i.e., altering ν^a_{ox} or ν^a_{red}), a shift in biomass composition affecting ATP or ECH demand (altering ν^a_{ana} or ν^e_{ana}), regulated ATP hydrolysis (breakdown to ADP), or some additional sink for ECH (e.g., CO_2 fixation as in ref. 34). We chose to model regulated ATP hydrolysis as this approach is simpler than others. Including an ECH sink, for example, is straightforward, but requires an additional terminal e^- acceptor or more complex flux-coupling (*SI Appendix*, section S5). Regardless of the underlying mechanisms, regulated ATP homeostasis appears to be a substantial cost for bacteria (35).

After adding a regulated ATP hydrolysis reaction (flux J_H), growth became feasible across the full range of $Z_{C,red}$ values (purple curve in Fig. 4*B*), and such homeostasis was not required for growth within the above limits. Yet, outside those limits, ATP homeostasis imposes a substantial cost: λ_{max} decreases roughly quadratically as $Z_{C,red}$ diverges from the preferred range (Fig. 4*B* and *SI Appendix*, Fig. S15). This result highlights that metabolic flexibility has a cost: Heterotrophs can arrange their metabolic stoichiometry to optimize growth on a preferred range of C sources and e^- acceptors, but switching to a substantially different metabolism (e.g., with a very different $Z_{C,red}$) requires some regulatory mechanism ensuring balanced flows of ATP, e^-, and C. Moreover, if the goal is to maintain fast growth without regulatory overhead, it is preferable to make biomass that is similar to nutrients in redox state, as λ_{max} arises when $Z_{C,B} \approx Z_{C,red}$ (Fig. 4*A*).

Thus far, our model offered two valuable lessons regarding physiology: i) carbon storage molecules enable a continuum of submaximal growth rates (Fig. 3*B*), and ii) flux-balancing regulatory mechanisms enable heterotrophic growth on carbon sources spanning a wide range of redox states (Fig. 4 *B* and *C*). This latter point might provide context for observations that light-harvesting proteins called rhodopsins are abundantly expressed by heterotrophic microbes across Earth's oceans (36), as these proteins provide a light-dependent source of ATP synthesis—an analog of alternative e^- flows for heterotrophy.

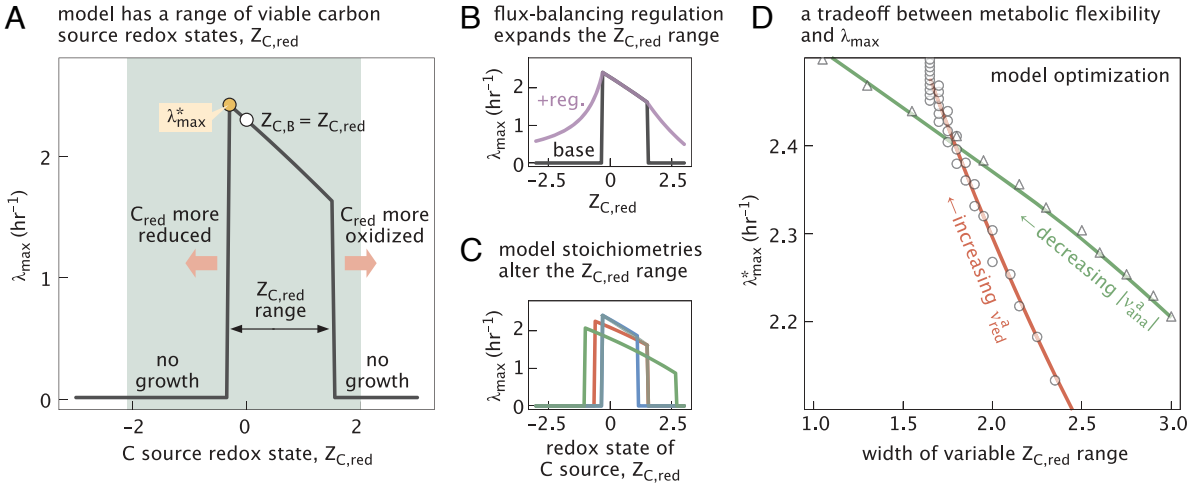

Fig. 4. Understanding the consequences of metabolic flexibility through the integrated model. Panels show results from optimization of the linearized model discussed in the text. (*A*) A model of heterotrophic respiration achieves its fastest growth rate when the C source is somewhat more reduced than biomass (yellow circle marking λ^*_{max}), consistent with the dual roles of C_{red} as a source of atoms for anabolism and e^- for energy extraction. The gray range marks the span of $Z_{C,red}$ values for organic C sources from ref. 3. If $Z_{C,red}$ and $Z_{C,B}$ were very different, growth was infeasible (where $\lambda_{max} = 0$). A model instantiation can therefore be characterized by a viable range of $Z_{C,red}$ values, as marked. The failure to grow outside this range is due to the inability to simultaneously balance flows of e^-, energy, and C, as discussed in the text. This baseline model has ATP stoichiometries $\nu^a_{ox} = 0.5$ ATP/C, $\nu^a_{red} = 1$ ATP/e^-, $\nu^a_{ana} = -0.3$ ATP/C, and $Z_{C,B} = 0$. (*B*) Adding a regulated ATP hydrolysis reaction (+reg.) enables the model to balance ATP and e^- carriers by hydrolysis of excess ATP. This permits growth with any $Z_{C,red}$ value, as shown by the purple curve. (*C*) Altering stoichiometric coefficients, e.g., of ATP production or consumption, can widen or shrink the range of $Z_{C,red}$ values that permit growth. A variant with $\nu^a_{ox} = 0.75$ is in blue, $\nu^a_{red} = 0.5$ in red, and $\nu^a_{ana} = -1$ in green. As such, regulated switching of metabolic pathways, which can alter ν^a_α, can also serve as a means of balancing flows of C, ATP, and e^-. (*D*) Varying the ATP yield of reduction (ν^a_{red}, red curve) or the ATP cost of anabolism ($-\nu^a_{ana}$, green curve) reveals a tradeoff between λ^*_{max} and metabolic flexibility. Higher ATP yields or lower ATP costs give the largest λ^*_{max} values, but over the narrowest range of viable $Z_{C,red}$ values. For detail, see *SI Appendix*, Fig. S15.

Metabolic Flexibility and Fast Growth Are Conflicting Goals. We saw that regulatory mechanisms are needed to permit growth when there is a large difference between the redox state of biomass carbon ($Z_{C,B}$) environmentally supplied C sources ($Z_{C,red}$). These mechanisms can take various forms, but inevitably take up "space" in biomass and consume resources like ATP to achieve flux balance. Such homeostasis is only required when $Z_{C,red}$ falls outside a preferred range that we calculate analytically from model parameters in *SI Appendix*, section S6. As shown in Fig. 4*C*, this range depends on the stoichiometric coefficients in the model, e.g., the ATP yield of reduction (ν^a_{red}) and the ECH requirements of anabolism (ν^e_{ana}; see *SI Appendix*, section S6). This emphasizes an equivalence between regulated pathway switching, which can alter stoichiometries, and other forms of flux-balancing, e.g., regulated ATP hydrolysis (Fig. 4*B*) or alternative e^- flows.

