Title

Proteomic allocation by E.coli during growth transitions, from strategy to mechanism

Author

Wu, Chenhao

Publication Date

2022

UNIVERSITY OF CALIFORNIA SAN DIEGO

Proteomic allocation by *E.coli* during growth transitions, from strategy to mechanism

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in

Physics

by

Chenhao Wu

Committee in charge:

Professor Terence Hwa, Chair
Professor Suckjoon Jun
Professor Elena F. Koslover
Professor Sergey Kryazhimskiy

2022

The dissertation of Chenhao Wu is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

DEDICATION

I dedicated my dissertation to my husband Daniel Lombardo and our parents. I’m grateful for their love and support. It was encouraging when my husband shared his experience in graduate study with me. Our cat has also been a great support to me, always comforting me with her warm belly, although unwillingly.

EPIGRAPH

Il n'ya qu'un héroïsme au monde：c'est de voir le monde tel qu'il est et de l'aimer
—Romain Rolland

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my advisor, Terence Hwa, for his support and guidance through these years. His advice for my research projects are very insightful and it has always been a great pleasure to discuss scientific problems with him. He largely changed the way I think and made me better in solving problems. Without his support and encouragement, I would not have been able to successfully complete my Ph.D. research.

I would also like to thank my other committee members, Suckjoon Jun, Sergey Kryazhimskiy, Elena F. Koslover. Talking with them always motivated me and encouraged me to move forward in my studies.

I would like to thank all the people who worked with me during these years, including current and previous members in Hwa's lab and other collaborator labs. They shared their experience in experiments with me, and discussed scientific problems with me. From daily errands to important research issues, they have helped me a lot . They have made my student life more colorful.

Chapter 2, is coauthored with Zhongge Zhang, Vadim Patsalo, James R. Williamson, Terence Hwa. The dissertation author was the primary investigator and author of this material.

Chapter 3, in full, has been submitted for publication in Nature Microbiology in 2022. The authors are Chenhao Wu, Zhongge Zhang, Amir Banaei-Esfahani, Hiroyuki Okano, Matteo Mori, Ruedi Aebersold, Christina Ludwig, Terence Hwa. The dissertation author was the primary investigator and author of this material.

Chapter 4 is coauthored with Terence Hwa. The dissertation author was the primary author of this chapter.

Chapter 5, in full, has been approved for publication by PNAS, 2022. The authors are Chenhao Wu, Rohan Balakrishnan, Nathan Braniff, Matteo Mori, Gabriel Manzanarez, Zhongge Zhang, Terence Hwa. The dissertation author and Rohan Balakrishnan were the primary investigators and authors of this material.

VITA

2014	Bachelor of Physics, Nanjing University, China
2014-2021	Teaching Assistant, University of California San Diego
2014-2021	Research Assistant, University of California San Diego
2022	Doctor of Philosophy in Physics, University of California San Diego

PUBLICATIONS

C. Wu, R. Balakrishnan, et al., Cellular perception of growth rate and the mechanistic origin of bacterial growth laws. bioRxiv, 2021.10.16.464649 (2021); PNAS (in press).

C. Wu, et al., Proteome reserves define the onset of enzyme recovery across biosynthesis pathways. Under review, Nature Microbiology (2022).

FIELDS OF STUDY

Major Field: Physics

Studies in Quantitative Biology

Professor Terrence Hwa

ABSTRACT OF THE DISSERTATION

Proteomic allocation by *E. coli* during growth transitions, from strategy to mechanism

by

Chenhao Wu
Doctor of Philosophy in Physics

University of California San Diego, 2022

Professor Terence Hwa, Chair

In different environments, bacteria are known to allocate their proteome differently and achieve different growth rates. Extensive quantitative studies on proteomic allocations in steady states have been done. Yet it's still not clear how bacteria manage to adjust their proteome responding to different environments through thousands of reactions and eventually reach different growth rates. In order to gain better understanding on the adaptation strategy and mechanism, in my dissertation, I studied the kinetic behavior of *E.coli* during various environmental changes.

Mathematical modeling were used to quantitatively capture the proteomic re-allocation and growth rate adaptation during transitions. Further studies on the kinetics of Guanosine tetraphosphate (ppGpp) and translational elongation rate during growth transition revealed the molecular basis of well-known ribosomal ‘growth law’.

In Chapter 2 and 3, we studied the growth transitions from ‘rich media’, a condition with ample amino acids or/and other nutrient supplies. We started with the case of methionine depletion in Chapter 2. MetE were found to be the major bottleneck in met downshift. By adapting the flux-controlled regulatory model established for carbon shift, we quantitatively captured the relationship between pre-shift MetE reserves and lag time before growth recovery. In Chapter 3, we studied the proteomic allocation during all AAs depletion. Quantitative proteomic analysis revealed a linear relationship between the onset time of enzymes across AAs biosynthesis pathways and their fractional reserve in pre-shift condition, indicating a ‘as-needed’ gene expression strategy during all AAs downshift. Combining flux-controlled global regulation and pathway specific end-production inhibition, we successfully captured the proteome recovery kinetics using only data collected in the pre- and post-shift steady-states.

Chapter 4 is focused on bacterial kinetic response to sub-lethal chloramphenicol (an antibiotic inhibiting translation) treatment. By using translational elongation rate as a flux sensor and the key signaling variable, we accurately predicted the kinetics of biomass accumulation and gene expressions.

In Chapter 5, we show that translational elongation rate is inversely proportional to ppGpp, an essential molecule regulating translational machinery, during transient and steady-states. We established ppGpp rate-sensing strategy and thus closed a key regulatory circuit linking ppGpp, growth rate, ribosomal content and translational rate together.

Chapter 1

Introduction

Bacterial growth rate or reproduction rate is central to their fitness in all environment. The basic picture of bacterial growth can be described as Fig. 1.1 : Bacteria take in nutrient source from the environment and convert them into biomass (proteins, RNA, lipids, etc.) for reproduction. This process needs the coordination of transporters, ribosomes and many other proteins as shown in Fig. 1.1. It's known that bacteria adjust their proteome composition according to the nutrient conditions, temperature, pH, etc. of the environment and achieve different growth rates accordingly. Previous studies have revealed various gene regulations involved in proteomic allocations at molecular level. However, the overall strategy and mechanism of how bacteria adapts its growth to different conditions remain unclear. This is also the major interest of this dissertation. In my dissertation, I focus on the growth of E.coli, which is one of the most well-studied living species. Its natural habitats are closely related to the living space of human beings, including soil, guts, ocean, etc.. The fluctuations of these environments make it essential for E.coli to quickly respond to the external changes. Many studies has been done related to its proteomic allocation in steady-state and even in transient.

1.1 Steady-state growth law

When providing constant growing environment, E.coli can eventually grow exponentially with a constant rate for hundreds of cycles before dying, which is referred to as steady-state growth. During steady-state growth, at population level, every component of biomass (proteins, RNA, etc.) accumulated exponentially with the same exponential rate. To describe this behavior, we let M_X

denote the abundance of X in a standard culture volume, e.g., 1ml of culture. X can be any specific molecule or any group of molecules in the cells. Then

$$\frac{dM_X}{dt} = \lambda M_X, \tag{1.1}$$

where in steady-state growth, λ is constant, representing shared exponent accumulation rate by all components of cells. We called it growth rate in the following notes. Therefore all the intensive properties inside cells, such as concentration, are constant during steady-state growth. Here we introduce an important concept called protein mass fraction, which can be written as

$$\phi_X = \frac{M_X}{M_{pro}}, \tag{1.2}$$

serves a good representation of the concentration of protein X in various growth conditions (1, 2). By simply changing nutrient conditions, e.x. carbon source, nitrogen source or add additional AAs, its doubling time can vary from more than 24 hrs to less than 30mins. As a result, even though steady-state growth is not common in E.coli's natural habitat, it's provide a great opportunity to understand the allocation of bacterial proteome across different conditions.

Since 1950s, many studies on exponential phase showed macromolecular composition has strong correlation with steady-state growth rate. One of the most important findings associated with proteomic allocation strategy is the linear relationship between the ribosome contents and growth rates while changing nutrient source (Fig. 1.2a), which applies to many different species (3–5). This physiological observation can be written as

$$\phi_R = \phi_{R0} + \lambda/\kappa_t \tag{1.3}$$

Here ϕ_R is the abundance of ribosomal related proteins in mass fraction and λ is the growth rate. κ_t and ϕ_{R0} represent the inverse of the slope and the offset in the observed linear relationship in Fig. 1.2a. The constant slope (κ_t) has been well accepted as the implication of constant saturated translational rate across varies nutrient condition since Maaleo (6). In other words, in order to achieve higher growth rate λ in better nutrient source, cells need higher ribosomal concentration due to the saturated translational rate. However, in recent studies, many measurements shows that the translational rate drops significantly in slow growth region, indicating that the proteome allocation strategy behind Eq. (1.3) is not as simple as it looks. In Chapter 5, we looked into Eq. (1.3) from the perspective of the regulatory mechanism of ribosomes. Also, we took the growth rate (or nutrient) sensing into consideration and eventually reproduce the linear growth law (Eq. (1.3)) mechanistically. It turns out that the linear relationship in Fig. 1.2a actually comes out of a set of nonlinear signaling and regulation equations of Guanosine tetraphosphate (ppGpp) signaling system and it's consistent with the drop of translational elongation rate in slow growth region.

Besides Eq. (1.3), such a simple linear relationship has also been found between different proteome components with growth rate under various limitations-- under carbon limitation, most catabolic proteins increase linearly with the decrease of the growth rate (Fig. 1.2b); Under chloramphenicol treatment (an antibiotic inhibit translation), ribosomal contents decrease with growth rate linearly, as opposite to that in nutrient limitation (Fig. 1.2c); under resource limitation by over-expressing useless proteins, the growth rate linearly decrease with the abundance of OE proteins (Fig. 1.2d). Although many regulatory systems related to those proteins (e.x., cAMP-CRP on catabolic proteins and ppGpp on ribosomal synthesis) have been well studies at molecular level, it's impossible to quantitatively describing those simple relationships based on molecular interactions due to numerous parameters involved. In 2010, Scotts et al (3) rationalized the set of linear relationships including Eq. (1.3) through a phenomenological coarse-grained flux-balance theory. To be detailed, the total protein content is coarse grained into three part (Fig. 1.3a)—growth rate independent part (Q-sector), translational related sector (R-sector) and growth rate dependent non-translational enzymes (P-sector, e.x., catabolic proteins), which added up to be 1—

$$\phi_R + \phi_P = 1 - \phi_Q = \phi_{max} \tag{1.4}$$

Since the protein concentration is found to be constant across various conditions (1, 2), Eq. (1.4) form an actual constrain for the R-sector and P-sector proteins. The increase concentration of R protein means the drop of concentration in P protein and vice versa. Proteins in R sector contribute to protein synthesis flux via ($\kappa_t \cdot (\phi_R - \phi_{R,0})$) according to Eq. (1.3) and proteins in P sector contribute to nutrient uptake flux via ($\kappa_n \cdot \phi_P$) where κ_n represents nutrient conditions. At balanced steady-state growth, the protein synthesis flux should balance with the nutrient uptake flux—

$$\kappa_t \cdot \left(\phi_R - \phi_{R,0}\right) = \kappa_n \cdot \phi_P = \lambda. \tag{1.5}$$

Using Eq. (1.4) and (1.5), we can then describe the set of linear relationships mentioned above. Intuitively, when E.coli grow in relative good nutrient condition (larger κ_n), in order to reach the flux balance under the constrain Eq. (1.5), cells need to increase ϕ_R and decrease ϕ_P comparing to the condition with poorer nutrient (smaller κ_n) since κ_t is constant (Comparing Fig. 1.3a to Fig. 1.3b). As a result, the growth rate increase. When chloramphenicol is added to repress translational capacity (decrease κ_t), cells also need to increase ϕ_R and decrease ϕ_P to keep the flux balance (Comparing Fig. 1.3c to Fig. 1.3a). But in this case, growth rate drops since $\lambda = \kappa_n \cdot \phi_P$ and κ_n is constant (same nutrient). For the over-expression (OE) case, useless protein took away a portion on the pie chart (ϕ_{max}) that R and P proteins can access (Comparing Fig. 1.3d to Fig. 1.3a). With constant κ_t and κ_n, growth rate decrease proportionally with the decrease of ϕ_P or ϕ_R, or with the increase of OE proteins.

This top-down flux balance model is very powerful in explaining proteomic allocation under various constrains. Starting with empirical results and without invoking the details of molecular interactions in E.coli, it facilitate quantitative understanding on the connection between bacterial growth and gene expression under various conditions. As a result, this method lays a great foundation for the proteomic studies that came out in the last ten years (2, 7, 8), including my dissertation also.

Meanwhile, the limitation of the steady-state study is obvious. It is hard to give further insights on how E.coli coordinate their proteome to various growth conditions. In other words, the regulation strategy and mechanism of adapting to different conditions was not easily revealed in the steady-state studies. In my dissertation study, instead of steady-state, I focused on growth adaptation and proteomic reallocation process of E.coli while transfer from one condition to the other. In Chapter 2&3, we looked into cells' responses to amino acid(s) downshift. In Chapter 4, we looked into the cellular adaptation while adding Cm. In Chapter 5, we dig further into bacterial regulation mechanism by studying carbon diauxic shift.

1.2 Bacterial response to environmental change

Due to the significance of quick adaptation to changing environment, bacterial behavior during various nutrient shifts, especially carbon shift, has been extensively studied (8–10). In the recent study on growth transition between co-utilized carbon sources, Erickson et al (8) extended the proteomic allocation model introduced above to kinetic regime and established a Flux-Controlled Regulation (FCR) model capturing the gene expression and biomass accumulation during carbon shift. The key feature of their work was that they used dimension reduction to bypass the molecular details of gene regulations and had protein synthesis under the regulation of translational activity instead.

To be detailed, in FCR model, all the internal metabolites, including ketoacids and amino acids, are coarse-grained into a single 'precursor' pool K. This precursor pool controls the synthesis of ribosomal proteins and catabolic proteins through cAMP-crp (11, 12), ppGpp (13–15) signaling system. Mathematically it was written as

$$\frac{dM_{Rb}}{dt} = \chi_{Rb}(K(m_1(t), m_2(t), \dots, m_i(t)))\frac{dM_{pro}}{dt}, \tag{1.6}$$

$$\frac{dM_C}{dt} = \chi_C(K(m_1(t), m_2(t), \dots, m_i(t)))\frac{dM_{pro}}{dt}. \tag{1.7}$$

Here χ_{Rb} and χ_C is called regulation function, representing the fraction of protein synthesis resources allocated to ribosomal and catabolic proteins. Their kinetics control the kinetics of transition process. When cells reach new steady-state, they should be the same as corresponding protein mass fractions due to ignorable degradation rate of proteins. At coarse-grained time scale, the influx of the precursor pool (nutrient uptake flux) was assumed to balance with the outflux of the precursor pool (protein synthesis flux) considering fast reactions between metabolites. Under the flux-balance constraint, the translational activity, defined as $\sigma = \frac{dM_{pro}}{dt}/M_{Rb}$, was set by the precursor pool. Using the idea of dimension reduction, σ is then used to set the regulation functions—

$$\chi_{Rb}(K(m_1(t), m_2(t), \dots, m_i(t))) = \hat{\chi}_{Rb}(\sigma) \tag{1.8}$$

$$\chi_C(K(m_1(t), m_2(t), \dots, m_i(t))) = \hat{\chi}_C(\sigma). \tag{1.9}$$

The form of $\hat{\chi}_{Rb}(\sigma)$ and $\hat{\chi}_C(\sigma)$ were determined by their steady-state relationships under carbon limitations, which is derived from the phenomenological growth law introduced above. Now by solving the kinetics of σ, the kinetics of proteomic allocation and growth adaptation during carbon shift can then be predicted without involving molecular details on gene regulations.

Except the work introduced above, many other studies on different type of carbon transition (for example, shift between hierarchical carbon substrates or carbon runout) have together lead to novel understanding on carbon uptake system, central carbon pathway and the role of central metabolites. However, most of the quantitative studies mainly focused on the bacterial growth in minimal conditions where only necessary carbon and nitrogen source are supplied while E.coli's natural habitat (e.x., human guts) commonly contains amino acids (AAs) and other nutrient supplies, at least transiently before they are exhausted. The condition with a multiple of amino acids (AAs) are called 'rich condition'. Due to the complexity of the composition of rich media, quantitative proteomic study on growth on and growth transition from rich medium is lacking. Here in Chapter 2 and 3, we quantitatively studied the growth transitions of amino acid(s) downshift. In Chapter 2, we looked into single amino acid downshift from rich media–met downshift, which caused 1hr lag. We identified MetE to be the major bottleneck and established a coarse-grained model based on the flux balance assumption in FCR model. It successfully capturing the relationship between the pre-shift MetE reserve and the lag time without requiring any measurement of cellular response specific to methionine shortage, even at the steady state level. In Chapter 3, we systematically studied proteomic allocation strategy in various rich conditions and during all amino acids downshift. We found that the main strategy for E.coli to reach higher growth rate in rich medium is to allocate resource from amino acid biosynthesis (AAB) to protein translation, which in turn caused 2-hr lag when all amino acids were washed away. The proteomic study during all AAs downshift revealed a striking linear relationship between the onset time of the enzymes across AAB pathway and their fractional reserves in pre-shift condition. The strategy of sequentially prioritizing the expression of various bottlenecks was not observed in the previous study of carbon downshift (8). Expanding the model of single AA depletion in Chapter 2 and incorporating the on-off allocation strategy introduced by Pavlov et al (16) to implement the sequential recovery across AAB pathways, we were able to quantitatively capture the kinetics of proteome reallocation process with only steady-state measurements.

With the success in quantitatively characterizing bacterial responses to nutrient changes by the top-down method, we further looked into other environmental changes. In Chapter 4, we studied the bacterial response to sub-lethal Chloramphenicol (Cm) treatment. Chloramphenicol, as introduced before, is an antibiotic inhibiting translational process. It's used in treating a number of bacterial infections. Considering the significant role of antibiotics and increased threats of antibiotic resistance due to wild usage of antibiotics, it's very important to quantitatively understanding the dynamic responses of bacteria to antibiotics like Cm. In our study, we quantified the dynamics of biomass accumulation and typical gene expressions while adding sub-lethal Cm in various nutrient conditions. We extended the FCR method to quantitatively capture the transition process and first used translational elongation rate as the central dynamic variable governing gene expressions. Compared to σ (translational activity per ribosome), translational elongation rate has a clear physiological meaning, leading us one step closer to understand the mechanism of growth rate control, which was the focus of Chapter 5.

Inspired by the success of dimension reduction methods in capturing various transition processes, we suspected that the translational rate could be actually used by E.coli as an control signal in its regulatory scheme. By looking into the glucose to glycerol shift in Chapter 5, we revealed that Guanosine tetraphosphate (ppGpp, a key signaling molecule regulating hundreds of genes) responds inversely to the translational elongation rate. Furthermore, the same relationship was also found in steady state. Combined with the downstream regulation function of ppGpp, we were able to close a circuit of growth rate control mechanism and reproduced the ribosomal growth law (Eq. 1.3) as we discussed above.

1.3 Figures

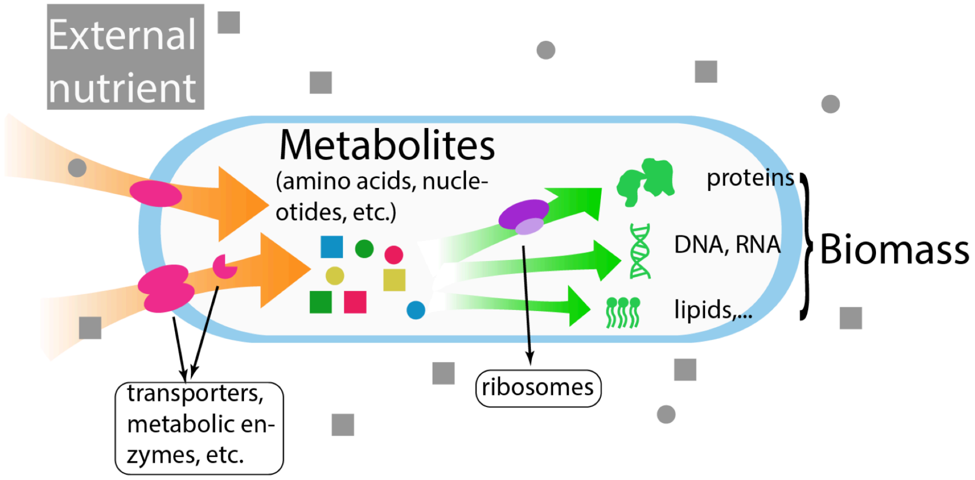


Figure 1.1. Cartoon illustration of the bacterial growing process. The blue rod-shape structure represents the membrane of bacteria, dividing intra-cellular environment from external environment. Bacteria growth can be divided into two parts. In the first part, the Bacteria take in external nutrients (gray squares, gray circles) and convert it into internal metabolites (such as amino acids, nucleotides, etc.) by transporters and many metabolic enzymes (e.x., amino acid biosynthesis enzymes). In the second part, bacteria uses the pool of metabolites to synthesize macro-molecules like protein, DNA, lipids, etc. to gain biomass by various enzymes/machines such as ribosomes.

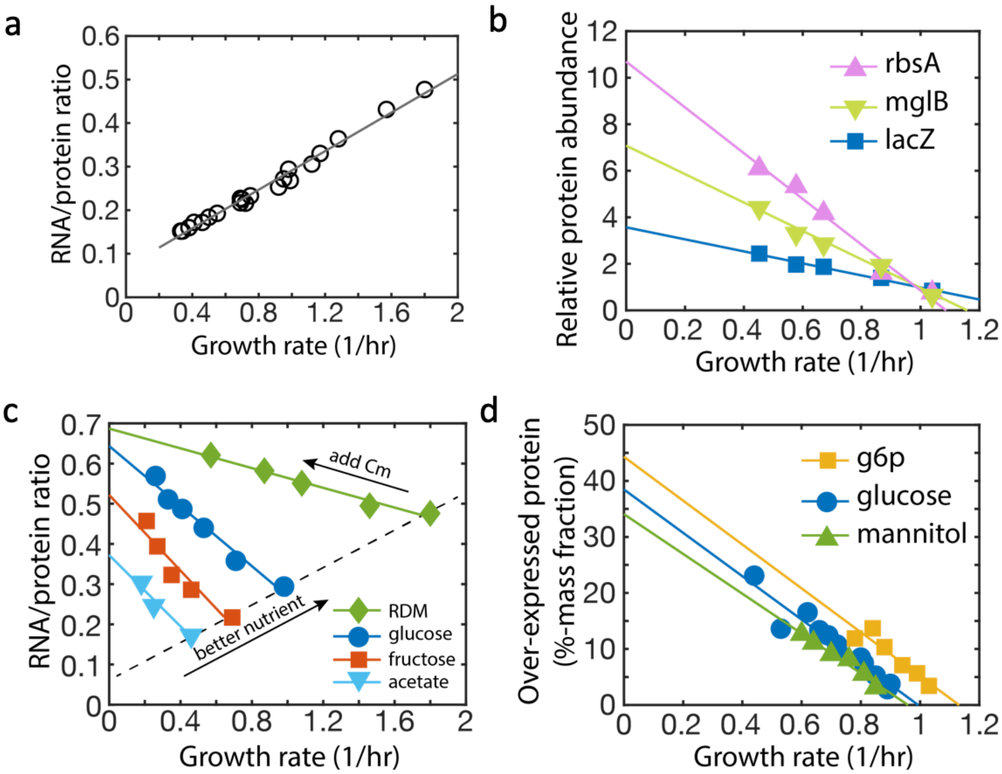


Figure 1.2. Steady-state relationship between gene expression and growth rate under various conditions. (a) Ribosomal contents (characterized by RNA/protein ratio) linearly increase with growth rates while changing nutrient conditions. **(b)** The relative expression levels of some cAMP-CRP regulated catabolic genes (rbsA, mglB, lacZ) increase while growth rates decrease under carbon limitation. Data came from (7). **(c)** when adding chloramphenicol (Cm), ribosomal content increase with the decrease of growth rate. Different symbols represent different nutrient conditions (RDM: rich defined medium from Teknova; the rest indicates the carbon sources in minimal media). Data in **(a)** and **(c)** come from (17). **(d)** Growth rates linearly decrease with the abundance of over-expressed useless protein (LacZ). Different symbols represent minimal media with different carbon sources. Data come from (18).

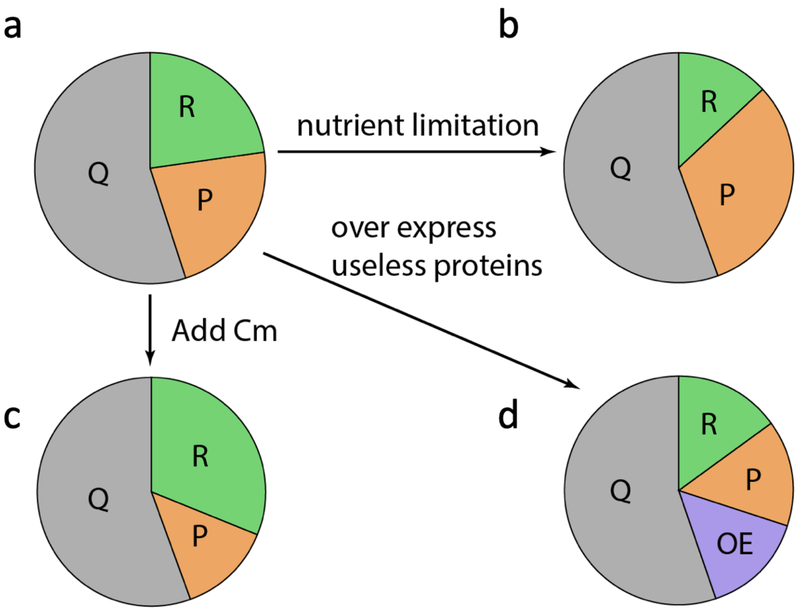


Figure 1.3. Illustration of proteome allocation under various condition based on the coarse-grained model in Scotts (3)**. (a)** cellular proteome is coarse-grained in three-parts: Q—growth rate independent part (marked by gray), R—translational related part (marked by green), P—growth rate dependent and non-translational enzymes, such as catabolic proteins (marked by orange). The area in the pie chart represents the mass fraction of corresponding protein groups. Since total protein concentration is relative constant in various conditions, the area in the pie chart also represent protein concentrations. Compared to the cells grow in poor nutrient condition **(b)**, cells grow in relative better nutrient condition **(a)** allocate more resource from P sector to R sector under flux balance constrain. **(c)** Under chloramphenicol (Cm) treatment, cell increase R sector and lead to drop in P sector. **(d)** When over expressing useless proteins, whose mass fraction is labeled as OE (marked by purple), R and P sector are reduced as results.

Chapter 2

MetE reserve shortens the lag time of met downshift

2.1 Abstract

Bacteria are known to experience nutrient fluctuation frequently in their nature habitat. While many extensively studies on how E.coli adapts to carbon changes have been done, quantitative understanding on how E.coli responds to amino acid depletion is lacking, despite amino acids are one of the most common nutrient source for E.coli in their natural habitat. In this work, we focus on met downshift from rich medium which lead to half the lag of all AA depletions. Consistent with the known inefficiency of MetE enzyme in met biosynthesis pathway (19, 20), we found that MetE is the major bottleneck in met downshift and the reserves of MetE in pre-shift condition strongly shortens the lag of met downshift. Adapting the top-down approach used in characterize carbon shift previous, we were able to quantitatively capture the dependence of lag time on pre-shift MetE reserve without invoking any fitting parameters. It gave an example on quantitatively characterizing transitions under rich conditions.

2.2 Introduction

Bacteria are known to grow in a broad range of conditions, and the rapidity of their adaptation to environmental fluctuation is an important component of their overall fitness. For enteric bacteria, a common nutrient environment both in the wild and in laboratory studies is the "rich medium" containing multiple amino acids (21). However, due to their more complex nature, quantitative understanding of downshift from rich condition is lacking. In this study, we focus on methionine depletion from rich media which we found causing half of the lag of all AAs downshift. Our data identified MetE to be the major bottleneck of met downshift. Based on that, We further established a quantitative model to characterize the lag dependence on MetE pre-shift expression level (reserves).

