nature reviews microbiology

https://doi.org/10.1038/s41579-022-00818-6

Review article

Check for updates

# Shaping bacterial gene expression by physiological and proteome allocation constraints

Matthew Scott [1] & Terence Hwa [2]

## Abstract

Networks of molecular regulators are often the primary objects of focus in the study of gene regulation, with the machinery of protein synthesis tacitly relegated to the background. Shifting focus to the constraints imposed by the allocation of protein synthesis flux reveals surprising ways in which the actions of molecular regulators are shaped by physiological demands. Using carbon catabolite repression as a case study, we describe how physiological constraints are sensed through metabolic fluxes and how flux-controlled regulation gives rise to simple empirical relations between protein levels and the rate of cell growth.

## Sections



[1]Department of Applied Mathematics, University of Waterloo, Waterloo, ON, Canada. [2]Department of Physics, University of California at San Diego, La Jolla, CA, USA. e-mail: mscott@uwaterloo.ca; thwa@ucsd.edu

## Introduction

The mechanics of protein expression in bacteria follow directly from the central dogma of molecular biology: a gene designated for expression is transcribed into mRNAs by RNA polymerases and subsequently translated into proteins by ribosomes. Regulatory mechanisms modulating each step of this process, including transcriptional and post-transcriptional control, have been studied in detail for numerous systems[1–5]. Yet, the central dogma is only one part of what is necessary to determine the concentrations of proteins in living cells (Fig. 1a).

Protein concentrations are affected not only by their synthesis, degradation, and dilution but also by the cytoplasmic volume. The bacterial cytosol is densely packed with proteins that catalyse biochemical reactions, anchor structural components and regulate gene expression as well as with macromolecular machinery responsible for transcription and translation[6]. Keeping protein concentrations high is obviously an advantage as it increases all the metabolic fluxes for the same number of proteins per cell; however, an overly crowded cytoplasm would slow down molecular diffusion and eventually limit key cellular processes[7–9]. Empirically, the biomass density is found to be approximately constant for several microorganisms grown under various conditions[6,10–13]. This constraint is not related to cell size and may arise from the homeostatic control of cytoplasmic water volume through intracellular osmolarity as the biomass density changes primarily upon variations in external osmolarity[13,14].

Because protein is the major component of biomass and the lengths of typical proteins are similar, the invariance of biomass density implies an approximate constancy of the total cellular protein density (also referred to as concentration)[15]. For example, in *Escherichia coli*, the protein concentration seems to vary by less than 10%[13,16]. This global constraint has a profound impact on the canonical notion of protein expression via the central dogma: suppose the transcriptional activities of all genes are doubled and there is no post-transcriptional regulation, then one would naively expect all protein concentrations to be doubled. However, this is not possible given the protein density constraint. Instead, if the concentrations of certain proteins are upregulated, then the concentration of other proteins needs to be reduced to keep the protein density approximately constant. This can be achieved by downregulating the synthesis of designated proteins; if not, the cytoplasmic water volume can then increase to implement the total density constraint, which results in an effective reduction in the concentration of all other proteins.

For growing cells, another important constraint is that the macromolecular machinery needed for RNA and protein synthesis is required at varying concentrations under different growth conditions[17–19]. In good nutrient conditions supporting fast growth, cells must maintain a high concentration of ribosomes and related translational machinery to satisfy the demand for a high flux of protein synthesis (because the speed of peptide elongation does not vary strongly across growth conditions[20]). This, coupled with the protein density constraint, immediately imposes the requirement that the concentration of non-ribosomal proteins must be reduced in fast growth. As a result, a growing bacterium is faced with a fundamental engineering problem of balancing the need for protein synthesis machinery to fuel biomass growth and the need to maintain sufficient concentrations of enzymes to generate fluxes of amino acids, nucleotides and energy to fuel protein synthesis and cell growth.

The focus of this Review is on the model bacterium *E. coli*, for which these constraints have been quantitatively characterized. Consider, for example, adjustments in the macromolecular composition of the cell when the growth rate is modulated by the carbon source (Fig. 1). With good carbon sources such as glucose (Fig. 1b), *E. coli* doubles quickly, which requires high metabolic fluxes mediated by high concentrations of ribosomes and enzymes for biosynthesis. Maintaining high concentrations requires high synthesis rates of these proteins. In poor nutrient conditions (Fig. 1c), the metabolic fluxes are reduced, with a consequent reduction in the concentrations of biosynthetic enzymes and ribosomes, and increased concentrations of catabolic proteins, including transporters.

The approximate constancy of protein density and the required allocation of macromolecular machinery for growing cells are two important examples of 'physiological constraints' that sculpt the outcome of direct regulatory interactions in the cell. In this Review, we describe several features of protein expression that seem surprising when viewed from the perspective of the central dogma modulated by regulators but become quite natural when viewed as consequences of these physiological constraints. We first discuss how physiological constraints are manifest in *E. coli*. We next review how physiological constraints impinge upon one of the fundamental metabolic regulatory networks: carbon catabolite repression (CCR). Finally, we describe how the quantitative framework developed in the context of steady-state CCR has been modified to successfully predict growth rate and proteome re-modelling during nutrient shifts. Throughout, we will focus on inviolable physiological constraints operating independently of possible notions of optimality. In fact, as we detail below, the regulation that *E. coli* uses to satisfy these operating constraints can be suboptimal in terms of, for example, growth-rate maximization. We will not discuss constraints operating on DNA synthesis and cell size; these have recently been reviewed in ref.[21].

## Physiological constraints

The constraints that exponential growth imposes on the protein synthetic machinery are revealed by tracking the cellular ribosome concentration[22]. The total cellular RNA concentration is proportional to the number of ribosomes across growth conditions[17], and therefore the total RNA per-total protein-mass abundance is a convenient proxy for ribosome concentration[22]. For an exponentially growing culture, the RNA to protein ratio exhibits a simple linear correlation with the growth rate, one of the 'bacterial growth laws' known since the 1960s[23] (Fig. 2a). This linear relation, also referred to as the 'R-line'[24], was understood early on[23,25]: to grow twice as fast requires the cell to synthesize proteins twice as quickly. Under moderate-to-fast growth conditions, when empirically the peptide elongation rate by the ribosome is nearly constant[20], the only way to double the protein synthesis rate is by doubling the ribosome concentration[23]. Along with the R-line, the constancy of protein density imposes a complementary constraint: if the cell allocates more of its protein synthesis flux towards ribosomal proteins, then the synthesis (and hence concentrations) of some other proteins must decrease. Therein lies the fundamental constraint on protein synthesis. An illustrative case is provided by the growth dependence of the concentration of an unregulated protein (which provides the starting point for analysing the growth dependence of more complex regulatory motifs[26]). In Fig. 2b, the activity-per-total protein-mass of a reporter enzyme (β-galactosidase) is used as a proxy for its abundance in terms of reporter concentration. When growth rate is modulated by changes in nutrient quality, the concentration of an unregulated protein exhibits the opposite growth-rate dependence from the ribosome concentration, decreasing as the growth rate increases[22,27] (Fig. 2b).

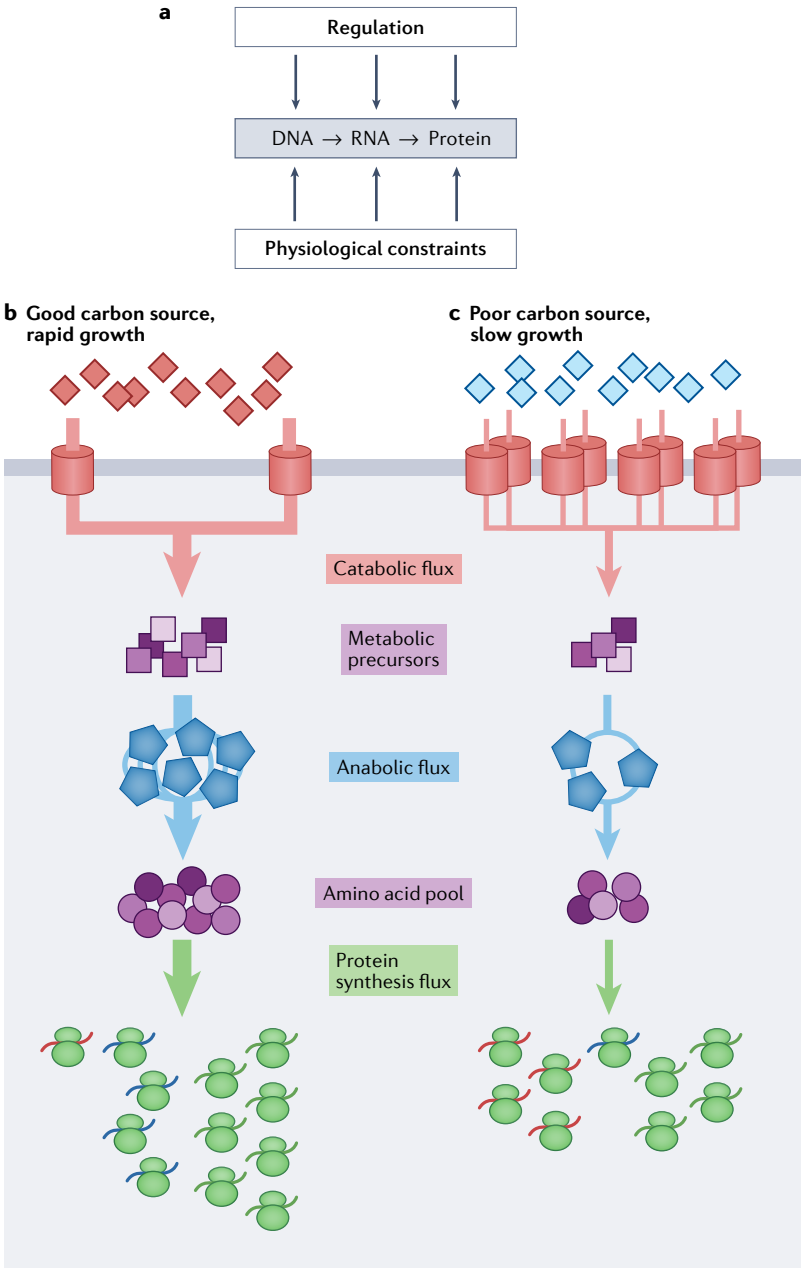



**Fig. 1 | Physiological constraints on gene expression. a**, Protein expression in bacteria follow directly from the central dogma of molecular biology: a gene designated for expression is transcribed into mRNAs by RNA polymerases and subsequently translated into proteins by ribosomes. Protein expression is constrained by direct regulation and physiological constraints arising from cell growth. **b**,**c**, The external carbon source (diamonds) is imported and processed to metabolic precursors (purple squares) via catabolic proteins (here represented by the red transporters), then converted into amino acids (purple circles) via anabolic enzymes (blue pentagons), and further assimilated into peptides via ribosomes (green ovals). The curly lines in the lower part of each panel denote the mRNAs, with colour corresponding to each protein type. In a good carbon source (part **b**), rapid exponential growth necessitates high metabolic fluxes (thick arrows). This requires a large concentration of ribosomes and anabolic enzymes; the concentration of catabolic proteins must be small to satisfy the constraint on total protein density. To maintain this proteome composition, more ribosomes must be allocated to synthesize ribosomes and anabolic enzymes than catabolic proteins, represented by a larger fraction of ribosomes translating the ribosomal and anabolic enzyme mRNAs. In a poor carbon source (part **c**), the growth rate, and hence the metabolic fluxes, are greatly reduced (thin arrows). The demand for anabolic enzymes and ribosomes is reduced, and their concentrations are therefore reduced whereas the concentration of catabolic proteins is increased to increase the carbon influx. This altered proteome composition is maintained by a larger fraction of ribosomes translating the catabolic enzyme mRNAs. The constraint on protein density is illustrated by having the same number of proteins (red, blue and green) in parts **b** and **c**.