Altering stoichiometries—e.g., by switching oxidative (37) or reductive pathways (38, 39)—affects the preferred $Z_{C,red}$ range. The range can be widened by decreasing the reductive ATP yield (ν^a_{red}), for example, but at the expense of decreasing λ_{max} (Fig. 4*D*). Increasing anabolic ATP demand ($-\nu^a_{ana}$) had a similar effect, highlighting that a tradeoff between flexibility and fast growth is embedded in the structure of heterotrophic metabolism.

Autotrophy and Fermentation Are More Constrained than Respiration. An advantage of redox formalism is that it naturally describes the full diversity of microbial metabolism, including heterotrophic respiration (Fig. 2*C*), oxygenic photosynthesis (Fig. 5*A*), chemolithoautotrophy (*SI Appendix*, Fig. S4), and fermentation (*SI Appendix*, Fig. S5). Indeed, in our framework, equations for photosynthesis are nearly identical to those for respiration, containing one additional mass balance constraint because autotrophs produce reduced carbon (C_{red}) intracellularly and use it to make biomass.

$$\frac{d[C_{red}]}{dt} = J_{red} - J_{ana} - \lambda[C_{red}]. \quad [6]$$

Due to this additional constraint, models of autotrophy are defined by at least four equations. As such, four variables (ϕ_{ox}, ϕ_{red}, ϕ_{ana}, ϕ_O) do not permit a continuum of growth rates $\lambda < \lambda_{max}$; it was natural to include alternative e^- flow as a fifth process in models of photosynthesis ($J_{AEF} \propto \phi_{AEF}$, Fig. 5*A*).

Additional constraints on autotrophy can only decrease λ_{max}. In our linearized model, the effect on λ_{max} depends on the C_{red} concentration via the dilution term $\lambda[C_{red}]$ (*SI Appendix*, Fig. S17), but this additional constraint reduces autotrophic λ_{max} values even when $[C_{red}]$ is negligible because $J_{red} \approx J_{ana}$ is required in this limit (see *SI Appendix*, Fig. S18 for a geometric explanation). To estimate the magnitude of this effect on λ_{max}, we sampled parameter values to construct comparable models of photosynthesis and respiration, observing a roughly threefold reduction in λ_{max} (Fig. 5*B*), with the exact difference depending on the sampling procedure (*SI Appendix*, Fig. S17). Our results therefore indicate that constraints absent from respiration could explain why the fastest-growing autotrophs have longer generation times than the fastest heterotrophs (Fig. 5*C*), even when grown in nutrient surfeit (40–43).

Similar to autotrophy, minimal models of fermentation contain a mass-balance constraint not present in respiration. Fermentations are heterotrophic metabolisms where the e^- acceptor is an intracellular product of the C source, e.g., the glycolytic product pyruvate is the e^- acceptor in lactic acid fermentation of glucose (*SI Appendix*, Fig. S5). Steady-state growth therefore requires balanced fluxes of pyruvate production and consumption. Yet, as shown in *SI Appendix*, section S5.3, this constraint is mathematically distinct from autotrophy and, as such, affects fermentative growth differently.

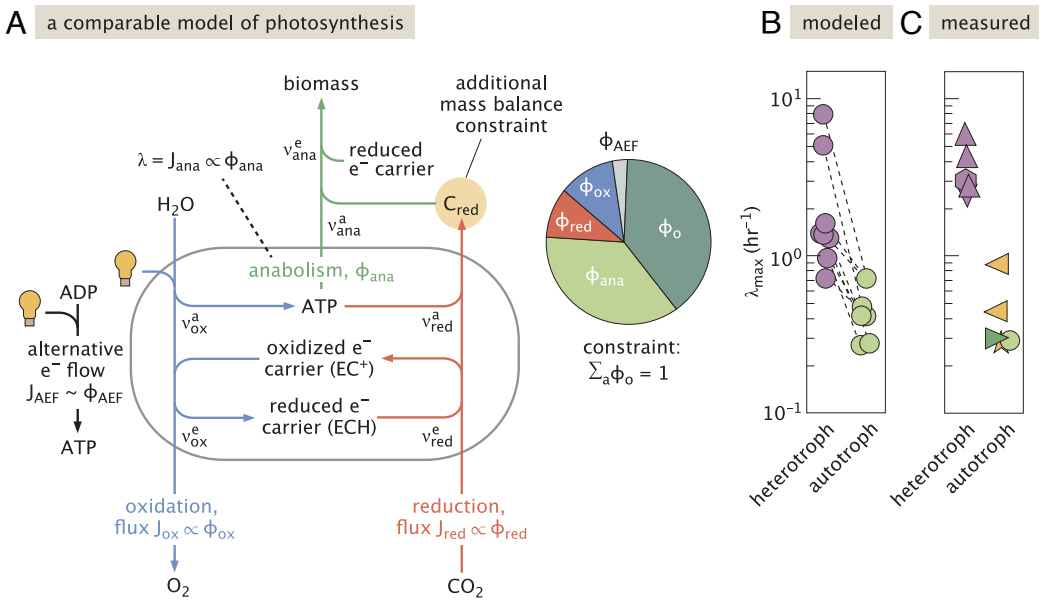

Fig. 5. A constraint imposed only on autotrophs reduces maximum growth rates. (*A*) A model of photosynthesis comprises nearly identical equations to respiration, but with key differences. Due to the use of light energy, phototrophs can balance ATP and redox carriers with new ATP synthesis, i.e., via alternative e^- flows ($J_{AEF} \propto \phi_{AEF}$). Second, because autotrophs produce and consume C_{red} intracellularly, fluxes of C_{red} production, consumption, and dilution must be balanced. Heterotrophs, in contrast, acquire C_{red} from outside the cell and can deplete it. (*B*) Comparing models of photosynthesis and respiration with identical kinetic parameters, we found that respiratory λ_{max} values exceeded photosynthetic ones by approximately threefold (compare green and purple dots). λ_{max} was calculated by optimization of the linearized model, as described in *SI Appendix*, section S4. (*C*) This result is qualitatively consistent with measurements. The fastest-growing photosynthetic microbes known grow 5 to 10 times more slowly than heterotrophic counterparts. Purple markers denote heterotrophs, green photoautotrophs, and yellow chemolithoautotrophs. Heterotrophic organisms: ▲ *V. natriegens*, ⬟ *E. coli* (γ-proteobacteria), ♦ *C. perfringens* (clostridia). Autotrophs: ⬟ *P. celeri*, ● *C. ohadii* (green algae), ► *S. elongatus* PCC 11801 (cyanobacterium), ◄ *T. crunogena* (γ-proteobacterial chemoautotroph).

Energy Economy Impacts the Redox State of Proteins. Above we saw that large differences between biomass and nutrient redox state ($|Z_{C,B} - Z_{C,red}|$) challenge heterotrophic growth (Fig. 4), raising the prospect that microbes may adapt the composition of their biomass over physiological and evolutionary timescales to reduce the magnitude of the biochemical task. Fig. 4*A* gave us a sense of the expected direction of adaptation: In nutrient surplus, microbial growth is improved by making biomass with $Z_{C,B} \approx Z_{C,red}$.