2.3 Results

2.3.1 MetE is the major bottleneck in methionine downshift

In order to characterize the kinetic of met downshift resembling the natural met depletion, we let the cells grew in the media with all AA supplies (18 AAs, no cysteine and tyrosine; see **Methods**) at the concentration the same as that in rich defined media (RDM, (22)) except methionine. Met concentration was lowered so that met was depleted in the middle of exponential growth (see **Methods**). Glycerol was always supplied as carbon source to ensure that cell only experience met shortage during the shift. Fig. 2.1a and 2.1b shows the growth curve during met downshift. Before the met depletion, cells grew exponentially with the growth rate the same as that measured in higher met concentration (same as RDM, the green bar in Fig. 2.1b). When met was depleted, cells experience an immediate growth halt as shown in Fig. 2.1a and 2.1b. The shift time (t=0) is determined by extrapolating the pre-shift growth curve to the OD_{600} where the growth pauses (the star in Fig. 2.1a). After about an hour, the growth started picking up and eventually reached the final steady-state. Here we defined 'lag time' as the lost time due to the adaptation to new condition as shown in Fig. 2.1a. The lag time of met downshift is about 1-hr, which is about half the lag time of all AA depletion (Fig. 2.2d)

Qualitatively, in response to met depletion, cells have to increase *de novo* met biosynthesis to make up for the missing met influx. However, in the pre-shift condition, the expression of met biosynthesis enzymes is largely repressed; see Fig. 2.1c which compares the expression levels of met biosynthesis enzymes measured in pre-shift rich media and minimal media. Therefore, in order to recover from met shortage, cells need to reallocate protein synthesis flux towards the synthesis of enzymes of the met pathway, mainly from ribosomal synthesis flux, which lead to the growth lag observed in Fig. 2.1a.

Among all the enzymes needed for met biosynthesis, the one catalyzing the last step of the met pathway, MetE, displayed the largest fold change in mass fraction between rich media and minimal media, by more than 10-fold comparing to other enzymes according to Fig. 2.1c. Also, the absolute abundance of MetE in minimal media is about 100 times larger than other enzymes on met pathway (Fig. S1). It's thus likely that MetE is the major bottleneck during met downshift. To validate that, we introduced a strain (HE588) with titratable expression of MetE in addition to the native metE gene (Fig. 2.2a, Methods). We pre-expressed MetE by supplying the inducer cTc in the pre-shift medium containing all AAs (same AA conc. as in Fig. 2.1a), and monitored the effect of pre-expression on the lag time upon met downshift (Methods). Figure 2.2b and 2.2c show the resulting growth curves and instantaneous growth rates at several cTc levels. Without cTc, the growth curve of HE588 (circles) was the same as that of the wild type strain (blue diamonds, see Fig. S2 for repeatability), indicating low background expression of MetE in HE588. For increasing cTc levels (squares and triangles), the lag time of met downshift clearly shortened, reaching < 20 mins at 50ng/ml cTc (Fig. 2.2d). These results thus validate our hypothesis that MetE is the major bottleneck for met downshift. In other words, met biosynthesis flux is limited by MetE during the shift and cells need to direct protein synthesis flux from ribosomal synthesis to MetE synthesis to recover growth.

2.3.2 Model of methionine recovery kinetics in met downshift

Based on this qualitative picture, we adapted a mathematical modeling approach introduced recently by Erickson et al. in the context of carbon diauxie shift (23) to describe the kinetics of growth recovery during met downshift. This approach allows us to quantitatively characterize the proteomic reallocation process without getting bogged down by unknown details of the regulations of the met enzymes; see illustration in Fig. 2.3a and elaboration below.

We focused on the flux of protein synthesis $J_R(t)$ (defined by Eq. (2.1) in Fig. 2.3a), measured in the number of AA polymerized per time, with $M_p(t)$ being the total protein mass (in unit of AA residues) in a fixed culture volume. After growth arrest abruptly set in due to met shortage (at $t = 0$), we let $J_R(t)$ be limited by the flux of methionine biosynthesis, $j_{met}(t)$, according to Eq. (2.2) in Fig. 2.3a where f_{met} represents the met frequency in the proteome. Since MetE is the major bottleneck in met synthesis, we can express $j_{met}(t)$ as the product of the mass of MetE ($M_{MetE}(t)$) and its specific catalytic rate k_{MetE}. We note that while in principle catalytic rates can be time dependent due to, e.g., allosteric regulation when the end-product is in surplus, we expect k_{MetE} to be at its maximum value, k^*_{MetE} through the shift as MetE is the growth bottleneck. Thus, $j_{met}(t)$ is written as Eq. (2.3) in Fig. 2.3a.

Combining Eqs. (2.2) and (2.3), we obtain

$$J_R(t) = k^*_{MetE} M_{MetE}(t)/f_{met} \tag{2.9}$$

as the dynamic relationship between protein synthesis flux and MetE abundance, throughout the duration of the shift and including the post-shift steady-state. In terms of "intensive variables", the instantaneous growth rate $\lambda(t)$ and the proteome fraction $\phi(t)$ defined as

$$\phi_x(t) \equiv \frac{M_x(t)}{M_p(t)}, \tag{2.10}$$

$$\lambda(t) \equiv \frac{d}{dt}\ln M_p(t) = \frac{J_R(t)}{M_p(t)}, \tag{2.11}$$

where Eq. (2.1) was used in the last equality. Eq. (2.9) can then be rewritten as

$$\lambda(t) = \frac{k_{MetE}^{*}\phi_{MetE}(t)}{f_{met}}. \tag{2.12}$$

Applying Eq. (2.12) to the post-shift steady state (i.e., exponential growth in minimal media) where $\lambda(t) \equiv \lambda^*$ is the final growth rate and $\phi_{MetE}(t) = \phi_{MetE}^{*}$ is the final abundance of MetE, these relations predict a linear relation between ϕ_{MetE}^{*} and λ^*, with slope given by k_{MetE}^{*}/f_{met}. Fig. S1 shows that indeed the abundance of MetE in minimal media, ϕ_{MetE}^{*}, is linearly proportional to λ^*, allowing us to fix the value of k_{MetE}^{*}/f_{met}; see **Appendix 2.9.1**.

Next, the rate of MetE synthesis and ribosomal proteins synthesis can be written as Eqs. (2.4) and (2.5) in Fig. 2.3a, based on the definitions of $\chi_{MetE}(t)$ and $\chi_{Rb}(t)$, which are the fraction of protein synthesis flux directed to synthesizing MetE and ribosomal proteins at time t, respectively. These functions are dictated by gene regulatory process. Molecularly, MetE expression is down-regulated by the met pool via the transcriptional activator MetR and repressor MetJ (24). Ribosome synthesis is regulated by ppGpp, which can sense single AA shortage (here, met) (14, 25–27). However, quantitative details of both regulatory processes are lacking.

To overcome the lack of regulatory details, we adapted the coarse-grained scheme of Flux-Controlled regulation (FCR) previously introduced (23). Translational activity of the ribosomes, $\sigma(t)$, defined by Eq. (2.6) in Fig. 3.2a where ϕ_{Rb}, the mass fraction of the ribosomal proteins, is taken to be the signal driving the regulation functions. For the carbon diauxie shifts studied in (23), FCR scheme was justified because the allocation of synthesis of ribosomal and carbon catabolic proteins are both regulated by the precursor pool (amino acids and keto-acids, which are in equilibrium with each other and activates ppGpp signaling system (14, 25–27) and cAMP-CRP signaling system (11, 12), respectively); and the same precursor pool sets $\sigma(t)$ at coarse-grained time scale. In met downshift, since the protein synthesis flux $J_R(t)$ is limited by $j_{met}(t)$, $\sigma(t)$ serves to represent the effect of the met pool, which controls the expression of ribosomal proteins by affecting ppGpp signaling and the expression of MetE through MetR/MetJ as mentioned above. Therefore, we let $\chi_{MetE}(t)$ and $\chi_{Rb}(t)$ be set by $\sigma(t)$, through regulation functions $\hat{\chi}_{MetE}(\sigma(t))$ and $\hat{\chi}_{Rb}(\sigma(t))$; see Eqs. (2.7) and (2.8) in Fig. 2.3a.

The form of these two regulation functions can in principle be obtained through their steady-state relationship $\chi^*_{Rb}(\sigma^*)$ and $\chi^*_{MetE}(\sigma^*)$ under met limitation assuming quasi-steady state during shift (23). Since in steady state the abundance of a protein x as measured in proteome fraction is equal to the fraction of flux synthesizing the protein, i.e., $\phi^*_x = \chi^*_x$, the form of the regulation functions can thus be obtained through the steady-state relationship between ϕ^*_x and σ^*, with the latter determined by steady-state growth rate λ^* and ribosomal protein abundance ϕ^*_{Rb} as shown in Eq. (2.8). Therefore from the empirical steady-state relationship between ϕ^*_{Rb} and λ^* under general nutrient limitation (met limitation is a special case of nutrient limitation), we can obtain the relationship between ϕ^*_{Rb} (χ^*_{Rb}) and σ^*, and hence $\hat{\chi}_{Rb}(\sigma) = \chi^*_{Rb}(\sigma^*)$ as detailed Fig. S2 and shown as the grey line in Fig. 2.3b.

The process of obtaining the form of $\hat{\chi}_{MetE}(\sigma)$ is not as straight forward. Following the idea used to construct $\hat{\chi}_{Rb}(\sigma)$ (Fig. S2), ideally we can construct $\hat{\chi}_{MetE}(\sigma)$ based on the relationship between growth rate λ^* and ϕ^*_{MetE} under steady-state met limitations since growth rate is directly linked with σ from above ($\lambda^* = \sigma^* \chi^*_{Rb}(\sigma^*)$). However, it is difficult to maintain a steady met limitation experimentally. And even if this could be done, this approach would only provide data over a narrow growth-rate range since the complete exclusion of met from the rich media only caused 20% growth reduction. Estimate of $\hat{\chi}_{MetE}(\sigma)$ obtained from data over a narrow range would not be very reliable when extended to the kinetic region where much lower growth rate is reached.

To overcome this problem, we exploited an idea introduced by Li et al (20) where the effect of growth reduction due to protein (mainly MetE) over-expression (3) was used to explain the change in steady-state growth rate λ^* and ϕ^*_{MetE} while growing in the media without met (as compared to the media with met). Fig. S3 shows that λ^* and ϕ^*_{MetE} in 'all-but-met' condition can be successfully captured by over-expression (OE) relationship. In that case, the question of obtaining λ^* and ϕ^*_{MetE} relationship under met limitation is converted to the question of obtaining λ^* and ϕ^*_{MetE} relationship under protein over-expression, which has been established before. Also, since the cost of protein over-expression applies across a wide growth range (3), we can extend the OE relationship between ϕ^*_{MetE} and λ^* (thus χ^*_{MetE} and σ^*) to the very slow growth rate region encountered during met downshift. The final form of $\hat{\chi}_{MetE}(\sigma)$ is shown as the orange line in Fig. 2.3b. The resulting forms of the regulatory functions $\hat{\chi}_{Rb}(\sigma)$ and $\hat{\chi}_{MetE}(\sigma)$ match our intuitive expectations that MetE is up-regulated and ribosomal proteins are down-regulated during met starvation when the translational activity σ drops.

Combining all the equations in **Fig. 2.3a**, we obtained the key differential equation describing the kinetics of σ (**Appendix 2.9.3**):

$$\frac{d\sigma(t)}{dt} = \sigma(t)\left(\frac{k^*_{MetE}}{f_{met}}\hat{\chi}_{MetE}\big(\sigma(t)\big) - \sigma(t)\hat{\chi}_{Rb}\big(\sigma(t)\big)\right), \tag{2.13}$$

which can be solved analytically after substituting the regulation functions of $\hat{\chi}_{MetE}(\sigma)$ and $\hat{\chi}_{Rb}(\sigma)$ shown in Fig. 3.2b (see **Appendix 2.9.3**). We can further solve the time course of ribosomal protein abundance $M_{Rb}(t)$ by Eq. (2.5) and (2.6), protein synthesis flux $J_R(t)$ by Eq. (2.6), MetE abundance $M_{MetE}(t)$ and met synthesis flux $j_{met}(t)$ by Eq. (2.2) and (2.3), biomass $M_p(t)$ by Eq. (2.1). With Eq. (2.10) and (2.11), we can then obtain the expression of instantaneous growth rate $\lambda(t)$ and the mass fractions of ribosomal proteins and MetE (**Appendix 2.9.3**). Their expressions depends on the pre- and post-shift steady-state growth rates and the abundance of the pre-expressed MetE. Using the pre- and post-shift steady-state growth rates obtained from met downshift experiments of WT strain (**Table 2.1**), we can predict the dependence of lag time on pre-expressed MetE abundance (**Fig. 2.3c**). (The effect of MetE pre-expression on pre-shift growth rate doesn't affect the prediction on lag time; see **Fig. 2.8**.) To directly test the prediction, we measured the pre-expressed ϕ_{MetE} of the metE titratable strain in **Fig. 2.2a** by mass spectrometry (**Methods**) under the same cTc concentration as used in **Fig. 2.2**. Plotting measured lag times and ϕ_{MetE} (symbols in **Fig. 2.3c**), we see excellent agreements with model prediction which contained no fitting parameters. Using the pre-expressed MetE corresponding to the measured lag times (**Table 2.2**), the model can then quantitatively explain the growth curve, instantaneous growth rate and other quantities as mentioned above through the course of met downshifts (solid lines in **Fig. 2.2b, 2.2c, Fig. 2.9**) without invoking any other fitting parameters.

2.4 Discussion

In this work, we studied methionine downshift from rich media which exhibited 1-hr lag. MetE are found to be the major bottleneck in met downshift since pre-expression of MetE dramatically shortened the lag time (Fig. 2.2). Based on that, we developed a coarse-grained model to quantitatively capture the enzyme and growth recovery kinetics during methionine downshifts. Different from the bottom-up approach starting from molecular interactions, our approach required only limited information on steady state proteome allocation, allowing us to overcome the challenge of obtaining the molecular interaction parameters.

Qualitatively, MetE synthesis is down-regulated by the met pool via the transcriptional regulators MetR and MetJ (24, 28). To directly model the regulation of MetE synthesis, we would need all the binding constants connecting the intracellular concentrations of met (and other metabolites of the met synthesis pathway) to MetR and MetJ, their effects on MetE expression and on the expressions of these regulators themselves. Besides MetE, the expression of many other proteins, particularly the translational proteins are also changed during met downshift, and direct modeling would require quantitative information regarding how tRNA charging and the ppGpp signaling pathway (14, 25–27) respond to the met shortage. In our approach, instead of describing the details of these molecular interactions, we let the translational capacity σ, set by the condition of flux balance, to represent the met pool and the kinetic signal governing the allocations of protein synthesis flux to MetE and to the translational machinery (χ_{MetE}, χ_{Rb}). Steady-state relationships between growth rates and proteomic allocations are used to obtain the forms of the regulatory functions, which can subsequently be extended to the kinetic regime through the time dependence of $\sigma(t)$. In this way, we can quantitatively predict the growth transition kinetics during met downshift using only steady-state information.

While the flux-based approach was originally developed in the context of carbon diauxie shift (23), the fact that we can modify it to quantitatively capture the very different AA downshifts reflect the generality of the approach. First, it shows that our approach is not limited to the simpler and extensively studied growth on carbon substrates. Second, the key metabolites involved in met downshift and the associated signaling and regulatory pathways are completely different from the ones involved in carbon shifts (keto-acids and cAMP (11, 12, 23)). Despite these differences, proteome allocation can be described by the same dynamic control variable—the translational capacity $\sigma(t)$. Thus, our approach is not dependent on details of the signaling systems involved when describing growth transition kinetics, leading us to expect that it can be applied more generally to other growth shifts without needing detailed knowledge of molecular signaling.

From both MetE pre-expression experiments (Fig. 2.2) and model predictions (Fig. 2.8), the cost of shortening the lag to 20mins is minor. It's corresponding to about 0.4% MetE over-expression in pre-shift condition and less than 5% drop in pre-shift growth rate. Considering that the natural habitat of E.coli (such as gut) usually have unstable nutrient supply, it seems more beneficial to the cells to have small amount of MetE (~0.4%) expressed all the time. However, it's not the strategy E.coli takes from our observation and the reason of keeping little MetE while met is supplied may be beyond the scope of protein cost. Considering met is the first amino acid needed for translation, it could be meaningful for cells to ensure lower met availability than other AAs in any conditions in order to avoid translational abortion. In our experiments, we only deplete methionine, while in nature, it's likely that met depletion happen together with other AAs' depletions (e.x., due to flux). In that case, it could be beneficial for cells keep little MetE when not needed. Further studies are needed to test that hypothesis, which is beyond the scope of this study.

2.5 Materials and Methods

2.5.1 Strain construction

The strain used in this study are either wild type E.coli K-12 NCM3722 strain (29, 30) or its derivative HE588 (metE titratable strain).

2.5.1.1 Construction of HE588

The metE structure gene was PCR amplified from NCM3722 genomic DNA using oligos Ptet-metE-F and metE-R (Table 2.3). The kanamycin resistance gene (km) plus its promoter (without flanked FRT scars) was amplified from pKD13 (PMID10829079) using oligos metKm-F ans Z-km-R. Note that the 33 nucleotides at the 3' end of the "metE" fragment are the same as the first 33 nucleotides at the 5' end of the km fragment. After gel purification, these two fragments were fused together by standard fusion PCR using both amplified products in equal ratios as templates and Ptet-metE-F and Z-km-R as primers. Using the Lamada-Red method as described in Datsenko and Wanner (2000. PMID10829079), the resultant fused product, "metE:km" (both genes transcribed in the same direction), was electroporated into the cells of strain EQ37 (PMID20064380; expressing the λ-Red recombinase) to replace the sequences from the +1 nucleotide to the +2389th nucleotide relative to the translational start point of lacZ. Note strain EQ37 carries Ptet driving lacZ at the lacZ locus and the lacY gene deletion. After electroporation, several Km resistant colonies were confirmed by PCR and DNA sequencing. The "rrnBT:Ptet-metE:km" cassette at the lacZ locus in the resultant strain was transferred to strain NCM3722-1R (carrying Ptet-tetR at the ycaD locus, (9)) by P1 transduction, yielding the metE titratable strain HE588. In this resultant strain, the metE expression is repressed by the auto-regulated TetR and the repression is released by a TetR inducer such as chlorotetracycline (cl-Tc). The more inducer added the greater the level of metE expression.

2.5.2 Growth medium

All the growth media used in this study were based on MOPS buffered media described in Cayley et al (31) with 50mM NaCl, 1.32mM KH2PO4, 0.4% glycerol and 10mM NH_4Cl.

The pre-shift rich media used in this work are based on the above media and with additions of 18 AAs (Cysteine is excluded because it oxidizes to cystine in aerobic growth. Tyrosine is excluded because of its small solubility). For simplicity, we refer it as 'all AA' condition. If not specified, amino acid concentrations are the same as those listed in the rich defined medium (22). In met downshift experiment, the met conc. in the batch is lowered to 10uM so that met can run out around OD_{600}=0.14 to trigger the shift.

2.5.3 Cell growth and medium shift

All the medium shift experiments were performed in a 37℃ water bath shaker shaking at 240 rpm. Shifts always start after cells reaching steady-state in pre-shift medium. To grow cells to exponential phase, we followed a standard cell growth round: seed culture, pre-culture and experimental culture. For seed culture, cells were inoculated to LB broth from fresh LB plate and grew for about 7hrs. Then cells were washed with corresponding Mops medium once and diluted to the same Mops medium (pre-culture) with starting OD_{600} around 10^-3. If chlortetracycline (cTc) is needed in this experiments, it was added to pre-culture too. After overnight pre-culture, cells were diluted to the same Mops medium (experimental culture) with starting OD_{600} below 0.01. For exponential growth measurements, we let cells grow to OD_{600} 0.5. More than 5 OD_{600} points were taken within the range of OD_{600} 0.05~0.5 for growth rate calculation.

In met downshift experiment, 200uM of met was used in the pre-culture and 10uM of met was used in experimental culture. While diluting cells from pre-culture to experimental culture, we first washed cells by centrifuging and resuspending cells with the experimental culture media.

The staring OD_{600} was also below 0.01. With 10uM met, the shift happens around $OD_{600} = 0.14$. Time 0 was determined by extrapolating the pre-shift growth curve to the OD_{600} of the first point after shift. Since for met downshift, growth ceases abruptly set in when met run out and OD_{600} was measured every 3 mins near met depletion, the uncertainty of time 0 is within 2 mins.

2.5.4 Growth rate calculation

For steady-state exponent growth rate, we simply fitted growth curve exponentially. If $R^2 < 0.999$, we didn't take it as reliable results.

To get instantaneous growth rate, we applied Savitzky-Golay filter (32) to N $\{t_i, Ln(OD_{600,i})\}$ data set. Recall that because of medium switch, OD_{600} was not continuously at t=0. Therefore we scaled the data in post-shift region before calculating growth rate to make the first OD_{600} after shift $(t < 5mins)$ equal to the value exponentially extrapolated from pre-shift data.

The window size we used includes 3 data points. First, we fitted the first 3 data points to the 2nd polynomial and substituted t of middle data point (t_2) into fitted equation to get its slope. Then we moved the window to the right and did the same thing for the 2nd to 4th data points to get the slope at t_3. Repeating moving windows and calculating slope, we got slopes at $t_2, t_3, \dots t_{N-1}$. Since we already did Ln operation on OD_{600}, the slopes we got are exponent instantaneous growth rates at $t_2, t_3, \dots t_{N-1}$. Cells are in steady-state at t_1 and t_N according to our experimental setting, so we only plotted data points from t_2 to t_{N-1} for instantaneous growth rate figures.

2.5.5 Mass spectrometry protocol

We used quantitative mass spectrometry to measure the MetE pre-expression levels under various cTc inductions. For each cTc concentration, at least $1OD_{600} \cdot ml$ sample was taken during steady-state growth in all AA condition (18 AAs) by centrifuging at max speed for 2mins. Samples were resuspended by pure water to 1ug/ul protein concentration and mixed with 15N labeled reference (MOPS glycerol minimal condition) with 1:1 mass ratio. Subsequent sample processing, including TCA precipitation, cysteine reduction, alkylation, tryptic digestion, and desalting procedures, was performed as previously described (17). Sciex 5600 TripleTOF were used to generate MS data. Following the data conversion and analysis steps described before (17, 23), we obtained the absolute MetE abundance, which is shown in **Fig. 2.3c**.

2.6 Acknowledgement

We are grateful to present and former members of the Hwa lab for helpful discussions through the course of this work. This research is supported by the NSF through Grant MCB-1818384 and the NIH through Grant R01GM109069.

Chapter 2, is coauthored with Zhongge Zhang, Vadim Patsalo, James R. Williamson, Terence Hwa. The dissertation author was the primary investigator and author of this material.

2.7 Figures

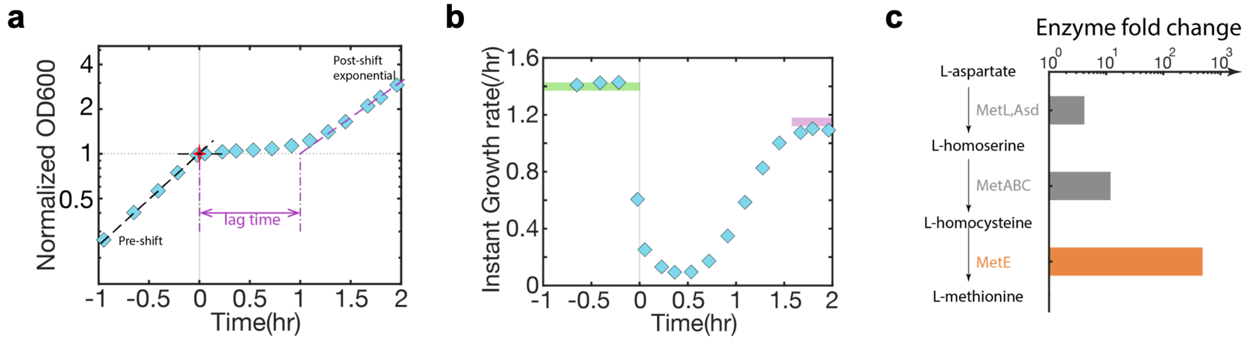


Figure 2.1. Growth kinetic of methionine downshift and methionine synthesis pathway. (a) growth curve and **(b)** instantaneous growth rate measured in methionine (met) downshift (blue diamonds; see **Methods**). The time when the shift starts (t=0) was determined by extrapolating the pre-shift growth curve to the OD_{600} where the growth pauses (the red star in **(a)**). Lag time is defined as the time lost in the transition, calculated by extrapolating exponential post-shift growth curve to the initial OD_{600} when the transition starts (purple lines in **(a)**). The green bar in **(b)** shows the steady-state growth rate (with 95% confidence interval) under the same nutrient condition as the pre-shift but with higher met conc. (0.2mM; same as that in rich defined media). The pink bar in **(b)** shows the steady-state growth rate (with 95% confidence interval) measured under the post-shift nutrient condition (‘all-but-met’). **(c)** Met synthesis pathway and enzymes involved in each step are shown on the left side. The bar plot on the right shows the fold increase in steady-state mass fraction of each enzyme(s) from the pre-shift rich condition (glycerol with 18 AAs; see **Methods**) to the minimal condition without AA (glucose only). The proteomic data used to calculate the fold changes were reported in the rich condition proteomic study in Chapter 3. We can see that the fold change for the met enzymes were all larger than 1, indicating that met enzymes are repressed when met was supplied. Notice that the post-shift steady-state of met downshift has a higher growth rate than the glucose minimal condition. It’s thus expected that cells need even more met enzymes in post-shift condition than in glucose minimal condition in order to support higher growth rate. Among all the met enzymes, MetE shows the largest fold change between rich and minimal condition.