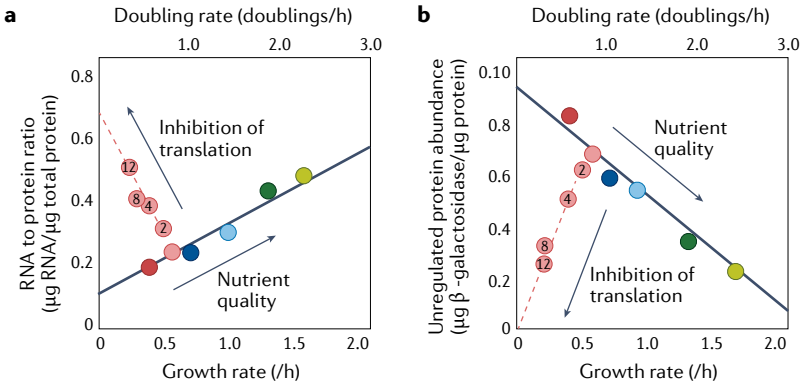



**Fig. 2 | Growth dependence of ribosomal and non-ribosomal proteins.** *Escherichia coli* cells are cultured to grow exponentially at different growth rates by using different nutrients (coloured symbols along the solid black lines) or by applying various sublethal concentrations of translation-inhibiting antibiotics for a given nutrient (dashed lines, micromoles of chloramphenicol shown inside the circles). **a**, The RNA to total protein ratio, which is proportional to the ribosome concentration, exhibits approximate linear dependence on the growth rate: the dependence is positive when changing nutrient quality (solid line) and negative when changing antibiotic concentration (dashed line). **b**, The concentration of an unregulated protein ($pTet_{01}$ driving lacZ (β-galactosidase) expression, using the activity to total protein ratio as a proxy) also exhibits approximate linear dependence on the growth rate but in the opposite directions from that of the RNA to protein ratio. Comparing both parts, the strong anti-correlation between the concentration of ribosomes and the concentration of a constitutive protein suggests a linear constraint (like a see-saw) operating between them. Parts **a** and **b** adapted with permission from ref.[22], AAAS.

Experimentally, there are various methods to investigate the coupling between growth rate and ribosome concentration. Early studies used changes in nutrient quality to modulate growth[23,28]; other modes of growth-rate modulation include physical methods (for example, changes in temperature[28,29] or osmolarity[9]), biochemical methods (for example, using subinhibitory levels of antibiotics[22]) and genetic methods (for example, auxotrophy[28] or expression of toxic[30] or unnecessary proteins[22]). Sub-lethal levels of translation-inhibiting antibiotics are a particularly useful method of growth modulation. In contrast to nutrient-limited growth with unperturbed translational elongation rate, ribosome-targeting antibiotics probe an 'orthogonal' scenario: translation-limited growth with fixed nutrient quality. Under conditions of translational inhibition, the ribosome concentration exhibits a strong (near-linear) negative correlation with growth rate. As with nutrient-modulated growth, the concentration of unregulated proteins exhibits the opposite growth-rate dependence from that of the ribosomes (Fig. 2). Translation-inhibited growth highlights the fundamental constraint on protein synthesis: whatever growth dependence is observed in the ribosomal concentration, the opposite growth dependence is observed in the concentration of an unregulated protein. The near-perfect anti-correlation between the ribosomal protein concentration and the concentration of a constitutive protein suggests a linear constraint operating between them. The bulk of non-ribosomal proteins are devoted to metabolic enzymes. Below, we consider CCR to illustrate the substantial effect that the protein-synthesis constraint can have on shaping the regulation of metabolic gene expression.

## Effects on CCR

Metabolic proteins are responsible for assimilating environmental nutrients to fuel biomass growth. Core metabolic tasks, including catabolism and anabolism, must be coordinated to ensure balanced growth[31]. CCR refers to the reduction in the expression of carbon catabolic proteins during steady-state growth on a good carbon source compared to a poor one[32]. The phenomenon of choosing a distinct carbon substrate for consumption in media with multiple carbon substrates, sometimes referred to interchangeably with catabolite repression[33,34], will be discussed further below in the context of hierarchical carbon utilization.

Early studies of CCR focused on the chemical origin of the substrate for the apparent suppression of specific catabolic enzymes in the presence of glucose[35] (Box 1). Later, CCR came to be understood as a manifestation of a more general strategy whereby the accumulation of a group of intermediate metabolites produced by carbon catabolism ('catabolites') signal a flux mismatch between carbon catabolism and consumption of these catabolites by anabolism[32]. Beyond catabolic proteins that are directly induced by their cognate substrate, in *E. coli*, a large fraction of the proteome is co-regulated to respond to changes in carbon availability even though many of these proteins carry no flux under these conditions[19]. Cyclic adenosine monophosphate (cAMP) antagonizes CCR[36,37]; with the subsequent discovery of the cAMP receptor protein (Crp) and its role in activating the expression of induced carbon catabolic gene expression[38], the direct regulatory mechanism responsible for catabolite repression in *E. coli* was resolved.

Yet, the physiological trigger of cAMP synthesis was never settled conclusively, and seemingly anomalous results kept the case from being closed: it has been long established that the synthesis of cAMP is inhibited by glucose transport[39] via the phosphoenolpyruvate-dependent carbohydrate phosphotransferase system (PTS)[40–43]; however, growth on PTS-independent carbohydrates and limitations in nitrogen or phosphorous also affected the cAMP pool and, consequently, the degree of CCR[44–48]. These anomalies, and subsequent work[49,50], brought into question the unique role of cAMP–Crp in the modulation of carbon catabolic protein expression[51] and began a (thus far unsuccessful) search for additional regulators[52].

### Growth on single carbon sources

By focusing to the carbon flux carried by catabolism and anabolism, in addition to constraints on protein synthesis, it becomes possible to disentangle the regulation-centric view of CCR from intrinsic protein-synthesis constraints on gene expression. Recent work supports the view that CCR acts primarily to coordinate metabolic flux and identifies keto

acids as the metabolic node responsible for conveying a flux mismatch between catabolism and anabolism[24].

Similar to the protein-synthesis constraints on ribosomal and non-ribosomal proteins (Fig. 2), the protein-synthesis constraints on catabolism and anabolism are analysed by independently varying carbon and nitrogen fluxes and then observing the resulting concentration of flux-carrying enzymes. The balance between carbon catabolic flux and anabolic flux can be shifted in several ways. Early studies focused on varying the composition of the growth medium[49,53] and using amino acid auxotrophs[50,52]; more recent work uses inducible promoters to control the expression of key metabolic enzymes. For example[24], the expression of lactose permease can be controlled to modulate carbon (lactose) influx in lactose-minimal media, and the expression of glutamate dehydrogenase (GDH) can be controlled to modulate anabolic flux via limiting amino acid synthesis (Fig. 3a). A proxy for the concentration of carbon catabolic proteins is (IPTG-induced) β-galactosidase (Fig. 3b) whereas a proxy for the concentration of anabolic proteins is glutamine synthetase (Fig. 3c).

Under conditions of lactose limitation, the concentration of β-galactosidase exhibits a negative correlation with growth rate and a zero-expression intercept at a specific growth rate of $\lambda_C \approx 1.2$ /h (Fig. 3b) in contrast with the zero-expression limit of an unregulated protein, which is at approximately 2.2 /h (Box 1, figure part b). The same growth-dependent behaviour in the concentration of carbon catabolic proteins is produced when the growth rate is modulated by changes in the carbon source (Box 1, figure part a). Other forms of carbon limitation (for example, titration of glycerol uptake) exhibit an identical response; therefore, we will refer to the negative correlation between the concentration of β-galactosidase and the growth rate as the 'carbon catabolism-limited response' of carbon catabolic genes. This negative linear correlation of carbon catabolic gene expression is synonymous with CCR, and has been referred to as the 'C-line'[24].

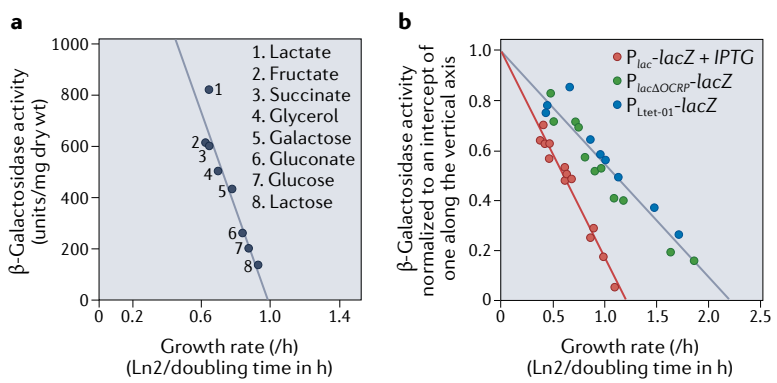
The expression of anabolic enzymes, using the reporter glutamine synthetase (*glnA*) as a proxy, exhibits a clear anti-correlation with those of carbon catabolic proteins. Under conditions of anabolic limitation, the concentration of anabolic proteins exhibits a negative linear correlation with growth rate (Fig. 3c) whereas carbon catabolic proteins exhibit a positive linear correlation with growth rate, passing through the origin (Fig. 3b). Under conditions of carbon catabolic limitation, the concentration of anabolic proteins exhibits a positive linear correlation with growth rate, passing through the origin (Fig. 3c). Subsequently, the carbon catabolism-limited and anabolism-limited response of the reporter enzymes was validated by proteomics measurements for a large number of catabolic and anabolic enzymes[19,27]. The ribosomal concentration (which is proportional to the RNA to protein ratio under both carbon catabolic and anabolic limitations) exhibits a positive

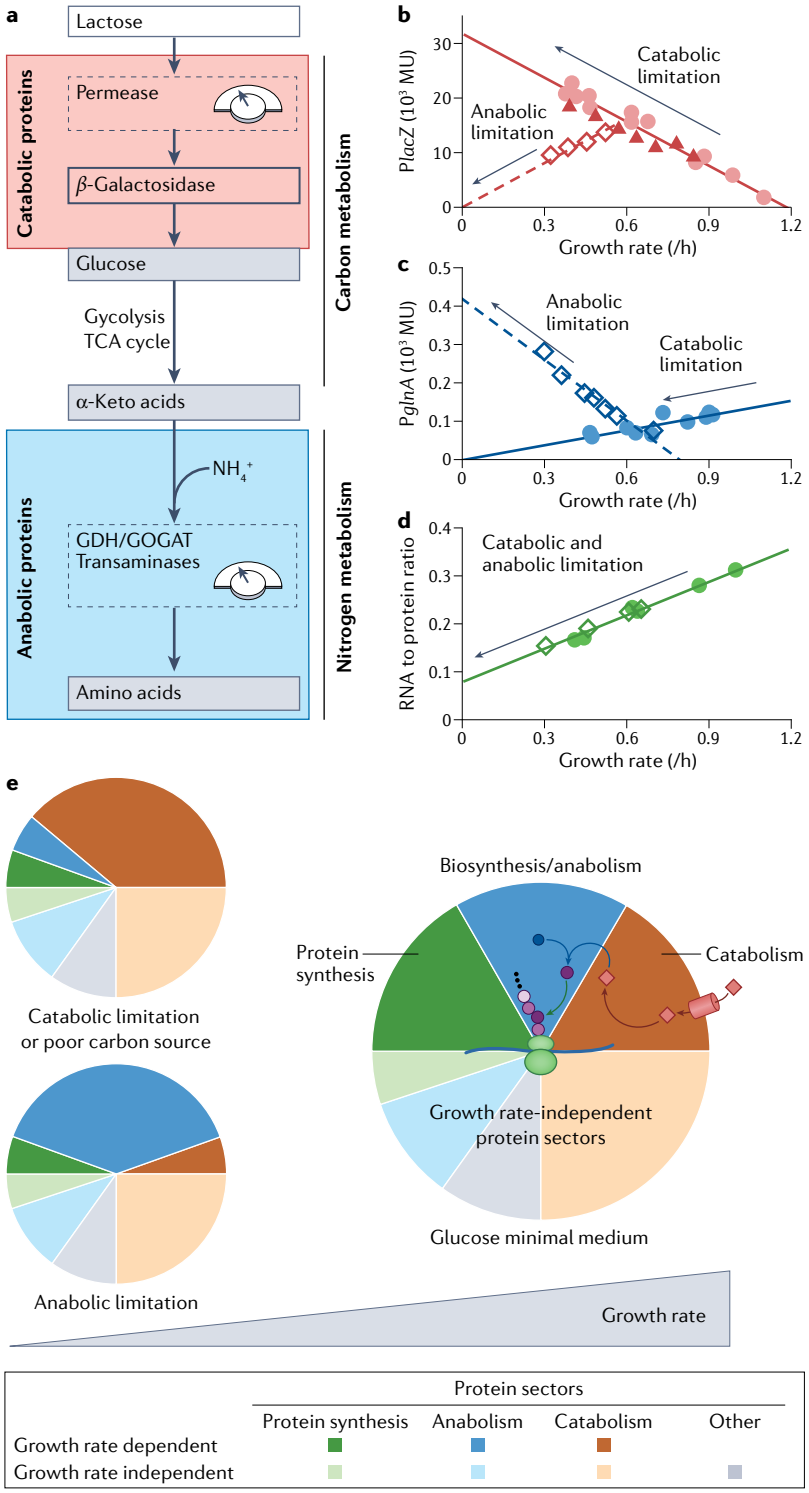
## Box 1

# The glucose effect and carbon catabolite repression

'Carbon catabolite repression' (CCR) was coined by Magasanik[32] to replace what had previously been referred to as the 'glucose effect,' whereby the presence of glucose in the growth medium was seen to correlate with a reduction in the synthesis of catabolic enzymes responsible for breaking down carbon substrates other than glucose. Carbon sources (such as glucose) that support rapid growth result in lower catabolic enzyme expression as compared to carbon sources (such as lactate) that support slower growth[49]. A scatter plot of the data (see the figure, part **a**) shows a striking negative linear correlation between the activity of β-galactosidase ($P_{lac}$-*lacZ*) and the growth rate; see also red symbols in the figure, part **b**, for β-galactosidase in a wild type *Escherichia coli* strain fully induced by IPTG.