Fig. 6*A* illustrates how λ_{max} varies as a function of $Z_{C,B}$ and $Z_{C,red}$. The largest values fall just above the line of equality ($Z_{C,B} = Z_{C,red}$), indicating that the λ-maximizing $Z_{C,B}$ value is somewhat more oxidized than the C source. This is due to the need to use a fraction of the e^- from C substrates to produce ATP. The λ-maximizing $Z_{C,B}$ can be expressed as

$$Z^*_{C,B} = K_Z Z_{C,red} + Z^\circ, \quad [7]$$

where the slope, K_Z, and intercept, Z°, are functions of all model parameters and typical K_Z values are positive (*SI Appendix*, section S6.2). We interpret the right-hand side of this equation as an effective environmental redox potential, one that includes $Z_{C,red}$, the potential of the e^- acceptor (via the reductive ATP yield, ν^a_{ox}), and the magnitudes of all intracellular fluxes. We also noted that the penalty of deviating from $Z^*_{C,B}$ by Δz is approximately quadratic in Δz (Fig. 4*B* and *SI Appendix*, Fig. S16). So adapting $Z_{C,B}$ to the local environment is desirable—it reduces excess expenditures of energy and catalytic capacity, maximizing λ in rich environments.

If microbial lineages have characteristic habitats—environments with typical $Z^*_{C,B}$ values—and experience surplus with sufficient frequency, we expect evolutionary selection to push $Z_{C,B}$ toward $Z^*_{C,B}$. Yet microbes frequently experience growth-limiting nutrient and/or energy deprivation in the wild—whether in soils (45, 46), sediments (47) or elsewhere—so it is worth clarifying why we nonetheless assume selection for faster growth in rich conditions. Selection and drift proceed over generations, not over absolute time (48). If microbes episodically experience nutrient surplus, fast growth can account for a sizable fraction of generations even if it represents a minuscule fraction of the time window considered. As such, selection for increased maximum growth rate is consistent with the widespread limitation of microbial growth.

We therefore hypothesized that the redox chemistry of microbial biomass should evolve to match their environments (49), becoming more reduced when $Z^*_{C,B} - Z^\circ > K_Z Z_{C,red}$ (Fig. 6*B*) or more oxidized when $Z^*_{C,B} - Z^\circ < K_Z Z_{C,red}$. Proteins, for example, can typically accept a variety of amino acid substitutions while still folding to achieve similar structures and biochemical activities (50). Indeed, across the tree of life, aminoacyl tRNA synthetases span nearly as wide a range of Z_C values as all globular proteins in the *E. coli* genome (*SI Appendix*, Fig. S20). Amino acid substitutions that move $Z_{C,B}$ toward optimality should therefore be both permitted and favored, for example substituting arginine ($Z_C = +1/3$) with lysine ($Z_C = -2/3$) in reducing conditions. As microbial genomes encode thousands of proteins, a generic selection pressure on $Z_{C,B}$ should affect all proteins, promoting substitutions moving in the same direction and resulting in positive correlations between Z_C values of proteins in the same genome (Fig. 6*B*).

Importantly, while our analysis presumes balanced growth—i.e., time-invariant intracellular concentrations—we do not require chemical equilibrium. Rather, our model describes intracellular processes that must all carry net forward flux for growth to occur. Directional fluxes are a hallmark of nonequilibrium systems. In contrast with equilibrium descriptions (49), the magnitudes of these intracellular fluxes alter the optimal biomass redox state ($Z^*_{C,B}$) in our model.

Positive Correlations in Protein C Redox State Across the Tree of Life. To test our hypothesis, we examined a set of ≈60,000 genomes from across the bacterial tree curated by the Genome Taxonomy Database [GTDB, (44)]. GTDB also identifies sequences for 120 single-copy genes in the *bac120* gene set. Using single-copy genes enabled comparison of protein Z_C values without the ambiguity arising from paralogs, i.e., genes with shared ancestry that have evolved different functions. We found that Pearson correlations (ρ) between *bac120* Z_C values were

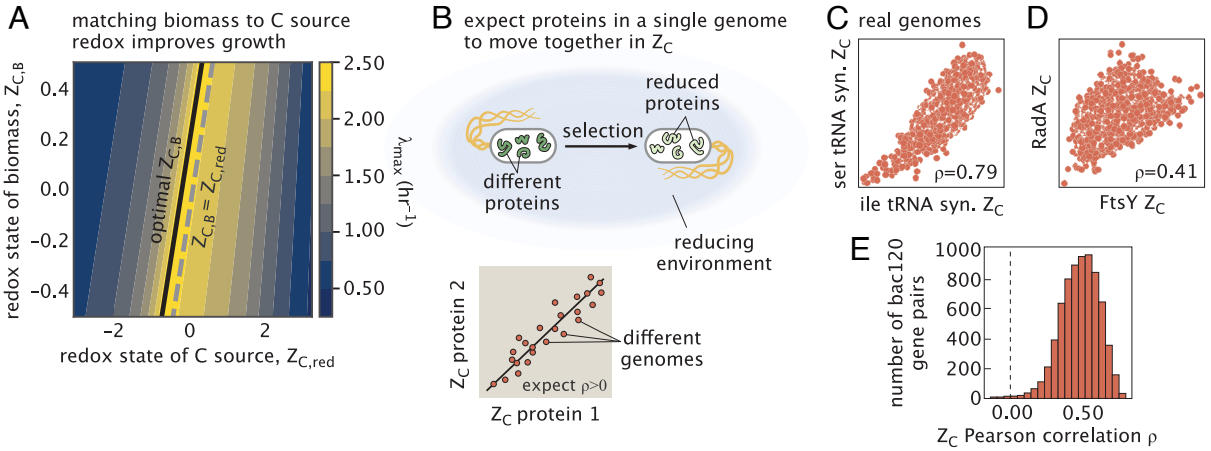

Fig. 6. Z_C values of conserved proteins are positively correlated, highlighting a redox-driven mode of protein evolution. (*A*) Optimization of our linearized model indicated that heterotrophic growth can be improved by matching the redox state of biomass ($Z_{C,B}$) to the carbon source ($Z_{C,red}$). The dashed gray line marks $Z_{C,B} = Z_{C,red}$ and the growth rate is maximized (black line) when the C source is somewhat more reduced than biomass. (*B*) If microbes are found in particular redox environments and/or have preferred metabolisms, we expect their biomass and, therefore, their proteins to take on the redox character of their environments. For example, organisms thriving in reducing environments should benefit from evolving more reduced proteins, as diagrammed. Proteins in the same genome should therefore exhibit positively correlated Z_C values. We calculated Z_C correlations for pairs of highly conserved single-copy genes across ≈60,000 genomes spanning the bacterial tree of life (44). Panel (*C*) documents strong positive correlation between Z_C values of serine and isoleucine tRNA synthetases. (*D*) Comparison of the signal recognition particle receptor, FtsY, and the DNA repair protein RadA. (*E*) Considering pairs of 120 conserved genes, Z_C correlations were almost uniformly positive, with an interquartile range of +0.41 to 0.59. See *SI Appendix*, section S1.6 and Fig. S22 for detail.

almost uniformly positive, with an interquartile range (IQR) of 0.41 to 0.59 and a 99% CI on the mean of 0.491 to 0.499 (Fig. 5 *C* and *F* and *SI Appendix*, Fig. S21).

Given their ubiquity, it is not surprising that *bac120* genes perform essential functions like transcription (6 genes), translation (70), genome replication (18), and protein secretion (6). So, in addition to recording the influence of environmental redox conditions on coding sequences, positive correlations might also reflect the prominent and mostly anabolic roles *bac120* genes play in all bacteria. Indeed, expression of many anabolic proteins—especially those involved in protein synthesis—correlates positively with λ (8, 21, 51, 52). Because increasing ribosome content results in the relative oxidation of biomass (Fig. 1*C*), one might expect anabolic enzymes to accrue reducing amino acid substitutions that compensate for oxidation, thereby preserving $Z^*_{C,B} \approx K_Z Z_{C,red} + Z^\circ$ as λ increases.