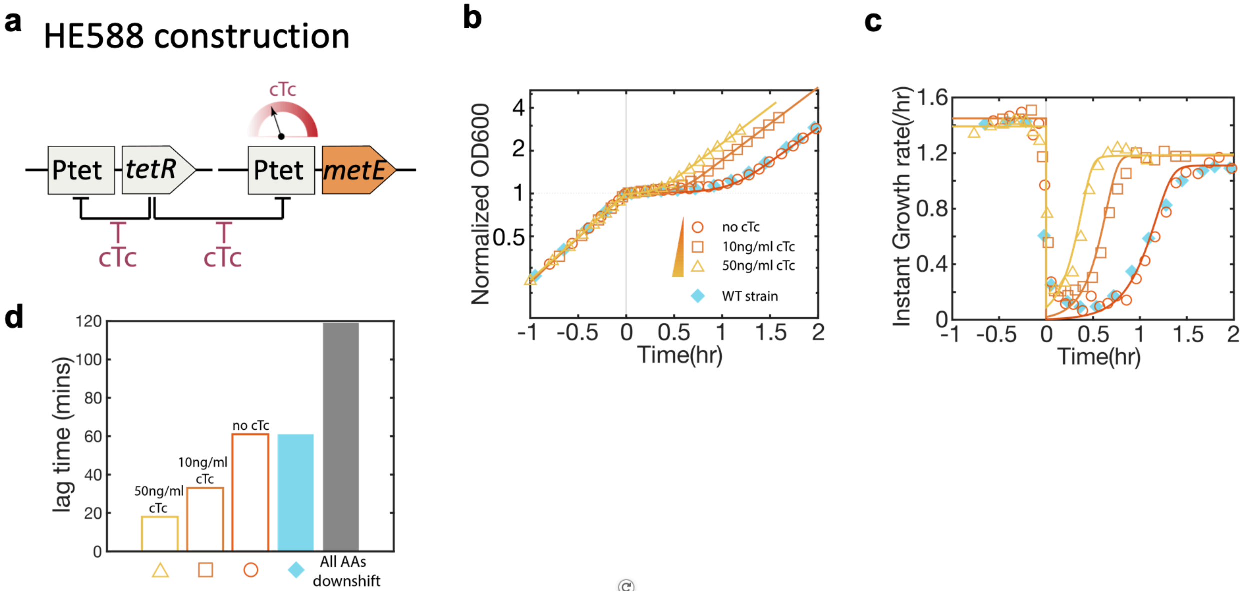
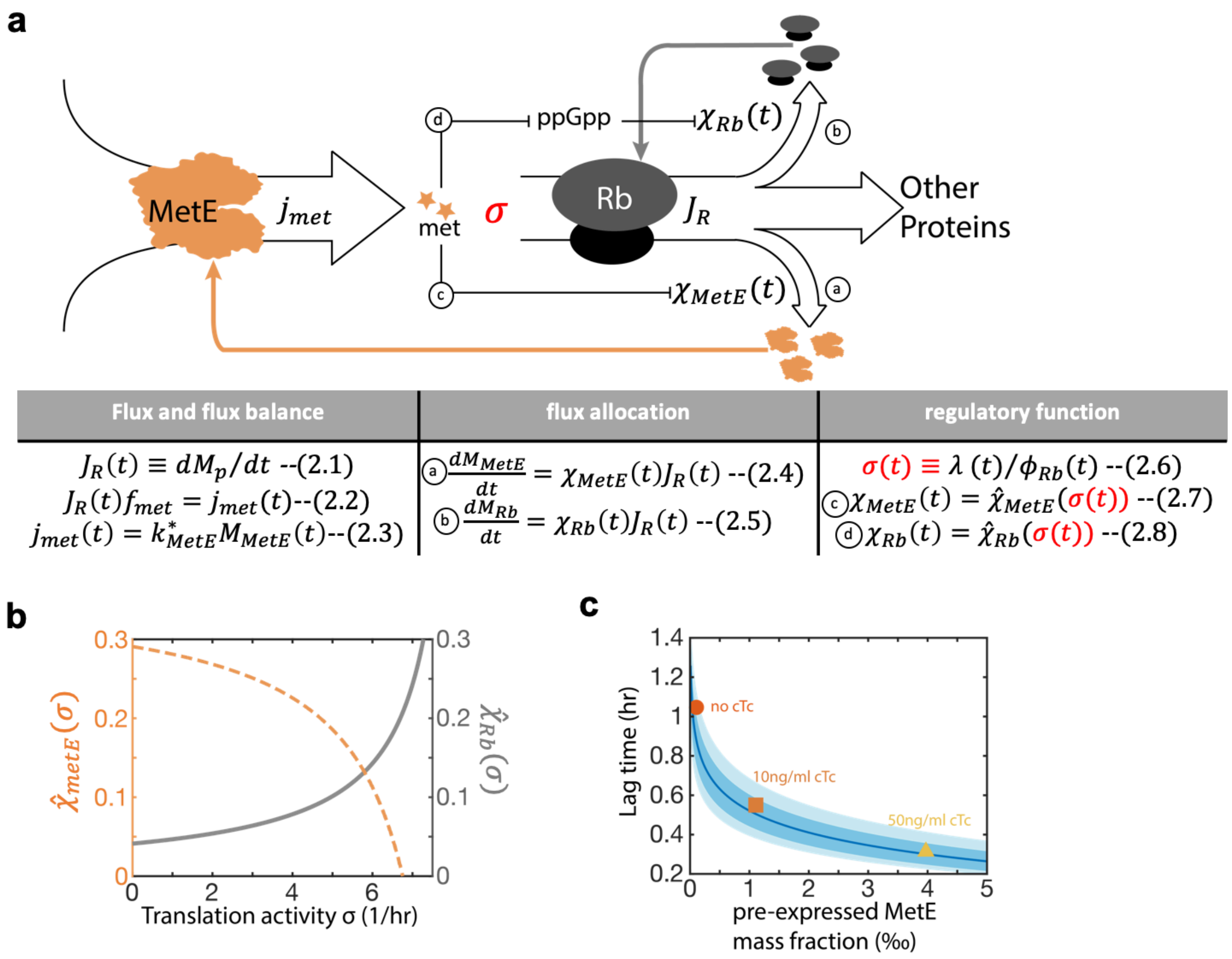


Figure 2.2. Met downshift with various MetE pre-expression. (a) Ptet-metE construction allows titrate MetE expression level with various cTc concentrations. (b) Growth curves and (c) instantaneous growth rates of met downshifts under different cTc conc. (open triangles: 50ng/ml cTc; open squares: 10ng/ml cTc; open circles: no cTc; blue diamond: wild type strain as shown in **Fig. 2.1**). Solid lines in **(b)** and **(c)** are model predictions under the corresponding cTc conc. (same color coding as the symbols). In the model, the pre-expressed MetE mass fractions was determined by the measured lag time according to predicted MetE-lag relationship (**Fig. 2.3b**). **(d)** shows the lag times (calculated according to **Fig. 2.1a**) of the met downshifts shown in **(b)** (same symbol as in **(b)**) and the lag time of all AAs downshift (grey bar, Chapter 3).



Flux and flux balance	flux allocation	regulatory function
$J_R(t) \equiv dM_p/dt$ --(2.1) $J_R(t) f_{met} = j_{met}(t)$--(2.2) $j_{met}(t) = k^*_{MetE} M_{MetE}(t)$--(2.3)	ⓐ $\frac{dM_{MetE}}{dt} = \chi_{MetE}(t) J_R(t)$ --(2.4) ⓑ $\frac{dM_{Rb}}{dt} = \chi_{Rb}(t) J_R(t)$ --(2.5)	$\sigma(t) \equiv \lambda(t)/\phi_{Rb}(t)$ --(2.6) ⓒ $\chi_{MetE}(t) = \hat{\chi}_{MetE}(\sigma(t))$ --(2.7) ⓓ $\chi_{Rb}(t) = \hat{\chi}_{Rb}(\sigma(t))$ --(2.8)

Figure 2.3. Model description of met downshift. (a) shows the schematic of cells' response to methionine shortage in the model. When methionine is depleted, methionine synthesis is required for growth, of which the biosynthesis flux j_{met} is limited by enzyme MetE (Eq. (2.3)). Under flux balance assumption at coarse-grained time scale (Eq. (2.2)), protein synthesis rate J_R (Eq. (2.1), or growth rate) is limited by MetE also as shown in Eq. (2.9) and Eq. (2.12). In other words, with given number of ribosomes (M_{Rb}), translation capacity σ drops as a response to methionine shortage by its definition (Eq. (2.6)). Regulation function χ_{metE} and χ_{Rb}, determining the allocation of J_R to MetE and ribosome synthesis respectively (Eq. (2.4) & Eq. (2.5)), respond to methionine pool via gene regulations. With σ being a reporter of methionine pool, χ_{metE} and χ_{Rb} can be written as a function of σ (Eq. (2.7) & Eq. (2.8)) based on steady-state relationships. The functions of $\hat{\chi}_{Rb}(\sigma)$ and $\hat{\chi}_{MetE}(\sigma)$ are constructed from steady-state relationship between protein abundance and growth rate (Fig. S3 and Fig. S4) and are plotted in **(b)**. **(c)** Model prediction of the dependence of the lag time on the pre-expressed MetE abundance (blue line). This prediction only rely on some phenomenological parameters describing steady-state growth laws (shown in Fig. S3 and Fig. S4) and the steady-state growth rates in the pre-shift and post-shift condition (shown in Fig. 2.1b as green and pink bar). No adjustable parameters are needed. The uncertainty of model prediction (shaded area: 68% and 95% confidence range) comes from the uncertainty in growth rates. Here we assumed the same pre-shift growth rate for different MetE pre-expression level. **Fig. S5** shows that the effect of MetE over expression on growth rate is indeed negligible for the range we showed. The measured lag times and pre-expressed MetE mass fractions in various cTc conditions are shown in **(c)** as solid symbols. They were well captured by the model..

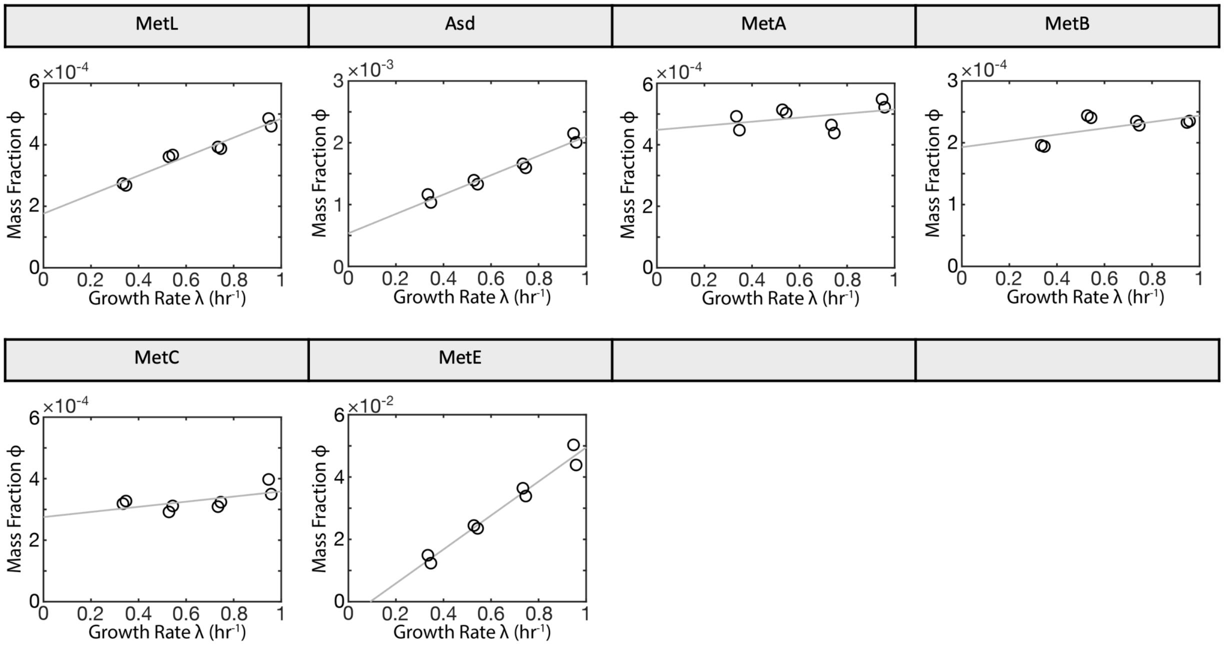


Figure 2.4. Steady-state mass fractions of enzymes on methionine synthesis pathway in minimal medium. Data points are from M1-M4, P1, P8-P10 conditions in (Chapter 3). Lines are linear fit of data.

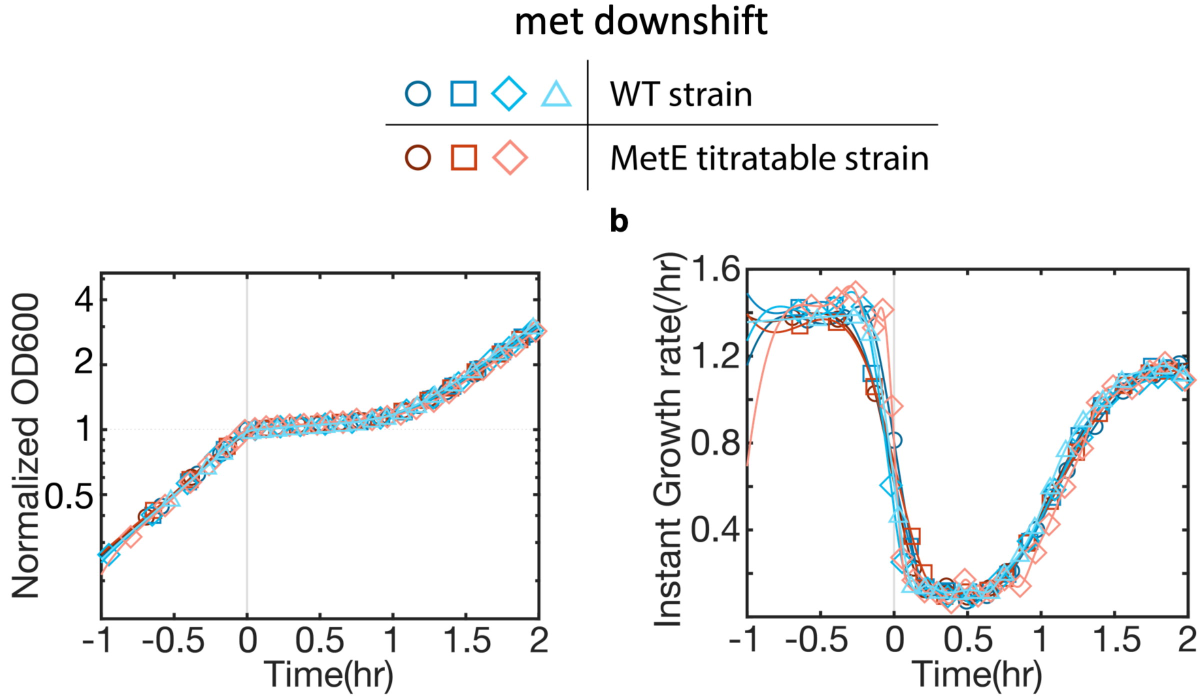


Figure 2.5. Met downshift growth curves **(a)** and instantaneous growth rates **(b)** of wild type (blue) and MetE titratable strain w/o cTc (red). Different symbols represent independent experiments. Lines here are guide of eyes.

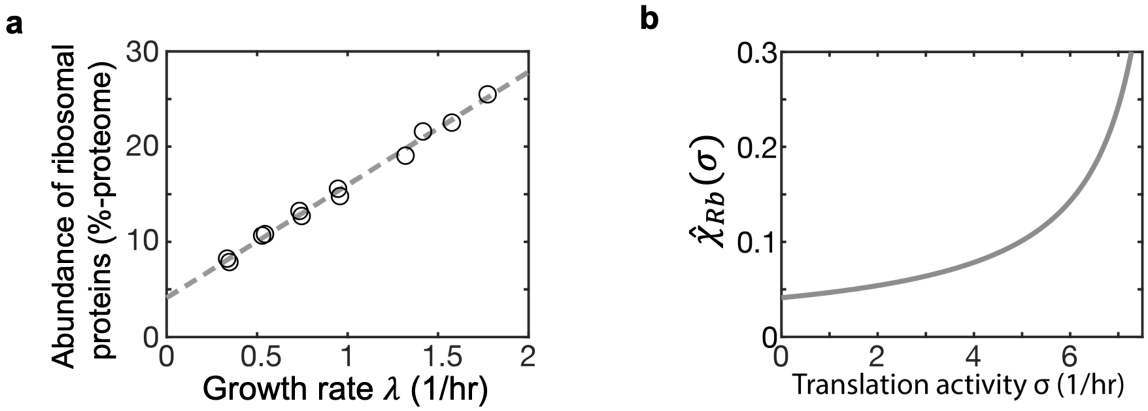


Figure 2.6: Construction of the regulatory function for ribosome synthesis, $\hat{\chi}_{Rb}(\sigma)$. (a) shows the abundance ϕ^*_{Rb} of ribosomal proteins at different growth rate (λ^*) for cultures growing exponentially in different nutrient sources. [Same data as that shown in **Fig. 3.6a**.] The data is well-captured by a linear fit (dashed line), i.e., $\phi^*_{Rb} = \phi_{Rb,0} + \lambda^*/\gamma$, with $\phi_{Rb,0} = 4.12\%$ and slope $\gamma = 8.43\ h^{-1}$. In steady state, we have $\chi^*_{Rb} \equiv \phi^*_{Rb}$, obtained by setting $\frac{d}{dt}\phi_{Rb} = 0$ using Eq. (2.1), (2.5) and (2.10) of the main text. Using the definition of σ in Eq. (2.6),

$$\sigma^* = \frac{\lambda^*}{\phi^*_{Rb}} = \frac{\lambda^*}{\phi_{Rb,0} + \lambda^*/\gamma} \tag{2.14}$$

we can invert Eq. (2.14) to obtain $\lambda^*(\sigma^*)$, and obtain the regulatory function as

$$\hat{\chi}_{Rb}(\sigma) = \chi^*_{Rb}(\sigma) = \phi^*_{Rb}\big(\lambda^*(\sigma)\big) = \frac{\phi_{Rb,0}}{1-\sigma/\gamma}. \tag{2.15}$$

The function obtained is sketched as the gray line in panel **(b)** and in **Fig. 2.3**.

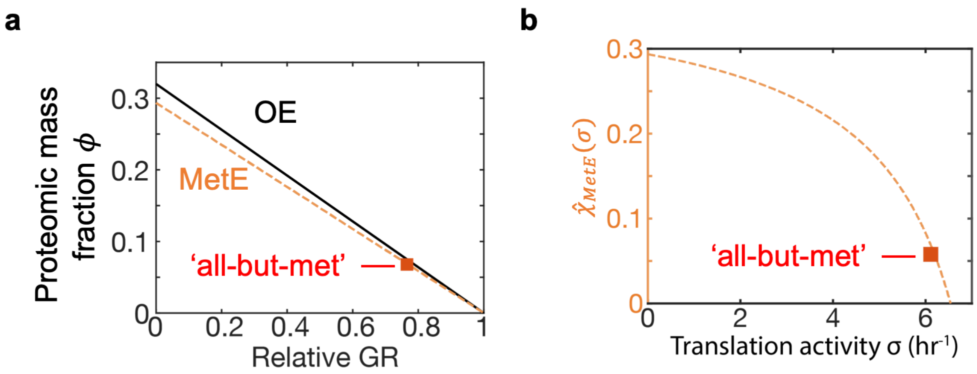


Figure 2.7: $\hat{\chi}_{metE}(\sigma)$ construction. Similar as obtaining $\hat{\chi}_{Rb}(\sigma)$ (Fig. 2.6), we first try to find the relationship between λ and ϕ_{MetE} under steady-state met limitation. As explained in the main text, to obtain the relationship, we convert the condition of met limitation to the condition of met enzyme over-expression as illustrated below.

Growth rate dependence on over-expressed proteins is established in Scotts 2010 (3). Over-expression (OE) growth law stated that OE limits resources for the synthesis of useful proteins, especially ribosomal proteins, leading to a growth rate drop (2). The relationship can be written as

$$\phi_{OE}^{*} = \phi_{max}(1 - \frac{\lambda^{*}}{\lambda_0}) \tag{2.16}$$

(black line in **(a)**), where ϕ_{OE}^{*} is the abundance of over-expressed proteins, λ_0 is the growth rate (GR) without OE and ϕ_{max} is a phenomenological parameter representing the maximum fraction that over-expressed protein can take (ϕ_{max} here is adjusted for NCM strains, see **Appendix 2.9.2** for details). Li et al (20) applied it to the condition lacking met supply in the media (referred as 'all-but-met' condition): the steady-state relationship between ϕ_{MetE}^{*} and λ^{*} in 'all-but-met' condition follows

$$\phi_{MetE}^{*} + \phi_{metother}^{*} = \phi_{max}(1 - \frac{\lambda^{*}}{\lambda_{met}}) \tag{2.17}$$

where λ_{met} is the growth rate with met supply ('all AA' condition), $\phi_{metother}^{*}$ is the mass fraction of other enzymes (except MetE) involved in met biosynthesis (basal expressions in 'all AA' condition are excluded in $\phi_{metother}^{*}$). Considering enzymes in met biosynthesis pathway are co-regulated, we can absorb $\phi_{metother}^{*}$ (10% of ϕ_{MetE}^{*}, see **Appendix 2.9.2**) into ϕ_{max} and Eq. (2.17) becomes

$$\phi_{MetE}^{*} = \phi'_{max}(1 - \frac{\lambda^{*}}{\lambda_{met}}) \tag{2.18}$$

which is the orange dash line in **(a)**. Comparing Eq. (2.18) to Eq. (2.16), MetE are considered as OE and the GR in 'all AA' condition is taken to be the growth rate without OE (MetE abundance in 'all AA' condition is negligible). In other words, Eq. (2.18) attributes the drop in steady-state growth rate when removing met from the media (create met limitation) to the burden of synthesizing MetE and linked the steady-state without met limitation (when met is supplied) and the steady-state with the most met limitation (no met supply). Thus, the relationship between λ and ϕ_{MetE} under met limitation is equivalent to the relationship between λ and ϕ_{MetE} under met enzyme over-expression, which is Eq. (2.18).

To double check whether the OE growth law can successfully capture the 'all-but-met' steady-state, we also plotted post-shift growth rate (measured) and MetE mass fraction (estimated by extrapolating the MetE-GR relationship in minimal media in **Fig. 2.4** to the measured GR in 'all-but-met' condition) in **(a)** (solid square). It is right on top of the orange line, meaning the post-shift 'all-but-met' steady-state is well described by OE growth law.

Substituting λ^{*} by $\sigma\phi_{Rb}^{*}$, ϕ_{Rb}^{*} by $\chi_{Rb}^{*}(\sigma)$ ($\chi_{Rb}^{*} = \phi_{Rb}^{*}$ under steady states) to Eq. (2.18), we then obtained the steady-state relationship between χ_{MetE}^{*} and σ for met limitation. Extending the form to the kinetic region (Eq. (2.19)), we got $\hat{\chi}_{MetE}(\sigma)$ shown in **(b)** and **Fig. 2.3b**.

$$\hat{\chi}_{MetE}(\sigma) = \chi_{MetE}^{*}(\sigma) = \phi'_{max}(1 - \frac{\sigma\hat{\chi}_{Rb}(\sigma)}{\lambda_{met}}) \tag{2.19}$$

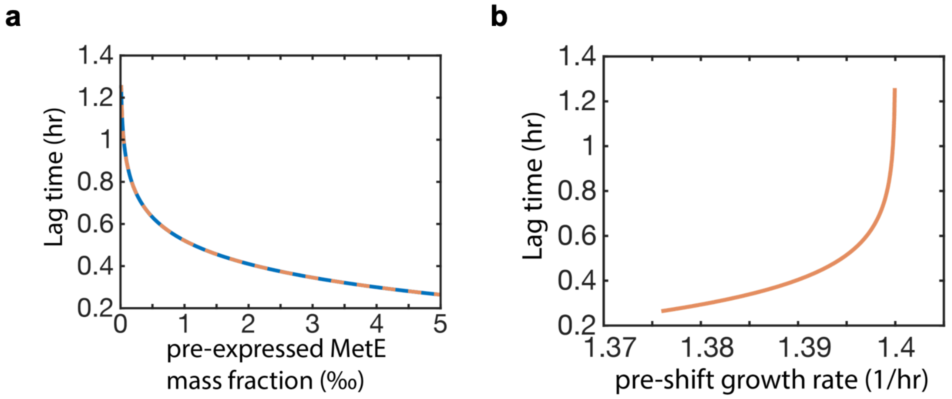


Figure 2.8: Model prediction when considering the effect of pre-expressed MetE on pre-shift growth rate. The blue line in **(a)** is the same as that shown in **Fig. 2.3c** where we assumed the same pre-shift growth rate for different MetE pre-expression levels. According to the over-expression growth law in (3), inducing MetE expression in the pre-shift condition will cause growth rate drop in pre-shift condition since MetE expression limited the resource for useful protein synthesis. Here we took it into consideration and let the pre-shift growth rate depend on the MetE pre-expression abundance. The relationship between pre-shift growth rate and MetE pre-expressed abundance is described by the OE growth law (the black line in **Fig. 2.7a**) with the maximum growth rate (no MetE pre-expression) being the measured WT growth rate in pre-shift condition. The model predictions are shown as orange lines in **(a)** and **(b)**. Notice that for the MetE range between 0 to 5‰, the predicted change in pre-shift growth rate is very small (**(b)**, less than 5%), while the predicted change lag time is significant. Due to the very small effect on growth rate, the prediction of lag time vs MetE stays the same (compare the orange line and the blue line in **(a)**).

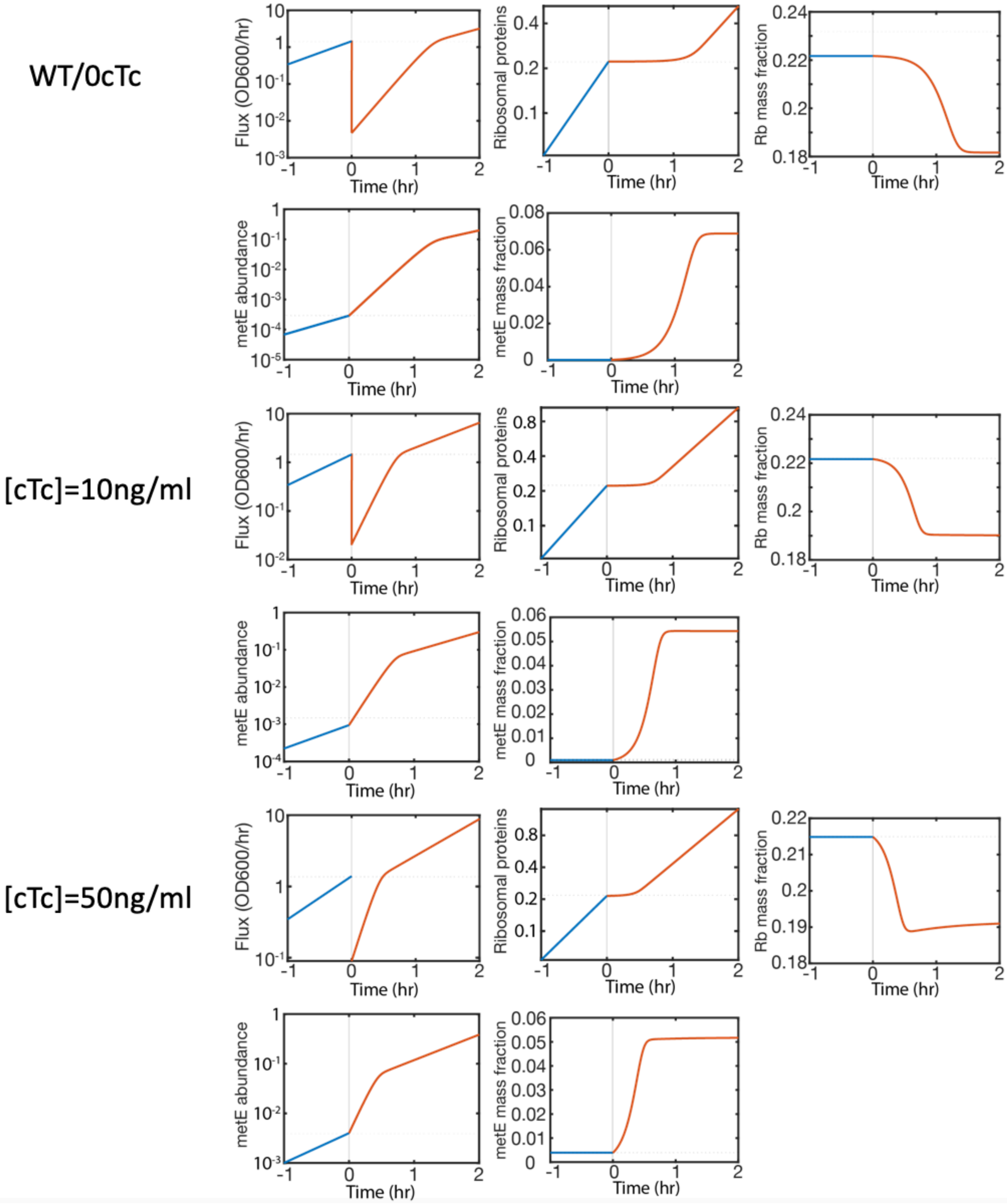


Figure 2.9: Model predictions of met downshift under various cTc concentrations. With measured pre-shift and post-shift steady-state growth rate and pre-expressed MetE abundance fitted to the lag time, our model can predict the kinetics of protein synthesis flux, ribosomal protein abundance and mass fraction, MetE abundance and mass fraction during met downshift at various cTc concentrations.

2.8 Tables

Table 2.1: parameter table for lag time vs MetE pre-expressed abundance (Fig. 2.3c) prediction.

