Proteome efficiency of metabolic pathways in *Escherichia coli* increases along the nutrient flow

Xiao-Pan Hu,[a,b#] Stefan Schroeder,[a,b] Martin J. Lercher[a,b#]

[a] Institute for Computer Science, Heinrich Heine University, Düsseldorf, Germany

[b] Department of Biology, Heinrich Heine University, Düsseldorf, Germany

[#]Address correspondence to Xiao-Pan Hu, xiao-pan.hu@hhu.de and Martin J. Lercher, martin.lercher@hhu.de

Abstract

Understanding the allocation of the cellular proteome to different cellular processes is central to unraveling the organizing principles of bacterial physiology. Proteome allocation to protein translation itself is maximally efficient, i.e., it represents the minimal allocation of dry mass able to sustain the observed protein production rate. In contrast, recent studies on bacteria have demonstrated that the concentrations of many proteins exceed the minimal level required to support the observed growth rate, indicating some heterogeneity across pathways in their proteome efficiency. Here, we systematically analyze the proteome efficiency of metabolic pathways, which together account for more than half of the *E. coli* proteome during exponential growth. Comparing the predicted minimal and the observed proteome allocation to different metabolic pathways across growth conditions, we find that the most costly biosynthesis pathways – those for amino acid biosynthesis and cofactor biosynthesis – are expressed for near optimal efficiency. Overall, proteome efficiency increases along the carbon flow through the metabolic network: proteins involved in pathways of nutrient uptake and central metabolism tend to be highly over-abundant, while proteins involved in anabolic pathways and in protein translation are much closer to the expected minimal abundance across conditions. Our work thus provides a bird's-eye view of metabolic pathway efficiency, demonstrating systematic deviations from optimal cellular efficiency at the network level.

Importance

Protein translation is the most expensive cellular process in fast-growing bacteria, and efficient proteome usage should thus be under strong natural selection. However, recent studies show that a considerable part of the proteome is unneeded for instantaneous cell growth in *E. coli*. We still lack a systematic understanding of how this excess proteome is distributed across different pathways as a function of the growth conditions. We estimated the minimal required proteome across growth conditions in *E. coli* and compared the predictions with experimental data. We found that the proteome allocated to the most

expensive internal pathways, including translation and the synthesis of amino acids and cofactors, are near the minimally required levels. In contrast, transporters and central carbon metabolism show much higher proteome levels than the predicted minimal abundance. Our analyses show that the proteome fraction unneeded for instantaneous cell growth decreases along the nutrient flow in *E. coli*.

Keywords

Resource allocation, proteome efficiency, growth rate, growth law, metabolic pathways, biosynthetic pathways, central carbon metabolism, glyoxylate shunt

Introduction

Proteins account for more than half of the cell dry mass in *E. coli* (1) and drive most biological processes. How and why proteome is allocated to different cellular processes and pathways is a vital question for understanding the principles behind bacterial physiology (2). Proteome allocation into different groups of genes is growth rate-dependent (3). When partitioning the proteome into specific, coarse-grained "sectors", the corresponding proteome fractions follow simple, empirical growth laws, increasing or decreasing linearly with the growth rate μ (4–7). For example, the proteome fraction allocated to the ribosome and ribosome-affiliated proteins (the R-sector (6)) scales as a linear function of growth rate under nutrient limiting conditions (4).

Why does the proteome composition scale with the growth rate? Protein is the most abundant and costly macromolecule in bacterial cells. It has thus been speculated that the proteome composition is adjusted to the specific growth condition to maximize the growth rate (8). If this were true, all protein concentrations would be at the minimal level required to sustain the observed cellular growth rate. This simple assumption has been widely used in computational models of cellular growth (9–15). However, even if proteome allocation had evolved to be maximally efficient, it is not obvious that this efficiency would simply maximize the instantaneous growth rate. Instead, it appears likely that proteome allocation has evolved to maximize cellular fitness in unpredictable, dynamic environments with varying nutrients and involving periods of famine and stresses (8). Indeed, recent experimental work indicates that the proteome is not expressed for maximal efficiency in unevolved *E. coli* strains, at least not in the naïve sense of maximizing the instantaneous growth rate. First, a large fraction of the expressed proteome is unneeded for the current environment, especially at low growth rates (16). Second, the growth rate can increase by approx. 20% in a few hundred generations in adaptive laboratory evolution experiments on minimal media (17), a process associated with reductions in the abundance of unused proteins (16). Finally, the fluxes through some cellular processes, e.g., nutrient transport and energy production, are not limited by specific proteins in these pathways at low growth rates (18). Thus, *E. coli* proteome allocation seems not to be globally optimized for maximizing the instantaneous growth rate.

On the pathway level, however, proteome allocation to at least one cellular process – protein translation – is optimized for maximal efficiency at the given protein synthesis rate (18–21). This indicates that while the global allocation of proteins is not always optimized for maximal growth rate, the proteome allocation to some cellular pathways is at a local optimum – i.e.,

the individual pathway utilizes the minimal protein mass required to support the observed pathway output. In contrast, proteome allocation to transporters scales contrary to the optimal demand with decreasing growth rate in *E. coli*: at increasingly lower growth rates, bacterial cells express more and more transporters for nutrients that are currently not available (18, 22).

Why do cells optimize resource allocation to certain pathways (translation) but not others (transporters)? From a cellular pathway topology perspective, transporters are located at the interface with the environment, while proteins for translation are located at the end of nutrient flow. Bacteria such as *E. coli* are living in constantly changing environments, but have only a very limited ability to sense external nutrient levels. Therefore, transporters should not only transport enough nutrients for cell growth under the current conditions, but also allow the cell to quickly import alternative substrates that become available in upcoming conditions. To maximize fitness across changing environments, it is plausible that bacteria growing on preferred nutrients should invest much of their resources into proteins required for instantaneous growth, while bacteria growing on unpreferred nutrients should allocate more resource to the preparation for future environments. Unlike transporters, translation proteins are located in the interior of cellular processes and rarely have direct connections to the environments. Moreover, in contrast to sensing the large number of potential nutrients and their combinations, sensing an increased or decreased demand of protein production is essentially a one-dimensional problem. Thus, the cell might have evolved a simple and efficient way to regulate the expression of the translation machinery to the minimal required level for instantaneous cell growth, rather than making them dependent on specific nutrients. Indeed, in *E. coli*, the ribosomal genes are mainly regulated by the concentration of a single molecule species, ppGpp (23, 24).

Based on these observations, we speculated that more generally, the proteomic efficiency of pathways might depend on their positions in the metabolic network. We hypothesized that proteome efficiency – defined as the ratio between minimally required and observed protein concentrations – increases along the carbon flow, from transporters at the network periphery to translation at the network core. In *E. coli* growing on minimal media, more than half of the proteome by mass is metabolic enzymes (22). Computational models can predict the optimally efficient proteome allocation to each metabolic pathway (9, 12, 14, 16), and quantitative proteomics data is available for *E. coli* growing on a wide range of minimal media with different carbon sources (22). To test our hypothesis, we exploit these resources to compare experimental data across diverse minimal media conditions (22) to the predicted optimal pathway expression at the observed growth rate. As expected, we find that pathways differ systematically in how much excess protein mass is allocated to them compared to the local optimum, with decreasing excesses over optimal allocation along the carbon flow from nutrient import to protein production.

Results and Discussion

Modeling proteome allocation with linear enzyme kinetics and growth-rate dependent biomass composition

To analyze local pathway efficiency, we first predict the local optima of all metabolic enzymes with an improved version of FBA with molecular crowding (9, 15). We modelled *E. coli* metabolism based on the constraint-based *i*ML1515 model (25). The standard model assumes

a constant composition of biomass across conditions. As the RNA/protein mass ratio (4) and the cell surface/volume ratio (26) can be expressed as functions of growth rate under the investigated conditions, we re-formulated the biomass function of *i*ML1515 with growth rate-dependent contents of RNA, protein, and cell envelope components (murein, lipopolysaccharides, and lipid) (see Methods, **Suppl. Fig. S1**, and **Suppl. Table S1**).