While both explanations are compatible with our model, we attempted to control for environmental effects, which should affect the entire genome, by calculating partial correlations controlling for the mean Z_C of protein-coding sequences in each genome. Partial correlations were attenuated but remained significantly biased toward positive values (IQR = 0.07 to 0.21, 99% CI = 0.135 to 0.141, *SI Appendix*, Fig. S22), indicating that both physiology and environment affect the redox state of protein carbon.

Redox Driven Protein Sequence Evolution. Proteins are typically considered to evolve toward improved individual function so long as the organism benefits from the improvement (50, 53). Based on the correlations documented in Fig. 6*F*, we propose an additional mode of protein evolution wherein mutations neutral or even weakly deleterious to the biochemical or structural functions of individual proteins might nonetheless be selected due to an organismal benefit—they cause $Z_{C,B}$ to better align with the redox chemistry of the host's typical environment (49). The affected proteins can be involved in any cellular process, but large and highly expressed proteins will affect $Z_{C,B}$ more.

Substantial changes to protein Z_C typically require more than one nucleotide substitution, however. To understand why, recall that single nucleotide substitutions tend to produce small changes in amino acid hydrophobicity metrics (54). As amino acid hydrophobicity correlates with Z_C, substitutions altering Z_C usually require multiple nucleotide changes (*SI Appendix*, Fig. S23). So the rate of mutations altering protein Z_C is lower than the average substitution rate, and larger selection coefficients are needed to drive these sequence changes to fixation (48).

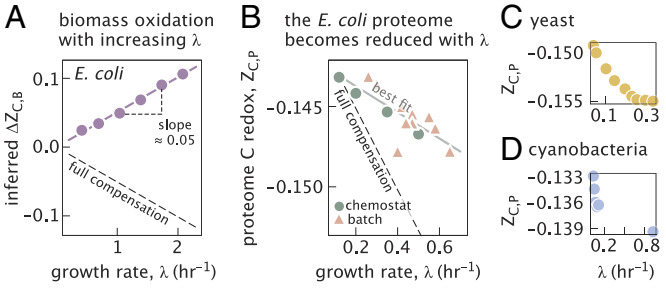

Fig. 7. Proteomes compensate for the oxidation of biomass during fast growth. (*A*) Faster-growing *E. coli* cells contain more RNA and less protein. We estimated the change in biomass C redox state, $\Delta Z_{C,B}$, due to measured λ-dependent changes in biomass composition [(14); see *SI Appendix*, section S1.4]. This calculation implied an oxidation of biomass of $\approx$0.05 Z_C units per unit λ increase (h^{-1} units), comparable to the SD of *E. coli* cytosolic proteins (σ = 0.04, *SI Appendix*, Fig. S20). Given the importance of $Z_{C,B}$ to attaining maximal growth rates in our model, we expected cells to compensate for this oxidation by making more reduced molecules during faster growth. (*B*) Quantitative proteomics in *E. coli* (21) show that the proteins expressed during fast growth are more reduced. The observed effect is of the same order as, but falls short of complete compensation for (dashed line), background oxidation estimated in (*A*). (*C* and *D*) Data from yeast (52) and cyanobacteria (51) show similar trends, suggesting a generic mechanism.

Faster Growth Induces Coordinated Shifts in the C Redox State of Proteomes. One principle arising from resource allocation logic is that, in conditions of surplus, λ is often limited by protein translation (4, 5, 7). Even in nutrient-limited conditions, higher λ typically coincides with greater ribosome activities and larger RNA:protein ratios (55, 56). Since ribosomes are $\approx$2/3 RNA by mass (10, 11) and RNA is about 1 e^-/C more oxidized than protein (3, 15), increased RNA content entails a relative oxidation of biomass at higher λ (Fig. 1*C*).

We estimated these biomass C redox changes by noting that slower-growing *E. coli* cultures are $\approx$10% RNA and $\approx$60% protein by mass ($\lambda \approx 0.4\ h^{-1}$), while fast growing cultures ($\lambda \approx 2\ h^{-1}$) are $\approx$20% RNA and $\approx$40% protein (14). Assuming $Z_C \approx +0.9$ for RNA and ≈ -0.15 for protein (15), $Z_{C,B}$ increases by $\approx$0.05 for each unit increase in λ (h^{-1} units; see *SI Appendix*, section S1.4). This value may seem small, yet, as noted above, deviations from optimal $Z_{C,B}$ lead to quadratic decreases in λ_{max} in our model. As translation cycle catalysts (ribosomes, tRNAs) are made of relatively oxidized polymers, we predicted that the remainder of biomass—proteins, lipids, storage polymers—should shift by a similar degree, but in the direction of reduction, to compensate (Fig. 7*A*).

In principle, $Z_{C,B}$ can be calculated from the elemental composition of biomass (3). While C:N:P ratios, or "Redfield ratios" (57), are commonly measured (58, 59), accurate C:H:O:S ratios are also required to calculate $Z_{C,B}$. We noted that quantitative proteomic surveys (21, 51, 52) report on the abundance and elemental composition of proteins and, therefore, the average redox state of protein C, $Z_{C,P}$. Consistent with our prediction, a comprehensive *E. coli* dataset (21) showed a systematic trend where the proteome was $\approx$0.005 units more reduced at the fastest growth rates measured (Fig. 7*B*). This work manipulated λ in chemostats and also by varying the C source in batch culture (21). These manipulations produced similar λ-$Z_{C,P}$ plots, suggesting that λ, rather than C source chemistry, is the primary determinant of $Z_{C,P}$. Assuming that all biomass constituents participate equally in compensating for RNA-induced oxidation, we estimated that the relative reduction of *E. coli* protein C ($\approx$0.01 Z_C units per λ unit) compensates for $\approx$30% of the total effect (*SI Appendix*, Fig. S25).

To see whether this observation is general, we examined additional *E. coli*, yeast (52) and cyanobacterial datasets (51). Though these microbes perform distinct metabolisms and are separated by billions of years of evolutionary divergence (60), all datasets displayed λ-dependent $Z_{C,P}$ changes of the same direction and magnitude (Fig. 7 *C* and *D* and *SI Appendix*, Fig. S24). This effect appeared to be diffuse and widely shared across the proteome because explaining $Z_{C,P}$ variance required a linear model tracking proteins with at least 5 distinct biological functions (*SI Appendix*, Fig. S26). Further, $Z_{C,P}$ values calculated from expressed proteins differed significantly from those estimated from coding sequences (*SI Appendix*, Fig. S27), indicating that i) regulation of gene expression substantially affects $Z_{C,P}$, and ii) genomes and metagenomes yield an incomplete picture of protein Z_C (15, 49).

These observations add to decades of literature documenting λ-dependent changes in microbial form and composition.

In *E. coli*, for example, fast-growing cells are larger (12), have greater mass (61), lower surface area to volume ratios (62), contain more ribosomes and, consequently, more RNA, less protein, and, typically, less lipid as a fraction of total biomass (5, 12–14). Similar trends are documented for other microbes, especially yeast (55). These rearrangements can be rationalized via resource allocation logic (4–9) and lead to the production of relatively oxidized (e^- poor) biomass during fast growth.