1*. λ_0 by definition is the growth rate without MetE expression. Since the expression level of MetE is very low in wild type strain under pre-shift 'all AA' condition, λ_0 is taken to be the measured wild-type growth rate under 'all AA' condition. The error comes from growth rate measurements (95% confidence interval).

2*. λ_i and λ_f are measured pre-shift and post-shift steady-state growth rate of wild-type strain respectively. Here we assume that MetE pre-expression do not affect pre-shift growth rate λ_i. The error comes from growth rate measurements (95% confidence interval).

$\gamma(1/hr)$	$\phi_{Rb,0}$	ϕ'_{max}	$\lambda_0^{1*}(1/hr)$	$\lambda_i^{2*}(1/hr)$	$\lambda_f^{2*}(1/hr)$
8.43	0.0412	0.29	1.4 ± 0.03	1.4 ± 0.03	1.15 ± 0.03

Table 2.2: parameter table of met downshift under various cTc concentration.

1*. $\phi_{MetE}(t=0)$ is the MetE pre-expression level in mass fraction. It's fitted to the lag time of each shift.

2*. λ_i and λ_f are measured pre-shift and post-shift growth rate of MetE titratable strain (HE588) under various cTc conc., obtained from the growth curves in **Fig. 2.2b**.

3*. Since the pre-shift growth rate listed here is a little different from that of wild-type strain (**Table 2.1**), we adjust the λ_0 to make sure it's not smaller than the pre-shift growth rates in any cTc conc..

cTc conc. (ng/ml)	0	10	50
$\phi_{metE}^{1*}(t=0)$	0.00029	0.00095	0.0039
$\lambda_i^{2*}(/hr)$	1.45	1.45	1.392
$\lambda_f^{2*}(/hr)$	1.11	1.182	1.194
$\lambda_0^{3*}(/hr)$	1.45	1.45	1.45
$\phi_{Rb,0}$	0.0412	0.0412	0.0412
$\gamma(/hr)$	8.43	8.43	8.43
ϕ_{max}	0.29	0.29	0.29

Table 2.3: Oligonucleotides used in this study.

Name	**Sequence**	**Use**
Ptet-metE-F	gcacatcagcaggacgcactgaccgaattcattaaagaggagaaaggtaccatgacaatattgaatcacaccctcg	Amplification of *metE*
metE-R	ctgtccattacccccgacgcaagttctgcg	Amplification of *metE*
metKm-F	gtgcaggcggcgcagaacttgcgtcgggggtaatggacagcaagcgaaccggaattgc	Amplification of *km*
z-Km-R	cgccttccagcgttcgacccaggcgttagggtcaatgcgggtcgcttcacttagaagaactcgtcaagaaggcgatag	Amplification of *km*
metE-ver-F1	tggcgctgggccacaaggtgaaacc	Verification of *metE* at the *lacZ* locus
metE-ver-R	ttcatcaatctgtacccactcgatg	Verification of *metE* at the *lacZ* locus
metE-ver-F2	tacaaaaatgccatgaactggcactg	Verification of *metE* at the *lacZ* locus
metE-ver-F3	atgctgacggggccggtgaccatac	Verification of *metE* at the *lacZ* locus

Table 2.4: Estimating the mass fraction change in MetE and other enzymes on met biosynthesis pathway from pre-shift 'all AA' condition to post-shift 'all-but-met' condition.

1*. Cys group include enzymes on cysteine biosynthesis pathway encoded by cysC, cysD, cysE, cysH, cysI, cysJ, cysK, cysM, cysN.

2*. ϕ_0 and ν_E are offset and inversed slope from linear fit of **Fig. 2.4**.

3*. If not specified, ϕ_i^{all} is measured mass fraction of each enzyme(s) in the pre-shift 'all AA' condition, and $\phi_f^{all-but-met}$ is estimated by $(\lambda_f/\nu_E + \phi_0)$ for each enzyme(s) with $\lambda_f = 1.15$/hr (the growth rate in 'all-but-met' condition)

4*. Asd is need for methionine, lysine, threonine and isoleucine biosynthesis in minimal medium. But in 'all-but-met' condition, lys, thr and ile are provided in the medium. So we calculated $\phi_f^{all-but-met}$ of Asd by $\left(\lambda_f/\nu_E \times \left(\frac{\eta_{met}}{\eta_{met}+\eta_{lys}+\eta_{thr}+\eta_{ile}}\right) + \phi_0\right)$, where η is amino acid molar composition of cells reported in Gene-wei Li et al (**Table 2.4**; (20)). Since the differences of η across various medium is small (**Table 2.4**; (20)). We used the η in glucose minimal medium for calculation.

5*. $\Delta\phi$ report the net change of mass fraction from 'all AA' condition to 'all-but-met' condition. From pre-shift to post-shift condition, the increase in MetE is 5.75% while the total increase in other met enzymes is 0.618%, which is about 10% of the MetE abundance.

Enzymes	$^{2*}\phi_0$ ($\times 10^{-3}$)	$^{2*}\nu_E$ ($\times 10^{-3}$)	$^{3*}\phi_i^{all}$ ($\times 10^{-3}$)	$^{3*}\phi_f^{all-but-met}$ ($\times 10^{-3}$)	$^{5*}\Delta\phi$ ($\times 10^{-3}$)
^{1*}cys group	5	1.56	7.8	12.3	6.18
MetL	0.18	32.4	0.183	0.53	
Asd	0.5	6.38	0.452	^{4*}0.74	
MetA	0.45	150	0	0.52	
MetB	0.19	196	0.035	0.25	
MetC	0.28	120	0.064	0.37	
MetE	-5	0.18	0.104	57.6	57.5

Table 2.5: Proteome amino acid molar compositions. Analyzed from data in Gene-wei Li et al.(20).

Composition	Rich defined medium w/ glucose (%)	Glucose minimal medium (%)	Genome (%)
Alanine	10.13	10.15	9.478
Cysteine	0.7192	0.8365	1.162
Aspartate	5.691	5.991	5.149
Glutamate	7.004	6.611	5.766
Phenylamine	3.382	3.398	3.902
Glycine	8.249	8.099	7.356
Histidine	1.919	1.950	2.267
Isoleucine	5.802	5.706	5.994
Lysine	6.830	6.325	4.421
Leucine	8.155	8.725	10.63
Methionine	2.725	2.632	2.818
Asparagine	3.972	4.213	3.970
Proline	3.664	3.778	4.415
Glutamine	3.726	3.973	4.442
Arginine	5.990	5.385	5.537
Serine	4.871	5.088	5.838
Threonine	5.481	5.511	5.408
Valine	8.256	7.782	7.057
Tryptophan	0.7947	0.9808	1.531
Tyrosine	2.638	2.864	2.859

2.9 Appendix

2.9.1 Obtain MetE catalytic rate k^*_{MetE}

According to the Eq. (2.12) in the main text, growth rate is the production of $\frac{k^*_{MetE}}{f_{met}}$ and MetE abundance ϕ_{MetE} during the whole shift. Therefore the value of $\frac{k^*_{MetE}}{f_{met}}$ can be fixed by the post-shift steady-state via

$$\frac{k^*_{MetE}}{f_{met}} = \frac{\lambda^*_f}{\phi^*_{MetE,f}}, \tag{2.20}$$

where λ_f^* is the steady-state growth rate in the post-shift media and $\phi_{MetE,f}^*$ is the steady-state mass fraction of MetE under the post-shift media. The value of $\phi_{MetE,f}^*$ is calculated by Eq. (2.18) in **Fig. 2.7** via λ_f^*. So

$$\frac{k_{MetE}^*}{f_{met}} = \frac{\lambda_f^*}{\phi'_{max}(1 - \frac{\lambda_f^*}{\lambda_0})}, \tag{2.21}$$

which only depend on the steady-state growth rates.

2.9.2. Estimation of ϕ'_{max} in $\hat{\chi}_{MetE}(\sigma)$ construction

As explained in Fig. 2.7, ϕ'_{max} was defined as ϕ_{max} absorbing the ratio of ϕ_{MetE}^* over ($\phi_{MetE}^* + \phi_{metother}^*$), where ϕ_{max} is a phenomenological parameter representing the maximum fraction that over-expressed protein can take, ϕ_{MetE}^* and $\phi_{metother}^*$ are the steady-state abundance of MetE and other enzymes in 'all-but-met' condition respectively. (Basal expressions in 'all AA' condition are excluded in ϕ_{MetE}^* and $\phi_{metother}^*$. For MetE, this value is almost zero). One important assumption behind this definition is that the change of ϕ_{MetE} and $\phi_{metother}$ are proportional to each other during the shift. Considering that in met pathway, enzymes are co-regulated by met pool directly or indirectly, it is reasonable to assume $\phi_{metother} = q \cdot \phi_{MetE}$ during the shift with q being a constant. Thus, ϕ'_{max} can be calculated by ϕ_{max} and q following the equation below.

$$\phi'_{max} = \frac{\phi_{max}}{1 + q} \tag{2.22}$$

In Scott et al (3), ϕ_{max} is calculated to be 0.48 from the measurements of MG1655 strain. Since we used NCM3722 strain in this work, we estimated $\phi_{max} \cong 0.32$ following the same way in Scott et al (3) but using the measurements of NCM3722 strain in Dai et al (17). It is qualitatively agreed with the quantification from LacZ over-expression measurements in Basan et al (18).

Since q is assumed to be constant through the shift, to estimate q, we only need to get ϕ_{MetE}^{*} and $\phi_{metother}^{*}$, which is the steady-state mass fractions of MetE and other enzymes involved in met biosynthesis (including cysteine biosynthesis enzymes since cysteine provide sulfate for met biosynthesis) in the post shift 'all-but-met' condition subtracting their mass fraction in the pre-shift 'all AA' condition (to exclude basal expressions since the basal expression of other enzymes are significant). **Table 2.3** listed measured mass fraction in 'all AA' condition (ϕ_i^{all}, Chapter 3) and estimated mass fraction in 'all-but-met' condition ($\phi_f^{all-but-met}$). The mass fractions in the post-shift condition ('all-but-met' condition) are estimated by extrapolating each enzyme's steady-state growth rate dependence (Figure 2.4) to the steady-state growth rate in the 'all-but-met' condition (around 1.15/hr). In other words, we linearly fitted steady-state relation between enzyme mass fraction ϕ and growth rate λ under minimal media (**Fig. 2.4**) to

$$\phi = \phi_0 + \lambda/\nu_E \tag{2.23}$$

with ϕ_0 and ν_E listed in **Table 2.3** and substitute the post-shift growth rate to get the post-shift mass fraction $\phi_f^{all-but-met}$. One special case is $\phi_f^{all-but-met}$ of enzyme Asd. Asd is an enzyme necessary for methionine, lysine, threonine and isoleucine synthesis in the minimal medium. But in the post-shift condition ('all-but-met'), lysine, threonine and isoleucine are already provided. Here we assumed that Asd expression level will be smartly adjusted to match the actual requirements. Referring to the cellular AA composition (**Table 2.4**, (20)), $\phi_f^{all-but-met}$ of Asd is approximated by $(\lambda_f^{all-but-met}/\nu_{E,asd} \times \left(\frac{\eta_{met}}{\eta_{met}+\eta_{lys}+\eta_{thr}+\eta_{ile}}\right) + \phi_{0,asd})$, where $\lambda_i^{all-but-met}$ is the pre-shift growth rate, $\nu_{E,asd}$ and $\phi_{0,asd}$ are ν_E and ϕ_0 in Eq. (2.23) for enzyme Asd, η_x is the cellular composition of AA x. With estimated mass fractions of pre-shift and post-shift steady-states listed in **Table 2.3**, we can get the estimated increase of other enzymes ($\phi^*_{metother}$) is about 10% of that of ϕ^*_{MetE} (last column in **Table 2.3**), that is $q = 10\%$. Thus $\phi'_{max} = \frac{\phi_{max}}{1+q} = 0.29$.

2.9.3 Analytical solution of growth kinetics during met downshift

In our model, σ is the signaling variable. Once we solve the kinetics of $\sigma(t)$, we can solve the time dependence of other variables since they solely depend on the change of $\sigma(t)$. From the definition of growth rate $\lambda(t)$ (Eq. (2.11)) and mass fraction ϕ (Eq. (2.10)), the definition of σ (Eq. (2.6)) can also be written as

$$\sigma(t) = \frac{J_R(t)}{M_{Rb}(t)}. \tag{2.24}$$

By taking the time derivative on both sides of Eq. (2.24), we can write the time differential equation of σ:

$$\frac{d\sigma(t)}{dt} = \frac{d}{dt}\left(\frac{J_R(t)}{M_{Rb}(t)}\right) = \frac{1}{M_{Rb}(t)}\frac{dJ_R(t)}{dt} - \frac{J_R(t)}{M_{Rb}^2(t)}\frac{dM_{Rb}(t)}{dt}. \tag{2.25}$$

Here we let $t \geq 0$, since we are only interested in the kinetics of $\sigma(t)$ after the shift starts. Substituting the flux balance constrain with J_R limited by MetE (Eq. (2.9)) and expressing the protein synthesis kinetics via regulation functions (Eq. (2.4) and Eq. (2.5)), we got

$$\frac{d\sigma(t)}{dt} = \frac{k^*_{MetE}}{f_{met}}\frac{J_R(t)}{M_{Rb}(t)}\hat{\chi}_{MetE}(\sigma(t)) - \frac{J_R^2(t)}{M_{Rb}^2(t)}\hat{\chi}_{Rb}(\sigma(t)). \tag{2.26}$$

Replacing $\frac{J_R(t)}{M_{Rb}(t)}$ with $\sigma(t)$, Eq. (2.26) then becomes

$$\frac{d\sigma(t)}{dt} = \sigma(t)\left(\frac{k^*_{MetE}}{f_{met}}\hat{\chi}_{MetE}\big(\sigma(t)\big) - \sigma(t)\hat{\chi}_{Rb}\big(\sigma(t)\big)\right), \tag{2.27}$$

which is the Eq. (2.13) in the main text.

From Eq. (2.27), we can see that the change of translational capacity σ is determined by the regulation functions of MetE and ribosomal proteins. The increase in the flux allocation towards MetE synthesis makes σ rises, while the increase in the flux allocation towards ribosomal proteins makes σ drops. It can be understood considering MetE supplies the met pool for translation process while ribosomes drains the met pool.

We can further simplify Eq. (2.27) via the expression of $\hat{\chi}_{MetE}\big(\sigma(t)\big)$ (Eq. (2.19)) and $\frac{k^*_{MetE}}{f_{met}}$ (Eq. (2.21)):

$$\frac{d\sigma(t)}{dt} = \sigma(t)(\frac{\lambda^*_f}{(1-\lambda^*_f/\lambda_{met})}\left(1-\frac{\sigma(t)\hat{\chi}_{Rb}(\sigma(t))}{\lambda_0}\right) - \sigma(t)\hat{\chi}_{Rb}\big(\sigma(t)\big)). \qquad (2.28)$$

Eq. (2.28) can be rearranged to

$$\frac{d\sigma(t)}{dt} = \sigma(t)\frac{\lambda^*_f}{(1-\lambda^*_f/\lambda_{met})}\left(1-\frac{\sigma(t)\hat{\chi}_{Rb}(\sigma(t))}{\lambda^*_f}\right). \qquad (2.29)$$

We can see that $\phi'_{max} = \frac{\phi_{max}}{1+q}$ in the expression of $\hat{\chi}_{MetE}\big(\sigma(t)\big)$ and $\frac{k_{MetE}}{f_{met}}$ canceled each other. The kinetics of σ is thus independent of the value of ϕ'_{max}, the maximum proteomic allocation MetE can reach.

Now we write out the dependence of $\hat{\chi}_{Rb}$ on σ (Eq. (2.15)) and let $\mu_f = \frac{\lambda^*_f}{(1-\lambda^*_f/\lambda_{met})}$ for simplification. After rearranging the differential equation of σ in order to separate $d\sigma$ and dt, we got

$$\frac{d\sigma}{\sigma}\frac{(1-\sigma/\gamma)}{1-\sigma\big(1/\gamma+\phi_{Rb,0}/\lambda^*_f\big)} = \mu_f dt. \qquad (2.30)$$

From the ribosomal growth law (**Fig. 2.6a**), $\big(1/\gamma+\phi_{Rb,0}/\lambda^*_f\big)$ is actually $1/\sigma^*_f$, where σ^*_f is the translational capacity in the post-shift steady-state. Then we rewrote Eq. (2.30) into

$$\frac{d\sigma}{\sigma}\frac{(1-\sigma/\gamma)}{1-\sigma/\sigma_f^*}=\mu_f dt. \tag{2.31}$$

Integrating both side, we got

$$\ln\left[\frac{\sigma}{\left(1-\sigma/\sigma_f^*\right)^{1-\sigma_f^*/\gamma}}\right]\Bigg|_{\sigma(0)}^{\sigma(t)}=(\mu_f t)|_0^t. \tag{2.32}$$

Therefore,

$$\ln\left[\frac{\sigma(t)}{\sigma(0)}\left(\frac{1-\sigma(0)/\sigma_f^*}{1-\sigma(t)/\sigma_f^*}\right)^{1-\sigma_f^*/\gamma}\right]=\mu_f t. \tag{2.33}$$

So far, the kinetics of the key signaling variable σ has been solved analytically. It only depends on post-shift steady-state growth rate (through μ_f and σ_f^*), some phenomenological parameter γ, $\phi_{Rb,0}$, λ_{met} that describe steady-state growth laws and the initial condition $\sigma(0)$ right after the shift. $\sigma(0)$ is determined by the pre-shift expression of MetE through Eq. (2.6) and Eq. (2.12):

$$\sigma(0)=\frac{\lambda(t=0+)}{\phi_{Rb}(t=0)}=\frac{k_{MetE}^*\phi_{MetE}(t=0)}{f_{met}\phi_{Rb}(t=0)}, \tag{2.34}$$

where $\phi_{Rb}(t=0)$ is calculated through Eq. (2.14) given pre-shift growth rate. $\phi_{MetE}(t=0)$ is taken as a variable in the prediction in Fig. 2.3c.

With $\sigma(t)$, we can predict the time dependence of protein synthesis and flux during the shift. Let us start from ribosomal proteins.

The protein synthesis process of ribosomal proteins are regulated by $\hat{\chi}_{Rb}(\sigma(t))$ through Eq. (2.5). We substituted J_R with σM_{Rb}, dt with $d\sigma/\dot{\sigma}$ and rearranged Eq. (2.5) into

$$\frac{dM_{Rb}}{M_{Rb}}=\frac{\hat{\chi}_{Rb}(\sigma)\sigma}{\dot{\sigma}}d\sigma. \tag{2.35}$$

Substituting Eq. (2.15) for $\hat{\chi}_{Rb}(\sigma)$ and Eq. (2.31) for $\dot{\sigma}$, we got

$$\frac{dM_{Rb}}{M_{Rb}} = \frac{\phi_{Rb,0}}{\mu_f \left(1 - \sigma/\sigma_f^*\right)} d\sigma. \tag{2.36}$$

The time dependence of M_{Rb} is then solved by integrating both sides of Eq. (2.36). It is shown as

$$M_{Rb}(t) = M_{Rb}(0) \left(\frac{1 - \sigma(0)/\sigma_f^*}{1 - \sigma(t)/\sigma_f^*}\right)^{\frac{\left(1 - \sigma_f^*/\gamma\right)\lambda_f^*}{\mu_f}}, \tag{2.37}$$

where $\phi_{Rb,0}$ is replaced by $(\lambda_f^*/\sigma_f^* - \lambda_f^*/\gamma)$ for consistency. $M_{Rb}(0)$ is the pre-expressed ribosomal protein abundance calculated from $M_p(0) \cdot \phi_{Rb}(t = 0)$. We can further simplify Eq. (2.37) by substituting Eq. (2.33) into Eq. (2.37):

$$M_{Rb}(t) = M_{Rb}(0) \left(\frac{\sigma(0)}{\sigma(t)}\right)^{\frac{\lambda_f^*}{\mu_f}} \exp\,(\lambda_f^* t). \tag{2.38}$$

Eq. (2.38) shows the short term and long term time dependence of ribosomal proteins. In short term, the change of M_{Rb} relies on the change of $\sigma(t)$ through $\left(\frac{\sigma(0)}{\sigma(t)}\right)^{\frac{\lambda_f^*}{\mu_f}}$. While σ settles to the post-shift steady-state, ribosomal proteins increase exponentially with the exponent λ_f^* for a long term.

With $M_{Rb}(t)$, protein synthesis flux $J_R(t)$ can be derived through $\sigma(t)M_{Rb}(t)$—

$$J_R(t) = J_R(0) \left(\frac{\sigma(t)}{\sigma(0)}\right)^{\frac{\lambda_f^*}{\lambda_0}} \exp\left(\lambda_f^* t\right). \tag{2.39}$$

And met biosynthesis flux is fixed through $J_R(t) f_{met}$ as

$$j_{met}(t) = j_{met}(0) \left(\frac{\sigma(t)}{\sigma(0)}\right)^{\frac{\lambda_f^*}{\lambda_0}} \exp\left(\lambda_f^* t\right). \tag{2.40}$$

Here the initial condition of $J_R(0)$ and $j_{met}(0)$ are also dependent on the pre-expressed MetE abundance $\phi_{MetE}(t = 0)$.

Eq. (2.9) connected $J_R(t)$ and $M_{MetE}(t)$ via $\frac{k^*_{MetE}}{f_{met}}$. Combining Eq. (2.39) and Eq. (2.9), we can get the time dependence of $M_{MetE}(t)$ through

$$M_{MetE}(t) = M_{MetE}(0)\left(\frac{\sigma(t)}{\sigma(0)}\right)^{\frac{\lambda_f^*}{\lambda_0}} \exp(\lambda_f^* t). \tag{2.41}$$

In the end we solve the time dependence of total protein abundance $M_p(t)$, whose time derivative is $J_R(t) = \sigma(t)M_{Rb}(t)$. We substituted the non-simplified form of $M_{Rb}(t)$ (Eq. (2.37)) and changed dt to $d\sigma$ through Eq. (2.31). Then $M_p(t)$ can be calculated by integrating the equation below:

$$\frac{\mu_f}{M_{Rb}(0)} dM_p = \frac{1-\frac{\sigma}{\gamma}}{1-\frac{\sigma}{\sigma_f^*}}\left(\frac{1-\sigma(0)/\sigma_f^*}{1-\sigma/\sigma_f^*}\right)^{\frac{(1-\sigma_f^*/\gamma)\lambda_f^*}{\mu_f}} d\sigma. \tag{2.42}$$

Replacing $M_{Rb}(0)$ with $M_p(0)\phi_{Rb}(0)$ where $\phi_{Rb}(0)$ is determined by pre-shift growth rate λ_i^* via ribosomal growth law (**Fig. 2.6a**) and using the definition of σ and μ_f, we can do the integral of Eq. (2.42) and get

$$
\begin{aligned}
&M_p(t) \\
&= M_p(0)\left[1 - \frac{\lambda_i^* \sigma_f^*}{\lambda_f^* \sigma_i^*} \frac{1 + \left(\frac{\lambda_{met}}{\lambda_f^*} - 1\right)\frac{\sigma(0)}{\gamma}}{1 + \left(\frac{\lambda_{met}}{\lambda_f^*} - 1\right)\frac{\sigma_f^*}{\gamma}}\right] \\
&+ M_p(0)\frac{\lambda_i^* \sigma_f^*}{\lambda_f^* \sigma_i^*} \frac{1 + \left(\frac{\lambda_{met}}{\lambda_f^*} - 1\right)\frac{\sigma(t)}{\gamma}}{1 + \left(\frac{\lambda_{met}}{\lambda_f^*} - 1\right)\frac{\sigma_f^*}{\gamma}} \left(\frac{1 - \sigma(t)/\sigma_f^*}{1 - \sigma(0)/\sigma_f^*}\right)^{(1-\sigma_f^*/\gamma)(\lambda_f^*/\lambda_{met}-1)} \qquad (2.43) \\
&= M_p(0)\left[1 - \frac{\lambda_i^* \sigma_f^*}{\lambda_f^* \sigma_i^*} \frac{1 + \left(\frac{\lambda_{met}}{\lambda_f^*} - 1\right)\frac{\sigma(0)}{\gamma}}{1 + \left(\frac{\lambda_{met}}{\lambda_f^*} - 1\right)\frac{\sigma_f^*}{\gamma}}\right] \\
&+ M_p(0)\frac{\lambda_i^* \sigma_f^*}{\lambda_f^* \sigma_i^*} \frac{1 + \left(\frac{\lambda_{met}}{\lambda_f^*} - 1\right)\frac{\sigma(t)}{\gamma}}{1 + \left(\frac{\lambda_{met}}{\lambda_f^*} - 1\right)\frac{\sigma_f^*}{\gamma}} \left(\frac{\sigma(t)}{\sigma(0)}\right)^{(\lambda_f^*/\lambda_{met}-1)} \exp\left(\lambda_f^* t\right).
\end{aligned}
$$

The form of $M_p(t)$ looks complicated, but we can tell that for the long term $M_p(t)$ is dominated by $\exp\left(\lambda_f^* t\right)$ and increase exponentially.

With mass of proteins and fluxes calculated above, other intensive properties, such as instantaneously growth rate $\lambda(t)$ and mass fraction $\phi(t)$ can be calculated from their definitions by divisions. The lag time T_{lag} can be calculated from the instantaneously growth rate via

$$T_{lag} = \int_0^\infty (1 - \lambda(t)\,/\lambda_f^*)dt. \qquad (2.44)$$

So that $e^{T_{lag}\cdot\lambda_f^*} = \lim_{t\to\infty} e^{t\cdot\lambda_f^*}/M_p(t)$, representing the lost fold changes in biomass due to transition.

Chapter 3

Proteome reserves define the onset of enzyme recovery across biosynthesis pathways

3.1 Abstract

Bacteria can grow in many different nutrient environments, and the rapidity of their adaptation is central to the overall fitness. *Escherichia coli* growing in “rich medium” primarily direct gene expression resources to ribosome biogenesis and cell growth. Upon transition to “poor medium”, a large number of biosynthesis genes must be turned on before growth can resume. Quantitative proteomic analysis reveals a continuous spectrum of growth bottlenecks in amino acid biosynthesis. A striking linear relation is revealed between the onset time of many enzymes across biosynthesis pathways and the fractional “reserve” of the corresponding enzymes kept by cells during growth in rich media when they are not needed. A coarse-grained kinetic model quantitatively captures the observed recovery kinetics of these enzymes across many pathways, as well as the adaptation of ribosome biogenesis and cell growth, based solely on snapshots of the proteome right before and long after the transition, without invoking any ad-hoc fitting parameters. Our results establish the coordinated adoption of an as-needed gene expression program across biosynthetic pathways, and elucidate the implementation of this program by flux-controlled global regulation and end-product inhibition.

3.2 Introduction

In different nutrient environments, bacteria are known to allocate their proteome differently (3). Previous proteomic research has focused on growth in minimal media with one or a few carbon and nitrogen sources (2, 7, 20, 33–35), and on transitions between these minimal media (9, 10, 23, 36). Much less is known for bacterial growth in different 'rich media' containing a multitude of amino acids and other nutrients, even though such nutrient conditions may be more common in natural habitats (e.g., the gut for enteric bacteria (37, 38)) than minimal medium, at least transiently until their exhaustion.

In this study, we characterized the proteome of *E. coli* cells growing in a number of 'rich media' as well as during growth transition from rich to minimal media, using a recently developed high-accuracy method of quantitative protein mass spectrometry (2). Our data reveals the reallocation of proteome between amino acid biosynthesis (AAB) enzymes and protein synthesis machineries as essential for attaining increased growth rate in rich media. As a consequence of this reallocation, *E. coli* suffers substantial growth lag when downshifting from rich media to minimal media. Detailed, quantitative experiments and mathematical modeling were performed to establish the origin of the growth lags, and to characterize the recovery kinetics of the individual AAB pathways.