We performed calculations using MOMENT (MetabOlic Modeling with ENzyme kineTics) (9, 27, 28), a version of flux balance analysis (FBA) with molecular crowding (15). Similar to other constraint-based approaches (12, 14), MOMENT estimates the enzyme concentration required to support a given flux v_i as $[E_i]= v_i / k_i$, where k_i is the effective turnover number of the enzyme. This effective turnover number was assumed to be constant across conditions, a zero-order approximation to the true growth rate-dependence (29). Maximal *in vivo* effective enzyme turnover number ($k_{app,max}$) represents turnover in the cellular environment better than *in vitro* estimates of enzyme turnover numbers (k_{cat}) (30, 31). We thus parameterized the reactions of the *i*ML1515 model with the $k_{app,max}$ from Ref. (31) by replacing the original k_{cat} (27) when $k_{app,max}$ was available. For reactions for which neither $k_{app,max}$ nor k_{cat} were available, we used enzyme turnover numbers predicted by machine learning (31) in the simulation (see **Suppl. Table S2** for the enzyme turnover numbers and their sources). Most enzymes in the metabolic model have experimentally estimated parameters ($k_{app,max}$ or k_{cat}); these account for ~70% of total enzyme by mass in the whole metabolic network, and for ~80% when excluding transport reactions (**Suppl. Fig. S2**).

With the growth rate-dependent biomass function and updated enzyme turnover numbers, we identified the minimal total mass concentration of enzymes and transporters (in units of gram per gram of dry weight, g/g_{DW}) that can support the observed growth rate on the given carbon source (see **Methods**; the predicted and measured concentrations of individual proteins are listed in **Suppl. Table S3**). Thus, our predictions do not reflect globally optimal resource allocation, but quantify the minimal proteome allocation into pathways required to sustain the observed growth rate (local optimality). Note that the calculation of required concentrations assumes that all enzymes are fully saturated with their products; this means that our estimates provide a lower bound of proteome allocation into pathways, which is expected to deviate increasingly from the actual demand at lower growth rates (29).

Proteome efficiency increases along nutrient flow in coarse-grained pathways

Following earlier work (16), we first compared the predicted minimal required proteome with experimental data across the whole metabolic network. As *E. coli* uses different central metabolic reactions for growth on glycolytic and gluconeogenic carbon sources and most of the proteome data in Ref. (22) were measured on glycolytic carbon sources, we focus on the proteome efficiency of metabolic pathways on glycolytic carbon sources here; results for gluconeogenic carbon sources are shown in **Suppl. Table S4**. We classified proteins into three groups on the basis of their experimental and predicted expression. An individual protein is labeled as:

- "*shared*" if its presence is predicted under local optimality and is confirmed in the experiment (these proteins were labeled "utilized" in Ref. (16));
- "*measured-only*" if it is found in the experiment but predicted to be absent (these proteins were labeled "un-utilized" in Ref. (16));
- "*predicted-only*" if its presence is predicted but not confirmed in the experiment.

The *predicted-only* proteins account for only a very small fraction of the total predicted proteins (<1%) in all studied pathways, except for nutrient transport and proteins without assigned pathways in this study ("others") (**Suppl. Fig. S3**). We thus do not include the *predicted-only* proteins in the following figures.

Metabolic enzymes account for a decreasing fraction of the proteome with growth rate, with observed proteome fractions ranging from 67% to 53% (**Suppl. Fig. S4**). In agreement with earlier work (16), we found that the total abundance of *shared* proteins – those required for maximally efficient growth – increases with growth rate, but far exceeds the predicted globally optimal abundance especially at lower growth rates (**Suppl. Fig. S4**).

To assess the pathway-specific proteome efficiency, we examined the following four aspects.

(1) For a given pathway, we summed the mass concentrations of all *shared* proteins – those that are predicted to be active and found experimentally – in each growth condition for both the observed proteins and for the locally optimal prediction. We then calculated the Pearson correlation coefficient r between the two combined mass concentrations across conditions (denoted as r_{pathway}). For locally optimal proteome allocation and if the assumption of constant enzyme saturation would hold, this correlation should approach r=1, independent of enzyme kinetic parameter values.

(2) The geometric mean fold-error (GMFE) of predicted vs. observed protein concentrations of the pathway's *shared* proteins (denoted as $\text{GMFE}_{\text{pathway}}$), calculated across proteins and growth conditions. The GMFE shows by which factor the observed concentrations deviate from predicted values on average.

(3) The experimentally observed mass fraction of *measured-only* proteins of the pathway in a given growth condition (denoted as $f_{\text{measured-only}}$). This is the proteome fraction that makes no contribution to growth according to our predictions.

(4) The squared Pearson's correlation coefficient between predicted and measured abundances across individual proteins in a given growth condition (denoted as $r_{\text{individual}}$). While measures (1)-(3) assess optimality at the pathway level, this last measure quantifies the relationships between proteins within the pathway: a correlation coefficient close to 1 indicates that all proteins are equally close to – or equally distant from – the optimal prediction. Note that in contrast to measure (1), the comparison across individual proteins relies strongly on the accuracy of the individual turnover numbers. As the latter are only known approximately, we expect these estimates to be noisy.

Table 1 shows the pathway proteome efficiency measures on glycolytic carbon sources, which are discussed in the following subsections.

Table 1. Proteome efficiency of pathways.

Pathway		Pathway expression (for *shared* proteins) (n = 14)[a]			*measured-only* fraction[b] (median across 14 conditions) ($f_{measured\text{-}only}$)	individual *shared* proteins; median across 14 conditions[c]		
		$r_{pathway}^2$	*p*	GMFE $_{pathway}$		$r_{individual}^2$	*p*	n^d
Measures (1)-(4)		(1)	(1)	(2)	(3)	(4)	(4)	
Biosynthesis		**0.84**	**4.8×10^{-6}**	**1.70**	**0.26**	**0.45**	**4.2×10^{-31}**	**226**
	Amino acid	0.77	3.7×10^{-5}	1.40	0.30	0.45	1.1×10^{-10}	72
	Nucleotide	0.67	3.7×10^{-4}	3.32	0.23	0.15	0.05	28
	Cell envelope	0.43	0.01	1.88	0.14	0.38	2.3×10^{-5}	40
	Cofactor	0.84	4.9×10^{-6}	1.24	0.11	0.59	4.1×10^{-15}	72
	Biosynthesis others	0.60	1.1×10^{-3}	2.91	0.25	0.46	5.4×10^{-5}	29
Central metabolism		**0.024**	**0.60**	**2.32**	**0.31**	**0.15**	**3.3×10^{-3}**	**56**
	Glycolysis	0.63	6.9×10^{-4}	2.21	0.08	0.35	0.05	11
	Pentose phosphate pathway	0.72	1.3×10^{-4}	1.30	0.39	0.32	0.24	6
	TCA cycle	(-) 0.43[e]	0.01	6.40	0.10	0.38	0.03	12
	Glyoxylate shunt	-	-	-	1	-	-	0
	Energy generation	(-) 0.02[e]	0.61	1.63	0.06	0.11	0.08	28
	Central metabolism others	0.44	9.4×10^{-3}	1.56	0.55	0.98	0.10	3
Transporters		**(-) 0.57[e]**	**1.9×10^{-3}**	**3.39**	**0.92**	**0.13**	**0.64**	**4**
***Others* (other metabolic enzymes)**		**0.004**	**0.84**	**1.79**	**0.91**	**0.16**	**0.03**	**30**
Total metabolism		**0.72**	**1.4×10^{-4}**	**1.79**	**0.52**	**0.35**	**1.7×10^{-30}**	**309**
Translation		**0.87**	**1.1×10^{-6}**	**1.35**	**0**	**0.98**	**0.08**	**3**

[a] Values reflect the local optimality of complete pathways across conditions. *n* = 14 indicates the number of glycolytic carbon sources analyzed.

[b] Mass fraction of *measured-only* (un-predicted but observed) proteins relative to all proteins in the pathway.