Beyond the observed changes in protein Z_C, we know of only one other example of biomass constituents becoming more reduced (e^- rich) during fast growth: the relative reduction of lipids with increased temperature or λ (63–65). Over the past two decades, several redox-based lipid biomarkers (e.g., TEX_{86}) have been empirically established as paleothermometers and are now widely used to infer ancient sea surface temperatures and climate over the past 100 million years (63). The mechanisms controlling lipid reduction remain debated, but recent work indicates that lipid biomarker chemistry is primarily controlled by growth rate (64). Our integrated model explains this relationship as optimizing microbial growth by compensating for the relative oxidation of biomass due to production of ribosomes during faster growth. In *SI Appendix*, Fig. S28, we calculate the Z_C of total *E. coli* lipids (66), finding λ-dependent changes of the same sign and similar magnitude as the proteome, suggesting a common mechanism.

Discussion

Unifying Redox Chemistry and Cellular Resource Allocation. Here, we developed a quantitative framework for building simple models coupling metabolic chemistry to the growth and persistence of cells. This framework aims to express the impressive flexibility and diversity of microbial metabolism by integrating two approaches: i) using redox chemistry to describe the flows of e^- and atoms producing energy and biomass (Fig. 2*A*) and ii) using principles of cellular resource allocation to understand how microbes apportion their biosynthetic activity to achieve certain metabolic rates (Fig. 2*B*).

We rely on redox chemistry to concisely describe (coarse-grain) metabolism as coupling oxidation and reduction reactions to load energy carriers (ATP) and reducing equivalents (ECH) for anabolism, as in Fig. 2*C*. One advantage of redox models is their chemical-specificity: Oxidations of glucose ($Z_C = 0$), pyruvate ($Z_C = +2/3$), and palmitate ($Z_C = -1.75$) have different ECH stoichiometry, and respiration, fermentation, and photosynthesis are represented by comparable, yet distinct equations. Coarse-grained models of cellular resource allocation, by contrast, lack such chemical specificity (4–9).

There is also a close relationship between redox reactions, thermodynamic potentials, and the principle of "microbial infallibility." This impressively predictive principle asserts that wherever there is energy to be gained from environmental nutrients, microbes "find a way" to sustain themselves (67–71), i.e., evolve a workable combination of metabolic couplings. Redox formalism enables the use of laboratory measurements, e.g., redox potentials, to calculate the energy content of different metabolic couplings and, in so doing, sets upper bounds on the ATP yields of diverse metabolisms (Box 1 and *SI Appendix*, section S2). Quantitative connections between metabolic chemistry and bioenergetics are useful in engineering (72) and environmental science, explaining why some substrates (e.g., O_2) are consumed before others [e.g., Fe(III)] in natural (73–76) and laboratory settings (77). A final advantage of redox formalism is that it naturally tracks flows of atoms, reducing equivalents, and energy carriers so that models can enforce the simultaneous conservation of energy and matter (*SI Appendix*, section S3) and interrogate the effects of imbalances (Fig. 4).

Despite their many attractive qualities, redox models omit a crucial aspect of growth, namely self-replication or autocatalysis. Biomass is not just the product of anabolism, but also composed of myriad enzymes catalyzing processes like nutrient transport, translation, and lipid biosynthesis. By apportioning anabolic activity to the production of catalysts for key cellular processes, resource allocation models concisely express the catalytic nature of biomass. Typically, such models subdivide protein translation (Fig. 2*B*). Indeed, proteins are the primary effectors in cells and the focus on proteins permits direct comparison with high-throughput studies of protein expression (21, 51, 52). Yet, a crucial step in developing a unified framework was to abandon the subdivision of protein translation and subdivide total anabolic activity instead (Fig. 2*C*). This was essential because redox descriptions of growth balance flows of atoms, energy, and e^- to produce all of biomass, not just protein.

Our integrated models can therefore enforce the simultaneous conservation of matter, energy, and biosynthetic activity. As these models are coarse-grained and chemically specific, they enabled us to reason about the constraints imposed by metabolic chemistry on growth. An overview of model assumptions and limitations is given in *SI Appendix*, section S1.2.

Here, we focused primarily on the limits of metabolic flexibility during respiratory growth (Figs. 3 and 4) and constraints unique to autotrophy (Fig. 5*A*). Despite efforts to find superlative photoautotrophs (41, 42), the fastest known heterotrophs grow much faster (Fig. 5*C*). We found that this discrepancy likely arises because autotrophs must balance intracellular production and consumption of reduced carbon, while heterotrophs can deplete environmentally supplied C_{red}. This constraint substantially reduces λ_{max} in our models (*SI Appendix*, Figs. S17 and S18). Indeed, photoheterotrophic growth of green algae—a mixed metabolism where cells photosynthesize and consume extracellular organics at the same time—is notably faster than growth by photosynthesis alone (42, 78). It appears, therefore, that the requirement to balance intracellularly produced C_{red} limits autotrophic growth rates both in theory and in practice.

Our exploration of heterotrophic respiration highlighted the need for regulated mechanisms of flux balancing, akin to alternative e^- flows in photosynthesis (Fig. 4). Regulated mechanisms of flux balancing enabled growth on a wide range of C sources; however, these mechanisms came at a cost, consuming energy and requiring that cells dedicate resources to the synthesis of enzymes and regulators. Reducing this expenditure by generating biomass that matches the environment produced faster maximum growth rates in the model (Fig. 6*A*), which led us to posit the existence of a generic selection pressure for microbes to match the redox state of biomass ($Z_{C,B}$) to their habitats (49). Consistent with this prediction, we found strong and widespread positive correlations between the Z_C values of conserved proteins (Fig. 6), revealing that proteins are subject to shared evolutionary pressures along a compositional dimension not directly related to their individual functions. Additional support for this unexpected mode of protein evolution is given by our finding that diverse microbes express comparatively reduced proteins during fast growth (Fig. 7), apparently using the proteome as an e^- acceptor to compensate for the relative oxidation of biomass due to the RNA content of ribosomes.

Context and Conclusions. Our work here is situated in a long history of modeling microbial physiology. Prior contributions

come in a variety of "flavors," modeling growth at various levels of detail and emphasizing different aspects of living matter. Here, we briefly discuss approaches related to, yet distinct from, our efforts here.

Stoichiometric constraint-based modeling (CBM) is an influential framework developed by chemical engineers. Like redox models, CBMs enumerate reactions in a system of interest, whether a single pathway, core metabolism, or a whole cell (79). The reaction network is described by a stoichiometric matrix, S, used to enforce steady-state mass balance of intracellular metabolites, i.e., an identifying a vector of reaction fluxes $\mathbf{v}$ for which $S \cdot \mathbf{v} = 0$ (see ref. 80 for a mathematical introduction). For larger networks, many such vectors exist, so CBMs are often numerically optimized to achieve some goal, e.g., identifying genetic knockouts that maximize fatty acids production (81). As CBMs preserve the relationship between genes, enzymes, and reactions, they write full chemical reactions rather than half-reactions. This makes it difficult to apply redox principles for coarse-graining as we did here. Indeed, microbial genomes often encode >1,000 enzymes (82), so a focus on genes is essentially incompatible with coarse-graining.

CBMs are, nonetheless, exquisitely chemically specific. Indeed, notable prior efforts to integrate chemical-specificity and resource allocation relied on genome-scale CBMs (83, 84). While these integrated models express the constraints of self-replication in a chemically specific setting and make predictions consistent with measurements, the detail of genome-scale models makes it difficult to extract principles and impractical to compare distinct metabolisms, as we did here for photosynthesis and respiration.