3.3 Results

3.3.1 Proteome allocation for steady state growth in rich media

We characterized the proteome of exponentially growing *E. coli* K-12 cells in a number of 'rich media' using a recently developed method for quantitative protein mass spectrometry (2). This method combines accurate relative quantification provided by DIA/SWATH (39, 40) and a new protein inference algorithm (xTop), together with the excellent absolute quantification provided by ribosome profiling (20), to yield accurate fractional abundances for about 2000 proteins in each condition. For comparison, we also characterized the proteome of the same cells growing on the same MOPS-based minimal medium with glucose as the sole carbon source, as well as for cells subjected to reduced glucose uptake; see **Table 3.1** for strain and **Table 3.2** for medium conditions.

The legend table in **Fig. 3.1** shows the growth rate in rich media and minimal glucose medium. As expected, the addition of casamino acid (CAA, containing mostly amino acid monomers) or rich-defined medium (RDM, containing amino acids, nucleotides, and vitamins (22)) to glucose minimal medium increased the steady-state growth rate, from ~1/hr to 1.6/hr and 1.8/hr respectively. Interestingly however, while the growth on CAA alone (1.3/h) well exceeded that on glucose alone (1/h), further addition of glucose (1.6/h) boosted the growth more than the addition of nucleotides and vitamins (RDM alone, 1.4/h). This underscores the dominant roles amino acids and additional carbon supplies play in rapid bacterial growth, with the supply of nucleotides and vitamins exerting relatively minor effects.

In **Fig. 3.1a**, we show the total abundances of the proteins in several key functional groups (41, 42) for wild type cells grown in glucose minimal medium (xxx bars) and in glucose supplemented by RDM (yyy bars). Here and throughout the text, the abundance of a protein X is reported as the mass fraction of that protein among the total cellular protein mass, denoted as ϕ_X, and the abundance of a functional group is the sum of the mass fraction of all proteins in that group (as specified in **Table 3.3**). The result in **Fig. 3.1a** shows the clear dominance of translation-related proteins for growth in rich medium and the co-dominance of amino-acid biosynthesis (AAB) enzymes in minimal medium. Looking into the data more closely by plotting abundances against the growth rate of each culture, **Fig. 3.1b** shows a linear increase in the abundance of translation-related proteins with increasing growth rates, by ~14%-proteome from glucose alone (black circle) to glucose supplemented by RDM (green square), with similar increases across the board for ribosomal proteins and auxiliary proteins such as elongation factors, and for tRNA synthetases (**Fig. 3.6**). Since translational capacity is nearly saturated in these growth conditions (17), the observed increase in the abundance of translational machinery is necessary to support the increase in protein synthesis flux in rich media.

Since the total cellular protein content changes in unison with cell volume for cells grown in different media (1, 2), the abundance measured in mass fraction is proportional to cellular protein concentration (with mass fraction of 1% corresponding to $\sim 3 \times 10^4$ proteins per μm^3 of cell volume for typical proteins 250aa in length). Thus, the increased concentration of translational proteins in rich media must be accompanied by decreased concentration of other proteins. Our data shows that this was mostly achieved by decreases in the allocation of AAB enzymes, which dropped from ~14% to < 4% of total protein mass when supplied with either CAA or RDM (**Fig. 3.1c**). The effect of nucleotide synthesis involves <1% proteome (**Fig. 3.1d**) and is negligible compared to AAB enzymes. Although a part of the proteome devoted to TCA cycle and glycolysis/gluconeogenesis (~2%-proteome total) were relieved when RDM was supplied in the media due to the reduced demand of carbons for AAB, their effect was offset by the increase in enzymes allocated to acetate overflow (2%-proteome) for ATP generations (**Fig. 3.1e-g**). Additionally, the ~20% difference in growth rate due to glucose supplementation to CAA or RDM, reflected in the reallocation of proteome between translational proteins and enzymes of TCA cycle, gluconeogenesis, etc. (compare open and filled symbols, **Fig. 3.1f-i**, **Fig. 3.7**), can be rationalized by the need to degrade amino acids (AA) for energy biogenesis in rich medium lacking a carbon source, since pyruvate and TCA intermediates are the common products resulting from AA degradation (**Fig. 3.2a**-**b**).

3.3.2 AA downshift from rich to minimal media leads to long lags

Data in **Fig. 3.1** indicate that the reallocation of the proteome from AA biosynthesis to translational machineries is a primary requirement for cells to gain fast growth in rich media. However, removal of AAB enzymes from the proteome makes it more difficult for cells to maintain growth once the supply of AA is exhausted, where *de novo* AA biosynthesis is needed (**Fig. 3.2a**). This could give rise to a substantial lag in growth recovery in growth transition from rich to minimal media.

To substantiate this growth-lag tradeoff, we characterized AA downshift kinetics by first growing cells in media with AAs (same concentration as in RDM, see **Table 3.4**) supplemented by glucose, and abruptly shifting the exponentially growing culture to the same medium but without any AAs (**Fig. 3.2b**; see details in **Methods**). A 2-hour lag was observed before growth recovery for a pre-shift mix of all 20 AAs, as well as for a mix of 18 AAs (**Table 3.4**) with cysteine and tyrosine excluded (**Fig. 3.8a**). Due to the instability of cysteine and poor solubility tyrosine, we exclude these two minor amino acids in the remainder of this study.

In principle, the lag observed could arise from a combination of bottlenecks involving *de novo* AA biosynthesis and the uptake of carbon substrate. We thus tested the effect of different carbon supplement on AA downshift kinetics. For a variety of carbon substrates which enter central metabolism from upper glycolysis by a variety of mechanisms (**Table 3.5**), the lag time were all ~2h (**Fig. 3.2c** and **Fig. 3.8b**) despite the boost in pre-shift GRs they provided (**Fig. 3.2d**). It is thus unlikely that the uptake of glycolytic carbon sources is strongly repressed before the shift.

We eliminated this possibility using AA downshift with glycerol, whose uptake system is well characterized (10): By repeating the AA downshift experiment with glycerol supplement using a mutant NQ399 (**Table 3.1**) harboring full unregulated expression of the glycerol uptake proteins GlpFK (43), we found these cells still exhibited the same 2h lag (**Fig. 3.8c-e**). Thus, the common 2h lag arose from internal growth bottleneck(s) related to abrupt AA removal from rich media. Below, we investigate this common 2hr lag for the downshift from 18AA to none using glycerol as the carbon supplement and refer to it generally as “AA downshift”.

3.3.3 Proteomic analysis during AA downshift

To gain a more quantitative look at the cellular recovery kinetics during AA downshift, we characterized the proteome after the shift using the DIA/SWATH method as for the steady-state measurements (2). **Fig. 3.9a-l** show the time courses of protein groups with panels a-h being the same groups shown in **Fig. 3.1**. The instantaneous growth rate is plotted in **Fig. 3.9m** for reference. As expected, large increase in AAB enzymes occurred during the shift (**Fig. 3.9b**). Translational proteins and nucleotide synthesis enzyme started decreasing after 1-hr (**Fig. 3.9a, c**) when growth recovery started, indicating their decrease mainly came from dilution rather than degradation. Somewhat surprisingly, glycerol uptake enzymes along with other transporters increased significantly (**Fig. 3.9i-k**), by about 3-fold, similar to the total of AAB enzymes, even though glycerol uptake was shown to be not a bottleneck during AA downshift (**Fig. 3.8d**).

The reason that the AAB enzymes did not show a larger increase is that the total abundance plotted in **Fig. 3.9b** masks the behavior of the enzymes in individual pathways, some of which can have much larger fold change than others and each AA is needed for growth in the post-shift minimal medium. We next plotted the abundances of individual AAB enzymes, grouped by the pathways (**Fig. 3.10**). The responses exhibited across these pathways were quite diverse, both in terms of fold changes and recovery kinetics: MetE showed >100-fold change (**Fig. 3.10a**) while enzymes in threonine synthesis pathway showed no more than two-fold change (**Fig. 3.10j**). Some enzymes started increasing in the early stage of the shift, see e.g., the methionine and arginine pathways (**Fig. 3.10a,b**), while other enzymes did not increase until 100 mins after the shift, see e.g., the serine and threonine pathways (**Fig. 3.10i, j**).

As the enzymes in the same pathway (shown in the same panel of **Fig. 3.10**) exhibited similar kinetics, attributed at least in part to the operon structure (e.g., for trp, his, leu), we summed up the abundances of enzymes in the same pathway to obtain a coarser grained view of the kinetic response (triangles in **Fig. 3.3a-n**). The time courses of the 14 enzyme groups were ordered according to their fold changes during the shift. Interestingly, a systematic change in the recovery kinetics can be seen from the plots: The groups at the top exhibited the largest fold change and increased the earliest (before 50 mins), while the groups at the bottom, with only 2x change, showed the slowest response (started increasing after 100 min). To quantitatively characterize the recovery kinetics, we estimated the onset time of each group (defined as the time when the abundance of each group increased by 25% above the pre-shift value; see **Fig. 3.11a**). The onset time clearly decreased with fold change in enzyme abundance (**Fig. 3.11b**). Plotting against the reciprocal of the fold change, describing the pre-shift enzyme abundance level relative to the final, we obtained a striking linear relation passing through the origin (**Fig. 3.3o**). This linear relation is robust to the precise definition of the onset time (**Fig. 3.11c**), and is moreover seen not only at the pathway level, but also for individual AAB enzymes (crosses in **Fig. 3.3p**, **Fig. 3.11d-e**). Correlations between the onset time and the pre- and post- shift abundances themselves or their differences are much weaker (**Fig. 3.11f-k**). Since the pre-shift enzyme abundance relative to the final represents the adequacy of enzyme reserve during pre-shift growth in rich medium (referred to below as “pre-shift reserve”), our data reveals a simple rule that the recovery onset time of AAB enzymes is set linearly by their pre-shift reserves .

3.3.4 Coarse-grained model of AA downshift kinetics

The correlation observed between the enzyme onset time and pre-shift reserve is reminiscent of qualitative predictions made in a theoretical study by Pavlov and Ehrenberg (PE) (16), who posited that allocating protein synthesis preferentially to enzyme groups with the smallest reserve would provide the fastest recovery. However, it is difficult to obtain concrete predictions on enzyme onset times from the PE model in order to compare to the data, since the PE model was assumed to be globally applicable to all enzymes, while the relationship observed between onset time and pre-shift reserve was specific for AAB enzymes, but not for the majority of other detected proteins (circles vs crosses, **Fig. 3.3p**). Below, we incorporated elements of the PE model into a coarse-grained model of flux-controlled proteome allocation (23) to examine quantitatively the onset of AAB enzymes, as well as their effect on the overall growth recovery.

Qualitatively, the problem cells face right after AA downshift can be described as follows: Cells need to make AAB enzymes to supply the missing AA influx. However, the overall protein synthesis flux is limited due to the lack of AAs. Therefore, it is necessary to reallocate the limited protein synthesis flux towards AAB enzymes right after the shift, away from the synthesis of ribosomal proteins which dominated before the shift (**Fig. 3.2a**). The generic kinetic equations for the n^{th} AAB enzyme group (of proteomic mass fraction $\phi_{A,n}(t)$) and ribosomal proteins (mass fraction $\phi_{Rb}(t)$) can be written as (**Appendix in Chapter 3**)

$$\frac{d\phi_{A,n}}{dt} = \lambda(t) \cdot \left(\chi_{A,n}(t) - \phi_{A,n}(t)\right), \tag{3.1}$$

$$\frac{d\phi_{Rb}}{dt} = \lambda(t) \cdot \left(\chi_{Rb}(t) - \phi_{Rb}(t)\right). \tag{3.2}$$

Here $\lambda(t) \equiv \frac{d}{dt}\ln M(t)$, is the instantaneous growth rate during the shift, with $M(t)$ being the total protein mass. It describes the magnitude of protein synthesis flux (dM/dt) normalized by the total protein mass. $\chi_{A,n}(t)$ and $\chi_{Rb}(t)$ are "allocation functions" describing the fraction of protein synthesis flux directed towards synthesizing the n^{th} AAB enzyme group and ribosomal proteins, respectively. In order to solve the recovery kinetics, it is necessary to specify the time dependence of growth rate $\lambda(t)$ and the allocation functions $\chi_{A,n}(t)$, $\chi_{Rb}(t)$ in terms of the dynamical variables, $\phi_{A,n}(t)$, $\phi_{Rb}(t)$.

We adopted the PE model (16), setting the instantaneous growth rate in terms of the abundances of the rate-limiting enzyme group(s), in this case, the AAB enzyme groups $\phi_{A,n}(t)$, as

$$\frac{\lambda(t)}{\lambda^*} = \min_n\left(\frac{\phi_{A,n}(t)}{\phi^*_{A,n}}\right) = \min_n q_n(t). \tag{3.3}$$

Here λ^* and $\phi^*_{A,n}$ represent the growth rate and the abundance of the n^{th} AAB enzyme group in the post-shift steady-state, respectively, and $q_n(t) \equiv \phi_{A,n}(t)/\phi^*_{A,n}$ is the enzyme abundance relative to the final abundance, i.e., the fractional enzyme reserve. As elaborated in **Fig. 3.12**, Eq. (3) is simply a statement that the protein synthesis flux (proportional to $\lambda(t)$) is governed by the abundance of the most bottlenecked enzyme group(s), a result which follows straightforwardly from the assumption that the catalytic rates of the rate-limiting enzyme groups are at maximum (i.e., uninhibited) as are the catalytic rates in the post-shift steady-state, and is consistent with allosteric inhibition of non-limiting AAB enzymes (44).

Specifying the time-dependent allocation functions $\chi_{A,n}(t),\ \chi_{Rb}(t)$ is more challenging. Molecularly, the synthesis of different enzyme groups are regulated by their corresponding AA products (**Fig. 3.4a**). According to the well-known mechanisms of end-products inhibition (44), the synthesis of growth-limiting AAB enzyme group(s) are expected to be up-regulated due to the low concentration of their AA products, while synthesis of non-growth limiting group(s) are expected to be repressed due to the accumulation of their corresponding AA products. To determine the protein synthesis flux allocated to each AAB group without requiring molecular details (which are mostly unknown at the quantitative level), we developed a two-stage allocation scheme as illustrated in **Fig. 3.4a**: In the first stage, we extended the scheme of flux-controlled regulation (FCR) introduced by Erickson et al (23) to describe the allocation between ribosome synthesis $\chi_{Rb}(t)$ and *total* AAB enzyme synthesis $\chi_{A,tot}(t)$, which is based molecularly on ppGpp-mediated regulation (45). Then in the second stage, we describe more refined sub-allocation of the AAB enzymes to each group, i.e., from $\chi_{A,tot}(t)$ to $\chi_{A,n}(t)$, which is based on end-product inhibition.

In the FCR scheme, the translational activity, defined as

$$\sigma(t) \equiv \lambda(t)/\phi_{Rb}(t), \tag{3.4}$$

is taken to represent the overall abundance of the AA pools while growth is limited by AA synthesis (Eq. (3)). Using $\sigma(t)$, we can express the allocation functions $\chi_{A,tot}(t)$ and $\chi_{Rb}(t)$ in terms of the *regulatory functions* $\hat{\chi}_{A,tot}(\sigma)$ and $\hat{\chi}_{Rb}(\sigma)$, as

$$\chi_{A,tot}(t) = \hat{\chi}_{A,n}\big(\sigma(t)\big), \tag{3.5}$$

$$\chi_{Rb}(t) = \hat{\chi}_{Rb}\big(\sigma(t)\big). \tag{3.6}$$

The forms of these regulatory functions, shown in **Fig. 3.4b**, can be determined from steady-state measurements as detailed in **Fig. 3.13** and **Fig. 3.14**. We see that under limited growth (small σ), ribosomal synthesis is repressed while AAB enzyme synthesis is induced as may be expected intuitively. Molecularly, the translational activity $\sigma(t)$ is represented (inversely) by the ppGpp signaling system, during growth transition and in steady-state (45, 46). Down-regulation of ribosomal synthesis by high ppGpp (small σ) is well documented (14, 25–27). Our model also expects up-regulation of AAB enzyme synthesis by ppGpp (**Fig. 3.4b**), which is supported by recent studies characterizing the transient effect of ppGpp on AAB synthesis (47).

In the second stage of our regulatory scheme, we specified the allocation of AAB enzyme synthesis towards the individual groups using another set of allocation functions:

$$\chi_{A,n}(t) \equiv \eta_n(t) \cdot \chi_{A,tot}(t). \tag{3.7}$$

The allocation functions $\eta_n(t)$ were chosen in accordance to the outcome of end-product inhibition (44), which has a simple mathematical implementation as described by Pavlov and Ehrenberg (16) and detailed in **Fig. 3.15**: At a given time t, the AAB groups can be divided into two classes based on whether a group limited protein synthesis. The non-limiting group (with $q_n(t) > \lambda(t)/\lambda^*$) received no allocation, i.e., $\eta_n(t) = 0$, while the limiting groups were allocated in proportional to their final abundance, $\phi^*_{A,n}$. If one orders the enzyme groups according to their relative pre-shift abundance level $q_n(0)$ such that $q_1(0) < q_2(0) < \cdots$, then the group-specific allocation imposed by end-product inhibition can be expressed as

$$\eta_n(t) = \begin{cases} 0 & n > \tilde{n}(t) \\ \phi^*_{A,n} / \sum_{m=1}^{\tilde{n}(t)} \phi^*_{A,m} & n \leq \tilde{n}(t) \end{cases} \tag{3.8}$$

where $\tilde{n}(t)$ is the number of enzyme groups being synthesized at time t, so that the groups with $n \leq \tilde{n}(t)$ are equally growth-limiting while those with $n > \tilde{n}(t)$ are not limiting.

3.3.5 Enzyme and growth kinetics of the two-stage model

Eqs. (1)-(8) completely specify our model of the recovery kinetics during AA downshift. With the initial and final enzyme abundance values $\phi_{A,n}(0)$ and $\phi^*_{A,n}$ fixed by the proteomic data (**Fig. 3.3**), the model predicts the kinetics of the ribosomes, each AAB enzyme groups, as well as the overall growth dynamics without any adjustable parameters. First of all, the allocation for ribosomal synthesis $\chi_{Rb}(t)$ is predicted to be repressed right after the shift, while the total allocation for AAB enzymes $\chi_{A,tot}(t)$ is upregulated as shown in **Fig. 3.5a**. The resulting changes in the abundances of these two groups of proteins, plotted as the black and green curves in **Fig. 3.5b**, capture the data (triangles) very well. The predicted allocation functions $\chi_{A,n}(t)$ for each of the 14 AAB groups are shown in **Fig. 3.5c**: Right after the shift, only the Met group (orange) was synthesized since methionine was most growth limiting (based on the pre-shift reserve of the Met enzymes shown in **Fig. 3.3a**). After ~20 minutes when methionine was no longer the only growth-limiting AA, the synthesis of other AAB enzymes would turn on one after another, in a multi-step recovery process. The resulting dynamics for each enzyme group is plotted in **Fig. 3.5d** for the relative abundance $q_n(t)$ using the same color scheme. Predictions of the time courses of the abundance of each enzyme group also capture the data well (**Fig. 3.3a-n**, lines and symbols), again without resorting to any adjustable parameters. From these solutions, we computed the onset time for each enzyme group. The results (**Fig. 3.5e**, crosses) are in very good agreement with the observed onset time (circles).

In addition to the kinetics of each enzyme group, our model also gave prediction on the growth kinetics itself during the shift (**Fig. 3.5f**), again without adjustable parameters. Interestingly, although our model captured the AAB enzyme onset time very well, it predicted a growth lag (blue line) which was about only half of the 2-hr lag observed (grey circles). To understand this discrepancy, we looked more closely into the recovery kinetics of each enzyme group (**Fig. 3.3a-n**) and noted that the group of Ser enzymes (**Fig. 3.3i**) was the first to show significant difference between measurements and model predictions. This can also be seen from the comparison of onset times (purple color, **Fig. 3.5e**). This "serine problem" turns out to be an artifact of the common recipe for "rich medium" whose composition has 10x higher serine content compared to the composition of typical proteins (**Fig. 3.16**). With normal AA composition, serine is depleted well before other AAs (48, 49). Indeed, using "casamino acid" (obtained as the hydrolysis product of the protein casein (50)) as the source of amino acid in pre-shift, we see that the growth kinetics is better captured by the model (**Fig. 3.5f**).

3.4 Discussion

In this study, we investigated the allocation of the proteome of *E. coli* grown in rich and minimal media, in steady state growth and during transition. The most significant changes are the tradeoff between AAB enzymes and the translational machinery, which allows cells to achieve higher protein synthesis flux, hence growth rate, in rich media where *de novo* AA biosynthesis is relieved (**Fig. 3.1**). However, drastically reducing the AAB enzymes causes problems when AA supply is abruptly depleted, manifested by a lag during AA downshift even when a separate carbon source is provided throughout the shift (**Fig. 3.2**).

Quantitative proteomic measurements during AA downshift revealed a diversity of time scales in the recovery kinetics of AAB enzymes (**Fig. 3.3**, **Fig. 3.10**). A linear relationship was established between the onset time and the fractional pre-shift reserve of AAB enzymes *across* pathways (**Fig. 3.3o**). The observed relationship is a manifestation of aa as-needed strategy of gene regulation for AAB enzymes, where those involved in the biosynthesis of the amino acid most needed for protein synthesis, methionine, were expressed first, followed by the enzymes for the next most needed AA, arginine, etc. In this context, we recall that a related 'just-in-time' gene expression strategy was proposed long ago for AA downshift (51), and was claimed as the optimal strategy for minimizing the amino acid recovery time while minimizing enzyme production cost (51). That strategy posits that *within* an unbranched pathway, the closer an enzyme is to the beginning of the pathway, the earlier its expression is turned on and the higher is its maximal expression level during AA downshift (**Fig. 3.17**). We examined four unbranched pathway segments for which the enzymes are not encoded in the same operon, three of which were claimed to follow the optimal strategy of Zaslaver et al (51) based on the expression kinetics of plasmid-borne GFP reporters. However, our proteomic data clearly refuted the adoption of such a strategy within a pathway (**Fig. 3.17c-f**). Instead, *E. coli* adopts just-in-time strategies across AAB pathways.

To quantitatively understand the linear recovery kinetics during AA downshift, we constructed a coarse-grained model with proteome allocation specified in two stages (**Fig. 3.4a**). In the first stage, we extended the framework of flux-controlled regulation introduced in Erickson et al (23) to AA biosynthesis, using the translational activity σ to represent the collective effect of

AA pools and to direct the dynamic allocation of protein synthesis flux between the AAB enzymes and ribosomal proteins. In the second stage, we allocated the synthesis of AAB enzymes into individual pathways, using the on-off strategy formulated by Pavlov & Ehrenberg (16) to capture the effect of end-product regulation of individual AAB pathways (44). With the only input parameters in the model being the growth rates and the proteomic data in the pre- and post-shift steady-states, our model was able to quantitatively capture the onset time of AAB enzyme groups during AA downshift (**Fig. 3.5e**) together with the actual kinetic profile (**Fig. 3.3a-n**), without invoking any adjustable parameters.

The overall success of our model raises the question of whether *E. coli* cells might have been configured to minimize the recovery time, since the on-off strategy employed in our model was shown theoretically by Pavlov and Ehrenberg (PE) to give the fastest recovery kinetics given an overall proteomic resource constraint. Using our model, we investigated the dependence of the expected recovery time on the allocation of AAB enzyme reserves to the different pathways in the pre-shift state (i.e., $q_n(0)$) for a fixed total abundance of AAB enzymes before the shift. The distribution of lag times obtained for 1000 randomly chosen sets of $q_n(0)$ values (**Fig. 3.18a,** blue bars) shows that that the mode of the distribution is nearly twice as long as the lag time with uniform $q_n(0)$ (dashed red line), where the AAB enzyme reserves are allocated in proportion to their final abundances. In fact, the lag time for the measured pre-shift reserve (**Fig. 3.18b**) falls on the long side of the mode of the lag-time distribution (solid red line). These results suggest that minimizing the recovery time of downshift to minimal media is not the primary concern for *E. coli* cells growing in rich media.

Additionally, in our model the allocation for ribosome synthesis ($\chi_{Rb}(t)$) was gradually turned on as AAB enzyme synthesis ($\chi_{A,tot}(t)$) was reduced (Fig. 3.5a, 3.5b), not abruptly switched on as prescribed by the on-off model. The portion of the proteome allocated to the ribosomes and AAB enzymes, which is the focus of this study, comprised < 30% of the total proteome (dashed line, Fig. 3.4b). The remainder of the proteome was allocated to such functions as carbon and AA transport, and general stress response (Fig. 3.9j-l), as well as other housekeeping functions carrying no metabolic fluxes in the growth conditions we studied. Many of these proteins are turned on before the AAB enzymes (grey circles, Fig. 3.3p). If we applied the PE model to all these proteins beyond the AAB enzymes, the cells would adapt much faster. Clearly, such a strategy is not adopted by E. coli. Instead, E. coli casts its response broadly, underscoring the importance of anticipating different nutrient types when AA supplies are exhausted in the native habitat. The kinetic responses exhibited by the various protein groups reflect the interplay of their respective regulatory mechanisms which are set by a variety of physiological and ecological reasons. The coarse-grained framework described in this work can be used to quantitatively capture the regulatory functions, and from them, quantitatively predict the recovery kinetics, based merely on proteome allocation in the pre- and post-shift steady state.

3.5 Methods

3.5.1 Strain construction

The strain used in this study are either wild type *E.coli* K-12 NCM3722 strain or its derivative, see **Table 3.1** for details.

3.5.2 Growth medium

Unless stated otherwise, all the growth media used in this study were based on MOPS buffered media described in Cayley et al (31). The medium contains 40mM Mops and 4mM

Tricine (adjust to pH 7.4 by NaOH), 0.1mM $FeSO_4$, 0.276 mM Na_2SO_4, 0.5 μM $CaCl_2$, 0.523 mM $MgCl_2$, and also micronutrients mixtures used in Neidhardt et al (22).

If not specified, the media used in steady-state proteomic measurements are MOPS buffered media with 50mM NaCl, 1.32mM KH_2PO_4, plus carbon source, nitrogen source and other nutrients shown on **Table 3.2**. Rich defined medium kit is obtained from Teknova M2105 which is based on Neidhardt et al (22).

For the proteomic measurements in AA downshift (18AA to none shift), the post-shift medium (referred as 'none') was MOPS buffered media with 50mM NaCl, 1.32mM KH_2PO_4, carbon source (0.4% glycerol) and nitrogen source (10mM ammonia chloride). The pre-shift medium was based on 'none' medium with additions of 18 amino acids (except cysteine and tyrosine). Amino acid concentrations are the same as those listed in the rich defined medium (22); see Supplementary Table 4. Cysteine is excluded because it oxidizes to cystine in aerobic growth. Tyrosine is excluded because of its small solubility. Among these 18 amino acids, asparagine and glutamine are prepared freshly for every experiments in case of fast degradation. Through this work, '18AA to none' shift has also been done with the carbon source other than 0.4% glycerol, e.x., 0.2% glucose, 0.2% lactose, 0.2% xylose or 0.2% fructose. If not specified, glycerol is taken to be the default carbon source for all the shift experiments shown in this work.

In some AA downshifts starts from less than 18AA (**Fig. 3.16b**), the post-shift medium was still 'none' medium. The pre-shift medium (usually referred as '18AA-but-xxx' with xxx being amino acid(s)) was based on the 'none' medium with the same 18 amino acids stated above except xxx (**Table 3.4**). The casamino acid (CAA) medium used in **Fig. 3.5f** was also based on 'none' medium, but with addition of 0.2% casamino acids (**Table 3.4**).

In the experiments requires cysteine and tyrosine, cysteine solution was prepared and added to the medium right before inoculating cells, and tyrosine was freshly dissolved in 0.1M KOH before added to the medium. The PH of the medium with tyrosine was still around 7.4.