[c] These columns reflect the local optimality compared across individual proteins within each pathway at a given growth condition; values are medians across the *n* = 14 glycolytic growth conditions.

[d] Number of proteins in each pathway or pathway set.

[e] Negative correlation coefficient $r_{pathway}$.

To test if the proteome efficiency of pathways increases with carbon flow, we first assigned the metabolic proteins in the *i*ML1515 model into four coarse-grained sets (see **Methods**, and **Suppl. Tables S5** and **S6** for pathway membership): (1) transporters, which shuttle metabolites across the outer or inner membrane; (2) central metabolism, which produces precursor

metabolites and energy for all other cellular processes; (3) biosynthesis pathways, which utilize precursors and energy generated by central metabolism to produce building blocks of macromolecules; (4) other enzymes, that is, all enzymes in the *i*ML1515 model not included in (1)-(3) (denoted as "*others*"; these proteins are not assigned to a specific position along the nutrient flow). The *i*ML1515 model does not include a representation of translation processes. To provide a more complete birds-eye view of nutrient flow, we also included in our analyses the proteome efficiency of the translation machinery (predicted and measured expression of ribosome, elongation factor Tu, and elongation factor Ts) from our previous work (19), which was based on the same proteomics data analyzed here (22).

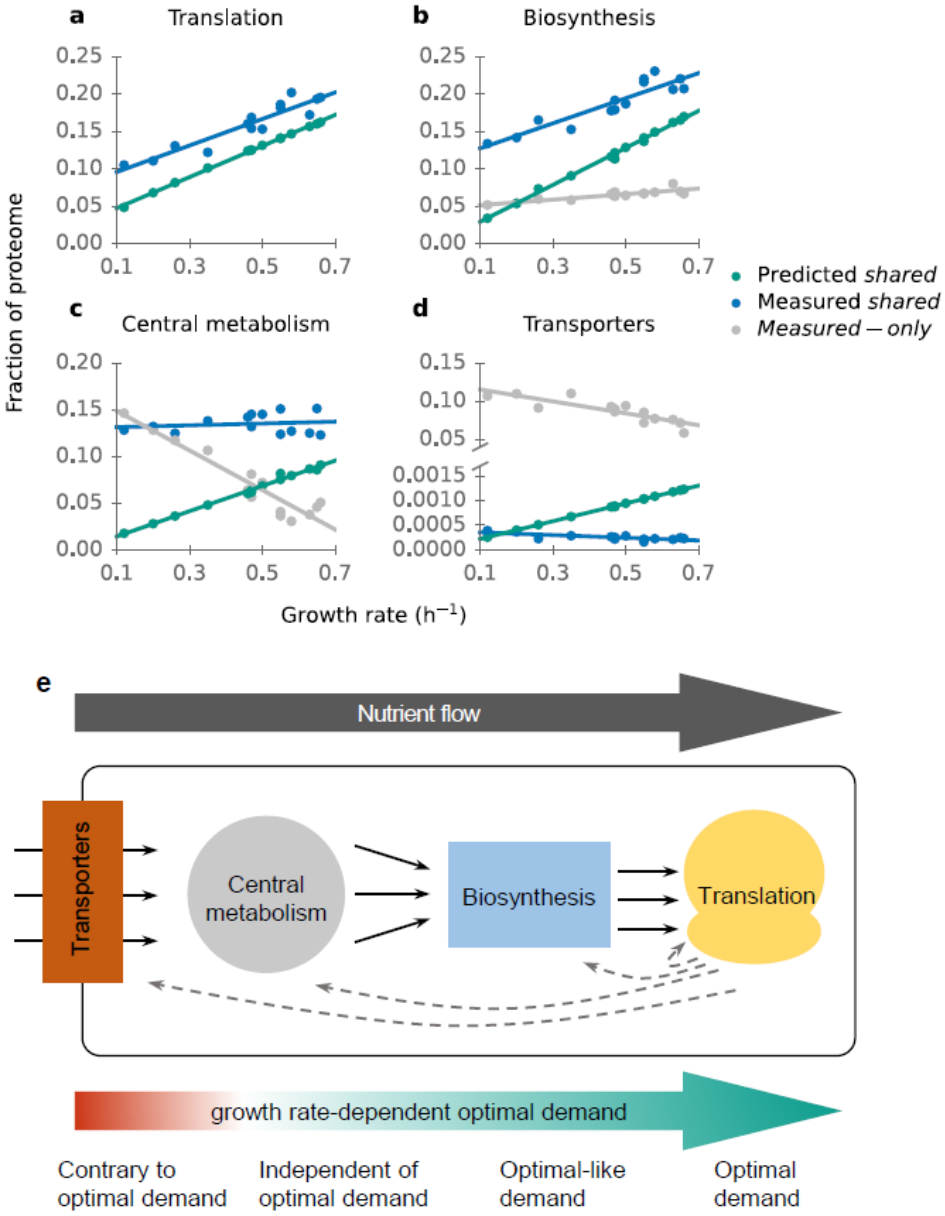

Fig. 1. Growth rate-dependent proteome efficiency increases along the nutrient flow. Predicted and observed proteome allocation to (**a**) translation machinery, (**b**) biosynthesis pathways, (**c**) central metabolism, and (**d**) transporters. (**e**) Schematic diagram of nutrient flow and proteome efficiency.

In these coarse-grained pathways, carbon and other nutrients flow from transporters to central metabolism to biosynthesis pathways to translation. For all four aspects assessed, the proteome efficiency gradually increases along the nutrient flow (**Fig. 1** and **Table 1**): $r_{pathway}$ increases from -0.75 to 0.93, $GMFE_{pathway}$ decreases from 3.39 to 1.35, $f_{measured\text{-}only}$ decreases from 0.92 to 0, and $r_{individual}^2$ increases from 0.13 to 0.98.

Proteome allocation to translation is near the optimal prediction (**Fig. 1a**, **Table 1**), with no expression of unneeded proteins ($f_{measured\text{-}only}$ = 0), a very high correlation between observed and predicted total investment across conditions ($r_{pathway}^2$ = 0.87), a mean deviation between predicted and observed individual protein concentrations of only 35% ($GMFE_{pathway}$ = 1.35), and a strong correlation between observed and predicted individual protein concentrations (median across the 14 glycolytic conditions: $r_{individual}^2$ = 0.98). The remaining discrepancy between measured and predicted data is largely caused by the presence of deactivated ribosomes and elongation factor Tu at the studied growth rates (19), which cannot be predicted by optimization.

Proteome allocation to biosynthesis pathways is quantitatively consistent with the predictions for *shared* proteins, i.e., those whose presence is both predicted and observed (**Fig. 1b**; $r_{pathway}^2$= 0.84; $GMFE_{pathway}$ = 1.70; $r_{individual}^2$ = 0.45). However, about a quarter of the biosynthesis protein mass present in the cell is not predicted ($f_{measured\text{-}only}$ = 0.26).

In central metabolism, the abundance of *shared* proteins is almost constant across growth rates in measured data, whereas it should increase with growth rate according to the predictions (**Fig. 1c**). Remarkably, the abundance of *measured-only* proteins is very high at low growth rates and decreases sharply with growth rate.

In stark contrast to all other pathways, the vast majority of transporters – more than 90% – are *measured-only*, i.e., the experimentally observed proteins are not part of the predicted optimal proteome ($f_{measured\text{-}only}$ = 0.92; **Fig. 1d**; see **Methods** for the treatment of carbon transporters). Moreover, proteome allocation to transporters decreases with growth rate in measured data (both *shared* and *measured-only*), whereas it increases with growth rate in the locally optimal predictions ($r_{pathway}$ = -0.75, p = 1.9×10^{-3}). We note that when the concentration of a substrate is the limiting factor for cell growth, the optimal expression of its transporter increases with decreasing growth rate (10). Here, to compare transporters across growth conditions, we excluded all carbon transporters used in the studied conditions. Since these carbon sources are the only different nutrients across growth conditions (22), the data shown here are the non-growth-limiting transporters and their abundance indeed scales contrary to optimal demand. The true deviation from optimality may be smaller than this estimate due to the existence of many alternative transporters (25) and due to inaccurate turnover number estimates for transporters; only 24 out of 774 transport reactions have experimental measured turnover numbers.