Environmental scientists often apply more compact models describing the effect of nutrient concentrations on microbial growth (85). This interest follows Albert Redfield's landmark observations that oceanic concentrations of inorganic phosphorus (P), nitrogen (N), oxygen, and carbon covary in the rough proportions one might expect from the oxidation of biomass with typical C:N:P ratios (57). Although the modern picture is more complex (59), these observations left many scientists wondering how such characteristic ratios arise (86) and how microbes cope with imbalances in the concentrations of multiple nutrients. Nutrient quota models posit that growth rates are controlled by intracellular concentrations (quotas Q), which have minimum values reflecting limits on cellular composition (85, 87–89). As protein, lipid, nucleic acids, and carbohydrates have characteristic C:N:P ratios (intimately related to their Z_C values), the elemental content of cells is tightly coupled to their macromolecular composition. This coupling has been understood for decades (12, 14, 90, 91) and integrated into quota models (85, 92, 93). Yet such models omit autocatalysis and the redox basis of bioenergetics, highlighting an opportunity to extend our framework to address questions in environmental science.

While in this work we track only carbon, energy, and electrons, our framework could be modified to consider N and P quotas, as well as the costs of accessing these nutrients (89, 93). Other topics of interest include nutrients that have a redox character [e.g., O_2 (94)] as well as intrinsic constraints on chemolithoautotrophy, fermentation, and heterotrophic carbon use efficiency, as these have implications for the origins of metabolism (95), anoxic heterotrophy on the early Earth (96), and vital attempts to predict microbial CO_2 emissions (19).

All of our results here derive from having integrated two potent, but disparate, views of cellular physiology—merging redox and resource allocation models. This integration couples intracellular processes—regulatory, structural, and catalytic—to the extracellular redox environment, whether in a lake, an animal gut, a tumor matrix, or a soil pore. How should the character of the chemical environment affect the metabolism and physiology of diverse microbes? Our integrated, quantitative approach opens the door to asking and answering such questions across domains of life.

Materials and Methods

Unified Models of Microbial Physiology. Model derivations, assumptions, parameter ranges, optimization, and model simulations are detailed in supplement. While our framework admits arbitrary equations describing the kinetics of metabolic processes, we mostly explored a linearized model in the main text. Analysis of nonlinear models is discussed in *SI Appendix*, section S7. Code for all analyses is available at github.com/flamholz/redox-proteome.

Data, Materials, and Software Availability. Source code for models and analyses are available on GitHub (https://github.com/flamholz/redox-proteome) (97). All other data are included in the manuscript and/or supporting information.

ACKNOWLEDGMENTS. We are grateful to E. Afik, L. Aristilde, A. Burlacot, J. Ciemniecki, A. Duarte, J. Goldford, S. Hirokawa, D. McRose, R. Murali, T. Roeschinger, and G. Salmon for useful discussions and to Y. M. Bar-On, G. Chure, S. Kuehn, D. LaRowe, and R. Milo for comments on the manuscript. Financial support from the NSF (PHY-1748958 to the Kavli Institute for Theoretical Physics to A.I.F. and A.G.), fellowships from the Jane Coffin Childs Memorial Fund and Burroughs Wellcome Fund (to A.I.F.), and the Govt of India's Ramalingaswami Fellowship (to A.G.); Additional support from the Gordon and Betty Moore Foundation (grant GBMF4513 to AG), the Caltech Center for Evolutionary Sciences (W.W.F.), NIH (1R01AI127850-01A1 to D.K.N.), the Rosen Center at Caltech and NIH MIRA grant 1R35 GM118043 (to R.P.).

Author affiliations: [a]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125; [b]Physics of Living Systems, Department of Physics, Massachusetts Institute of Technology, Cambridge MA 02139; [c]International Centre for Theoretical Sciences, Tata Institute of Fundamental Research, Bengaluru 560089, India; [d]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125; and [e]Division of Physics, Mathematics and Astronomy, California Institute of Technology, Pasadena, CA 91125

1. C. G. Friedrich, B. Friedrich, B. Bowien, Formation of enzymes of autotrophic metabolism during heterotrophic growth of *Alcaligenes eutrophus*. *Microbiology* **122**, 69–78 (1981).
2. F. M. Harold, *The Vital Force: A Study of Bioenergetics* (W H Freeman & Co, 1987).
3. D. E. LaRowe, P. Van Cappellen, Degradation of natural organic matter: A thermodynamic analysis. *Geochim. Cosmochim. Acta* **75**, 2030–2042 (2011).
4. D. Molenaar, R. van Berlo, D. de Ridder, B. Teusink, Shifts in growth strategies reflect tradeoffs in cellular economics. *Mol. Syst. Biol.* **5**, 323 (2009).
5. M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, T. Hwa, Interdependence of cell growth and gene expression: Origins and consequences. *Science* **330**, 1099–1102 (2010).
6. M. Basan *et al*., Overflow metabolism in *Escherichia coli* results from efficient proteome allocation. *Nature* **528**, 99–104 (2015).
7. N. M. Belliveau *et al*., Fundamental limits on the rate of bacterial growth and their influence on proteomic composition. *Cell Syst.* **12**, 924–944.e2 (2021).
8. R. Balakrishnan *et al*., Principles of gene regulation quantitatively connect DNA to RNA and proteins in bacteria. *Science* **378**, eabk2066 (2022).
9. G. Chure, J. Cremer, An optimal regulation of fluxes dictates microbial growth in and out of steady state. *eLife* **12**, e84878 (2023).
10. S. Reuveni, M. Ehrenberg, J. Paulsson, Ribosomes are optimized for autocatalytic production. *Nature* **547**, 293–297 (2017).
11. V. Ramakrishnan, Ribosome structure and the mechanism of translation. *Cell* **108**, 557–572 (2002).
12. M. Schaechter, O. Maaloe, N. O. Kjeldgaard, Dependency on medium and temperature of cell size and chemical composition during balanced grown of *Salmonella typhimurium*. *J. Gen. Microbiol.* **19**, 592–606 (1958).
13. F. C. Neidhardt, B. Magasanik, Studies on the role of ribonucleic acid in the growth of bacteria. *Biochim. Biophys. Acta* **42**, 99–116 (1960).
14. H. Bremer, P. P. Dennis, Modulation of chemical composition and other parameters of the cell at different exponential growth rates. *EcoSal Plus* **3**, 1–49 (2008).
15. J. M. Dick, M. Yu, J. Tan, A. Lu, Changes in carbon oxidation state of metagenomes along geochemical redox gradients. *Front. Microbiol.* **10**, 120 (2019).