Regulation by cyclic adenosine monophosphate (cAMP)–cAMP receptor protein (Crp) is necessary to produce the distinct growth dependence of carbon catabolic gene expression. If the Crp-binding site in the *lac* promoter is scrambled (see the figure, part **b**, $P_{lac\Delta OCRP}$), then the growth dependence of the expressed gene reverts to the canonical unregulated growth dependence (see the figure, part **b**, $P_{Ltet\text{-}O1}$; compare with the solid line in Fig. 2b). To facilitate comparison, the activities in Box figure part **b** were scaled so that a linear fit through the data intercepts the vertical axis at 1. Data in part **a** from ref.[49]; data in part **b** from ref.[24].






linear correlation with growth rate (Fig. 3d) irrespective of growth limitation by catabolism or anabolism.

Using the protein mass fraction $\phi_i$ of an individual protein as a measure of cellular abundance (which we adopt henceforth) provides several advantages over concentration. First, in exponential growth, the protein mass fraction of a given protein is equal to the fraction of ribosomes actively translating that protein (thereby manifesting the macromolecular machinery constraint; Supplementary Box 1). Second, all protein fractions must sum to one, thereby conveniently enforcing the constant-protein-density constraint.

Using a series of growth perturbations and protein mass spectrometry, it was found that individual proteins could be grouped into

**Fig. 3 | Modulation of carbon and nitrogen flux reveals protein synthesis constraints on catabolic and anabolic proteins.** In a minimal growth medium without externally supplied amino acids, the bacterium must synthesize amino acids via amination of carbon precursors (keto acids)[131] (part **a**). These precursors are supplied through the carbon catabolic enzymes, glycolysis, the tricarboxylic acid (TCA) cycle and various biosynthesis pathways. To minimize substrate-specific effects, carbon influx was modulated by titrating the expression of the uptake system (lactose permease for growth on lactose[24]) or by changing the carbon source in minimal medium. The nitrogen flux was modulated by titrating glutamate dehydrogenase (GDH) in a glutamine oxoglutarate aminotransferase (GOGAT)-deleted background[24], or by titrating GOGAT in a GDH-deleted background[27]. The flux control points are denoted by dashed boxes. Proxies for the carbon catabolic and anabolic protein concentrations are β-galactosidase activity (*lacZ*; panel **b**) and glutamine synthetase activity (*glnA*; panel **c**), respectively. These proxies quantitatively capture the behaviour of many other catabolic and anabolic proteins as validated by later proteomic work[19,27]. Carbon catabolic and anabolic enzyme concentrations exhibit obvious anti-correlation and near-linear growth dependence under various growth perturbations, whereas the ribosomal proteins exhibit a positive linear correlation with growth rate irrespective of the metabolic limitation that is used (RNA to protein ratio; part **d**). The red line in part **b** is sometimes called the 'C-line'[24], and can be taken as a defining feature of carbon catabolite repression (Box 1). Carbon flux was modulated by a change in carbon source (filled circles) or by titrating lactose permease (filled triangles); nitrogen flux was modulated by titrating GDH in a GOGAT-deleted background (open diamonds). The correlations among the abundances of catabolic, anabolic and ribosomal proteins can be understood quantitatively if abundance is measured in units of protein mass fraction (part **e**). The constancy of protein density and the allocation constraint on the protein synthetic machinery are both captured by a coarse-grained partitioning of the proteome. For simplicity, only four sectors are shown: growth-rate dependent ribosomal (and ribosome affiliated) sector (green), biosynthetic/anabolic (blue) and catabolic (red) sectors, along with their associated growth-rate independent basal expression (pale sectors), and a growth-rate independent sector (grey)[27]. The near-linear response of the growth-dependent sectors is rationalized by invoking a simple flux balance: external nutrients are converted to carbon precursors by catabolic proteins (red arrows) and converted into amino acid precursors (purple circles) by biosynthetic proteins (blue arrow) at a rate matched with amino acid consumption by protein synthesis (green arrow). Catabolic limitation leads to increased expression of catabolic proteins (top left), whereas anabolic limitation leads to increased expression of anabolic proteins (bottom left). Parts **b**, **c** and **d** adapted from ref.[24], Springer Nature Limited.

a small number of sectors sharing a common response to growth rate change[27] (Fig. 3e) (this was confirmed in a recent large-scale study[19]). The cellular abundance of each individual protein is expressed as a protein mass fraction $\phi_i$ composed of a growth-independent (basal) component $\phi_i^0$ and a growth-dependent component $\Delta\phi_i(\lambda)$, where $\lambda$ denotes the exponential growth rate:

$$\phi_i = \phi_i^0 + \Delta\phi_i(\lambda)$$

With the proteome partitioned into proteome sectors, it becomes clear that, if proteins in one sector increase in mass fraction, then other sectors must decrease to accommodate the change (Fig. 3e). Thus, the combination of constraints on total protein synthesis flux and on total cellular protein concentration leads to a constraint on the composition of individual protein concentrations (as represented by the pie chart, Fig. 3e), which we call the 'proteome allocation constraint.'

Irrespective of the method of growth perturbation, all growth-dependent protein sectors exhibit a linear dependence on growth rate, and enzymes with correlated expression are found to largely share common functionalities (for example, protein synthesis, carbon catabolism amino acid biosynthesis)[27]. This behaviour is rationalized by assuming that the flux carried by enzymes in that sector is proportional to the sum of the growth-dependent fractions of those enzymes[27,54]:

$$\Delta\Phi_\alpha = \frac{\sum_{i\subset\alpha}\Delta\phi_i}{\text{enzymes in sector }\alpha},$$

$$\text{flux through enzymes contained in proteome sector }\alpha = \kappa_\alpha\Delta\Phi_\alpha(\lambda)$$

where $\kappa_\alpha$ is a growth rate-independent factor characterizing the effective catalytic constant for the coarse-grained activity of that sector. It is important to note that, although all proteins within a given sector $\alpha$ are co-expressed, not all proteins in that sector may be carrying flux in any particular growth condition[19]. Proteomic data[19,27], which supports the hypothesis that the enzyme-catalysed reaction flux is proportional to the growth-dependent protein mass fraction $\Delta\Phi_\alpha$, provides a quantitative, empirical framework for coupling flux balance analysis with proteome allocation constraints[55].

Assuming a flux balance between carbon catabolism and anabolism, as required by exponential growth, the linearity of the growth-dependent response of each metabolic sector implies that the flux is proportional to the growth-dependent protein fractions. Denoting the growth-dependent mass fraction of carbon catabolic proteins by $\Delta\Phi_C$ and the growth-dependent mass fraction of anabolic proteins by $\Delta\Phi_A$, the flux balance constraint is expressed succinctly as:

$$\kappa_C\,\Delta\Phi_C = \kappa_A\,\Delta\Phi_A,$$

where $\kappa_C$ and $\kappa_A$ are proportionality constants that quantify the overall catalytic efficiency of each sector. Beyond the linearity of the response, the second striking feature of the data shown in Fig. 3b,c is the anti-correlation between the catabolic and anabolic protein sectors, implying a second proteome allocation constraint operating in addition to the overarching constraint illustrated in Fig. 2.

Altogether, the data suggest a hierarchical scheme of proteome allocation: first, the amino acid demands of protein synthesis determine the allocation of the growth-dependent sectors of the proteome between ribosomal proteins (setting the amino acid consumption flux) and metabolic proteins (setting the amino acid supply flux[56]; Fig. 2). Within the metabolic proteins, a second balance is struck between carbon catabolism and anabolism.

## Coordination of flux balance and proteome allocation constraints

Proteome allocation constraints lead to simple, linear dependence of protein mass fractions on the steady-state growth rate. Yet, the constraints themselves reveal little about the underlying regulation responsible for this apparent simplicity. Much is known about the molecular regulators that control the expression of these large proteome sectors: guanosine tetraphosphate (ppGpp)[57] for ribosomal proteins and cAMP–Crp[40] for carbon catabolic proteins. In this section, we discuss

these signalling pathways in the context of proteome allocation constraints and flux balance.

The ribosomal protein mass fraction (denoted by $\Delta\Phi_R$ in Fig. 4) is set by the flux of ribosome biogenesis, which is inhibited by ppGpp through the synthesis of ribosomal RNA[58]. The synthesis rate of ppGpp itself responds negatively to the rate of translational elongation[59], with translation rate acting as an integrated sensor of the availability of all 20 amino acids. An accumulation of the amino acid pools relieves repression of ribosomal RNA transcription, leading to feedforward activation of ribosome synthesis. Subsequently, the increased rate of amino acid consumption reduces amino acid accumulation (Fig. 4). In this way, the amino acid pools balance protein synthesis and amino acid biosynthesis fluxes.

Flux balance between carbon catabolism and biosynthesis is maintained by the regulatory complex cAMP–Crp, which activates the expression of carbon catabolic genes[38], whose abundance is denoted by $\Delta\Phi_C$ in Fig. 4. The readout of a flux mismatch seems to be the keto acid pool (α-ketoglutarate, pyruvate and oxaloacetate), with a build-up of keto acids attenuating the synthesis of cAMP[24]. Consequently, an increase in the keto acid pools produces an effective feedback inhibition of carbon catabolic genes via downregulation of the activator cAMP–Crp (Fig. 4). Less is known about the regulation of biosynthesis enzymes, whose abundance is denoted by $\Delta\Phi_A$ in Fig. 4. The abundance of biosynthetic enzymes is reduced under carbon-limited growth[19,24,27]; this could arise from end-product inhibition mediated by the pool of individual amino acids and/or tRNAs[60], inhibition due to the lack of carbon precursors[24,61,62], or passive effects of global regulation[63].

Despite the divergent molecular mechanisms to achieve flux balance, there are two key regulatory motifs that enable *E. coli* to use the concentration of a signalling molecule to sense and control flux (Box 2).

A persistent feature of the proteome fractions is the linear growth-rate dependence under various modes of growth limitation. Flux-based regulation provides a mechanism to generate a linear relation between the flux and the enzyme concentration. Consider, for example, the feed-forward scheme illustrated in Box 2, figure part a. If the concentration $[Z]$ of an active enzyme $Z$ is directly proportional to the concentration $[S]$ of a signalling molecule $S$ (that is, $[Z] = c[S]$) and if the flux $v$ through enzyme Z has a Michaelis–Menten dependence on the signalling molecule concentration $[S]$ (that is, $v = k_{cat}[Z][S]/([S] + K_s)$), then a linear relation between flux and the total enzyme concentration is obtained over a range of fluxes (as long as the flux is not too small):

$$[Z] = av + b$$

where $a \propto k_{cat}^{-1}$ is the slope and $b \propto cK_s$ is the flux-independent offset. This relation would readily give rise to a linear dependence of the enzyme concentration on the growth rate $\lambda$ because many metabolic fluxes are proportional to the growth rate[64].

The simplest way to achieve this regulatory scheme is by making signalling molecule S the substrate of enzyme Z or the substrate of another reaction in the same linear pathway as Z (and hence sharing the same flux). The latter has been demonstrated explicitly for the glycolytic enzyme PykF[65] (Box 2, figure part c), for which the linear dependence between flux and enzyme abundance is expected to hold so long as the enzyme is biased to operate in the regime of excess substrate $[S] \geq K_s$ (this is the case for many metabolic enzymes[54,66]). A more elaborate implementation of this regulatory scheme is known for ribosome biogenesis (Box 2, figure part e). In this case, charged tRNAs are the ribosome substrates[8,67], and ppGpp is the signal that integrates the different substrates. An effective Michaelis–Menten relation between the substrates and the flux, assumed in past phenomenological studies[8,20,67,68], is shown to be molecularly regulated by ppGpp via its response to the translational elongation rate, on the one hand, and its regulation of rRNA and tRNA, on the other[59].