A large mass fraction of the proteins that cannot be assigned to one of the pathways described above (*others*) is also not expected to be present in the cell according to our predictions ($f_{measured\text{-}only}$ = 0.91; **Suppl. Fig. S5a**). About 40% of this unexpected protein mass is related to degradation pathways. At the same time, the abundance of *shared* proteins is similar to the predictions ($GMFE_{pathway}$ = 1.79).

In sum, proteome efficiency increases along the nutrient flow in the four coarse-grained

pathways (**Fig. 1e**). Transporters represent the metabolic interface of the cell to the environment. In the absence of external sensors, the expression of a transporter for a potential nutrient is a necessary condition for its detection by the cell; thus, non-optimal transporter expression serves an important cellular function unrelated to steady-state growth. Central metabolism acts as a hub that connects all other pathways. When nutrients are transported into the cell, they either directly enter central metabolism, or they first need to be degraded by catabolism. For this reason, optimal proteome allocation to central metabolism is strongly environment-dependent. Just as is the case for transporters, keeping a certain fraction of central metabolism enzymes in standby for environmental changes will thus be beneficial in transitions between physiological states. Moreover, the optimal expression of central metabolism proteins would require detailed, environment-dependent regulation, which may be difficult to achieve without substantial cellular investment into sensing and regulation. In contrast, optimal resource allocation into translation and the biosynthesis (anabolic) pathways, which synthesize building blocks for the cell, is largely independent of nutrients across minimal environments, and depends almost exclusively on the growth rate. Their optimal regulation is thus a one-dimensional problem that requires only a sensor for growth rate itself, and can be implemented relatively easily. Consistent with this speculation, biosynthesis and translational genes are regulated by fewer transcriptional factors than transporters and central metabolic genes (**Suppl. Fig. S6**). At the same time, our observations are consistent with a reserve of unused biosynthesis enzymes at low growth rates (**Fig. 1a and 1b**), which can benefit the cell in fluctuating conditions (32, 33).

The most expensive biosynthesis pathways are consistent with optimality

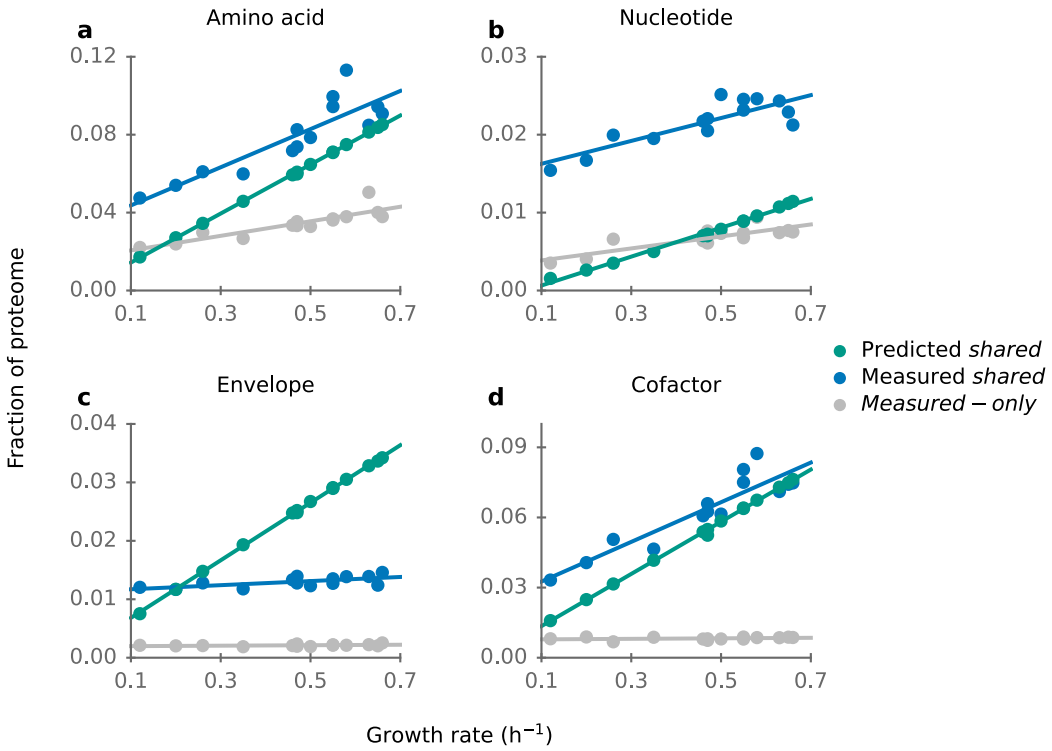

Fig. 2. Experimentally observed and predicted proteome fractions of biosynthesis pathways across glycolytic carbon sources. See **Suppl. Fig. S5b** for biosynthetic proteins not covered here.

To find if proteome efficiency varies in biosynthesis, we further divided biosynthesis pathways into five sets of pathways: amino acid biosynthesis; nucleotide biosynthesis; cofactor biosynthesis; cell envelope component biosynthesis; and all other biosynthesis enzymes. The predicted proteome fractions of these pathways are almost linear functions of the growth rate (**Fig. 2**), as mostly the same reactions are expected to be used for biosynthesis across the studied minimal conditions.

A large fraction of the proteome is allocated to amino acid biosynthesis pathways at high growth rates on minimal media (about 15%, **Fig. 2a**). Similar to the situation for translation, proteome allocation to amino acid biosynthesis pathways is strongly correlated with predictions (**Fig. 2a**; $r_{pathway}^2$ = 0.77; $GMFE_{pathway}$ = 1.40; $r_{individual}^2$ = 0.45; **Table 1**). However, in contrast to translation, a sizeable proteome fraction for amino acid biosynthesis is invested into proteins not predicted to be active ($f_{measured\text{-}only}$ = 0.30).

For nucleotide biosynthesis pathways, predicted and observed abundances of *shared* proteins are also strongly correlated in ($r_{pathway}^2$ = 0.67), but their magnitudes differ by more than 3-fold ($GMFE_{pathway}$ = 3.32; **Fig. 2b**). Moreover, the expression of individual enzymes in this pathway cannot be explained well by the predictions ($r_{individual}^2$ = 0.15).

Cell envelope biosynthesis pathways encompass lipid, peptidoglycan, and lipopolysaccharide (LPS) biosynthesis. While predicted and observed expression of *shared* enzymes in these pathways show a statistically significant correlation ($r_{pathway}^2$=0.43; **Table 1**), the slopes of their growth rate dependences differ markedly. The observed proteome allocation is almost constant across growth conditions; in contrast, the predicted proteome allocation increases proportionally with growth rate (**Fig. 2c**). It is noteworthy that this disagreement does not stem from an incorrect assumption of constant biomass composition across conditions: our model explicitly accounts for the changing biomass fractions of cell envelope components (**Methods**), which are in particular due to changes in cell size. Theoretically, the predicted optimal proteome allocation should provide a lower limit on the required proteome investment; that predictions substantially exceed observed proteome allocation for cell envelope biosynthesis at faster growth suggests that one or more enzymes were assigned turnover numbers that are much lower than the true values.

Similar to amino acid biosynthesis pathways, cofactor biosynthesis pathways are also highly abundant at high growth rates (about 10% of the total proteome, **Fig. 2d**). Proteome allocation to cofactor biosynthesis pathways is highly consistent with the optimal predictions ($r_{pathway}^2$ = 0.84; $GMFE_{pathway}$ = 1.24; $f_{measured\text{-}only}$ = 0.11; $r_{individual}^2$ = 0.59).