16. P. G. Falkowski, T. Fenchel, E. F. Delong, The microbial engines that drive earth's biogeochemical cycles. *Science* **320**, 1034–1039 (2008).
17. J. P. Amend, E. L. Shock, Energetics of overall metabolic reactions of thermophilic and hyperthermophilic Archaea and bacteria. *FEMS Microbiol. Rev.* **25**, 175–243 (2001).
18. D. H. Rothman, D. C. Forney, Physical model for the decay and preservation of marine organic carbon. *Science* **316**, 1325–1328 (2007).
19. W. R. Wieder, G. B. Bonan, S. D. Allison, Global soil carbon projections are improved by modelling microbial processes. *Nat. Clim. Chang.* **3**, 909–912 (2013).
20. P. W. Crockford, Y. M. Bar On, L. M. Ward, R. Milo, I. Halevy, The geologic history of primary productivity. *Biol. Curr.* **33**, 4741–4750.e5 (2023).
21. A. Schmidt *et al*., The quantitative and condition-dependent *Escherichia coli* proteome. *Nat. Biotechnol.* **34**, 104–110 (2016).
22. R. Milo, R. Phillips, *Cell Biology by the Numbers* (Garland Science, 2015).
23. P. Mitchell, Chemiosmotic coupling in oxidative and photosynthetic phosphorylation. *Biol. Rev. Camb. Philos. Soc.* **41**, 445–502 (1966).
24. S. Manzoni, P. Taylor, A. Richter, A. Porporato, G. I. Ågren, Environmental and stoichiometric controls on microbial carbon-use efficiency in soils. *New Phytol.* **196**, 79–91 (2012).
25. E. J. Zakem, M. F. Polz, M. J. Follows, Redox-informed models of global biogeochemical cycles. *Nat. Commun.* **11**, 5680 (2020).
26. J. Preiss, Bacterial glycogen synthesis and its regulation. *Annu. Rev. Microbiol.* **38**, 419–458 (1984).
27. Z. T. Wang, N. Ullrich, S. Joo, S. Waffenschmidt, U. Goodenough, Algal lipid bodies: stress induction, purification, and biochemical characterization in wild-type and starchless *Chlamydomonas reinhardtii*. *Eukaryot. Cell* **8**, 1856–1868 (2009).
28. A. J. Klok, D. E. Martens, R. H. Wijffels, P. P. Lamers, Simultaneous growth and neutral lipid accumulation in microalgae. *Bioresour. Technol.* **134**, 233–243 (2013).
29. T. Holme *et al*., Continuous culture studies on glycogen synthesis in *Escherichia coli* B. *Acta Chem. Scand.* **11**, 763–775 (1957).
30. H. Lees, J. R. Postgate, The behaviour of Azotobacter chroococcum in oxygen- and phosphate-limited chemostat culture. *J. Gen. Microbiol.* **75**, 161–166 (1973).
31. J. D. Linton, R. E. Cripps, The occurrence and identification of intracellular polyglucose storage granules in methylococcus NCIB 11083 grown in chemostat culture on methane. *Arch. Microbiol.* **117**, 41–48 (1978).
32. D. I. Arnon, Conversion of light into chemical energy in photosynthesis. *Nature* **184**, 10–21 (1959).
33. A. Burlacot, Quantifying the roles of algal photosynthetic electron pathways: A milestone towards photosynthetic robustness. *New Phytol.* **240**, 2197–2203 (2023).
34. J. B. McKinlay, C. S. Harwood, Carbon dioxide fixation as a central redox cofactor recycling mechanism in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11669–11675 (2010).
35. W. H. Lin, C. Jacobs-Wagner, Connecting single-cell ATP dynamics to overflow metabolism, cell growth, and the cell cycle in *Escherichia coli*. *Curr. Biol.* **32**, 3911–3924.e4 (2022).
36. L. Gómez-Consarnau *et al*., Microbial rhodopsins are major contributors to the solar energy captured in the sea. *Sci. Adv.* **5**, eaaw8855 (2019).
37. A. Flamholz, E. Noor, A. Bar-Even, W. Liebermeister, R. Milo, Glycolytic strategy as a tradeoff between energy yield and protein cost. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 10039–10044 (2013).
38. M. Bekker, S. de Vries, A. Ter Beek, K. J. Hellingwerf, M. J. T. de Mattos, Respiration of *Escherichia coli* can be fully uncoupled via the nonelectrogenic terminal cytochrome *bd*-II oxidase. *J. Bacteriol.* **191**, 5510–5517 (2009).
39. V. B. Borisov *et al*., Aerobic respiratory chain of *Escherichia coli* is not allowed to work in fully uncoupled mode. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 17320–17324 (2011).
40. H. W. Jannasch, C. O. Wirsen, D. C. Nelson, L. A. Robertson, Thiomicrospira crunogena sp. nov., a colorless, Sulfur-Oxidizing bacterium from a Deep-Sea hydrothermal vent. *Int. J. Syst. Bacteriol.* **35**, 422–424 (1985).
41. D. Jaiswal *et al*., Genome features and biochemical characteristics of a robust, fast growing and naturally transformable cyanobacterium *Synechococcus elongatus* PCC 11801 isolated from India. *Sci. Rep.* **8**, 16632 (2018).
42. H. Treves *et al*., Metabolic flexibility underpins growth capabilities of the fastest growing alga. *Curr. Biol.* **27**, 2559–2567.e3 (2017).
43. A. Krishnan *et al*., Picochlorum celeri as a model system for robust outdoor algal growth in seawater. *Sci. Rep.* **11**, 11649 (2021).
44. D. H. Parks *et al*., A complete domain-to-species taxonomy for bacteria and archaea. *Nat. Biotechnol.* **38**, 1079–1086 (2020).
45. T. A. Caro, J. McFarlin, S. Jech, N. Fierer, S. Kopf, Hydrogen stable isotope probing of lipids demonstrates slow rates of microbial growth in soil. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2211625120 (2023).
46. S. J. Blazewicz *et al*., Taxon-specific microbial growth and mortality patterns reveal distinct temporal population responses to rewetting in a California grassland soil. *ISME J.* **14**, 1520–1532 (2020).
47. F. Beulig, H. Røy, C. Glombitza, B. B. Jørgensen, Control on rate and pathway of anaerobic organic carbon degradation in the seabed. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 367–372 (2018).
48. T. Ohta, The nearly neutral theory of molecular evolution. *Annu. Rev. Ecol. Syst.* **23**, 263–286 (1992).
49. J. M. Dick, D. Meng, Community- and genome-based evidence for a shaping influence of redox potential on bacterial protein evolution. *mSystems* **8**, e0001423 (2023).
50. M. A. DePristo, D. M. Weinreich, D. L. Hartl, Missense meanderings in sequence space: A biophysical view of protein evolution. *Nat. Rev. Genet.* **6**, 678–687 (2005).
51. T. Zavřel *et al*., Quantitative insights into the cyanobacterial cell economy. *eLife* **8**, e42508 (2019).
52. J. Xia *et al*., Proteome allocations change linearly with the specific growth rate of *Saccharomyces cerevisiae* under glucose limitation. *Nat. Commun.* **13**, 2819 (2022).
53. L. Noda-Garcia, W. Liebermeister, D. S. Tawfik, Metabolite-Enzyme coevolution: From single enzymes to metabolic pathways and networks. *Annu. Rev. Biochem.* **87**, 187–216 (2018).
54. D. Haig, L. D. Hurst, A quantitative measure of error minimization in the genetic code. *J. Mol. Evol.* **33**, 412–417 (1991).
55. E. Metzl-Raz *et al*., Principles of cellular resource allocation revealed by condition-dependent proteome profiling. *eLife* **6**, e28034 (2017).