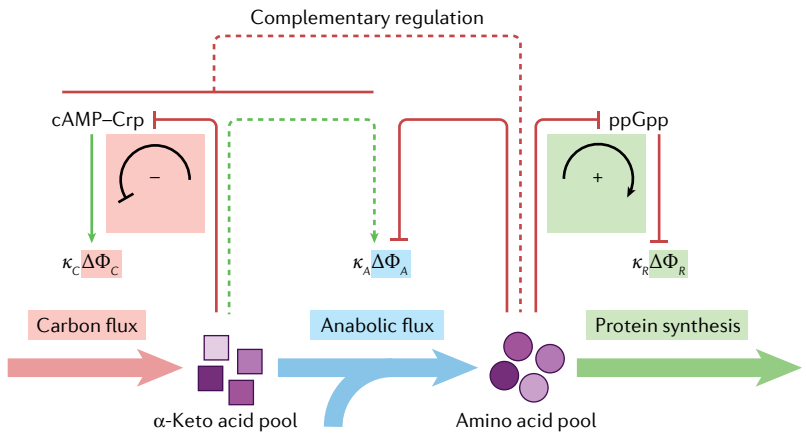



**Fig. 4 | Coordination of catabolic and anabolic flux via cAMP–Crp signalling.** Two notable features of the growth dependence in the various protein fractions in Fig. 3b–d are the near-linear response and the strong anti-correlation. The linearity can be rationalized by assuming that the flux mediated by each sector is proportional to protein mass[54], where the proportionality constant $\kappa_i$ is a measure of the catalytic efficiency of each sector. The anti-correlation suggests the existence of complementary regulation to enforce the proteome allocation constraints. The direct regulatory interactions are well characterized (solid lines). Accumulation of α-keto acids results in the downregulation of carbon catabolic proteins ($\Delta\Phi_C$) by decreasing the activity of the global activator cyclic adenosine monophosphate (cAMP)–cAMP receptor protein (Crp) (negative feedback loop indicated in the red box)[62]. Accumulation of amino acids leads to upregulation of ribosomal proteins ($\Delta\Phi_R$) by decreasing the level of the alarmone ppGpp which represses ribosome biogenesis (positive feedforward loop indicated in the green box)[58,132]. The biosynthetic proteins ($\Delta\Phi_A$) are directly regulated by end-product inhibition via individual amino acids[60] (red solid arrow). In addition to these direct mechanisms, proteome allocation constraints necessitate anti-correlated, complementary regulation (complementary regulation is denoted by the dashed lines).

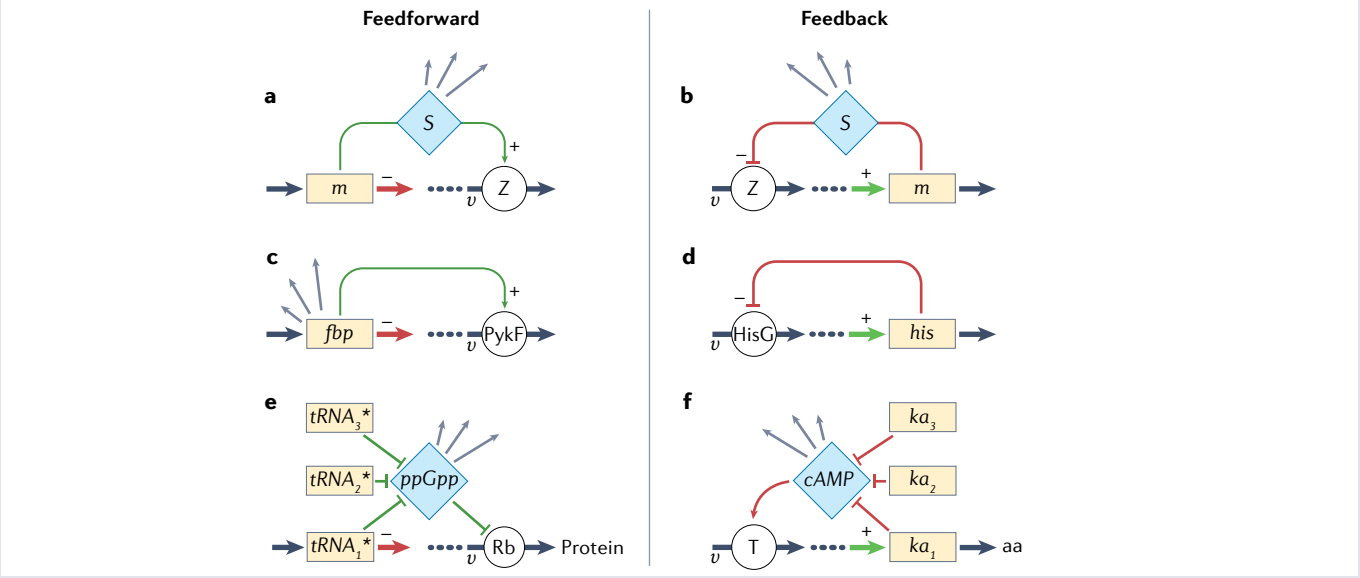
## Box 2

# Implementation of flux-controlled regulation

Flux-controlled regulation shares common features (see the figure, parts **a** and **b**):

- The abundance of a metabolite pool *m* is affected by the pathway flux *v* modulated by an enzyme *Z*.
- The activity of enzyme *Z* uniquely determines the pathway flux *v*.
- A signalling molecule *S* determines the activity of the enzyme *Z*.
- The concentration of the signalling molecule *S* responds to the metabolite pool *m* to provide stable feedback.

Here, we describe implementations of the core motifs. The metabolite *fbp* is an allosteric activator of the downstream enzyme PykF and acts as a glycolytic flux sensor[133] (see the figure, part **c**). The flux *v* through the pathway consumes and hence depletes *fbp* (thick red arrow), but the activation of PykF by *fbp* (green regulatory link) ensures flux balance. End-product inhibition is another common flux-responsive regulatory strategy. In the histidine biosynthesis pathway (see the figure, part **d**), histidine is an allosteric inhibitor of the upstream enzyme HisG[134]. If the flux *v* of histidine synthesis (thick green arrow) does not balance the incorporation into protein synthesis, negative feedback (red regulatory link) acts to restore the balance. In the figure, parts **c** and **d**, the metabolite pool functions as its own signalling molecule. If the flux depends upon a range of metabolites, then a signalling metabolite *S* is necessary to integrate their combined effect on the pathway flux as is the case in the regulation of protein synthesis (see the figure, part **e**) and carbon catabolic protein expression (see the figure, part **f**). Ribosomes convert charged tRNA (*tRNA**) to proteins (see the figure, part **e**, thick red arrow). The regulatory link[58] from the charged tRNA to ribosome abundance involves a signalling molecule[135], guanosine tetraphosphate (ppGpp), which detects the shortage of any one of the charged tRNAs by monitoring the ribosome elongation rate[59]. The double-negative regulation results in an overall positive connection between charged tRNA and ribosome abundance (green regulatory links). Keto acid pools (*ka*) report on total carbon flux *v* (see the figure, part **f**, thick green arrow), and are supplied by the carbon catabolic enzymes denoted by the transporter T. The keto acids inhibit the synthesis of cyclic adenosine monophosphate (cAMP)[24], which activates the expression of many catabolic enzymes[38], resulting in an overall negative connection between keto acid pools and catabolic enzyme abundance (red regulatory arrows). Signalling molecules, such as ppGpp[92,136] and cAMP[31,40], are global physiological signals that are used to regulate numerous other operons (grey arrows), effectively coupling the expression of many distal genes to the pathway flux *v*, driving coherent global changes observed in proteome dynamics. aa, amino acids.




Under conditions of severe growth limitation, the linear relation generated by the simple scheme above would break down; however, the relation between ribosome concentration and growth rate in *E. coli* is still found to be approximately linear when grown in very poor carbon sources[20] or in carbon-limited chemostat culture[69]. The maintenance of this linear relation has recently been suggested[59] to involve the additional regulation of ribosomal translational activity by using ribosome hibernating factors to make the majority of ribosomes inactive at very slow growth[70]. By contrast, in nitrogen-limited growth[69], it seems that tRNA charging (primarily glutamine) becomes rate limiting whereas, in phosphate-limited growth[69], the decreased synthesis of nucleic acid itself reduces ribosome abundance. How nitrogen and phosphate limitations lead to their corresponding ribosome abundances is not understood at the same quantitative level as carbon limitation.

The right column of Box 2 shows a similar regulatory motif implemented by a feedback circuit (Box 2, figure panel b). Qualitatively, this regulatory motif is established for a broad range of metabolic systems, including amino acid biosynthesis (Box 2, figure panel d) and catabolite repression (Box 2, figure panel f), although the details of how the catabolite repression circuit senses flux and leads to the C-line (Fig. 3b) have yet to be quantitatively established.

### Growth on multiple carbon sources

Microorganisms exhibit preferential utilization when presented with several substitutable carbon sources[33]. Although this was demonstrated by Monod in the context of kinetic growth transitions[71], the phenomenon of preferential nutrient usage can be elucidated quantitatively in steady-state growth[72,73]. For *E. coli*, it was observed that most pairs of glycolytic carbon sources are hierarchically utilized (that is, carbon sources are metabolized one at a time) whereas pairs of gluconeogenic carbon sources are simultaneously utilized (that is, they are metabolized concurrently) as are combinations of the two types[72]. The observed substrate utilization pattern has been rationalized in terms of the proteome cost involved in the conversion between glycolytic and gluconeogenic substrates[74] or more generally in terms of elementary growth modes[75] (recently reviewed in ref.[76]).

For those combinations that are simultaneously utilized, the uptake rate of each substrate is generally not equal: it is reduced in the presence of a co-utilized substrate, yet the steady-state growth rates on the two substrates exceed that on either substrate alone. The same inverse linear relation between carbon catabolic enzyme expression and growth rate (C-line) is observed for growth on various combinations of co-utilized carbon substrates, with the same zero-expression intercept at a growth rate of $\lambda_C = 1.2\ h^{-1}$[72] (Fig. 3b). The common intercept implies that, although the growth rate on multiple substrates increases with respect to a single substrate, there is a 'speed limit' imposed at a growth rate of $\lambda_c$. The co-utilization phenomenon can be accounted for quantitatively using the same flux-matching scheme as in the single-carbon case (now enabling for carbon influx from multiple sources). Using only the carbon-limited speed limit $\lambda_c$, the resulting model can predict the growth rate on multiple substrates based on growth rates on individual substrates[72].

For the combinations of substrates that give rise to hierarchical usage, the preferred substrate is usually the one with a higher single-substrate growth rate. This preferred utilization can be converted to simultaneous utilization if the uptake rate of the preferred substrate is reduced below the uptake rate of the other substrate[73], indicating that the onset of preferred utilization is again controlled by a flux-controlled regulation scheme. Detailed analysis[73] identified the involvement of cAMP–Crp together with another substrate-specific regulator.

### Growth transition kinetics

To control the kinetics of growth transitions, the cell needs to monitor its own state of growth, which involves thousands of reactions. Which molecular species should the cell monitor to diagnose its own state of protein synthesis? A useful strategy, both for the scientists studying bacterial response and for the bacterium managing its own growth, is to capture key kinetic variables that reflect the coarse-grained dynamics of the cell[59,77].

The feedback–feedforward control schemes mediated by ppGpp and cAMP-Crp outlined in the previous sections (and shown schematically in Fig. 4) exhibit a common flux-controlled regulatory motif (Box 2). In both cases, flux mismatch is measured by the accumulation of signalling molecules (α-keto acids in the case of CCR and charged tRNAs in the case of ribosome biogenesis). The result of accumulation of the signalling molecule is a reduction of the supply flux along with a simultaneous increase in the consumption flux. A transformative effect of the cAMP and ppGpp signalling pathways is for the cell to condense, or coarse-grain, the large variety of individual metabolites (for example, the different species of keto acids for cAMP signalling[24] and the different amino acid–tRNA species for ppGpp signalling[59]) onto the concentrations of just two molecular species, cAMP and ppGpp (Box 2, figure panels e, f). In this way, the metabolites no longer directly participate in downstream regulation but rather participate through their signalling molecules.