In sum, proteome efficiency varies substantially across biosynthesis pathways. While observed proteome investment only increases by roughly two-fold for amino acid, nucleotide, and cofactor biosynthesis and shows almost no increase in envelope and other biosynthesis pathways, predicted investment increases by almost a factor of 5.5 (which is the fold-change of growth rate across the examined conditions). At lower growth rates, we expect decreasing enzyme saturation (29) and thus a progressively stronger underestimation of the required proteome by the model; accordingly, **Figs. 2a** and **2d** appear to be highly consistent with an optimal expression of the *shared* proteins of amino acid and cofactor biosynthesis pathways. On the other hand, proteome allocation to nucleotide, envelope, and other biosynthesis pathways (**Suppl. Fig. S5b**) appears to be sub-optimal.

Central metabolism: precursor metabolite and energy generation pathways appear not to be regulated for optimality

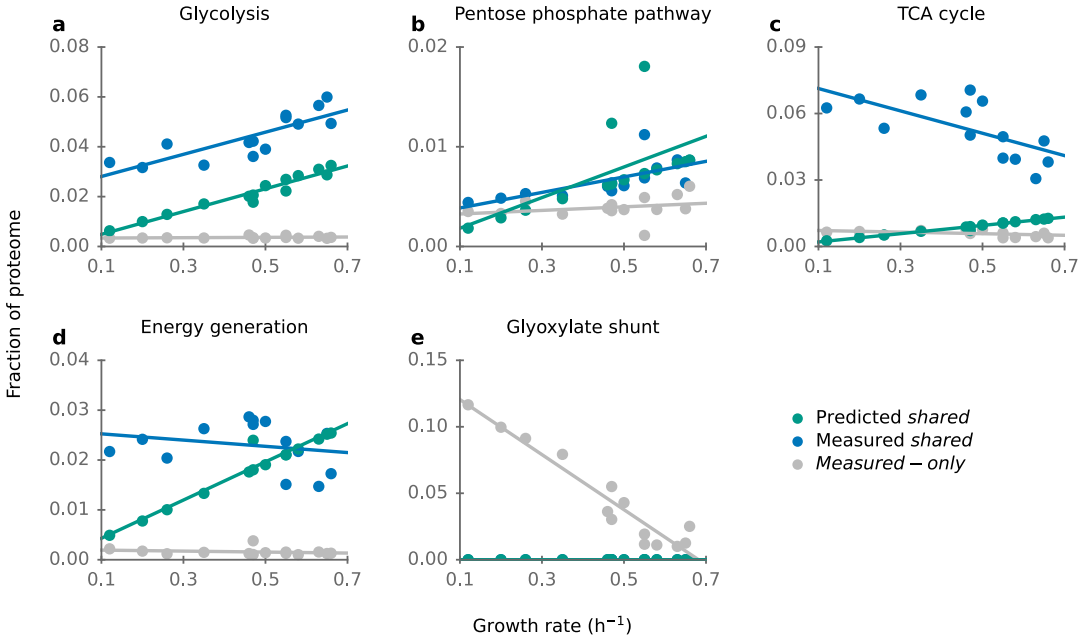

Fig. 3. **Experimentally observed and predicted proteome fractions of central metabolic pathways.** See **Suppl. Fig. S5c** for central metabolic proteins not covered here.

The enzymes of central metabolism show little systematic variation with growth rate, and their abundance is at most weakly correlated with the predicted concentrations ($r_{pathway}^2$ =0.024; $GMFE_{pathway}$=2.32). To examine if individual pathways show a stronger agreement between observations and predictions, we examined six central metabolic pathways: glycolysis; pentose phosphate pathway; TCA cycle; energy generation pathways, comprising the electron transport chain and ATP synthase; glyoxylate shunt; and other central metabolic enzymes.

Proteome allocation to glycolysis increases markedly with growth rate and is strongly correlated with predicted values (**Fig. 3a**; $r_{pathway}^2$ = 0.63; $f_{measured\text{-}only}$ = 0.08). However, protein levels are substantially higher than predicted (GMFE = 2.21). A potential reason for this discrepancy is that most of the reactions in glycolysis are reversible, while the simple approximation for enzyme activity used here (k_{cat}) cannot capture the demand of enzymes close to thermodynamic equilibrium (34). Moreover, many of the enzymes in glycolysis are regulated allosterically (35), and may hence act at lower activities than assumed in the simulations.

The pentose phosphate pathway also shows significant signs of partial optimality: the measured abundance of *shared* proteins is close to and strongly correlated with the predictions (**Fig. 3b**; $r_{pathway}^2$ = 0.72; $GMFE_{pathway}$ = 1.3). However, *measured-only* proteins account for 39% of the pathway proteome.

Enzyme expression in the TCA cycle is decidedly non-optimal. The expression of *shared* enzymes decreases with growth rate, while predictions indicate it should increase (**Fig. 3c**; $r_{pathway}$ = -0.65). In addition, enzyme abundance is massively higher than predicted across all growth rates ($GMFE_{pathway}$ = 6.4). At the same time, *measured-only* proteins account for only

a very small fraction of the pathway ($f_{\text{measured-only}}$ = 0.10), and the abundances of individual proteins are also correlated with measured data ($r_{\text{individual}}^2$ = 0.38, p = 0.03).

The proteome fraction allocated to energy generation pathways – comprising the electron transport chain and ATP synthase – is almost independent of the growth rate, while predictions increase with growth rate (**Fig. 3d**). Similar to the TCA cycle, *measured-only* proteins make up only a small fraction of the pathway (6%). *E. coli* fully oxidizes carbon sources to CO_2 at low growth rates under aerobic conditions (aerobic respiration), while at high growth rates it only partially oxidizes some carbon sources – in particular glucose and fructose – resulting in the excretion of acetate (aerobic fermentation, leading to overflow metabolism). Along with the metabolic switch from aerobic respiration to aerobic fermentation, the TCA cycle is gradually down-regulated (36). In our predictions, aerobic fermentation is more efficient than aerobic respiration for all conditions, so that only aerobic fermentation was active in the predictions. However, even with a model that predicted the switch to fermentation, our conclusions would likely not change; this is because the switch would not affect lower growth rates, and because the predicted demand into the TCA cycle would only change slightly.

We were surprised to find that the proteins of the glyoxylate shunt (comprising AceA, AceB, and GlcB) are highly abundant at low growth rates (~12% of the proteome at μ = 0.12 h^{-1}; **Fig. 3e**), with a proteome fraction almost twice that of its alternative pathway, the TCA cycle (**Fig. 3c**). This high abundance at low growth rates does not appear to be specific to the BW25113 strain, as it is mirrored in the MG1655 strain (**Suppl. Fig. S7a**) (3, 37). Fluxomics data shows that across many conditions with low growth rates, flux into the glyoxylate shunt is roughly equal to the flux into the TCA cycle (38–43) (**Suppl. Fig. S7b**). In contrast, the model predicts the glyoxylate shunt to be inactive except in growth on acetate.

In sum, proteome allocation to the pathways of central metabolism is not well explained by optimal proteome efficiency alone, at least not as far as can be discerned with the type of model employed here. This is particularly true for the metabolic switches from aerobic respiration to aerobic fermentation and from the glyoxylate shunt to the TCA cycle.

Utilization of alternative pathways cannot be explained by optimal proteome efficiency

With increasing growth rate, metabolic fluxes may shift between alternative pathways. For example, energy production from glucose switches from aerobic respiration to aerobic fermentation (overflow metabolism) (36). Consistent with previous studies (38, 40), we found that with increasing growth rate, flux gradually transitioned from the PEP-glyoxylate cycle to the TCA cycle (**Suppl. Fig. S7**).

Neither aerobic respiration nor the glyoxylate shunt are used in the predicted flux distributions. In constraint-based models, overflow metabolism emerges when a previously redundant, additional growth-limiting constraint becomes active (44). While there is evidence that overflow metabolism is rooted in a limit on proteome investment into catabolic enzymes (36, 45), this effect cannot be reproduced in mechanistic models without corresponding empirical adjustments. For example, one way of enforcing aerobic fermentation is to impose a decrease in proteome usage and an increase in energy production with increasing growth rate (36, 46); another is to allocate a constant empirical mass of proteins to energy production (47).