3.5.3 Cell growth and medium shift

All the steady-state measurements and medium shift experiments were performed in a 37℃ water bath shaker shaking at 240 rpm. Shifts always started after cells reaching steady-state in pre-shift medium. To grow cells to exponential phase, we followed a standard cell growth round: seed culture, pre-culture and experimental culture. For seed culture, cells were inoculated to LB broth from fresh LB plate and grew for about 7hrs. Then cells were washed with corresponding Mops medium once and diluted to the same Mops medium (pre-culture) with starting OD_{600} around 10^-3. If inductions (e.x., 3-methylbenzyl alcohol, IPTG) were needed in this experiments, they were added to pre-culture too. After overnight pre-culture, cells were diluted to the same MOPS medium (experimental culture) with starting OD_{600} below 0.01. For exponential growth measurements, we let cells grow to OD_{600} 0.5. More than 5 OD_{600} points were taken within the range of OD_{600} 0.05~0.5 for growth rate calculation. For medium shift experiments, we let cells grow to about OD_{600} 0.3, then washed cells twice with pre-warmed post-shift medium and diluted it to OD_{600} 0.1 in pre-warmed post-shift medium. Whole medium transition process was kept within 5mins and we called it time 0 when cells were transferred to post-shift medium. A control shift from 18AA w/ glycerol to the same media was done and no lag or growth rate change was observed in the control shift. We took at least 5 OD_{600} points in pre-shift medium starting from OD_{600} 0.05. After medium transition, we took OD_{600} points every 5-20mins according to lag time until OD_{600} reached 0.5.

3.5.4 Growth rate calculation

For steady-state exponent growth rate, we simply fitted growth curve exponentially. If $R^2 < 0.999$, we didn't take it as reliable results.

To get instantaneous growth rate, we applied Savitzky-Golay filter (32) to N $\{t_i, Ln(OD_{600,i})\}$ data set. Recall that because of medium switch, OD_{600} was not continuously at t=0. Therefore we scaled the data in post-shift region before calculating growth rate to make the first OD_{600} after shift $(t < 5mins)$ equal to the value exponentially extrapolated from pre-shift data.

The window size we used includes 3 data points. First, we fitted the first 3 data points to the 2nd polynomial and substituted t of middle data point (t_2) into fitted equation to get its slope. Then we moved the window to the right and did the same thing for the 2nd to 4th data points to get the slope at t_3. Repeating moving windows and calculating slope, we got slopes at $t_2, t_3, \dots t_{N-1}$. Since we already did Ln operation on OD_{600}, the slopes we got are exponent instantaneous growth rates at $t_2, t_3, \dots t_{N-1}$. Cells were in steady-state at t_1 and t_N according to our experimental setting, so we only plotted data points from t_2 to t_{N-1} for instantaneous growth rate figures.

3.5.5 Measurements of proteins using mass spectrometry

We used quantitative mass spectrometry to characterize the abundance of E. coli proteomes at various growth conditions (**Table 3.2**).

For steady-state growing cells (sample ID M1-M4, P1, P8-P10, N5-N8), we followed the cell growth protocol stated above and took samples around $OD_{600} = 0.4$. For AA downshift experiments (sample ID S1-S5) executed following the protocol above, we took one sample (S1) at pre-shift steady-state around $OD_{600} = 0.3$. During the shift, we took samples at 15, 50,100, 200mins after the shift started. We collected at least 1 $OD_{600} \cdot ml$ cells for each sample by centrifuging at 13200 RPM for 2mins.

Quantitative measurements of E. coli proteomes and determination of protein mass fractions (**Table_proteomicdata.xls**) were carried out as described previously (2). Briefly, E. coli cell pellets were lysed with 2% sodium deoxycholate and digested with LysC and trypsin. An iRT peptide mix (Biognosys) was added to all samples for retention time alignment. Tryptic peptides were measured in SWATH mode (64 variable windows) on a TripleTOF 5600 mass spectrometer (Sciex). The DIA/SWATH data was analyzed using OpenSWATH (www.openswath.org). Protein mass fractions were computed from the detected peptide intensities using the xTop algorithm and a final scaling of xTop protein intensities with ribosome profiling data (2) (http://www.ebi.ac.uk/pride/archive/projects/PXD028559).

3.6 Data availability

The proteomic raw and data analysis files are publicly accessible through the ProteomeXchange Consortium via the PRIDE partner repository: http://www.ebi.ac.uk/pride/archive/projects/PXD028559 (reviewer username: reviewer_pxd028559@ebi.ac.uk; password: Q5KSl2Ry). The E. coli spectral library used for DIA/SWATH data analysis has been published previously (2) and is available via SWATHAtlas: http://www.peptideatlas.org/PASS/PASS01421.

3.7 Acknowledgments

We are grateful to Xiongfeng Dai, Vadim Patsalo, and Rohan Balakrishnan for helpful discussions and technical assistance during the course of this work. This research is supported by the NSF Grant MCB 1818384 through TH. CL has been supported by EPIC-XS funded by the Horizon 2020 program of the European Union (project number 823839). RA and CL acknowledge the support of the European Research Council (Proteomics4D: AdvG grant 670821 and Proteomics v3.0: AdvG-233226).

Chapter 3, in full, has been submitted for publication in Nature Microbiology in 2022. The authors are Chenhao Wu, Zhongge Zhang, Amir Banaei-Esfahani, Hiroyuki Okano, Matteo Mori, Ruedi Aebersold, Christina Ludwig, Terence Hwa. The dissertation author was the primary investigator and author of this material.

3.8 Figures and Figure legends

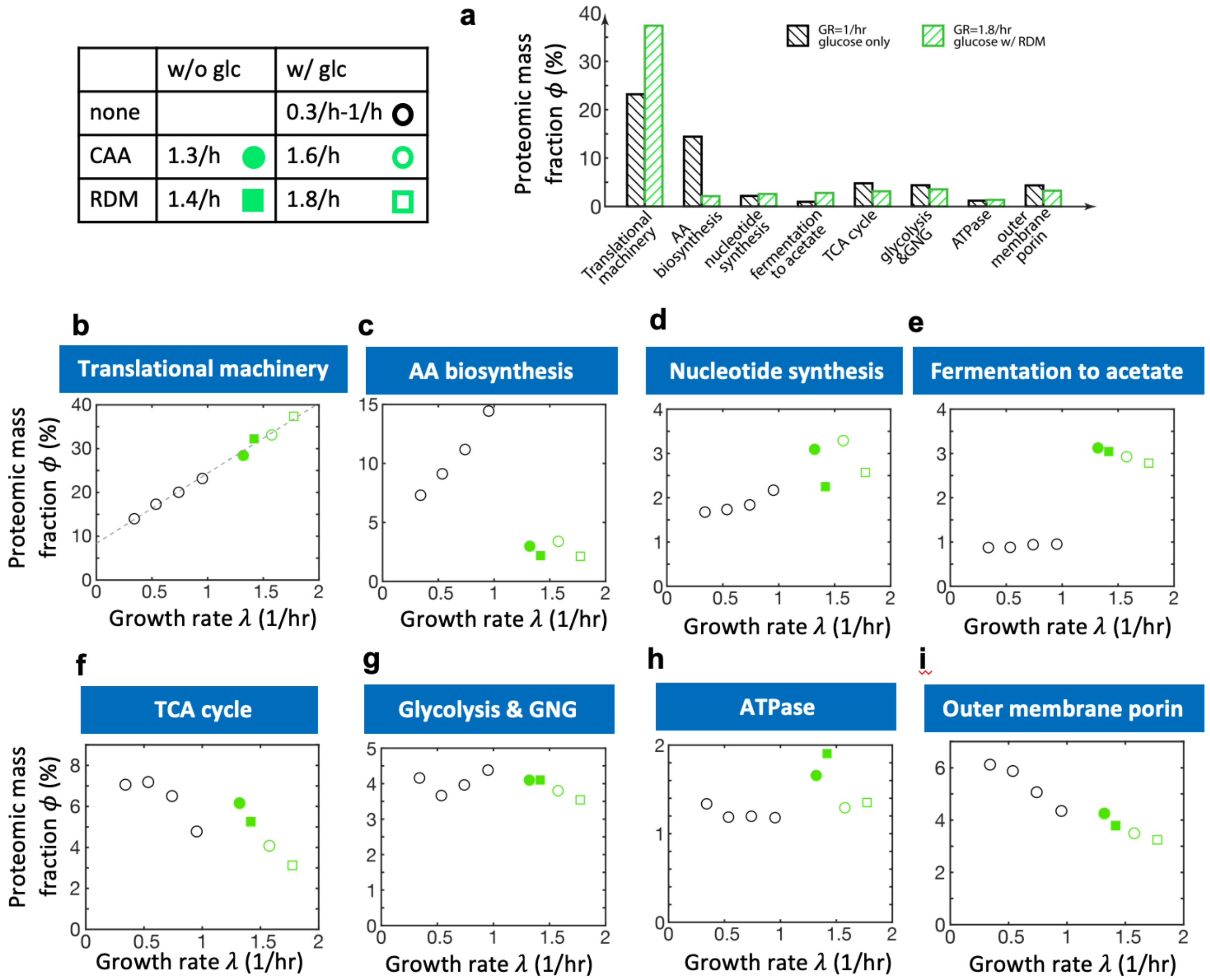


Figure 3.1: Steady-state protein abundance vs growth rates. The legend table indicates symbols used for the different growth conditions studied (see **Table 3.2**): Rich-defined medium (RDM) with glucose (open green square) and without glucose (filled green square), casamino acid (CAA) with glucose (open green circle) and without glucose (filled green circle), and minimal medium with glucose as the sole carbon source for wild type cells and mutants (open black circle) with limited expression of glucose uptake protein PtsG (10, 17); see **Table 3.1** for strain details. The values in each table entry indicates the steady-state growth rate of that culture. **(a)** Proteomic allocation comparison of WT strain between glucose minimal media (black bars) and glucose w/ RDM (green bars). The mass fraction changes in translational machinery and AAB enzymes are dominant compare to others. **(b)-(i)** Total abundance of proteins in various protein groups, as obtained from mass spectrometry (Methods) and reported as fraction of total protein mass ("proteome fraction"), is plotted against the growth rate of each culture. Classification of proteins into groups are specified in **Table 3.3** according to Refs. (41, 42).

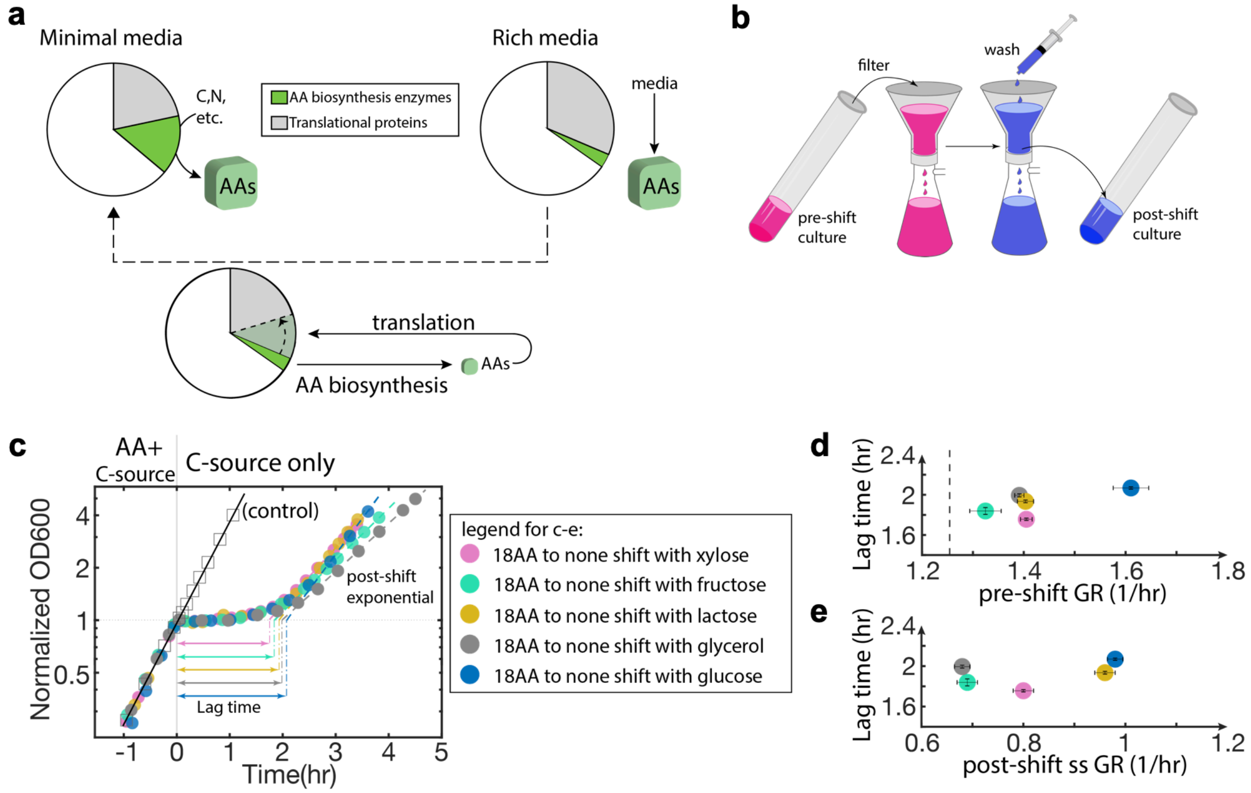
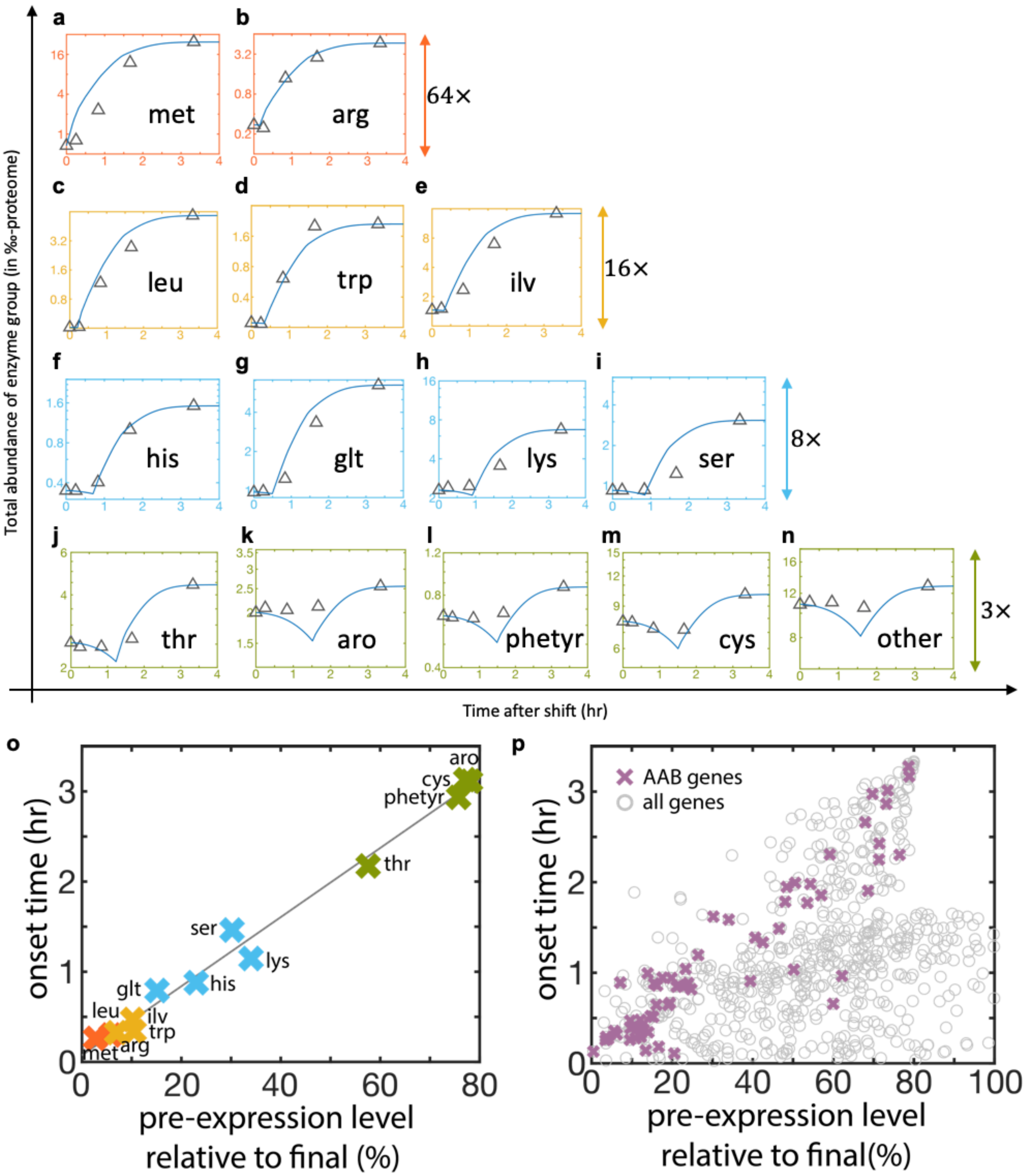


Figure 3.2: (a) Remodeling of the proteome needed during AA downshift. Pie charts shows the steady-state proteomic allocation in glucose minimal media and in CAA supplemented with glucose. Grey sector represents translational proteins, green sector represents AA biosynthesis (AAB) enzymes, and the white sector represents other proteins. In minimal media, cells need to synthesize AAs using AAB enzymes, which are expressed at low levels in rich media. When the environments suddenly change from rich media to minimal media (dash line), cells suffer from the shortage of both AAs and AAB enzymes. Limited AAB enzymes provide limited AA fluxes, which limit the synthesis of all proteins including the AAB enzymes. **(b)** shows a schematic diagram of the medium shift performed in this study, with the entire duration of the shift taking < 5 min. **(c)** Growth curves for transition from 18AA (**Table 3.4**) to none, for WT strain, while maintaining one of the carbon sources: glucose, lactose, fructose, xylose, glycerol throughout the shift. Time 0 is the downshift time. “Lag time” is defined by extrapolating the exponential growth of post-shift steady state to the OD of the culture at the time of shift. The open squares shows the results for control experiment from 18AA w/ glycerol to the same medium, showing that the filter-wash process in **(b)** does not affect cell growth. **(d)-(e)** scatter plots of lag times against pre-shift and post-shift steady-state growth rates for the shifts shown in **(c)** using the same color scheme. The dash line in panel d shows the growth rate of the culture on 18AA without additional carbon source.

Figure 3.3: AAB enzymes are sorted into fourteen groups according to the pathways they are in. The total abundance of each group, expressed in term of the fraction of total protein mass, is plotted as triangles in panels **(a)-(n)** at time t=0, 15 min, 50 min, 100 min, 200 min after shift. See **Fig. 3.10** for the abundances of individual enzymes in each group. The data are plotted in log scales, with the color of the axis representing the fold-change from the minimal to maximal values on the y-axis (orange: 64 ×, yellow: 16 ×, blue: 8 ×, green: 3 ×). The total abundance before the shift is plotted at time 0. Panel **(o)** shows the onset time of each enzyme group, as defined in **Fig. 3.12a**, plotted against the pre-shift abundance level relative to final post-shift level (the reciprocal of the fold change), referred to as the "pre-shift reserve". The gray line is the linear fit and the correlation coefficient is 0.994. The marker color in **(o)** matches the axis colors used in panels **(a)-(n)**. Panel (p) shows the onset time vs "pre-shift reserve" for all the up-regulated genes during the shift (circles), in which AAB genes are marked by crosses. The correlation coefficient of AAB genes between onset time and "pre-shift reserve" is 0.904. Panel **(p)** shows that the AAB enzymes have among the longest onset time, suggesting that they are expressed as needed while many other proteins are pre-programmed to increase in expression soon after the shift, regardless of the need. The blue lines in **(a)-(n)** are the predicted enzyme kinetics according to the model described in the text. For some enzyme groups arising later (e.g., cys, aro), model predicted a dip during recovery due to the dilution from cell growth. Note that tyrosine and cysteine were provided in the pre-shift growth medium in the experiment.

a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
p
met
arg
leu
trp
ilv
his
glt
lys
ser
thr
aro
phetyr
cys
other
64×
16×
8×
3×
Total abundance of enzyme group (in ‰-proteome)
Time after shift (hr)
onset time (hr)
pre-expression level relative to final (%)
pre-expression level relative to final(%)
AAB genes
all genes

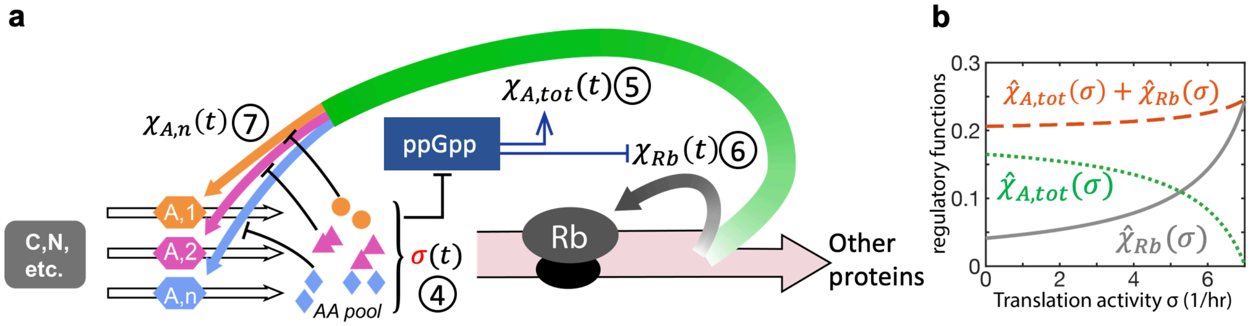


Figure 3.4: Model of dynamic proteome allocation for AA downshift. (**a**) Upon removal of AAs from the medium, cells synthesize AAs *de novo* from carbon, nitrogen substrates using AA biosynthesis (AAB) enzymes, indicated as A.1, A.2, etc. The AAs synthesized (indicated as circles, triangles, and diamonds) are consumed by ribosomes (Rb) for protein synthesis. Limited AAs synthesis fluxes lead to limited AAs pool and reduced ribosome activity, denoted by $\sigma(t)$ and defined mathematically by Eq. (4). $\sigma(t)$ serves here as a "signal" reflecting the AA pools, analogous to the role played by the (inverse of) the alarmone ppGpp (23), and directs the allocation of protein synthesis flux to AAB enzymes and ribosomal proteins, indicated by the thick green and grey arrows, respectively. The latter are described by the allocation functions, $\chi_{A,tot}(t)$ and $\chi_{Rb}(t)$, respectively, and set by the regulatory functions $\hat{\chi}_{Rb}(\sigma)$ and $\hat{\chi}_{A,tot}(\sigma)$ according to Eqs. (5) and (6). The protein synthesis flux directed to AAB enzymes is further allocated to the enzymes on each pathways by the pathway-specific allocation function $\eta_n(t)$ through Eq. (7). As explained in detail in **Fig. 3.15**, $\eta_n(t)$ is set according to the end-product inhibition so that the protein synthesis flux is directed to the enzyme groups limiting growth at the current time. **(b)** shows the forms of the regulatory functions $\hat{\chi}_{Rb}(\sigma)$ and $\hat{\chi}_{A,tot}(\sigma)$, solid gray and dotted green lines, respectively, constructed through the relation between the steady-state growth rate and the abundances of the corresponding protein sectors as detailed in **Fig. 3.13** and **Fig. 3.14**. Note that the sum of $\hat{\chi}_{Rb}(\sigma)$ and $\hat{\chi}_{A,tot}(\sigma)$, shown as the red dash line, is only 20~25% of the total proteome. Substantial flux is also directed to the synthesis of other proteins such as stationary phase proteins and generic transporters; see **Fig. 3.9j-l**. However, these other proteins do not carry AA flux and therefore need not be included in our model.

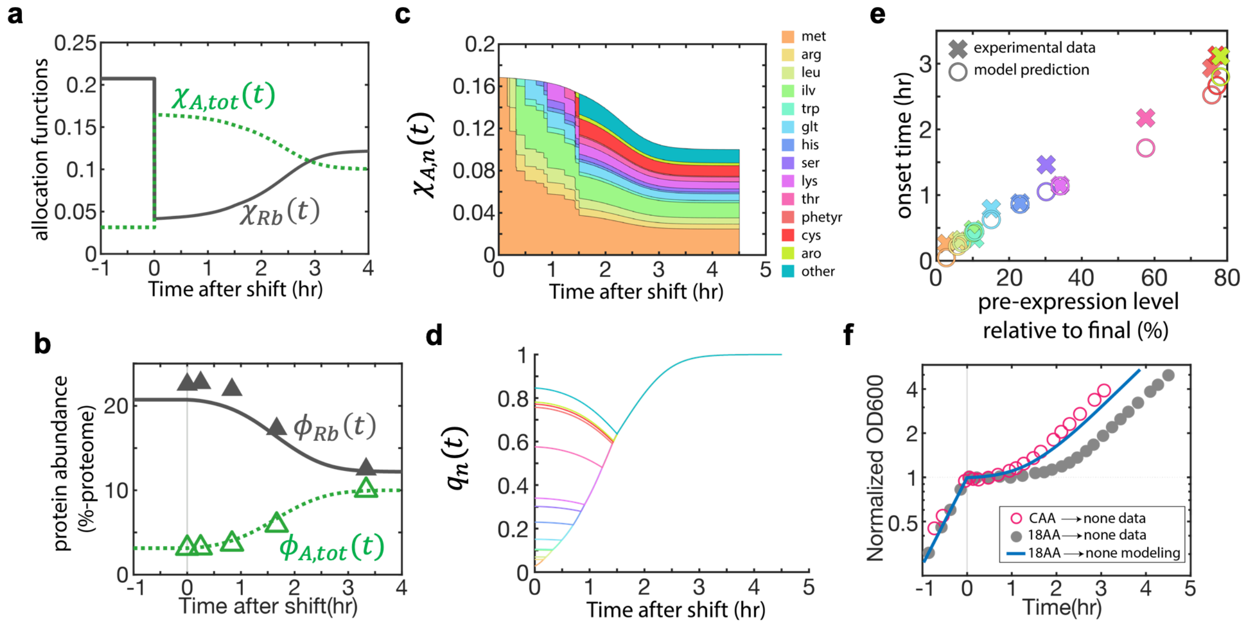


Figure 3.5: (a) Time courses of the allocation functions $\chi_{Rb}(t)$ (solid grey line) and $\chi_{A,tot}(t)$ (dotted green line), obtained according to the solution of the complete model defined by Eqs. (1)-(8). **(b)** Time courses of ribosomal protein abundance and total AAB enzyme abundance for model predictions (lines) and experimental data (triangles). Green represents AAB enzyme, while gray represents ribosomal proteins. **(c)** Time courses of the predicted regulatory functions (stacked) for each AAB group. The areas with different colors represent the allocation functions for the different AAB groups as described by the legend on the right, ordered from the smallest to the largest relative pre-expression level $q_n(0)$. The top curve represent the time course of $\chi_{A,tot}$ (same as the green curve in panel d) since it's the sum of $\chi_{A,n}$. **(d)** Time courses of the relative abundance $q_n(t) = \phi_{A,n}(t)/\phi_{A,n}^*$ for each AAB group n. Eventually all the $q_n(t)$ reach one, indicating cells reach the final steady-state expression levels. **(e)** compares the model prediction of onset time (circles) with experimental data (crosses). The color scheme used in **(d)** and **(e)** is the same as that in **(c)**. **(f)** Model predicted '18AA-to-none' growth curve (blue line) is compared with measured '18AA-to-none' growth curve (filled gray circles) and measured 0.2% casamino acid to none growth curve (open red circles). Model prediction better captured the lag time of 0.2% casamino acid to none shift.