However, this qualitative coarse-graining picture does not immediately lend itself to quantitative prediction because it remains largely unknown how cAMP and ppGpp synthesis (and turnover) respond to changes in the large number of metabolites they monitor. A simple coarse-grained framework was developed[78] that enables quantitative predictions of kinetic transitions (for example, variations of the classic diauxic shift) to be made based upon the gross topology of cAMP–ppGpp signalling (Box 2, figure panels e, f) and their effect on proteome allocation in the steady state (C-line and R-line in Fig. 3b,d), without the need to know any details of the signalling pathways. Instead of describing the complex regulatory interactions that link the concentrations of keto acids and amino acids to cAMP and ppGpp signalling, which control the expression of carbon catabolic and ribosomal protein mass (Fig. 5, $M_C$ and $M_R$), the authors postulated[78] that flux-controlled regulation is driven by the translational activity $\sigma$, defined as the ratio between the protein synthesis flux, $J_R$, and the ribosomal protein mass, $M_R$. Growth kinetics are then determined by a simple dynamic loop involving the effect of the present allocation of the proteome ($M_C$ and $M_R$) on translational activity and the effect of translational activity on future allocation of $M_C$ and $M_R$ via the regulatory functions $\chi_C$ and $\chi_R$ (Fig. 5). The form of the regulation functions, which are implemented by the signalling molecules cAMP and ppGpp and not known quantitatively, was bypassed in their treatment by using the relation between the regulatory functions $\chi_{C,R}$ and the translational activity $\sigma$ in the steady state (easily obtained via empirical steady state correlations) and by assuming that the same relation determines the regulation function in the kinetic regime, where translational activity itself changes in time[78].

The mechanistic justification of using translational activity to drive the regulation of ribosome biogenesis $\chi_R$ has since been substantiated through the discovery that the ppGpp pool responds directly to the peptide elongation rate[59]: because the translational activity $\sigma$ so-defined is a product of the peptide elongation rate and the fraction of active ribosomes and the ppGpp pool regulates the active ribosome fraction[58,59], it follows that the translational activity $\sigma$ is uniquely dependent on the ppGpp pool and acts as its proxy. In light of the relation between ppGpp and translational activity, we can understand the tremendous degree of 'dimensional reduction' performed by Erickson et al.[78], where the kinetics of bacterial growth transitions are determined by a few parameters despite the formal involvement of a large number of variables and parameters (metabolite concentrations and kinetic constants characterizing each reaction) as a consequence of coarse-graining by the cell itself, where ppGpp senses the availability of all the amino acids and tRNAs by simply sensing the rate of peptide elongation[59,77].

We note that a similar understanding of the connection between translational activity and regulation of catabolic protein synthesis $\chi_C$ is currently still lacking. The red feedback arrow in Fig. 5 does not mean that ppGpp activates the catabolic sector — the catabolic sector is

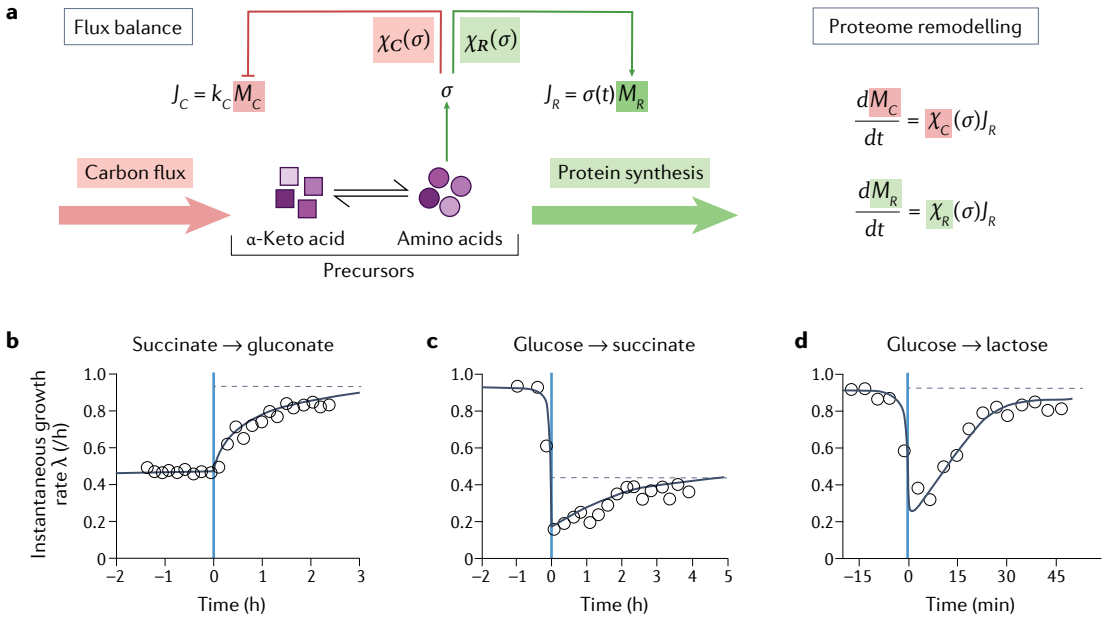



**Fig. 5 | Flux-controlled regulation and proteome remodelling during growth transitions. a**, The steady-state growth-dependence of the various protein fractions enables quantitative characterization of the kinetics of growth transitions. The schematic network shown in Fig. 4 can be further simplified to emphasize the main (direct) regulatory loops. α-Keto acids and amino acids are collected into a pool of 'precursors' that affect the magnitude of a single coarse-grained variable, the ribosome's translational activity $\sigma$. The carbon uptake flux $J_C$ is proportional to the carbon catabolic protein mass $M_C$. It is matched to the protein synthesis flux $J_R$, which is proportional to the ribosomal protein mass $M_R$. The ratio of the flux $J_R$ and the mass $M_R$ defines the translational activity $\sigma$, which is the central dynamical variable regulating the expression of carbon catabolic proteins ($M_C$, via cyclic adenosine monophosphate (cAMP)–cAMP receptor protein (Crp)), and ribosomal proteins ($M_R$, via guanosine tetraphosphate (ppGpp))[78]. These direct regulatory effects are implemented through two regulatory functions, $\chi_C(\sigma)$, the fraction of ribosomes actively translating carbon catabolic proteins, and $\chi_R(\sigma)$, the fraction of ribosomes actively translating ribosomal proteins. The quantitative form of the regulatory functions $\chi_C(\sigma)$ and $\chi_R(\sigma)$ are determined by the observed steady-state correlations of these quantities. In this way, the dynamic response of a bacterial culture to a shift in carbon substrate can be fully predicted by solving a single differential equation for the translational activity $\sigma(t)$. From the solution, the dynamic remodelling of the proteome, that is, changes in the cellular mass of proteins for ribosomes ($M_R(t)$), carbon catabolism ($M_C(t)$) and anabolism ($M_A(t)$), can be obtained by integrating the regulatory functions $\chi_R(\sigma(t))$ and $\chi_C(\sigma(t))$ with the anabolic protein mass $M_A(t)$ obtained from the constraint on total protein synthesis. For transitions between simultaneously utilized carbon sources, the time course of growth kinetics and proteome remodelling can be predicted for both upshifts and downshifts, using only the above mentioned information on steady-state growth in the pre-shift and post-shift medium. **b**,**c**, The relaxation of the instantaneous growth rate $\lambda = \frac{1}{M}\frac{dM}{dt}$ during two exemplary shifts (circles are data, solid black lines are predictions with no fitting parameters). Small modification of the same approach can also be used to capture the growth kinetics for shift between hierarchically utilized carbon sources. **d**, Result of the classic diauxie between glucose and lactose, whereby a single fitting parameter is introduced to set the time point of the *lac* operon activation after glucose depletion. Dashed lines denote the final growth rate on the second substrate. Parts **b**, **c** and **d** adapted from ref.[78], Springer Nature Limited.

clearly regulated via cAMP–Crp and cAMP synthesis is affected by the keto acid pools[24]. We hypothesize that a coarse-graining scheme similar to that which relates amino acids to ppGpp (via the peptide elongation rate[59]) relates keto acids to cAMP. However, because keto acids and amino acids are reversibly connected through transamination[61,79], a relation exists between the cAMP and ppGpp pools when growth is limited by carbon influx (as during diauxic shifts), which leads to a connection between catabolic allocation $\chi_C$ and translational activity $\sigma$ exploited in ref.[78].

Under these key assumptions, the dynamics of a nutrient shift are fully specified with the only additional information being the steady-state growth rate in the pre-shift and post-shift medium ($\lambda_i$ and $\lambda_f$, respectively). Furthermore, both nutrient downshifts and nutrient upshifts are accommodated within the same framework (distinguished only by whether $\lambda_i$ is larger or smaller than $\lambda_f$).

The simplest nutrient shift is achieved by exhausting one of the simultaneously utilized carbon sources. In that case, all requisite enzymes and transporters are present throughout and their cellular mass abundances are known from the pre-shift steady state. Figure 5b,c shows the predictability of this flux-controlled kinetic model on the adaptation dynamics of the instantaneous growth rate for an exemplary upshift and downshift; similar predictability is observed in the dynamics of proteome remodelling[78]. An unexpected feature of the growth-shift kinetics for simultaneously utilized carbon sources is that, despite the expression of all the required processing enzymes and cognate transporters prior to the shift, it takes several hours to reach the post-shift steady state. The success of the theoretical framework (Fig. 5b,c) suggests several contributions of the proteome allocation constraints on the prolonged adaptation timescale. First, flux-controlled regulation leads to coordinated response of large sectors of the proteome. Induction of a single operon is very rapid but remodelling the proteome following a nutrient shift requires altering the degree of expression of hundreds of genes (for example, carbon catabolic genes that comprise up to 30% of the proteome for a switch to poor-carbon conditions)

## Glossary

**Anabolic enzymes**
Enzymes responsible for biosynthesis, including amino acid and nucleotide synthesis.

**Carbon catabolic proteins**
Proteins responsible for the transport and breakdown of extracellular carbon sources. Operationally, these are genes regulated by cAMP–Crp.

**C-line**
The negative correlation between catabolic enzyme expression and growth rate in minimal media when growth rate is modulated by carbon source.

**Diauxic growth**
Multiple stages of exponential growth as carbon sources are preferentially utilized. The time to switch between carbon sources can take several hours.

**Nutrient quality**
The exponential growth rate can be modulated by changing the carbon source or nitrogen source, or enriching the medium with amino acids and nucleotides.

**Per-total protein-mass abundance**
The concentration constraint (biomass/cell water volume) allows a direct conversion between the concentration ((amount of a particular molecule)/(cell water volume)) and its abundance relative to total protein mass ((amount of a particular molecule)/(total protein mass)) assuming that the biomass contains a fixed fraction of protein.

**Protein density**
The buoyant density ((biomass and water mass)/cell volume) is independent of the growth rate under isotonic conditions. A constant density constraint (biomass/cell volume) therefore implies a constant concentration constraint (biomass/cell water volume). The cellular volume has strong growth-rate dependence; consequently, we do not use protein-number-per-cell as an abundance measure in this Review.

**Protein mass fraction**
The total number of a particular protein is proportional to its protein mass. Using the per-total protein-mass abundance defined above, the protein mass fraction (protein mass of a particular protein/total protein mass) is therefore a direct measure of concentration.

**Proteome**
The set of all expressed proteins in a given growth condition.

**Proteome sectors**
Subsets of the proteome that exhibit similar growth-rate dependence under various growth conditions.

**R-line**
The positive correlation between the abundance of protein-synthetic machinery (chiefly ribosomes) and growth rate when growth rate is modulated by nutrient quality.

**Unregulated protein**
Protein expression not subject to any transcriptional or post-transcriptional regulation.

and greatly extends the adaptation time. Second, beyond adjusting synthesis rates, the cell must remove proteins that are 'overexpressed' with respect to the post-shift growth condition. The inherent stability of most bacterial proteins means that a decrease in mass fraction must be achieved by dilution through cell doubling. For a downshift, this involves replacing a fraction of the ribosomal protein and anabolic proteins by carbon catabolic proteins, and vice versa for an upshift. Depending upon the nature of shifts, additional internal metabolic bottlenecks may further lengthen the adaptation time[80,81].