The PEP-glyoxylate cycle, which contains the glyoxylate shunt, represents an alternative route to the TCA cycle (38). Compared to the TCA cycle, the PEP-glyoxylate cycle produces an additional NADH instead of one NADPH (38). Since NADPH is a common cofactor in anabolic pathways in *E. coli*, it was suggested that the cell should choose the pathway which can produce more NADPH (the TCA cycle) at high growth rates (38). However, the interconversion between NADPH and NADH is a very common process in *E. coli* (48), and it is not clear how the small difference in pathway output (1 NADPH vs. 1 NADH) could explain the massive resource allocation (~ 12% of the proteome) into the glyoxylate shunt at low growth rates. Recent studies showed that overexpression of glyoxylate shunt enzymes can reduce the lag time when *E. coli* experiences a transition from a glycolytic carbon source to a gluconeogenic carbon source (49, 50). However, it is still challenging to develop mechanistic models that explain the growth rate-dependent expression of alternative pathways and lag times from first principles.

Conclusions

In this study, we systematically assessed proteome efficiency at the pathway level in *E. coli*. Overall, we found that the proteome efficiency of pathways increases along the nutrient flow, from transporters to central metabolism to biosynthesis pathways to translation. We note that this gradient is analogous to a gradient of genomic stability observed on much longer time scales, with central reactions being more stable over evolutionary time than reactions at the interface to the environment (51), which we found here to also be less efficient. Above, we showed that proteome allocation is near the optimal demand for the most expensive biosynthesis pathways, including translation as well as amino acid and cofactors biosynthesis pathways; the same pathways are located in the interior of the cellular biosynthetic network. In contrast, about half of the metabolic pathways by mass show a growth rate dependence contrary to that expected for optimal demand, including the TCA cycle, glyoxylate shunt, and transporters; typically, these pathways are located at the periphery of the cellular network. We hypothesize that these patterns of local optimality and sub-optimality arise from two tradeoffs and their interactions: on the one hand, the tension between maximal instantaneous growth and the cell's ability to quickly and efficiently transition its physiological state in response to environmental change; and on the other hand, the tension between the benefits of precise and optimal control of cellular resource allocation and the resource investment required for the corresponding control systems. Quantifying these tradeoffs and their joint influence on cellular physiology will require an enhanced, quantitative understanding of the evolutionarily relevant patterns of environmental changes as well as of the costs and effectiveness of regulatory strategies available to bacteria such as *E. coli*.

Methods

Growth rate-dependent biomass composition

The original biomass composition in the *i*ML1515 model is very similar to that of the *i*AF1260 model, formulated for a doubling time of 40 min or μ = 1.04 h^{-1} (52). However, biomass composition varies across growth rates. The two most significant changes are those of the RNA/protein mass ratio and the cell volume, which determines the surface/volume ratio (S/V). Both ratios can be expressed as functions of the growth rate; accordingly, we estimated the growth rate-dependent biomass fraction of RNA, protein, and cell envelope components (including murein, lipopolysaccharides, and lipid) as functions of the growth rate, as described

below.

We first fitted experimental data for the RNA/protein mass ratio ($\frac{m_{RNA}}{m_{protein}}$) (4, 53) and the surface/volume ratio (S/V) (26) to linear functions of the growth rate (**Suppl. Fig. S1**), resulting in the relationships

$$\frac{m_{RNA}}{m_{protein}}(\mu) = 0.223\mu + 0.08 \quad , \tag{1}$$

$$\frac{S}{V}(\mu) = -0.1895\mu + 7.952 \quad . \tag{2}$$

Assuming that the biomass contribution of cell envelope components ($m_{envelope}$) is proportional to the surface/volume ratio gives

$$\frac{m_{envelope}(\mu = \mu_1)}{m_{envelope}(\mu = \mu_2)} = \frac{\frac{S}{V}(\mu = \mu_1)}{\frac{S}{V}(\mu = \mu_2)} \quad . \tag{3}$$

The growth rate-dependent biomass fraction of cell envelope components ($m_{envelope}$) can then be estimated by equation (3) given equation (2) and $m_{envelope}$ at μ = 1.04 h^{-1}. The relative composition of murein, lipopolysaccharides, and lipid was assumed to be constant.

The biomass fractions of cellular components other than RNA, protein, and cell envelope components (m_{others}) were assumed to be independent of the growth rate. The sum of RNA and protein is given by:

$$m_{RNA} + m_{protein} = 1 - m_{others} - m_{envelope} \tag{4}$$

Combining equation (1) and (4), the content of RNA and protein can be calculated for all conditions (**Suppl. Fig. S1**). The relative contributions of individual nucleotides to total RNA and of individual amino acids to total protein were assumed to be growth rate-independent. The resulting growth rate-dependent biomass compositions are listed in **Suppl. Table S1**.

Implementation of MOMENT

To perform flux balance analysis with molecular crowding, we used ccFBA (27), which implements the MOMENT algorithm (9) with an improved treatment of co-functional enzymes (28). For enzymes for which maximal *in vivo* effective enzyme turnover numbers ($k_{app,max}$) were available from Ref. (31), we used these to replace the original *in vitro* k_{cat} values (see **Suppl. Table S2** for the turnover numbers used).

Instead of maximizing the growth rate at a given nutrient condition, we solved the complementary optimization problem that estimates the minimal required proteome (C) able to support the observed growth rate on the given carbon source. However, as the objective function in ccFBA is the growth rate, we used an indirect procedure for the soluiton. In the constraint-based type of model employed here, there is a linear relationship between proteome investment and predicted growth rate, $C = a\mu + b$ for two constants *a,b*. Note that due to a non-zero non-growth-related maintenance energy term included in the model, *b*>0. The constants *a* and *b* can be determined by any two pairs of proteome budget and growth rate.

For each experimental condition with observed growth rate μ' according to Ref. (22), we first estimated the biomass composition at μ'. At this biomass composition, we then predicted the growth rates at C=0.1 g/g_{DW} ($C_{0.1}$) and C=0.2 g/g_{DW} ($C_{0.2}$), denoted as $\mu_{0.1}$ and $\mu_{0.2}$, respectively. a and b were then calculated from $\mu_{0.1}$, $\mu_{0.2}$, $C_{0.1}$, and $C_{0.2}$. The total minimal required proteome (C') at the observed growth rate was then read out as $C' = a\mu + b$.

For a given protein i, its minimal demand at the observed growth rate μ' ($p_{i,\mu'}$) in units of g/g_{DW} can be expressed as

$$p_{i,\mu'} = \frac{C'}{C_{0.1}} p_{i,\mu_{0.1}} \quad (5)$$

with $p_{i,\mu_{0.1}}$ the minimal demand for protein i at $C_{0.1}$.

With the protein content in dry mass at μ' ($m_{protein,\mu'}$) estimated in equation 4, the proteome fraction of protein i at μ' ($m_{i,\mu'}$) can be written as

$$m_{i,\mu'} = \frac{p_{i,\mu'}}{m_{protein,\mu'}}. \quad (6)$$

Pathway membership

Proteins were characterized as transporters if the corresponding genes are assigned to transport processes according to the *i*ML1515 annotation (25). The carbon source is the only nutrient that differs between the minimal media used in the proteomics experiments (22). To make the transporters comparable across conditions, we thus excluded inner and outer membrane transporters for all carbon sources used in the studied conditions (22) and analyzed only the transporters for other metabolites.

We used the pathway ontology in EcoCyc (54) (downloaded on 13. January 2021) to assign the enzyme members for other metabolic pathways.

Proteins are labeled as biosynthetic enzymes based on the EcoCyc pathway ontology annotation "biosynthesis" (54). Pathways in this category are: (1) Amino acid biosynthesis ("Amino Acid Biosynthesis" in EcoCyc), (2) nucleotide biosynthesis ("Nucleoside and Nucleotide Biosynthesis"), (3) cofactors ("Cofactor, Carrier, and Vitamin Biosynthesis"), and (4) cell envelope components ("Cell Structure Biosynthesis and Fatty Acid and Lipid Biosynthesis"), including lipid, peptidoglycan, and LPS. All other biosynthetic enzymes are merged into (5) other biosynthetic pathways. See **Suppl. Table S5** for the corresponding hierarchy levels in the EcoCyc pathway ontology.