56. S. H. J. Li *et al*., *Escherichia coli* translation strategies differ across carbon, nitrogen and phosphorus limitation conditions. *Nat. Microbiol.* **3**, 939–947 (2018).
57. A. C. Redifield, "On the proportions of organic derivatives in sea water and their relation to the composition of plankton" in *James Johnstone Memorial*, R. J. Daniel, Ed. (University Press of Liverpool, 1934), pp. 176–192.
58. C. C. Cleveland, D. Liptzin, C: N: P stoichiometry in soil: is there a "redfield ratio" for the microbial biomass? *Biogeochemistry* **85**, 235–252 (2007).
59. A. C. Martiny *et al*., Strong latitudinal patterns in the elemental ratios of marine plankton and organic matter. *Nat. Geosci.* **6**, 279–283 (2013).
60. S. Kumar *et al*., TimeTree 5: An expanded resource for species divergence times. *Mol. Biol. Evol.* **39**, msac174 (2022).
61. B. R. K. Roller *et al*., Single-cell mass distributions reveal simple rules for achieving steady-state growth. *MBio* **14**, e0158523 (2023).
62. L. K. Harris, J. A. Theriot, Surface area to volume ratio: A natural variable for bacterial morphogenesis. *Trends Microbiol.* **26**, 815–832 (2018).
63. S. Schouten, E. C. Hopmans, E. Schefuß, J. S. Sinninghe Damsté, Distributional variations in marine crenarchaeotal membrane lipids: A new tool for reconstructing ancient sea water temperatures? *Earth Planet. Sci. Lett.* **204**, 265–274 (2002).
64. S. J. Hurley *et al*., Influence of ammonia oxidation rate on thaumarchaeal lipid composition and the TEX86 temperature proxy. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7762–7767 (2016).
65. H. C. Holm *et al*., Global ocean lipidomes show a universal relationship between temperature and lipid unsaturation. *Science* **376**, 1487–1491 (2022).
66. A. G. Marr, J. L. Ingraham, Effect of temperature on the composition of fatty acids in *Escherichia coli*. *J. Bacteriol.* **84**, 1260–1267 (1962).
67. E. F. Gale, *The Chemical Activities of Bacteria* (Academic Press, New York, 1951).
68. M. Alexander, Biochemical ecology of soil microorganisms. *Annu. Rev. Microbiol.* **18**, 217–252 (1964).
69. M. Alexander, Biodegradation of chemicals of environmental concern. *Science* **211**, 132–138 (1981).
70. E. Broda, Two kinds of lithotrophs missing in nature. *Z. Allg. Mikrobiol.* **17**, 491–493 (1977).
71. M. A. O'Malley, D. A. Walsh, Rethinking microbial infallibility in the metagenomics era. *FEMS Microbiol. Ecol.* **97**, fiab092 (2021).
72. B. E. Rittmann, P. L. McCarty, *Environmental Biotechnology: Principles and Applications* (McGraw-Hill Education, Columbus, OH, 2018).
73. D. R. Champ, J. Gulens, R. E. Jackson, Oxidation-reduction sequences in ground water flow systems. *Can. J. Earth Sci.* **16**, 12–23 (1979).
74. P. N. Froelich *et al*., Early oxidation of organic matter in pelagic sediments of the eastern equatorial Atlantic: Auboxic diagenesis. *Geochim. Cosmochim. Acta* **43**, 1075–1090 (1979).
75. W. H. Patrick Jr., R. E. Henderson, Reduction and reoxidation cycles of manganese and iron in flooded soil and in water solution. *Soil Sci. Soc. Am. J.* **45**, 855–859 (1981).
76. H. L. Ehrlich, *Ehrlich's Geomicrobiology*, H. L. Ehrlich, D. K. Newman, A. Kappler, Eds. (CRC Press, London, England, ed. 2, 2021).
77. I. Babcsányi, F. Meite, G. Imfeld, Biogeochemical gradients and microbial communities in winogradsky columns established with polluted wetland sediments. *FEMS Microbiol. Ecol.* **93**, fix089 (2017).
78. G. Lalibertè, J. de la Noüie, Auto-, hetero-, and mixotrophic growth of Chlamydomonas Humicola (chlorophyceae) on acetate. *J. Phycol.* **29**, 612–620 (1993).
79. A. Bordbar, J. M. Monk, Z. King, B. O. Palsson, Constraint-based models predict metabolic and associated cellular functions. *Nat. Rev. Genet.* **15**, 107–120 (2014).
80. W. Gottstein, B. G. Olivier, F. J. Bruggeman, B. Teusink, Constraint-based stoichiometric modelling from single organisms to microbial communities. *J. R. Soc. Interface* **13**, 20160627 (2016).
81. S. Ranganathan *et al*., An integrated computational and experimental study for overproducing fatty acids in *Escherichia coli*. *Metab. Eng.* **14**, 687–704 (2012).
82. J. M. Monk *et al*., iML1515, A knowledgebase that computes *Escherichia coli* traits. *Nat. Biotechnol.* **35**, 904–908 (2017).
83. A. Goelzer *et al*., Quantitative prediction of genome-wide resource allocation in bacteria. *Metab. Eng.* **32**, 232–243 (2015).
84. M. Mori, T. Hwa, O. C. Martin, A. De Martino, E. Marinari, Constrained allocation flux balance analysis. *PLoS Comput. Biol.* **12**, e1004913 (2016).
85. R. W. Sterner, J. J. Elser, *Ecological Stoichiometry: The Biology of Elements from Molecules to the Biosphere* (Princeton University Press, Princeton, NJ, 2003).
86. S. Sharoni, I. Halevy, Nutrient ratios in marine particulate organic matter are predicted by the population structure of well-adapted phytoplankton. *Sci. Adv.* **6**, eaaw9371 (2020).
87. M. R. Droop, Vitamin B_{12} and marine ecology. IV. The kinetics of uptake, growth and inhibition in *Monochrysis lutheri*. *J. Mar. Biol. Assoc. U. K.* **48**, 689–733 (1968).
88. M. R. Droop, The nutrient status of algal cells in continuous culture. *J. Mar. Biol. Assoc. U. K.* **54**, 825–855 (1974).
89. C. A. Klausmeier, E. Litchman, S. A. Levin, A model of flexible uptake of two essential resources. *J. Theor. Biol.* **246**, 278–289 (2007).
90. T. R. Parsons, K. Stephens, J. D. H. Strickland, On the chemical composition of eleven species of marine phytoplankters. *J. Fish. Res. Bd. Can.* **18**, 1001–1016 (1961).
91. P. J. Harrison, P. A. Thompson, G. S. Calderwood, Effects of nutrient and light limitation on the biochemical composition of phytoplankton. *J. Appl. Phycol.* **2**, 45–56 (1990).
92. B. Shuter, A model of physiological adaptation in unicellular algae. *J. Theor. Biol.* **78**, 519–552 (1979).
93. C. A. Klausmeier, E. Litchman, T. Daufresne, S. A. Levin, Optimal nitrogen-to-phosphorus stoichiometry of phytoplankton. *Nature* **429**, 171–174 (2004).
94. O. Ulloa, D. E. Canfield, E. F. DeLong, R. M. Letelier, F. J. Stewart, Microbial oceanography of anoxic oxygen minimum zones. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 15996–16003 (2012).
95. W. Martin, M. J. Russell, On the origin of biochemistry at an alkaline hydrothermal vent. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **362**, 1887–1925 (2007).
96. N. Lane, J. F. Allen, W. Martin, How did LUCA make a living? Chemiosmosis in the origin of life *BioEssays* **32**, 271–280 (2010).
97. A. I. Flamholz, A. K. Goyal, Source code. Zenodo. https://doi.org/10.5281/zenodo.14427140. Deposited 12 December 2024.