The model of flux-controlled regulation described for shifts between simultaneously utilized substrates can be readily adapted to describe diauxic growth between hierarchically utilized substrates. Figure 5d is an example of hierarchical utilization between glucose and lactose. In this case, glucose and lactose support almost identical steady-state growth rates, yet glucose in the medium results in a near-complete repression of the lactose catabolic genes. Although hierarchical utilization requires additional regulatory mechanisms (in this case, 'inducer exclusion'[82,83]), it seems that cAMP–Crp continues to be the primary flux sensor[73]. The kinetics of hierarchical utilization of glucose and lactose can be accurately described by invoking a single fitting parameter that specifies the time at which the *lac* operon is de-repressed, reflecting a threshold in glucose influx below which lactose starts to be taken up[73]. Thereafter, as with the simple nutrient shifts, the kinetics of hierarchical utilization conform to what is predicted from the coarse-grained model (Fig. 5d). Much faster transition times can be attained if the synthesis of growth-limiting enzymes are prioritized[81,84]. In the canonical shift from glucose to lactose (both of which support doubling times of about 40 min), at full induction, *lac* proteins occupy approximately 1–2% of the protein content of the cell. A targeted response to the shift would be expected to take less than a minute (1–2% of 40 min), and empirically can be achieved in less than 2 min (ref.[85]). Instead, it appears that *E. coli* adopts the versatile, generalist strategy of flux-controlled regulation which, though independent of the details of how metabolic flux is generated, nevertheless carries with it the burden of global proteome remodelling, extending the adaptation time. Instead of upregulating the required 1–2% of the catabolic proteins needed to respond to a carbon downshift, the cell upregulates the entire sector of catabolic proteins, most of which carry no flux[19,81]. This surprising behaviour possibly reflects the lack of dedicated sensors and regulators to detect the instantaneous nutrient landscape and 'compute' the optimal cause of action[86]. It is also possible that, in natural environments, the remaining nutrients are distributed across a variety of different types rather than in one type as usually done in diauxic shift studies[71].

## Conclusions and outlook

The bacterial cell can be viewed in (at least) two different ways: as a membrane-enclosed autocatalytic loop of ribosomes synthesizing ribosomes (along with the necessary metabolic enzymes fuelling the substrates needed for the ribosomes to do their work) or, alternatively, as a finely tuned collection of metabolic enzymes orchestrated by a chemical soup of signalling molecules designed to transform external nutrients into a replicating cell. Although both views are simultaneously accurate and useful, they constrain one another[22], and irrespective of whether the object of focus is protein synthesis or metabolism, the physiological constraints operating in the background unavoidably sculpt the gene expression landscape and consequently shape the behaviour of the system.

In this Review, we focused on the correlation between ribosomal protein abundance and growth rate (R-line; Figs. 2a and 3d), which imposes complementary regulation on the expression of all other non-ribosomal proteins. Among these, the expression of carbon catabolic proteins exhibits further strong correlation with growth rate (C-line; Box 1 and Fig. 3b), a phenomenon synonymous with CCR. Together, these proteome allocation constraints were captured by a phenomenological model of bacterial growth that quantitatively predicts the dynamics of growth transitions among carbon sources without adjustable parameters[78].

Underlying the growth-dependent proteome allocation constraints are nested loops of flux-controlled regulation responsible

for coordinating the flux through large sectors of the proteome (Fig. 4 and Box 2). Flux-sensing regulatory motifs are particularly useful in mediating the interface between large networks because they are not sensitive to the details of how the flux is generated, providing a generalized strategy that can accommodate metabolic innovation without adjudicating on a case-by-case basis. The plug-and-play functionality of flux-controlled regulation has the benefit of not being tailored to any single growth environment but, of course, this generality may suffer from occasional non-optimal response under exotic growth conditions[86,87]. Nevertheless, the success of the phenomenological approach in the study of CCR points to the importance of proteome allocation constraints and flux-controlled regulation in studies of gene expression.

A number of recent theoretical studies[88–90] have suggested that existing observations on bacterial proteome allocation (including the R-line and C-line), support the notion that the proteome is optimally allocated to maximize the growth rate. Closer quantitative inspection of the existing data necessitates a broader definition of optimality that can accommodate the substantial non-flux-carrying protein sectors (both the growth-rate independent proteome sector[54] and the growth-dependent catabolic proteins that carry no flux[19,81]).

Although the proteome allocation constraints discussed in this Review provide a quantitative framework to predict the interdependence between the growth rate and gene expression, many open questions remain and the applicability of this approach beyond *E. coli* is largely unexplored.

The coordination of ribosome allocation and proteome allocation constraints with growth rate change requires a mechanism to sense the growth rate; in *E. coli*, ppGpp has that role[59,91]. The alarmone ppGpp exerts its regulatory effect primarily through modulation of transcription[57,92]. In other species, what are the molecular signals that sense growth-rate change, and transform this information into the appropriate proteome allocation? The best-characterized physiological constraints have come from nutrient-limited growth — either balanced exponential growth or transitions between balanced growth states. What about growth under stress, when the growth-limiting factor is not any nutrient component but other environmental parameters such as temperature, pH and osmolarity? Quite a lot of information is available on regulatory factors operating under these conditions[93–95]. However, little is known, at the quantitative level, about the physiological constraints imposed by these environmental factors nor about their effects on growth reduction. Our focus on proteome allocation constraints has not touched upon changes in cell size. The broader question of how cell size homeostasis and DNA replication[96] is coordinated with proteome partitioning and protein-synthesis constraints remains unknown. Current ideas are centred around the accumulation of a protein (or proteins) to initiate cell division[12,96–98].

What about non-growing or stationary conditions? Are they similar to a very slowly growing state[99] or fundamentally different? How do physiological constraints affect gene expression by cells in the stationary phase[100]? What about survival through the stationary phase[101]? Recent work on the effect of pre-shift growth conditions on survival[102] provides a rare example of the physiological study of stationary cells but much work is left to be done.

The coordination of ribosome abundance with growth rate[57] and carbon catabolic repression[33,34] are ubiquitous responses in bacteria. Are the protein synthesis constraints active in other organisms the same as those found in *E. coli* and, if so, are these constraints responding to similar flux-matching signals? Certainly, the molecular details are different among distant species. In the model firmicute *Bacillus subtilis*, for example, guanosine (penta)tetraphosphate ((p)ppGpp) seems to modulate ribosomal RNA operon activity indirectly by lowering the concentration of GTP rather than by direct binding to the RNA polymerase as in *E. coli*[103]. Furthermore, CCR in *B. subtilis* is achieved by transcriptional repression rather than via cAMP–Crp activation of the catabolic promoters[33,34]. Despite these differences in implementation, are the resulting control structures responding to similar cues to generate a similar flux balance among proteome sectors?

*B. subtilis* and *E. coli* are both capable of rapid growth — are similar proteome allocation constraints operating in slow-growing bacteria? Are proteome allocation constraints relevant in slow-growing species, or is the protein synthesis machinery no longer a growth-limiting resource under these conditions[104]? Despite the proteins responsible for (p)ppGpp signalling and the regulation of ribosome biogenesis being largely conserved among bacteria[105], there are varied molecular implementations depending upon whether the bacteria are copiotrophic or oligotrophic. Bacteria adapted to nutrient-rich environments (copiotrophs) tend to synthesize ppGpp in response to a number of individual starvation cues, exhibiting OR-logic combinatorial control, whereas bacteria adapted to chronic starvation (oligotrophs) tend to synthesize ppGpp in response to a combination of cues, typically in addition to amino acid starvation, exhibiting AND-logic combinatorial control[106]. Are proteome allocation constraints (and other physiological constraints) likewise shaped by the bacterial lifestyle? Are they found in extremophiles and archaea? Quantifying physiological constraints in photosynthetic organisms is complicated by the additional degrees of freedom afforded to growth modulation via light intensity and duration. Nevertheless, it seems that simple empirical relationships exist, for example, linking cellular glycogen to growth rate in cyanobacteria[107].

Microorganisms continue to be an important vector for biomanufacturing and synthetic circuits design. Proteome allocation constraints impose a hard upper limit on the mass fraction that can be occupied by a heterologous protein[12,22], irrespective of whether the protein is produced endogenously or via an orthogonal synthesis pathway[108]. Is it possible to engineer the host[109] or the growth medium to increase the heterologous protein limit? Or, alternatively, are there suitable microorganisms that have a larger propensity for heterologous production? The implementation of synthetic genetic circuits is limited by a lack of predictability, largely due to unaccounted crosstalk between the host and the synthetic construct[110,111] (or among different modules of the construct[112]). Including physiological constraints in mechanistic models of gene expression improves predictability[111,113,114], particularly when the growth rate changes. Designing flux-controlled regulation that couples the synthetic circuit to the host physiology could improve performance by ensuring flux matching among modules.

Beyond bacteria and archaea, unicellular eukaryotes are subjected to far more complex protein-expression regulation and compartmentalization. Yet, they exhibit apparent catabolite repression[115] and other metabolic transitions, such as overflow metabolism[116], that are also manifested in bacteria. Work has been done quantifying some physiological constraints in *Saccharomyces cerevisiae*[117–120], fungi[121,122] and mammalian cell lines[123] but, for the most part, eukaryotic organisms, particularly higher eukaryotes, remain under-investigated. Generally, much less is known quantitatively about issues specific to eukaryotes, for example, the role of protein turnover and secretion on proteome allocation[124,125]. The existence of a constraint on biomass density and hence on protein concentration is another unknown.

Indeed, even defining the intracellular protein concentration is non-trivial in eukaryotes due to the existence of many intracellular compartments, in particular, vacuoles, which can take up substantial intracellular volume but do not directly affect the cytosolic protein concentration[124,126].

Turning to communities of interacting microorganisms, can we define meta-physiological constraints? These may be particularly straightforward to define in model ecosystems comprised of a few species[127–129]. As a simple example, in a two-species Lotka–Volterra-type system, where both the densities of the predator and prey oscillate, the ribosome abundance in the predator is expected to oscillate with the predator growth rate whereas, in the prey, the ribosome abundance is expected to exhibit no time dependence (as the density change of the prey is due to changes in predation not growth). Can empirical correlations of this kind be used to trace putative predator–prey interactions in environmental samples?

Finally, the proteome allocation constraints we have focused on in this Review operate on the scale of days; how do the characterizations of these constraints change over evolutionary timescales? One of the outstanding mysteries surrounding physiological constraints in *E. coli* is the substantial allocation of the proteome to non-flux carrying genes[19,81]. Does this strategy arise from a trade-off between optimizing current need against a contingency for future uncertainty? If so, how then is the magnitude of the non-flux carrying proteome determined? Recent work predicting the emergence of antibiotic resistance[130] suggests that physiological constraints persist over short adaptation periods. Is it possible to use coarse-grained phenomenological parameters to infer underlying physiological changes during other evolutionary adaptation scenarios? Will this provide a wider outlook than a regulator-centric view? Although simple to state, and often simple to quantify, proteome allocation constraints have far-ranging implications for how we understand gene regulation in microorganisms.