Enzymes are designated as being involved in precursors and energy generation according to the EcoCyc pathway ontology annotation "Generation of Precursor Metabolites and Energy". Pathways in this category are: (1) glycolysis, (2) Pentose Phosphate Pathways, (3) TCA cycle, (4) glyoxylate bypass (EcoCyc does not list a pathway for the glyoxylate shunt; the three genes classified as glyoxylate shunt are aceA, aceB, and glcB), (5) energy production ("Electron Transfer Chains and ATP biosynthesis"), and (6) other enzymes.

Treatment of enzymes involved in the nucleotide salvage pathway: In the range of studied growth rates, the transcription of mRNA accounts for more than half of the total RNA transcription (1). The half-life of mRNA is very short (~ 5.5 min) (55) compared to the doubling time, and degraded mRNA will be reused through the nucleotide salvage pathway. However,

our model only predicts the expression of *de novo* biosynthesis pathways. To make the prediction comparable with the observed data, the nucleotide salvage pathway was thus excluded from "nucleotide biosynthesis pathway".

Transcriptional regulation data

Experimental datasets of RegulonDB v10.9 (56) were used for counting the number of transcription factors regulating each protein.

Acknowledgments

This work was supported by the Volkswagen Foundation under the "Life?" initiative, and by the German Research Foundation (DFG) through grant CRC 1310. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

We thank Hugo Dourado, Deniz Sezer, and Peter Schubert for helpful discussions.

XPH conceived and designed the study and performed analysis. SS performed the analysis of transcriptional factor regulation. MJL supervised the study. XPH and MJL interpreted the results and wrote the manuscript.

The authors declare that no competing interests exist.

References

1. Bremer H, Dennis PP. 2008. Modulation of chemical composition and other parameters of the cell at different exponential growth rates. EcoSal Plus 3:1–49.

2. Basan M. 2018. Resource allocation and metabolism: the search for governing principles. Curr Opin Microbiol 45:77–83.

3. Peebo K, Valgepea K, Maser A, Nahku R, Adamberg K, Vilu R. 2015. Proteome reallocation in *Escherichia coli* with increasing specific growth rate. Mol BioSyst 11:1184–1193.

4. Scott M, Gunderson CW, Mateescu EM, Zhang Z, Hwa T. 2010. Interdependence of cell growth and gene expression: origins and consequences. Science 330:1099–1102.

5. You C, Okano H, Hui S, Zhang Z, Kim M, Gunderson CW, Wang YP, Lenz P,

568 Yan D, Hwa T. 2013. Coordination of bacterial proteome with metabolism by
569 cyclic AMP signalling. Nature 500:301–306.

570 6. Klumpp S, Scott M, Pedersen S, Hwa T. 2013. Molecular crowding limits
571 translation and cell growth. Proc Natl Acad Sci U S A 110:16754–16759.

572 7. Hui S, Silverman JM, Chen SS, Erickson DW, Basan M, Wang J, Hwa T,
573 Williamson JR. 2015. Quantitative proteomic analysis reveals a simple strategy of
574 global resource allocation in bacteria. Mol Syst Biol 11:784.

575 8. Bruggeman FJ, Planqué R, Molenaar D, Teusink B. 2020. Searching for principles
576 of microbial physiology. FEMS Microbiol Rev 44:821–844.

577 9. Adadi R, Volkmer B, Milo R, Heinemann M, Shlomi T. 2012. Prediction of
578 microbial growth rate versus biomass yield by a metabolic network with kinetic
579 parameters. PLoS Comput Biol 8:e1002575.

580 10. Molenaar D, van Berlo R, de Ridder D, Teusink B. 2009. Shifts in growth
581 strategies reflect tradeoffs in cellular economics. Mol Syst Biol 5:323.

582 11. Dourado H, Lercher MJ. 2020. An analytical theory of balanced cellular growth.
583 Nat Commun 11:1226.

584 12. Goelzer A, Muntel J, Chubukov V, Jules M, Prestel E, Nölker R, Mariadassou M,
585 Aymerich S, Hecker M, Noirot P, Becher D, Fromion V. 2015. Quantitative
586 prediction of genome-wide resource allocation in bacteria. Metab Eng 32:232–
587 243.

588 13. Goelzer A, Fromion V, Scorletti G. 2011. Cell design in bacteria as a convex
589 optimization problem. Automatica 47:1210–1218.

590 14. O'Brien EJ, Lerman J a, Chang RL, Hyduke DR, Palsson BØ. 2013. Genome-
591 scale models of metabolism and gene expression extend and refine growth
592 phenotype prediction. Mol Syst Biol 9:693.

593 15. Beg QK, Vazquez A, Ernst J, de Menezes MA, Bar-Joseph Z, Barabási A-LA-L,
594 Oltvai ZN. 2007. Intracellular crowding defines the mode and sequence of
595 substrate uptake by *Escherichia coli* and constrains its metabolic activity. Proc
596 Natl Acad Sci U S A 104:12663–12668.

597 16. O'Brien EJ, Utrilla J, Palsson BO. 2016. Quantification and classification of *E.*
598 *coli* proteome utilization and unused protein costs across environments. PLoS
599 Comput Biol 12:e1004998.

600 17. Ibarra RU, Edwards JS, Palsson BO. 2002. *Escherichia coli* K-12 undergoes
601 adaptive evolution to achieve in silico predicted optimal growth. Nature 420:186–
602 189.

603 18. Belliveau NM, Chure G, Hueschen CL, Garcia HG, Kondev J, Fisher DS, Theriot
604 JA, Phillips R. 2021. Fundamental limits on the rate of bacterial growth and their
605 influence on proteomic composition. Cell Syst 12:924-944.e2.

606 19. Hu X-P, Dourado H, Schubert P, Lercher MJ. 2020. The protein translation
607 machinery is expressed for maximal efficiency in *Escherichia coli*. Nat Commun
608 11:5260.

609 20. Hu X-P, Lercher MJ. 2021. An optimal growth law for RNA composition and its
610 partial implementation through ribosomal and tRNA gene locations in bacterial
611 genomes. PLOS Genet 17:e1009939.

21. Lalanne J-B, Li G-W. 2021. First-principles model of optimal translation factors stoichiometry. eLife 10:e69222.

22. Schmidt A, Kochanowski K, Vedelaar S, Ahrne E, Volkmer B, Callipo L, Knoops K, Bauer M, Aebersold R, Heinemann M. 2016. The quantitative and condition-dependent *Escherichia coli* proteome. Nat Biotechnol 34:104–110.

23. Potrykus K, Murphy H, Philippe N, Cashel M. 2011. ppGpp is the major source of growth rate control in *E. coli*. Environ Microbiol 13:563–575.

24. Wu C, Balakrishnan R, Braniff N, Mori M, Manzanarez G, Zhang Z, Hwa T. 2022. Cellular perception of growth rate and the mechanistic origin of bacterial growth law. Proc Natl Acad Sci 119:e2201585119.

25. Monk JM, Lloyd CJ, Brunk E, Mih N, Sastry A, King Z, Takeuchi R, Nomura W, Zhang Z, Mori H, Feist AM, Palsson BO. 2017. iML1515, a knowledgebase that computes *Escherichia coli* traits. Nat Biotechnol 35:904–908.

26. Si F, Li D, Cox SE, Sauls JT, Azizi O, Sou C, Schwartz AB, Erickstad MJ, Jun Y, Li X, Jun S. 2017. Invariance of Initiation Mass and Predictability of Cell Size in *Escherichia coli*. Curr Biol 27:1278–1287.

27. Desouki A. 2016. Algorithms for improving the predictive power of flux balance analysis. Doctoral dissertation. Heinrich Heine University Düsseldorf, Düsseldorf.

28. Heckmann D, Lloyd CJ, Mih N, Ha Y, Zielinski DC, Haiman ZB, Desouki AA, Lercher MJ, Palsson BO. 2018. Machine learning applied to enzyme turnover numbers reveals protein structural correlates and improves metabolic models. Nat Commun 9:5252.