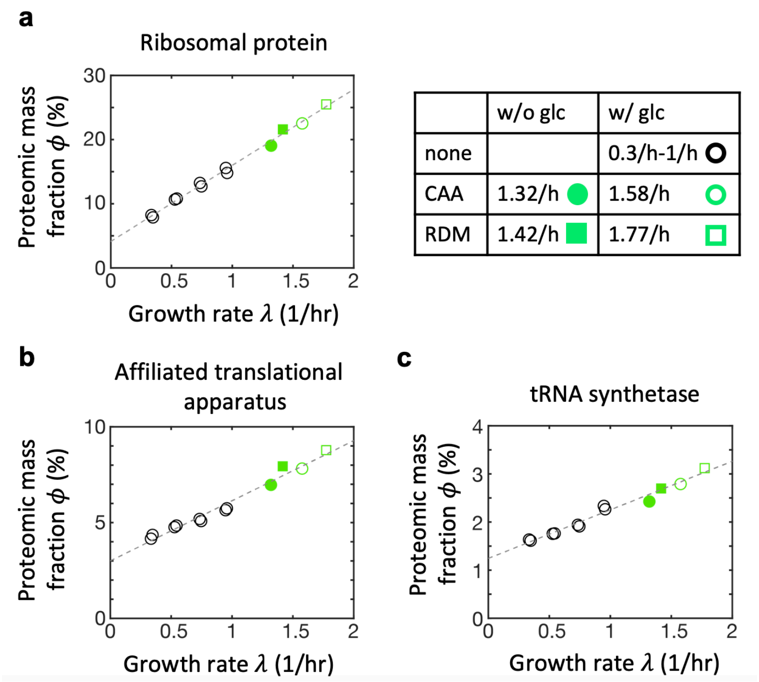
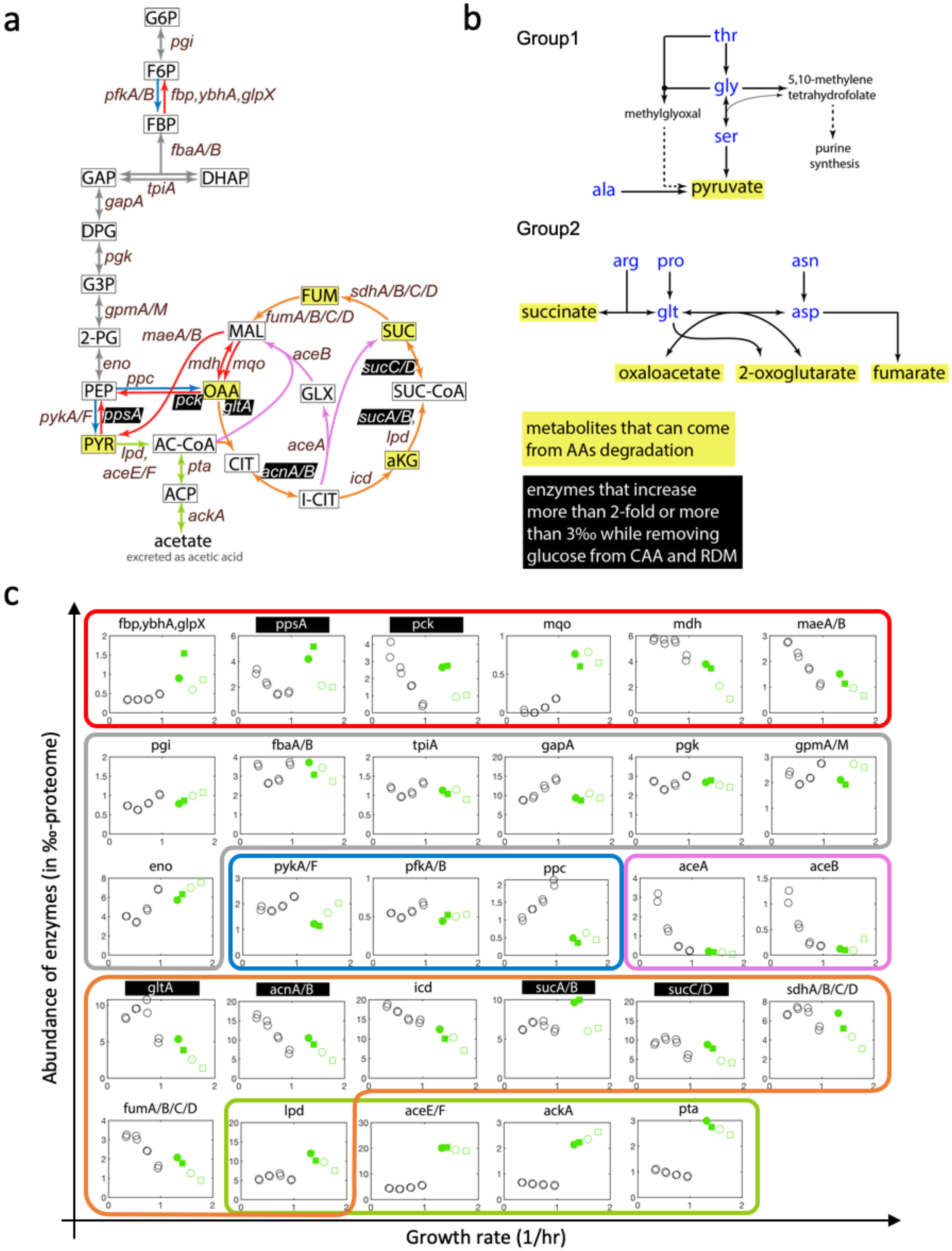


Figure 3.6: Translational machinery in **Fig. 3.1** is composed of **(a)** ribosomal proteins, **(b)** affiliated translational apparatus, and **(c)** tRNA synthetase, whose abundances are plotted against the growth rate under carbon limitations (black open circles) and four rich conditions (green symbols as shown in the legend table). See **Table 3.2** for the details of growth conditions and **Table 3.3** for gene classifications. The same symbols are used in **Fig. 3.1** and **Fig. 3.7**.

Figure 3.7: Metabolic pathways related to growth on amino acids. **(a)** Enzymes and metabolites involved in Central Carbon Metabolism (CCM). Metabolites are boxed, with yellow background indicating those which can be produced from AA degradation. Reactions are represented by colored arrows with the genes encoding the corresponding enzymes written next to it in Italic. The color of the arrows corresponds to the box color panel **(c)**, with gray indicating reversible reactions involved in both glycolysis and gluconeogenesis, red indicating gluconeogenesis only, blue indicating glycolysis only, orange indicating TCA cycle, pink indicating glyoxylate shunt, and green indicating pyruvate fermentation to acetate. **(b)** AA degradation pathways, including those producing CCM metabolites such as pyruvate (Group1), and metabolites in TCA cycle (Group2). **(c)** The abundances of CCM enzymes in steady state growth, plotted against the growth rate. The symbols are the same as those used in **Fig. 3.6**. The gene names shown with a black background (in both (a) and (c)) are enzymes that increased by more than 2-fold or by more than 3‰ in CAA or RDM without glucose supplement compared to those with glucose supplement (i.e., increased in filled green symbols compared to open green symbols). As can be seen from panel (a), these genes with increased enzyme levels in the absence of glucose supplement are mostly adjacent to the metabolite that are products of AA degradation (those with yellow background). This suggests that the AA degradation products are funneled through CCM to supply the cell's non-AA carbon needs (e.g., energy biogenesis and the biosynthesis of lipids and cell wall components) via gluconeogenesis (enzymes encoded by *ppsA* and *pck*) and via TCA (enzymes encoded by *gltA*, *acnAB*, *sucABCD*).


a
G6P
pgi
F6P
pfkA/B
fbp,ybhA,glpX
FBP
fbaA/B
GAP
DHAP
tpiA
gapA
DPG
pgk
G3P
gpmA/M
2-PG
eno
PEP
ppc
pykA/F
ppsA
PYR
lpd,
aceE/F
AC-CoA
pta
ACP
ackA
acetate
excreted as acetic acid
maeA/B
MAL
mdh
mqo
OAA
pck
gltA
FUM
fumA/B/C/D
sdhA/B/C/D
SUC
aceB
GLX
sucC/D
SUC-CoA
sucA/B,
lpd
aceA
CIT
acnA/B
I-CIT
icd
aKG
b
Group1
thr
gly
5,10-methylene
tetrahydrofolate
methylglyoxal
ser
purine
synthesis
ala
pyruvate
Group2
arg
pro
asn
succinate
glt
asp
oxaloacetate
2-oxoglutarate
fumarate
metabolites that can come from AAs degradation
enzymes that increase more than 2-fold or more than 3‰ while removing glucose from CAA and RDM
c
Abundance of enzymes (in ‰-proteome)
Growth rate (1/hr)
fbp,ybhA,glpX
ppsA
pck
mqo
mdh
maeA/B
pgi
fbaA/B
tpiA
gapA
pgk
gpmA/M
eno
pykA/F
pfkA/B
ppc
aceA
aceB
gltA
acnA/B
icd
sucA/B
sucC/D
sdhA/B/C/D
fumA/B/C/D
lpd
aceE/F
ackA
pta

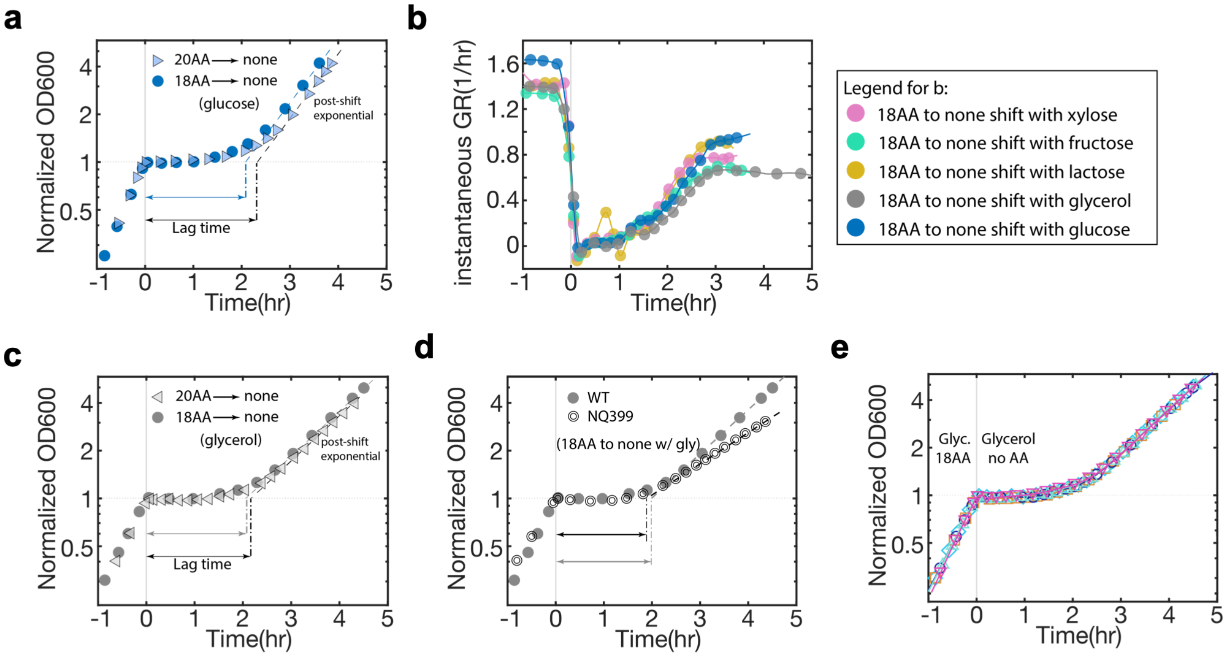


Figure 3.8: Amino acid downshifts with supplement of various carbon sources. (a) Comparison of ‘20AA-to-none’ shift to ’18AA-to-none’ shift (excluding cysteine and tyrosine due to poor solubility; see **Methods** and **Table 3.4**), both supplemented by 0.2% glucose before and after the shift, shows minor effects due to exclusion of cysteine and tyrosine. **(b)** Instantaneous growth rate of ‘18AA-to-none’ shift supplemented by various carbon sources shows nearly identical temporal profile. These carbon sources enter upper glycolysis at various points and use a variety of different uptake mechanisms (see **Table 3.5**). **(c)** AA downshift experiments of panel **(a)** is repeated with glucose replaced by 0.4% glycerol as the carbon supplement. The growth transition is nearly the same for both 20AA-to-none and 18AA-to-none shifts. **(d)** Comparison of the growth curve of ’18AA-to-none’ transition supplemented by 0.4% glycerol between WT strain and a mutant (NQ399, see **Table 3.1**) harboring the full expression of glycerol uptake proteins GlpFK. 1mM IPTG and 1mM 3MBA is added to the medium before and after the shift for the shift by NQ399 so that *glpFK* expression is fully induced throughout the shift. The observed lag time of NQ399 is the same as the lag time of WT strain, meaning that glycerol uptake is not the bottleneck during ‘18AA-to-none’ shift. **(e)** Repeatability of ’18AA-to-none’ shift supplemented by glycerol. Symbols of different colors represent results of independent experimental runs.

Figure 3.9: The total abundance of proteins from individual pathways or functional groups during the '18AA-to-none' shift supplemented by glycerol are shown in panels **(a)-(l)**. **(m)** Instantaneous growth rate following the shift. Membership of each protein group is shown in **Table 3.3**. All protein abundances are reported in mass fraction (%-proteome). Samples were collected at 15 min, 50 min, 100 min, 200 min after the shift. The pre-shift data points are shown at time 0.

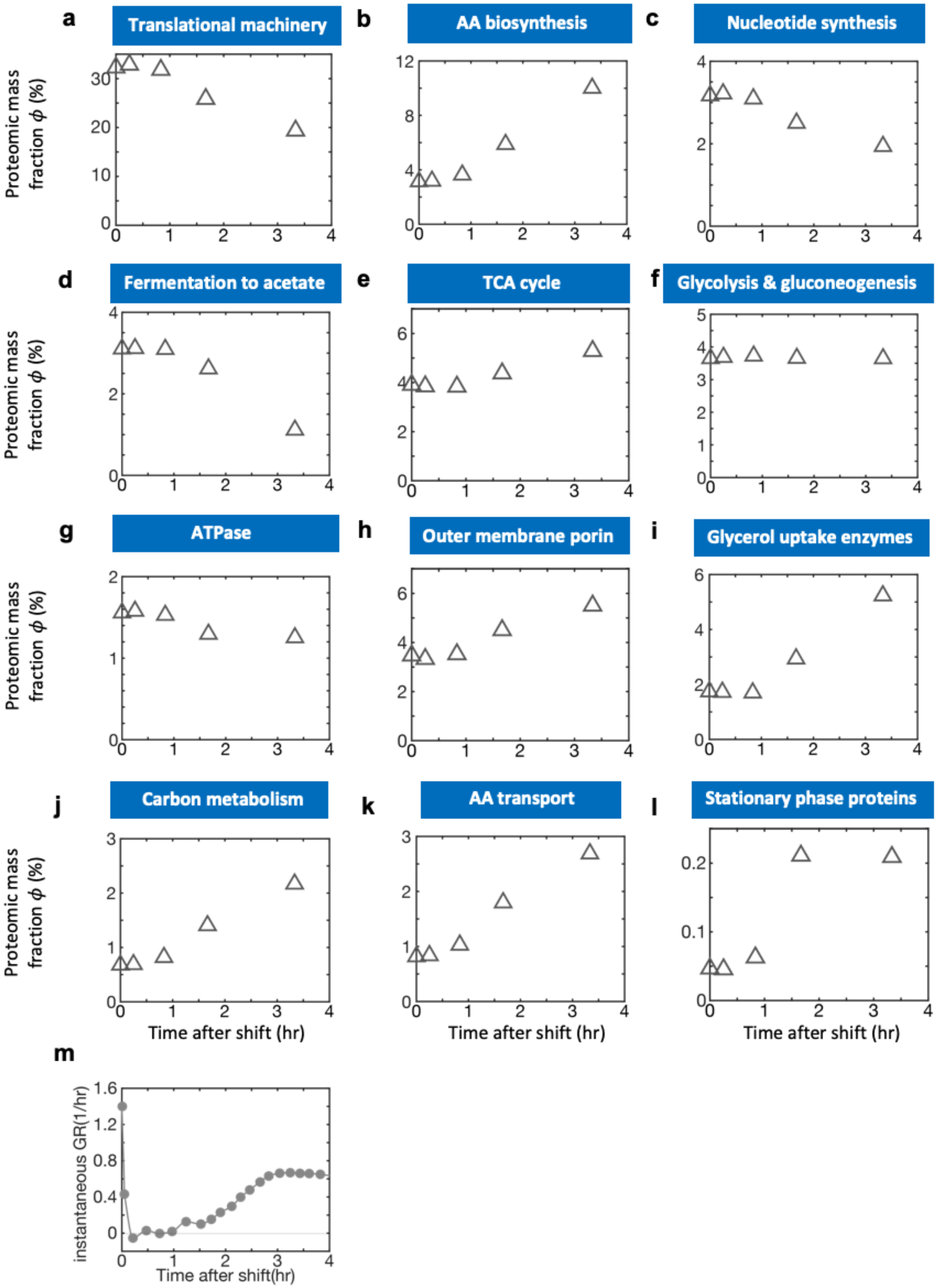

a
Translational machinery
b
AA biosynthesis
c
Nucleotide synthesis
d
Fermentation to acetate
e
TCA cycle
f
Glycolysis & gluconeogenesis
g
ATPase
h
Outer membrane porin
i
Glycerol uptake enzymes
j
Carbon metabolism
k
AA transport
l
Stationary phase proteins
m
Proteomic mass fraction ϕ (%)
Time after shift (hr)
instantaneous GR(1/hr)
Time after shift(hr)
0
1
2
3
4
10
20
30
8
12
6
5
0.1
0.2
0.4
0.8
1.2
1.6

Figure 3.10: The abundance time courses for individual AA biosynthesis enzymes during '18AA-to-none' shift supplemented by glycerol. The enzymes in the same pathway are plotted in the same plot. For those enzymes not detected at some time points, we used abundance values 10x smaller than the lowest abundance detected among all enzymes in the pathway. Most of these replacements are in the arginine pathway **(b)**. In each panel, the lines in the legend box next to the gene name indicates the operon structure. The same line style means the genes are on the same operon. E.x., for panel **(m)**, cysC and cysN are on the same operon, cysHIJ are on the same operon, while cysDEKM are each on different operons. The genes in different groups don't share operons except serC-aroA and aroF-tyrA. In each panel, the similar color of the recovery curves indicating similar kinetic behavior. E.x., in **(i)**, serA and serC (both in red) have similar kinetics, while serB (in blue) is different from them.

Generally, we can see that the similar kinetic behavior in **(c-d,f,j)** are all attributed to operon structure. Similar kinetic behavior are also seen in panel **(g,i,m-n)** although the genes are not on the same operon. To be detail, in the met group **(a)**, changes in the abundance of MetE dominated the entire group. In the arg group **(b)**, most enzymes showed large increase between 15 min and 50 min after the shift, except ArgA, ArgD, ArgH (show in bluish colors). The total abundance of this group was mainly determined by the abundance of ArgG and ArgI. In the ilv group **(e)**, four enzymes (in red) exhibited moderate increase in the first 50 min, while the rest (in blue) increased only after 50 min. Since IlvC is much more abundant than other enzymes, the group kinetics was largely determined by this enzyme alone. In the lys group **(h)**, three enzymes (in blue, DapA, DapE, DapF) were approximately constant throughout the shift. The rest (in red) started increasing after 50 min, most of which showed only small increases during the shift, except for LysC, with a big jump between 50 min and 100 mins. But since LysC is not the most abundant enzyme in the lys group, it did not significantly affect the group kinetics itself. In the aro group **(k)**, most of the enzymes (in blue) stayed constant or dropped during the shift. Three enzymes (in red) clearly increased during the shift, among which AroG was the most significant. In the phetyr group **(l)**, decrease in TyrA could likely be attributed to the fact that tyrosine was not provided in the pre-shift medium, so that TyrA was fully expressed before the shift and was then shutoff (with abundance decreasing due to dilution) during the growth recovery phase.

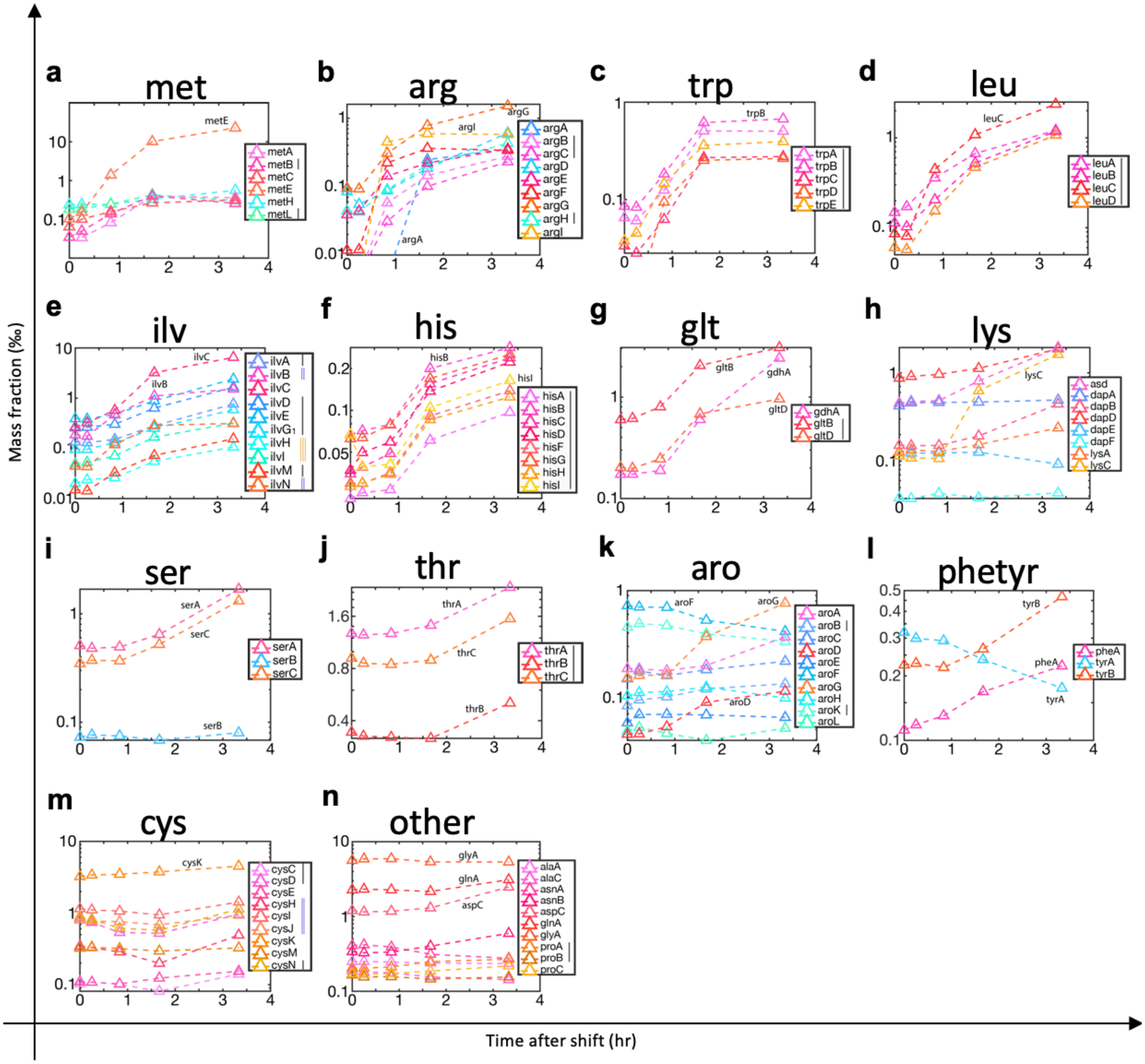

Mass fraction (‰)
Time after shift (hr)
a
met
metE
metA
metB
metC
metE
metH
metL
b
arg
argG
argI
argA
argB
argC
argD
argE
argF
argG
argH
argI
c
trp
trpB
trpA
trpB
trpC
trpD
trpE
d
leu
leuC
leuA
leuB
leuC
leuD
e
ilv
ilvC
ilvB
ilvA
ilvB
ilvC
ilvD
ilvE
ilvG₁
ilvH
ilvI
ilvM
ilvN
f
his
hisB
hisI
hisA
hisB
hisC
hisD
hisF
hisG
hisH
hisI
g
glt
gltB
gdhA
gltD
gdhA
gltB
gltD
h
lys
lysC
asd
dapA
dapB
dapD
dapE
dapF
lysA
lysC
i
ser
serA
serC
serB
serA
serB
serC
j
thr
thrA
thrC
thrB
thrA
thrB
thrC
k
aro
aroF
aroG
aroD
aroA
aroB
aroC
aroD
aroE
aroF
aroG
aroH
aroK
aroL
l
phetyr
tyrB
pheA
tyrA
pheA
tyrA
tyrB
m
cys
cysK
cysC
cysD
cysE
cysH
cysI
cysJ
cysK
cysM
cysN
n
other
glyA
glnA
aspC
alaA
alaC
asnA
asnB
aspC
glnA
glyA
proA
proB
proC
0
1
2
3
4
0.01
0.05
0.1
0.2
0.3
0.4
0.5
0.8
1.6
1
10

Figure 3.11: Relation between onset time and expression levels. (a) The total abundance of the trp group relative to the pre-shift total abundance is plotted at different times after the shift (same data as **Fig. 3.3d**. Between the measured data points, we used linear interpolation (gray dash lines) to connect the data points. The green arrow shows the fold-change in enzyme abundance at the final time point (200 min) from the pre-shift level for the trp pathway. The red arrow indicates the onset time, defined as the time when the enzyme abundance exceeds the pre-shift value by 25%. **(b)** The onset time is plotted against the fold change for each pathway. The marker color is the same as the axis color used in **Fig. 3.3**, grouped according to the fold-change. **(c)** The onset time is plotted against the inverse of fold change, which is the pre-expression abundance relative to the final abundance of each enzyme group. The crosses in green are the same as the data in panel **(b)**. The others are results obtained by using different thresholds for the definition of onset time, blue crosses for 10% increase over pre-shift value, and red crosses for 50% increase. Approximately linear relations are seen between the plotted quantities regardless of the threshold value used in the definition of the onset time. **(d)** Same plot as panel **(b)**, but for individual AAB enzymes. **(e)** Same plot as panel **(c)** (25%), for individual AAB enzymes. The correlation coefficient is label as R. **(f-h)** show plots of the onset time against the pre-shift, the post-shift abundances and the increase from pre- to post- shift for individual AAB enzymes, respectively. The correlation coefficient between the onset time and the pre-shift abundance (or the log of abundance) is 0.48 (or 0.62). The correlation coefficient between the onset time and the post-shift abundance (or the log of abundance) is -0.12 (or -0.02). The correlation coefficient between the onset time and the increase abundance from pre- to post-shift (or the log of abundance) is -0.21 (or -0.33). They are all much weaker compared to panels **(e)**. Due to the definition of the onset time in **(a)**, the enzymes with fold change less than 1.25 or undetected in the pre-shift condition do not have defined onset times, and are thus excluded in panels **(d)-(g)**. **(i-k)** show plots of the onset time with the pre-shift, the post-shift abundance and the increase abundance from pre- to post- shift for all enzymes of each pathway, respectively. The marker color in **(i-k)** are the same as those used in panel **(b)**.