Published online: 14 November 2022

## References


1. Neidhardt, F. C., Ingraham, J. L. & Schaechter, M. *Physiology of the Bacterial Cell: A Molecular Approach* (Sinauer Associates, 1990).
2. Bervoets, I. & Charlier, D. Diversity, versatility and complexity of bacterial gene regulation mechanisms: opportunities and drawbacks for applications in synthetic biology. *FEMS Microbiol. Rev.* **43**, 304–339 (2019).
3. Dorman, C. J. *Structure and Function of the Bacterial Genome* (Wiley-Blackwell, 2020).
4. Henkin, T. M. & Peters, J. E. *Snyder & Champness Molecular Genetics of Bacteria*. 5 edn (ASM Press, 2020).
5. Phillips, R. *The Molecular Switch: Signaling and Allostery* (Princeton University Press, 2020).
6. van den Berg, J., Boersma, A. J. & Poolman, B. Microorganisms maintain crowding homeostasis. *Nat. Rev. Microbiol.* **15**, 309–318 (2017).
7. Zhang, G. et al. Global and local depletion of ternary complex limits translational elongation. *Nucleic Acids Res.* **38**, 4778–4787 (2010).
8. Klumpp, S., Scott, M., Pedersen, S. & Hwa, T. Molecular crowding limits translation and cell growth. *Proc. Natl Acad. Sci. USA* **110**, 16754–16759 (2013).
9. Dai, X. et al. Slowdown of translational elongation in Escherichia coli under hyperosmotic stress. *mBio* https://doi.org/10.1128/mBio.02375-17 (2018).
10. Woldringh, C. L., Binnerts, J. S. & Mans, A. Variation in Escherichia coli buoyant density measured in Percoll gradients. *J. Bacteriol.* **148**, 58–63 (1981).
11. Kubitschek, H. E. Buoyant density variation during the cell cycle in microorganisms. *CRC Crit. Rev. Microbiol.* **14**, 73–97 (1987).
12. Basan, M. et al. Inflating bacterial cells by increased protein synthesis. *Mol. Syst. Biol.* **11**, 836 (2015).
13. Oldewurtel, E. R., Kitahara, Y. & van Teeffelen, S. Robust surface-to-mass coupling and turgor-dependent cell width determine bacterial dry-mass density. *Proc. Natl Acad. Sci. USA* https://doi.org/10.1073/pnas.2021416118 (2021).
14. Cayley, S., Lewis, B. A., Guttman, H. J. & Record, M. T. Characterization of the cytoplasm of Escherichia coli K-12 as a function of external osmolarity: implications for protein-DNA interactions in vivo. *J. Mol. Biol.* **222**, 281–300 (1991).
15. Milo, R. What is the total number of protein molecules per cell volume? A call to rethink some published values. *BioEssays* **35**, 1050–1055 (2013).
16. Balakrishnan, R. et al. Principles of gene regulation quantitatively connect DNA to RNA and proteins in bacteria. *bioRxiv* https://doi.org/10.1101/2021.05.24.445329 (2021).
17. Bremer, H. & Dennis, P. P. Modulation of chemical composition and other parameters of the cell at different exponential growth rates. *EcoSal* https://doi.org/10.1128/ecosal.5.2.3 (2008).
18. Schmidt, A. et al. The quantitative and condition-dependent Escherichia coli proteome. *Nat. Biotechnol.* **34**, 104–110 (2016).
19. Mori, M. et al. From coarse to fine: the absolute Escherichia coli proteome under diverse growth conditions. *Mol. Syst. Biol.* **17**, e9536 (2021).
20. Dai, X. et al. Reduction of translating ribosomes enables Escherichia coli to maintain elongation rates during slow growth. *Nat. Microbiol.* **2**, 16231 (2016).
21. Jun, S., Si, F. W., Pugatch, R. & Scott, M. Fundamental principles in bacterial physiology-history, recent progress, and the future with focus on cell size control: a review. *Rep. Prog. Phys.* **81**, 80 (2018).
22. Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z. & Hwa, T. Interdependence of cell growth and gene expression: origins and consequences. *Science* **330**, 1099–1102 (2010).
23. Neidhardt, F. C. & Magasanik, B. Studies on the role of ribonucleic acid in the growth of bacteria. *Biochim. Biophys. Acta* **42**, 99–116 (1960).
24. You, C. et al. Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature* **500**, 301–306 (2013).
25. Maaloe, O. in *Gene Expression Biological Regulation and Development* (ed Goldberger, R. F.) 487–542 (Plenum Press, 1979).
26. Klumpp, S., Zhang, Z. & Hwa, T. Growth rate-dependent global effects on gene expression in bacteria. *Cell* **139**, 1366–1375 (2009).
27. Hui, S. et al. Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Mol. Syst. Biol.* **11**, 784 (2015).
28. Schaechter, M., Maaloe, O. & Kjeldgaard, N. O. Dependency on medium and temperature of cell size and chemical composition during balanced grown of Salmonella typhimurium. *J. Gen. Microbiol.* **19**, 592–606 (1958).
29. Mairet, F., Gouze, J. L. & de Jong, H. Optimal proteome allocation and the temperature dependence of microbial growth laws. *NPJ Syst. Biol. Appl.* **7**, 14 (2021).
30. Kaspy, I. et al. HipA-mediated antibiotic persistence via phosphorylation of the glutamyl-tRNA-synthetase. *Nat. Commun.* **4**, 3001 (2013).
31. Chubukov, V., Gerosa, L., Kochanowski, K. & Sauer, U. Coordination of microbial metabolism. *Nat. Rev. Microbiol.* **12**, 327–340 (2014).
32. Magasanik, B. Catabolite repression. *Cold Spring Harb. Symposia Quant. Biol.* **26**, 249–256 (1961).
33. Deutscher, J. The mechanisms of carbon catabolite repression in bacteria. *Curr. Opin. Microbiol.* **11**, 87–93 (2008).
34. Gorke, B. & Stulke, J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nat. Rev. Microbiol.* **6**, 613–624 (2008).
35. Epps, H. M. & Gale, E. F. The influence of the presence of glucose during growth on the enzymic activities of Escherichia coli: comparison of the effect with that produced by fermentation acids. *Biochem. J.* **36**, 619–623 (1942).
36. Ullmann, A. & Monod, J. Cyclic AMP as an antagonist of catabolite repression in Escherichia coli. *FEBS Lett.* **2**, 57–60 (1968).
37. Perlman, R. & Pastan, I. Cyclic 3′5-AMP: stimulation of beta-galactosidase and tryptophanase induction in E. coli. *Biochem. Biophys. Res. Commun.* **30**, 656–664 (1968).
38. Zubay, G., Schwartz, D. & Beckwith, J. Mechanism of activation of catabolite-sensitive genes: a positive control system. *Proc. Natl Acad. Sci. USA* **66**, 104–110 (1970).
39. Saier, M. H. Jr, Feucht, B. U. & Hofstadter, L. J. Regulation of carbohydrate uptake and adenylate cyclase activity mediated by the enzymes II of the phosphoenolpyruvate: sugar phosphotransferase system in Escherichia coli. *J. Biol. Chem.* **251**, 883–892 (1976).
40. Kolb, A., Busby, S., Buc, H., Garges, S. & Adhya, S. Transcriptional regulation by cAMP and its receptor protein. *Annu. Rev. Biochem.* **62**, 749–795 (1993).
41. Postma, P. W., Lengeler, J. W. & Jacobson, G. R. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol. Rev.* **57**, 543–594 (1993).
42. Saier, M. H. Jr. Regulatory interactions involving the proteins of the phosphotransferase system in enteric bacteria. *J. Cell. Biochem.* **51**, 62–68 (1993).
43. Deutscher, J., Francke, C. & Postma, P. W. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol. Mol. Biol. Rev.* **70**, 939–1031 (2006).
44. Epstein, W., Rothman-Denes, L. B. & Hesse, J. Adenosine 3′:5′-cyclic monophosphate as mediator of catabolite repression in Escherichia coli. *Proc. Natl Acad. Sci. USA* **72**, 2300–2304 (1975).
45. Hogema, B. M. et al. Catabolite repression by glucose 6-phosphate, gluconate and lactose in Escherichia coli. *Mol. Microbiol.* **24**, 857–867 (1997).
46. Bettenbrock, K. et al. Correlation between growth rates, EIIACrr phosphorylation, and intracellular cyclic AMP levels in Escherichia coli K-12. *J. Bacteriol.* **189**, 6891–6900 (2007).
47. McFall, E. & Magasanik, B. Effects of thymine and of phosphate deprivation on enzyme synthesis in Escherichia coli. *Biochim. Biophys. Acta* **55**, 900–908 (1962).
48. Clark, D. J. & Marr, A. G. Studies on the repression of beta-galactosidase in Escherichia coli. *Biochim. Biophys. Acta* **92**, 85–94 (1964).
49. Mandelstam, J. The repression of constitutive beta-galactosidase in Escherichia coli by glucose and other carbon sources. *Biochem. J.* **82**, 489–493 (1962).
50. Magasanik, B. & Neidhardt, F. C. Inhibitory effect of glucose on enzyme formation. *Nature* **178**, 801–802 (1956).