634 29. Dourado H, Mori M, Hwa T, Lercher MJ. 2021. On the optimality of the enzyme–
635 substrate relationship in bacteria. PLOS Biol 19:e3001416.

636 30. Davidi D, Noor E, Liebermeister W, Bar-Even A, Flamholz A, Tummler K,
637 Barenholz U, Goldenfeld M, Shlomi T, Milo R. 2016. Global characterization of
638 in vivo enzyme catalytic rates and their correspondence to in vitro *k*cat
639 measurements. Proc Natl Acad Sci U S A 113:3401–3406.

640 31. Heckmann D, Campeau A, Lloyd CJ, Phaneuf PV, Hefner Y, Carrillo-Terrazas M,
641 Feist AM, Gonzalez DJ, Palsson BO. 2020. Kinetic profiling of metabolic
642 specialists demonstrates stability and consistency of in vivo enzyme turnover
643 numbers. Proc Natl Acad Sci 117:23182–23190.

644 32. Mori M, Schink S, Erickson DW, Gerland U, Hwa T. 2017. Quantifying the
645 benefit of a proteome reserve in fluctuating environments. Nat Commun 8:1225.

646 33. Korem Kohanim Y, Levi D, Jona G, Towbin BD, Bren A, Alon U. 2018. A
647 Bacterial Growth Law out of Steady State. Cell Rep 23:2891–2900.

648 34. Bar-Even A, Flamholz A, Noor E, Milo R. 2012. Rethinking glycolysis: On the
649 biochemical logic of metabolic pathways. Nat Chem Biol 8:509–517.

650 35. Diether M, Nikolaev Y, Allain FH, Sauer U. 2019. Systematic mapping of protein-
651 metabolite interactions in central metabolism of *Escherichia coli*. Mol Syst Biol
652 15:e9008.

653 36. Basan M, Hui S, Okano H, Zhang Z, Shen Y, Williamson JR, Hwa T. 2015.
654 Overflow metabolism in *Escherichia coli* results from efficient proteome
655 allocation. Nature 528:99–104.

656 37. Valgepea K, Adamberg K, Seiman A, Vilu R. 2013. *Escherichia coli* achieves
657 faster growth by increasing catalytic and translation rates of proteins. Mol Biosyst
658 9:2344-2358.

659 38. Fischer E, Sauer U. 2003. A novel metabolic cycle catalyzes glucose oxidation
660 and anaplerosis in hungry *Escherichia coli*. J Biol Chem 278:46446–46451.

661 39. Nanchen A, Schicker A, Sauer U. 2006. Nonlinear dependency of intracellular
662 fluxes on growth rate in miniaturized continuous cultures of *Escherichia coli*.
663 Appl Environ Microbiol 72:1164–1172.

664 40. Gerosa L, Haverkorn Van Rijsewijk BRB, Christodoulou D, Kochanowski K,
665 Schmidt TSB, Noor E, Sauer U. 2015. Pseudo-transition analysis identifies the
666 key regulators of dynamic metabolic adaptations from steady-state data. Cell Syst
667 1:270–282.

668 41. Haverkorn van Rijsewijk BRB, Nanchen A, Nallet S, Kleijn RJ, Sauer U. 2011.
669 Large-scale 13C-flux analysis reveals distinct transcriptional control of
670 respiratory and fermentative metabolism in *Escherichia coli*. Mol Syst Biol 7:477.

671 42. Ishii N, Nakahigashi K, Baba T, Robert M, Soga T, Kanai A, Hirasawa T, Naba
672 M, Hirai K, Hoque A, Ho PY, Kakazu Y, Sugawara K, Igarashi S, Harada S,
673 Masuda T, Sugiyama N, Togashi T, Hasegawa M, Takai Y, Yugi K, Arakawa K,
674 Iwata N, Toya Y, Nakayama Y, Nishioka T, Shimizu K, Mori H, Tomita M. 2007.
675 Multiple high-throughput analyses monitor the response of *E. coli* to perturbations.
676 Science 316:593–597.

677 43. Rui B, Shen T, Zhou H, Liu J, Chen J, Pan X, Liu H, Wu J, Zheng H, Shi Y. 2010.
678 A systematic investigation of *Escherichia coli* central carbon metabolism in

679 response to superoxide stress. BMC Syst Biol 4:122.

680 44. de Groot DH, Lischke J, Muolo R, Planqué R, Bruggeman FJ, Teusink B. 2020.
681 The common message of constraint-based optimization approaches: overflow
682 metabolism is caused by two growth-limiting constraints. Cell Mol Life Sci
683 77:441–453.

684 45. Schuster S, Boley D, Möller P, Stark H, Kaleta C. 2015. Mathematical models for
685 explaining the Warburg effect: a review focussed on ATP and biomass production.
686 Biochem Soc Trans 43:1187–1194.

687 46. Zeng H, Yang A. 2019. Modelling overflow metabolism in *Escherichia coli* with
688 flux balance analysis incorporating differential proteomic efficiencies of energy
689 pathways. BMC Syst Biol 13:3.

690 47. Chen Y, Nielsen J. 2019. Energy metabolism controls phenotypes by protein
691 efficiency and allocation. Proc Natl Acad Sci U S A 116:17592–17597.

692 48. Sauer U, Canonaco F, Heri S, Perrenoud A, Fischer E. 2004. The soluble and
693 membrane-bound transhydrogenases UdhA and PntAB have divergent functions
694 in NADPH metabolism of *Escherichia coli*. J Biol Chem 279:6613–6619.

695 49. Basan M, Honda T, Christodoulou D, Hörl M, Chang Y-F, Leoncini E, Mukherjee
696 A, Okano H, Taylor BR, Silverman JM, Sanchez C, Williamson JR, Paulsson J,
697 Hwa T, Sauer U. 2020. A universal trade-off between growth and lag in
698 fluctuating environments. 7821. Nature 584:470–474.

699 50. Balakrishnan R, de Silva RT, Hwa T, Cremer J. 2021. Suboptimal resource
700 allocation in changing environments constrains response and growth in bacteria.

Mol Syst Biol 17:e10597.

51. Pál C, Papp B, Lercher MJ. 2005. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. Nat Genet 37:1372–1375.

52. Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V, Palsson BØ. 2007. A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. Mol Syst Biol 3:121.

53. Dai X, Zhu M, Warren M, Balakrishnan R, Patsalo V, Okano H, Williamson JR, Fredrick K, Wang Y-P, Hwa T. 2016. Reduction of translating ribosomes enables *Escherichia coli* to maintain elongation rates during slow growth. Nat Microbiol 2:16231.

54. Keseler IM, Mackie A, Santos-Zavaleta A, Billington R, Bonavides-Martínez C, Caspi R, Fulcher C, Gama-Castro S, Kothari A, Krummenacker M, Latendresse M, Muñiz-Rascado L, Ong Q, Paley S, Peralta-Gil M, Subhraveti P, Velázquez-Ramírez DA, Weaver D, Collado-Vides J, Paulsen I, Karp PD. 2017. The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. Nucleic Acids Res 45:D543–D550.

55. Bernstein J a, Khodursky AB, Lin P-H, Lin-Chao S, Cohen SN. 2002. Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. Proc Natl Acad Sci U S A 99:9697–9702.

56. Santos-Zavaleta A, Salgado H, Gama-Castro S, Sánchez-Pérez M, Gómez-Romero L, Ledezma-Tejeida D, García-Sotelo JS, Alquicira-Hernández K,

724 Muñiz-Rascado LJ, Peña-Loredo P, Ishida-Gutiérrez C, Velázquez-Ramírez DA,
725 Del Moral-Chávez V, Bonavides-Martínez C, Méndez-Cruz C-F, Galagan J,
726 Collado-Vides J. 2019. RegulonDB v 10.5: tackling challenges to unify classic
727 and high throughput knowledge of gene regulation in *E. coli* K-12. Nucleic Acids
728 Res 47:D212–D220.

729