51. Ullmann, A. Catabolite repression: a story without end. *Res. Microbiol.* **147**, 455–458 (1996).
52. Wanner, B. L., Kodaira, R. & Neidhardt, F. C. Regulation of lac operon expression: reappraisal of the theory of catabolite repression. *J. Bacteriol.* **136**, 947–954 (1978).
53. Magasanik, B. & Neidhardt, F. C. The effect of glucose on the induced biosynthesis of bacterial enzymes in the presence and absence of inducing agents. *Biochim. Biophys. Acta* **21**, 324–334 (1956).
54. Dourado, H., Mori, M., Hwa, T. & Lercher, M. J. On the optimality of the enzyme-substrate relationship in bacteria. *PLoS Biol.* **19**, e3001416 (2021).
55. Mori, M., Hwa, T., Martin, O. C., De Martino, A. & Marinari, E. Constrained allocation flux balance analysis. *PLoS Comput. Biol.* **12**, e1004913 (2016).
56. Scott, M., Klumpp, S., Mateescu, E. M. & Hwa, T. Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Mol. Syst. Biol.* **10**, 747 (2014).
57. Hauryliuk, V., Atkinson, G. C., Murakami, K. S., Tenson, T. & Gerdes, K. Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat. Rev. Microbiol.* **13**, 298–309 (2015).
58. Paul, B. J., Ross, W., Gaal, T. & Gourse, R. L. rRNA transcription in Escherichia coli. *Annu. Rev. Genet.* **38**, 749–770 (2004).
59. Wu, C. et al. Cellular perception of growth rate and the mechanistic origin of bacterial growth laws. *Proc. Natl Acad. Sci. USA* **119**, e2201585119 (2022).
60. Umbarger, H. E. Amino acid biosynthesis and its regulation. *Annu. Rev. Biochem.* **47**, 532–606 (1978).
61. Reitzer, L. Nitrogen assimilation and global regulation in Escherichia coli. *Annu. Rev. Microbiol.* **57**, 155–176 (2003).
62. Huergo, L. F. & Dixon, R. The emergence of 2-oxoglutarate as a master regulator metabolite. *Microbiol. Mol. Biol. Rev.* **79**, 419–435 (2015).
63. Kochanowski, K. et al. Global coordination of metabolic pathways in Escherichia coli by active and passive regulation. *Mol. Syst. Biol.* **17**, e10064 (2021).
64. Varma, A. & Palsson, B. O. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110. *Appl. Environ. Microbiol.* **60**, 3724–3731 (1994).
65. Kochanowski, K. et al. Functioning of a metabolic flux sensor in Escherichia coli. *Proc. Natl Acad. Sci. USA* **110**, 1130–1135 (2013).
66. Bennett, B. D. et al. Absolute metabolite concentrations and implied enzyme active site occupancy in Escherichia coli. *Nat. Chem. Biol.* **5**, 593–599 (2009).
67. Hu, X. P., Dourado, H., Schubert, P. & Lercher, M. J. The protein translation machinery is expressed for maximal efficiency in Escherichia coli. *Nat. Commun.* **11**, 5260 (2020).
68. Marr, A. G. Growth rate of Escherichia coli. *Microbiol. Rev.* **55**, 316–333 (1991).
69. Li, S. H. et al. Escherichia coli translation strategies differ across carbon, nitrogen and phosphorus limitation conditions. *Nat. Microbiol.* **3**, 939–947 (2018).
70. Prossliner, T., Gerdes, K., Sorensen, M. A. & Winther, K. S. Hibernation factors directly block ribonucleases from entering the ribosome in response to starvation. *Nucleic Acids Res.* **49**, 2226–2239 (2021).
71. Monod, J. in *Selected Papers in Molecular Biology by Jacques Monod* (eds Lwoff, A. & Ullmann, A.) (Academic Press, 1978).
72. Hermsen, R., Okano, H., You, C., Werner, N. & Hwa, T. A growth-rate composition formula for the growth of E.coli on co-utilized carbon substrates. *Mol. Syst. Biol.* **11**, 801 (2015).
73. Okano, H., Hermsen, R., Kochanowski, K. & Hwa, T. Regulation of hierarchical and simultaneous carbon-substrate utilization by flux sensors in *Esherichia coli*. *Nat. Microbiol.* **5**, 206–215 (2020).
74. Wang, X., Xia, K., Yang, X. & Tang, C. Growth strategy of microbes on mixed carbon sources. *Nat. Commun.* **10**, 1279 (2019).
75. de Groot, D. H., Hulshof, J., Teusink, B., Bruggeman, F. J. & Planque, R. Elementary Growth Modes provide a molecular description of cellular self-fabrication. *PLoS Comput. Biol.* **16**, e1007559 (2020).
76. Okano, H., Hermsen, R. & Hwa, T. Hierarchical and simultaneous utilization of carbon substrates: mechanistic insights, physiological roles, and ecological consequences. *Curr. Opin. Microbiol.* **63**, 172–178 (2021).
77. Hwa, T. in *The Physics of Living Matter: Space, Time and Information* (eds Gross, D., Sevrin, A. & Shraiman, B.) 87–98 (World Scientific Publishing Co., 2020).
78. Erickson, D. W. et al. A global resource allocation strategy governs growth transition kinetics of Escherichia coli. *Nature* **551**, 119–123 (2017).
79. Yuan, J., Fowler, W. U., Kimball, E., Lu, W. & Rabinowitz, J. D. Kinetic flux profiling of nitrogen assimilation in Escherichia coli. *Nat. Chem. Biol.* **2**, 529–530 (2006).
80. Basan, M. et al. A universal trade-off between growth and lag in fluctuating environments. *Nature* https://doi.org/10.1038/s41586-020-2505-4 (2020).
81. Balakrishnan, R., de Silva, R. T., Hwa, T. & Cremer, J. Suboptimal resource allocation in changing environments constrains response and growth in bacteria. *Mol. Syst. Biol.* **17**, e10597 (2021).
82. Lengeler, J. W. in *Regulation of Gene Expression in Escherichia coli* (eds Lin, E. C. C. & Lynch, A. S.) Ch. 11, 231–254 (Chapman and Hall, 1996).
83. Magasanik, B. in *The Lactose Operon* (eds Beckwith, J. & Zipser, D.) 189–219 (Cold Spring Harbor Laboratory, 1970).
84. Pavlov, M. Y. & Ehrenberg, M. Optimal control of gene expression for fast proteome adaptation to environmental change. *Proc. Natl Acad. Sci. USA* **110**, 20527–20532 (2013).
85. Riley, M., Pardee, A. B., Jacob, F. & Monod, J. On the expression of a structural gene. *J. Mol. Biol.* **2**, 216–225 (1960).
86. Bren, A. et al. Glucose becomes one of the worst carbon sources for E.coli on poor nitrogen sources due to suboptimal levels of cAMP. *Sci. Rep.* **6**, 24834 (2016).
87. Towbin, B. D. et al. Optimality and sub-optimality in a bacterial growth law. *Nat. Commun.* **8**, 14123 (2017).
88. Muller, S., Regensburger, G. & Steuer, R. Enzyme allocation problems in kinetic metabolic networks: optimal solutions are elementary flux modes. *J. Theor. Biol.* **347**, 182–190 (2014).
89. Bruggeman, F. J., Planque, R., Molenaar, D. & Teusink, B. Searching for principles of microbial physiology. *FEMS Microbiol. Rev.* https://doi.org/10.1093/femsre/fuaa034 (2020).
90. Dourado, H. & Lercher, M. J. An analytical theory of balanced cellular growth. *Nat. Commun.* **11**, 1226 (2020).
91. Potrykus, K., Murphy, H., Philippe, N. & Cashel, M. ppGpp is the major source of growth rate control in E. coli. *Environ. Microbiol.* **13**, 563–575 (2011).
92. Sanchez-Vazquez, P., Dewey, C. N., Kitten, N., Ross, W. & Gourse, R. L. Genome-wide effects on Escherichia coli transcription from ppGpp binding to its two sites on RNA polymerase. *Proc. Natl Acad. Sci. USA* **116**, 8310–8319 (2019).
93. Hengge-Aronis, R. Recent insights into the general stress response regulatory network in Escherichia coli. *J. Mol. Microbiol. Biotechnol.* **4**, 341–346 (2002).
94. Richter, K., Haslbeck, M. & Buchner, J. The heat shock response: life on the verge of death. *Mol. Cell* **40**, 253–266 (2010).
95. Imlay, J. A. The molecular mechanisms and physiological consequences of oxidative stress: lessons from a model bacterium. *Nat. Rev. Microbiol.* **11**, 443–454 (2013).
96. Si, F. et al. Mechanistic origin of cell-size control and homeostasis in bacteria. *Curr. Biol.* **29**, 1760–1770.e7 (2019).
97. Zheng, H. et al. General quantitative relations linking cell growth and the cell cycle in Escherichia coli. *Nat. Microbiol.* **5**, 995–1001 (2020).
98. Colin, A., Micali, G., Faure, L., Cosentino Lagomarsino, M. & van Teeffelen, S. Two different cell-cycle processes determine the timing of cell division in Escherichia coli. *eLife* https://doi.org/10.7554/eLife.67495 (2021).
99. Cooper, S. On the fiftieth anniversary of the Schaechter, Maaloe, Kjeldgaard experiments: implications for cell-cycle and cell-growth control. *Bioessays* **30**, 1019–1024 (2008).
100. Gefen, O., Fridman, O., Ronin, I. & Balaban, N. Q. Direct observation of single stationary-phase bacteria reveals a surprisingly long period of constant protein production activity. *Proc. Natl Acad. Sci. USA* **111**, 556–561 (2014).
101. Kaplan, Y. et al. Observation of universal ageing dynamics in antibiotic persistence. *Nature* **600**, 290–294 (2021).
102. Biselli, E., Schink, S. J. & Gerland, U. Slower growth of Escherichia coli leads to longer survival in carbon starvation due to a decrease in the maintenance rate. *Mol. Syst. Biol.* **16**, e9478 (2020).
103. Krasny, L. & Gourse, R. L. An alternative strategy for bacterial ribosome synthesis: Bacillus subtilis rRNA transcription regulation. *EMBO J.* **23**, 4473–4483 (2004).
104. Muller, A. L. et al. An alternative resource allocation strategy in the chemolithoautotrophic archaeon Methanococcus maripaludis. *Proc. Natl Acad. Sci. USA* https://doi.org/10.1073/pnas.2025854118 (2021).
105. Atkinson, G. C., Tenson, T. & Hauryliuk, V. The RelA/SpoT homolog (RSH) superfamily: distribution and functional evolution of ppGpp synthetases and hydrolases across the tree of life. *PLoS One* **6**, e23479 (2011).
106. Boutte, C. C. & Crosson, S. Bacterial lifestyle shapes stringent response activation. *Trends Microbiol.* **21**, 174–180 (2013).
107. Zavrel, T. et al. Quantitative insights into the cyanobacterial cell economy. *eLife* https://doi.org/10.7554/eLife.42508 (2019).
108. Costello, A. & Badran, A. H. Synthetic biological circuits within an orthogonal central dogma. *Trends Biotechnol.* **39**, 59–71 (2021).
109. Kim, J., Darlington, A., Salvador, M., Utrilla, J. & Jimenez, J. I. Trade-offs between gene expression, growth and phenotypic diversity in microbial populations. *Curr. Opin. Biotechnol.* **62**, 29–37 (2020).
110. Cardinale, S. & Arkin, A. P. Contextualizing context for synthetic biology–identifying causes of failure of synthetic biological systems. *Biotechnol. J.* **7**, 856–866 (2012).
111. Brophy, J. A. & Voigt, C. A. Principles of genetic circuit design. *Nat. Methods* **11**, 508–520 (2014).
112. Qian, Y., Huang, H. H., Jimenez, J. I. & Del Vecchio, D. Resource competition shapes the response of genetic circuits. *ACS Synth. Biol.* **6**, 1263–1272 (2017).
113. Weisse, A. Y., Oyarzun, D. A., Danos, V. & Swain, P. S. Mechanistic links between cellular trade-offs, gene expression, and growth. *Proc. Natl Acad. Sci. USA* **112**, E1038–E1047 (2015).
114. Braniff, N., Scott, M. & Ingalls, B. Component characterization in a growth-dependent physiological context: optimal experimental design. *Processes* **7**, 23 (2019).
115. Ronne, H. Glucose repression in fungi. *Trends Genet.* **11**, 12–17 (1995).
116. Compagno, C., Dashko, S. & Piskur, J. in *Molecular Mechanisms in Yeast Carbon Metabolism* (eds Compagno, C. & Piskur, J.) 1–21 (Springer, 2014).
117. Kafri, M., Metzl-Raz, E., Jonas, F. & Barkai, N. Rethinking cell growth models. *FEMS Yeast Res.* https://doi.org/10.1093/femsyr/fow081 (2016).
118. Boer, V. M., Crutchfield, C. A., Bradley, P. H., Botstein, D. & Rabinowitz, J. D. Growth-limiting intracellular metabolites in yeast growing under diverse nutrient limitations. *Mol. Biol. Cell* **21**, 198–211 (2010).
119. Metzl-Raz, E. et al. Principles of cellular resource allocation revealed by condition-dependent proteome profiling. *eLife* https://doi.org/10.7554/eLife.28034 (2017).
120. Hackett, S. R. et al. Systems-level analysis of mechanisms regulating yeast metabolic flux. *Science* https://doi.org/10.1126/science.aaf2786 (2016).
121. Brown, C. M. & Rose, A. H. Effects of temperature on composition and cell volume of Candida utilis. *J. Bacteriol.* **97**, 261–270 (1969).

122. Alberghina, F. A., Sturani, E. & Gohlke, J. R. Levels and rates of synthesis of ribosomal ribonucleic acid, transfer ribonucleic acid, and protein in Neurospora crassa in different steady states of growth. *J. Biol. Chem.* **250**, 4381–4388 (1975).
123. Kochanowski, K. et al. Systematic alteration of in vitro metabolic environments reveals empirical growth relationships in cancer cell phenotypes. *Cell Rep.* **34**, 108647 (2021).
124. Hecht, K. A., O'Donnell, A. F. & Brodsky, J. L. The proteolytic landscape of the yeast vacuole. *Cell Logist.* **4**, e28023 (2014).
125. Tyo, K. E., Liu, Z., Magnusson, Y., Petranovic, D. & Nielsen, J. Impact of protein uptake and degradation on recombinant protein secretion in yeast. *Appl. Microbiol. Biotechnol.* **98**, 7149–7159 (2014).
126. Armstrong, J. Yeast vacuoles: more than a model lysosome. *Trends Cell Biol.* **20**, 580–585 (2010).
127. Hays, S. G., Yan, L. L. W., Silver, P. A. & Ducat, D. C. Synthetic photosynthetic consortia define interactions leading to robustness and photoproduction. *J. Biol. Eng.* **11**, 4 (2017).
128. Chuang, J. S., Frentz, Z. & Leibler, S. Homeorhesis and ecological succession quantified in synthetic microbial ecosystems. *Proc. Natl Acad. Sci. USA* **116**, 14852–14861 (2019).
129. Amarnath, K. et al. Stress-induced cross-feeding of internal metabolites provides a dynamic mechanism of microbial cooperation. *bioRxiv* https://doi.org/10.1101/2021.06.24.449802 (2021).
130. Pinheiro, F., Warsi, O., Andersson, D. I. & Lassig, M. Metabolic fitness landscapes predict the evolution of antibiotic resistance. *Nat. Ecol. Evol.* https://doi.org/10.1038/s41559-021-01397-0 (2021).
131. Reitzer, L. Biosynthesis of glutamate, aspartate, asparagine, L-alanine, and D-alanine. *EcoSal* https://doi.org/10.1128/ecosalplus.3.6.1.3 (2004).
132. Goldman, E. & Jakubowski, H. Uncharged tRNA, protein synthesis, and the bacterial stringent response. *Mol. Microbiol.* **4**, 2035–2040 (1990).
133. Kotte, O., Zaugg, J. B. & Heinemann, M. Bacterial adaptation through distributed sensing of metabolic fluxes. *Mol. Syst. Biol.* **6**, 355 (2010).
134. Winkler, M. E. & Ramos-Montanez, S. Biosynthesis of histidine. *EcoSal* https://doi.org/10.1128/ecosalplus.3.6.1.9 (2009).
135. Irving, S. E., Choudhury, N. R. & Corrigan, R. M. The stringent response and physiological roles of (pp)pGpp in bacteria. *Nat. Rev. Microbiol.* **19**, 256–271 (2021).
136. Magnusson, L. U., Farewell, A. & Nystrom, T. ppGpp: a global regulator in Escherichia coli. *Trends Microbiol.* **13**, 236–242 (2005).


## Acknowledgements

This Review was shaped by extended discussion with numerous colleagues and collaborators over the years. It grew from early discussions with Eduard Mateescu and Stefan Klumpp, and with Hans Bremer, Lazlo Csonka, Antoine Danchin, Patrick Dennis, Peter Geiduschek, Sydney Kustu, Bill Loomis, Elio Schaechter, and Dalai Yan. Many insightful ideas came from colleagues whose proteomic data underlies the pie charts shown in the figures: Ruedi Aebersold, Gene-wei Li, Christina Ludwig, and especially Vadim Patsalo, Josh Silverman and Jamie Williamson. Our current understanding of proteome allocation constraints would not have been possible without the input of Rosalind Allen, Frank Bruggeman, Mans Ehrenberg, Suckjoon Jun, Meriem el Karoui, Karl Kochanowski, Martin Lercher, Fernanda Pinheiro, Uwe Sauer, Bas Teusink, Yiping Wang, and current and former members of the Hwa laboratory, especially Rohan Balakrishnan, Markus Basan, David Erickson, Tony Hui, Matteo Mori, Hiroyuki Okano, Severin Schink, Chenhao Wu, Conghui You and Zhongge Zhang. Support for the Hwa laboratory has been provided by the NIH (R01GM095903, R01GM109069), the NSF (PHY105873, MCB 1818384) and the Simons Foundation (330378). M.S. was supported by NSERC (2016-03658).

## Author contributions

The authors contributed equally to all aspects of the article.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41579-022-00818-6.

**Correspondence** should be addressed to Matthew Scott or Terence Hwa.

**Peer review information** *Nature Reviews Microbiology* thanks Kerwyn Casey Huang, who co-reviewed with Handuo Shi; Joshua Rabinowitz; and Uwe Sauer for their contribution to the peer review of this work.