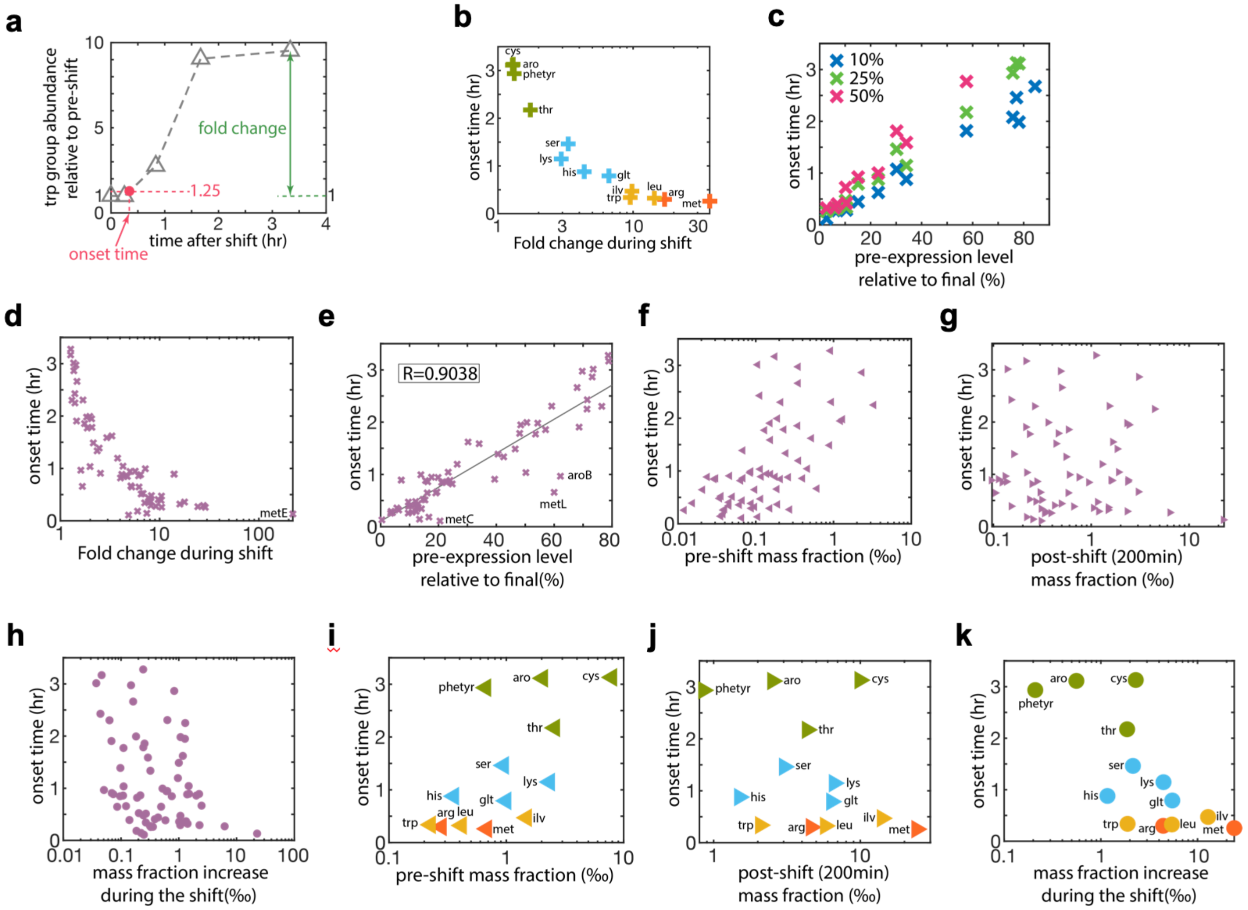

a
trp group abundance relative to pre-shift
time after shift (hr)
fold change
1.25
onset time
b
onset time (hr)
Fold change during shift
cys
aro
phetyr
thr
ser
lys
his
glt
ilv
trp
leu
arg
met
c
onset time (hr)
pre-expression level relative to final (%)
10%
25%
50%
d
onset time (hr)
Fold change during shift
metE
e
R=0.9038
onset time (hr)
pre-expression level relative to final(%)
aroB
metL
metC
f
onset time (hr)
pre-shift mass fraction (‰)
g
onset time (hr)
post-shift (200min) mass fraction (‰)
h
onset time (hr)
mass fraction increase during the shift(‰)
i
onset time (hr)
pre-shift mass fraction (‰)
phetyr
aro
cys
thr
ser
lys
his
glt
ilv
trp
arg
leu
met
j
onset time (hr)
post-shift (200min) mass fraction (‰)
phetyr
aro
cys
thr
ser
his
lys
glt
trp
arg
leu
ilv
met
k
onset time (hr)
mass fraction increase during the shift(‰)
aro
cys
phetyr
thr
ser
lys
his
glt
trp
arg
leu
ilv
met

Figure 3.12. Growth rate during AA downshift. If f_n represents the frequency of the n$^{\text{th}}$ AA, then the consumption flux of this AA by protein synthesis is given by $f_n \frac{dM}{dt}$, where $M(t)$ is the total protein mass. The synthesis flux of this AA is given by $k_n(t) M_{A,n}(t)$, where k_n is the specific catalytic rate of the pathway and $M_{A,n}$ is the total enzyme mass associated with the synthesis of this AA. The condition of flux balanced, obtained by equating the two fluxes, can be written as

$$\lambda(t) f_n = k_n(t) \phi_{A,n}(t), \tag{3.9}$$

using the definition of the instantaneous growth rate $\lambda(t) \equiv \frac{d}{dt} \ln M$ and proteome fraction $\phi_{A,n}(t) \equiv M_{A,n}(t)/M(t)$ introduced in the main text. In steady-state growth, Eq. (3.9) can be written as

$$\lambda^* f_n = k_n^* \phi_{A,n}^*, \tag{3.10}$$

where the asterisks indicate the steady-state values of the variables. We take f_n to be constant and the catalytic rates $k_{A,n}^* \geq k_{A,n}(t)$ to be the maximum catalytic rate, since during the period of growth recovery, some of the rates could be reduced due to allosteric inhibition.

In the illustration above, we plotted the quantity $k_n^* \phi_{A,n}(t)/f_n$, which is a measure of the maximum synthesis flux of the n$^{\text{th}}$ AA weighted by the frequency of this AA. These weighted frequencies allow flux comparisons across different AA groups. In steady state where $\phi_{A,n}(t) = \phi_{A,n}^*$, the weighted flux must be the same for all AAs according to Eq. (3.10). This is indicted by the black line. During the transition period after shift, the enzyme abundances $\phi_{A,n}(t)$ are below the steady state level by different amounts, resulting in differences in the maximum fluxes as illustrated by the dashed purple line at some time t. The overall flux of protein synthesis is limited by the lowest flux among these groups (here it is group 3 in the figure, with $k_3(t) = k_3^*$), indicated by the green line. This can be written mathematically as

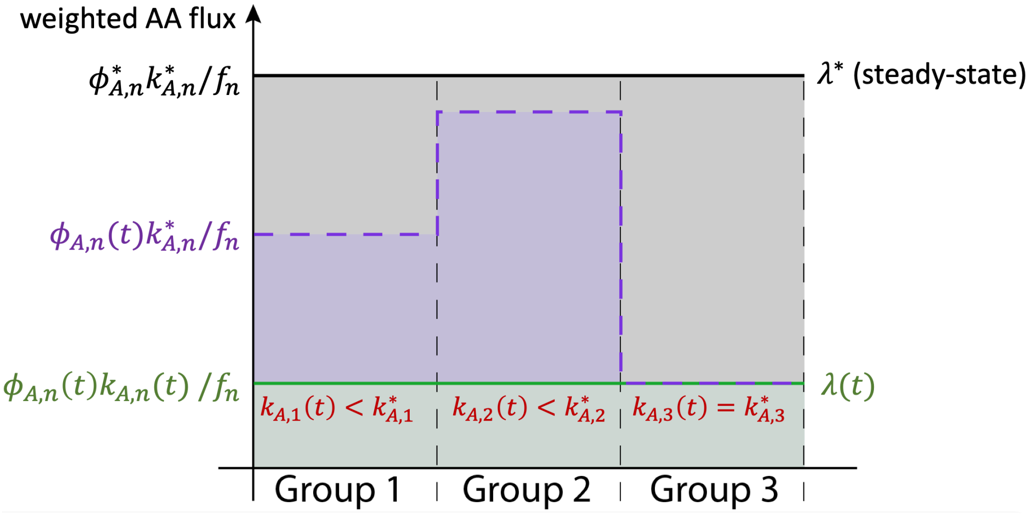
$$\lambda(t) = \min_{\mathrm{n}} \left[k_n^* \phi_{A,n}(t)/f_n \right], \tag{3.11}$$

which reduces to

$$\lambda(t) = \lambda^* \min_{\mathrm{n}} \left[\phi_{A,n}(t)/\phi_n^* \right] \tag{3.12}$$

using Eq. (3.10). Eq. (3.12) is Eq. (3.3) of the main text.

Note that in order for all AA groups to have the same weighted flux at all time during the recovery period, the catalytic rates of the non-growth limiting groups (here $k_1(t)$, $k_2(t)$) must be reduced from their maximum values k_1^*, k_2^*, respectively. This is expected to occur via end-product inhibition since those AAs with maximum synthesis flux above the minimum would synthesize more AAs than needed and those AAs in surplus are known to inhibit their own synthesis allosterically (44).


weighted AA flux
$\phi^*_{A,n} k^*_{A,n} / f_n$
λ^* (steady-state)
$\phi_{A,n}(t) k^*_{A,n} / f_n$
$\phi_{A,n}(t) k_{A,n}(t) \, / f_n$
$\lambda(t)$
$k_{A,1}(t) < k^*_{A,1}$
$k_{A,2}(t) < k^*_{A,2}$
$k_{A,3}(t) = k^*_{A,3}$
Group 1
Group 2
Group 3

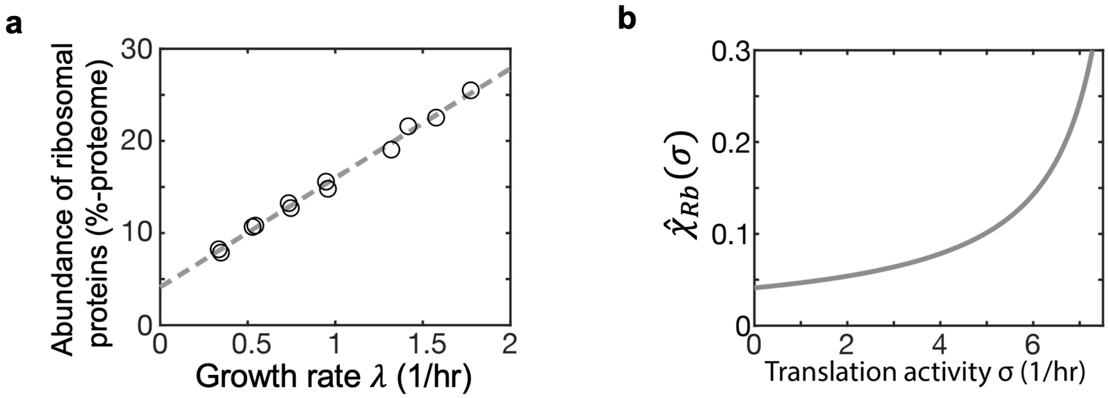


Figure 3.13: Construction of the regulatory function for ribosome synthesis, $\hat{\chi}_{Rb}(\sigma)$. (a) shows the abundance ϕ^*_{Rb} of ribosomal proteins at different growth rate (λ^*) for cultures growing exponentially in different nutrient sources. [Same data as that shown in **Fig. 3.6a**.] The data is well-captured by a linear fit (dashed line), i.e., $\phi^*_{Rb} = \phi_{Rb,0} + \lambda^*/\gamma$, with $\phi_{Rb,0} = 4.12\%$ and slope $\gamma = 8.43\ h^{-1}$. In steady state, we have $\chi^*_{Rb} \equiv \phi^*_{Rb}$, obtained by setting $\frac{d}{dt}\phi_{Rb} = 0$ in Eq. (3.2) of the main text. Using the definition of σ in Eq. (3.4),

$$\sigma^* = \frac{\lambda^*}{\phi^*_{Rb}} = \frac{\lambda^*}{\phi_{Rb,0} + \lambda^*/\gamma} \tag{3.13}$$

we can invert Eq. (3.13) to obtain $\lambda^*(\sigma^*)$, and obtain the regulatory function as

$$\hat{\chi}_{Rb}(\sigma) = \chi^*_{Rb}(\sigma) = \phi^*_{Rb}\big(\lambda^*(\sigma)\big) = \frac{\phi_{Rb,0}}{1 - \sigma/\gamma}. \tag{3.14}$$

The function obtained is sketched as the gray line in panel **(b)** and in **Fig. 3.4b**.

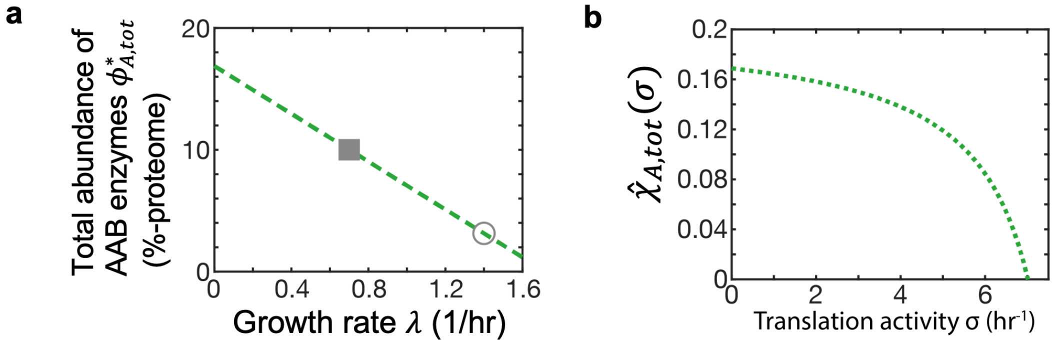
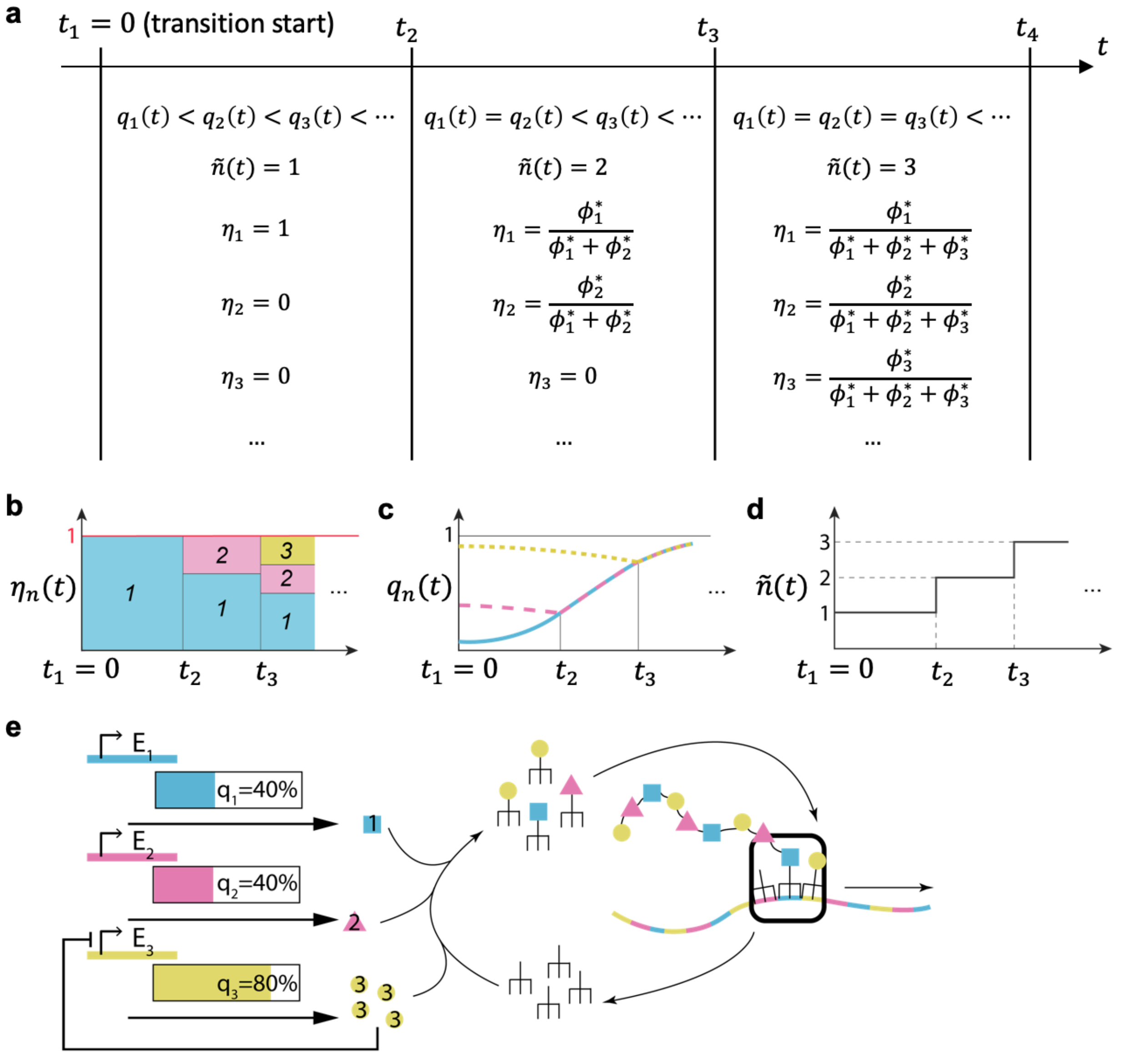


Figure 3.14: Construction of the global regulatory function for AAB enzymes, $\hat{\chi}_{A,tot}(\sigma)$. The form of $\hat{\chi}_{A,tot}(\sigma)$ is constructed similarly to that described in **Fig. 3.13** for $\hat{\chi}_{Rb}(\sigma)$. **(a)** Total abundance of AAB enzymes. We started with the steady state abundance $\phi^*_{A,tot}$ of the AAB enzymes in different growth conditions: the value in the pre-shift steady-state is indicated by the open circle, and that of the post-shift steady-state is indicated by the filled square. We connected these two points by a straight line (green dash line), which can be written as $\phi^*_{A,tot} = \phi^{max}_{A,tot} - \alpha_A \lambda^*$ with $\phi^{max}_{A,tot} = 16.5\%$ and $\alpha_A = 0.95\ h$. Using Eq. (3.13) in **Fig. 3.13** to relate λ^* to σ^*, we obtain

$$\hat{\chi}_{A,tot}(\sigma) = \phi^*_{A,tot}\big(\lambda^*(\sigma)\big) = \phi^{max}_{A,tot} - \alpha_A \cdot \sigma \hat{\chi}_{Rb}(\sigma), \tag{3.15}$$

where $\hat{\chi}_{Rb}(\sigma)$ is given by Eq. (3.14). The resulting function is plotted as the dotted line in panel **(b)**.
From **(b)**, we see that when the translational activity σ is low (reflecting limitation by the shortage of AAs), AAB enzymes are expected to be up-regulated. This is consistent with our qualitative expectation that due to the global regulation of ppGpp, cells allocate more protein synthesis to AAB enzymes after AA downshift. Quantitatively, the expression for $\hat{\chi}_{A,tot}(\sigma)$ arise from the hypothesized linear relation between $\phi^*_{A,tot}$ and λ^* shown in **(a)**. This linear relation can be rationalized in two ways: First, it can be taken as a result of passive regulation from ribosomal protein synthesis regulations, i.e., as ribosomal protein abundance decrease for decreasing growth rate (**Fig. 3.13a**), other protein sectors including AAB enzymes increase in abundance. Alternatively, panel **(a)** can be viewed as describing the expected reduction in growth rate if the total abundance of the AAB enzymes are dialed up. A linear relation between these two quantities is justified by earlier findings that the growth rate declined linearly upon increases in the over-expression of "useless" proteins (3, 7).

Figure 3.15: On-off allocation strategy. To allocate protein synthesis to the growth bottlenecks, Pavlov and Ehrenberg (16) formulated an "on-off strategy" which is easy to implement mathematically: Consider enzyme groups whose indices n are ordered by their relative pre-expression abundances $q_n(0) = \phi_{A,n}(0)/\phi^*_{A,n}$, so that $q_1(0) < q_2(0) < q_3(0) < \cdots$. Define $\tilde{n}(t)$ to be the number of enzyme groups that are limiting growth and are hence being synthesized at time t. **(a)** shows several key time points t_n indicating the time when the n^{th} enzyme group becomes growth-limiting, as defined by the condition $q_n\,(t) = q_{n-1}(t)$ for $n > 1$. At such a time t_n, the value of $\tilde{n}(t)$ is incremented by 1. Specifically, in the first stage ($t_1 < t < t_2$ with $t_1 = 0$), $q_1(t)$ is the minimum and only enzyme group 1 is growth-limiting, thus $\tilde{n} = 1$. Here all the available protein synthesis flux are used to synthesize group 1 enzymes. Thus the allocation function is $\eta_1 = 1, \eta_{n>1} = 0$. As a result $q_1(t)$ increases in time. It reaches the value $q_2(t)$ at $t = t_2$. For $t_2 < t < t_3$, enzyme groups 1 and 2 are both growth-limiting. $\tilde{n} = 2$, and protein synthesis flux is allocated to both group 1 and 2. The partition between η_1 and η_2 is proportional to the mass fractions of these two groups in final steady-state, i.e.. ϕ^*_1 and ϕ^*_2, so that $q_1(t) = q_2(t)$ holds throughout this time period. Similarly, when $q_1(t)$ and $q_2(t)$ reach $q_3(t)$ at $t = t_3$, cells start synthesizing group 3 ($\tilde{n} = 3$), with protein synthesis flux allocated in proportional to $\phi^*_1 : \phi^*_2 : \phi^*_3$. In general, $\eta_n(t)$ is given by Eq. (8) of the main text. Panels **(b)-(d)** provide an example of 3 enzyme groups with their mass faction in the final steady-state given by $\phi^*_1 = 2\phi^*_2 = 2\phi^*_3$. **(b)** $\eta_n(t)$; **(c):** $q_n(t)$; **(d)**: $\tilde{n}(t)$. In **(b)** and **(c)**, blue indicate group 1, pink indicates group 2, yellow indicates group 3. **(e)** The dynamics of a toy model for the three enzyme groups shown in panels **(b)-(d)**, at a time point t between t_2 and t_3. Enzyme groups E_1, E_2 and E_3 synthesize amino acid $1, 2$ and 3 respectively. The bars below the enzymes shows the expression level relative to the final (q_n) at time t. Amino acid $1, 2$ and 3 are "consumed" by the protein synthesis process with the help of tRNA as shown on the right side of the plot. The usage of AAs depend on the codons on the mRNA (blue, red and yellow color bars on mRNA). Here for simplicity we assumed equal codon frequencies. Since the level of E_3 is higher at this time, more AA$_3$ is synthesized than what can be utilized (since protein synthesis is limited by AAs 1 and 2). Since the expression of AAB enzymes is negatively regulated by their corresponding AA products through pathway-specific regulations (44) the surplus in AA$_3$ suppresses the expression of E_3, thereby channeling the flux of protein synthesis preferentially into the synthesis of E_1 and E_2 as panels **(b)-(d)** illustrates.


a
$t_1 = 0$ (transition start)
t_2
t_3
t_4
t
$q_1(t) < q_2(t) < q_3(t) < \cdots$
$\tilde{n}(t) = 1$
$\eta_1 = 1$
$\eta_2 = 0$
$\eta_3 = 0$
...
$q_1(t) = q_2(t) < q_3(t) < \cdots$
$\tilde{n}(t) = 2$
$\eta_1 = \frac{\phi_1^*}{\phi_1^* + \phi_2^*}$
$\eta_2 = \frac{\phi_2^*}{\phi_1^* + \phi_2^*}$
$\eta_3 = 0$
...
$q_1(t) = q_2(t) = q_3(t) < \cdots$
$\tilde{n}(t) = 3$
$\eta_1 = \frac{\phi_1^*}{\phi_1^* + \phi_2^* + \phi_3^*}$
$\eta_2 = \frac{\phi_2^*}{\phi_1^* + \phi_2^* + \phi_3^*}$
$\eta_3 = \frac{\phi_3^*}{\phi_1^* + \phi_2^* + \phi_3^*}$
...
b
1
$\eta_n(t)$
1
2
1
3
2
1
...
$t_1 = 0$
t_2
t_3
c
1
$q_n(t)$
...
$t_1 = 0$
t_2
t_3
d
3
2
1
$\tilde{n}(t)$
...
$t_1 = 0$
t_2
t_3
e
E_1
q_1=40%
1
E_2
q_2=40%
2
E_3
q_3=80%
3
3
3
3

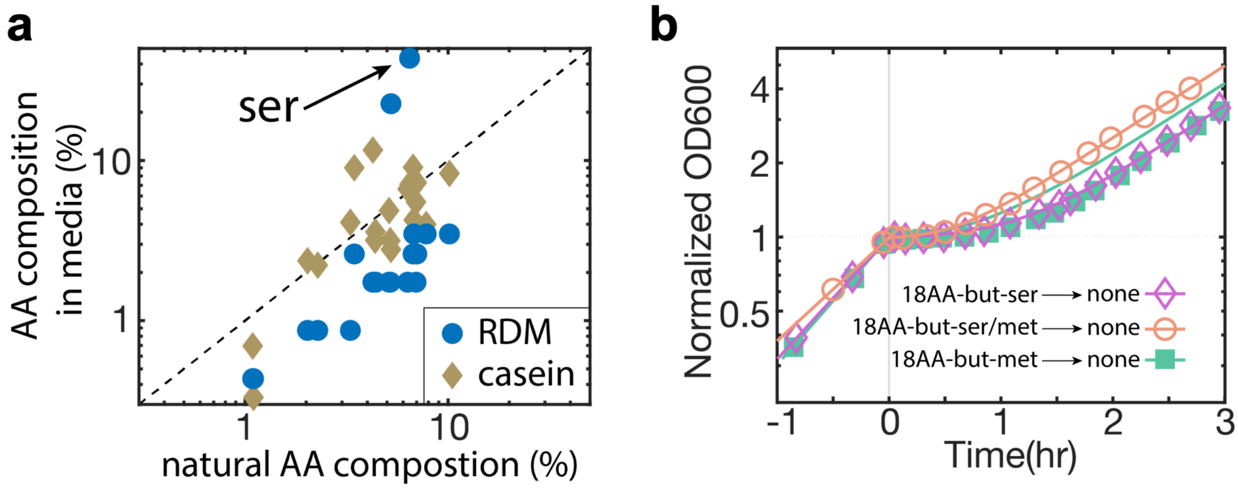


Figure 3.16: **(a)** Scatter plots comparing the normalized AA molar composition of RDM (blue circles, see **Table 3.4** for the recipe (22)) and casein (brown diamond, (50)) with the natural AA frequency across proteins (52). Serine (arrow) is over-represented by ~10x in RDM. **(b)** Growth curves for AA downshifts with additional AAs excluded from the pre-shift medium: ser excluded (open diamonds), ser and met excluded (open circles), met excluded (filled squares); see **Table 3.4** for AA concentrations. The corresponding model predictions are shown as lines with the corresponding colors. To model '18AA-but-ser' to 'none' shift, we excluded ser enzymes from the AAB groups while taking the $q_n(0)$ of the other AAB enzyme groups to be the same as those in the 'all' condition. Similar idea applies to the modeling of the other shifts.

Our model successfully captured the growth kinetics when ser was excluded in the pre-shift medium (open symbols), but not when ser was present (solid symbols, see also **Fig. 3.5f** for '18AA-to-none' shift). Our model did not quantitatively capture the growth kinetics of '18AA-to-none' as seen in **Fig. 3.5f** despite its success with the enzyme kinetics. Since the tradeoff between the synthesis of AAB enzymes and ribosome is the starting point of our model (Eqs. (3.1) and (3.2)), while the determination of growth by ribosome allocation is well established (Eq. (3.4), (23)), the discrepancy between the observed and predicted growth curves likely originated outside of the tradeoff between AAB enzymes and ribosomes. Close comparison of our data and model prediction pointed to prolonged delay in serine biosynthesis (**Fig. 3.3i**). In support of this, we found that AA downshifts without serine shortage were quantitatively captured by our model as shown in panel (**b**).

Looking further into the "serine problem", we noted that the common recipe of AA composition for the "rich medium" we used (RDM, Ref. (22)) contained >10-fold higher serine than the natural AA composition in proteins (blue circles in **(a)**). This recipe was presumably formulated to counteract the very high serine consumption rate by *E. coli* (48, 49) so that serine would never run out in steady state studies. For cells growing in media with AA composition more representative of the natural AA frequencies, e.g., "casamino acid" (CAA) obtained from hydrolysis of casein (brown diamonds in **(a)**), serine is known to be depleted well before any other AAs (49). We thus expected downshift of cells grown in CAA to exhibit growth kinetics similar to that seen in 18AA-ser shift (magenta diamonds in **(b)**), which was well-described by our model. This was indeed seen to be the case as shown in **Fig. 3.5f** (open red circles). Thus, the "serine problem" is likely an artifact of artificial medium composition that is not likely encountered in natural AA downshifts. In this light, the extra delay in the onset of serine biosynthesis enzymes observed from growth in high serine could result from a built-in protection mechanism against surge of serine (which would be toxic (53–55)) but not usually encountered in natural AA downshifts.

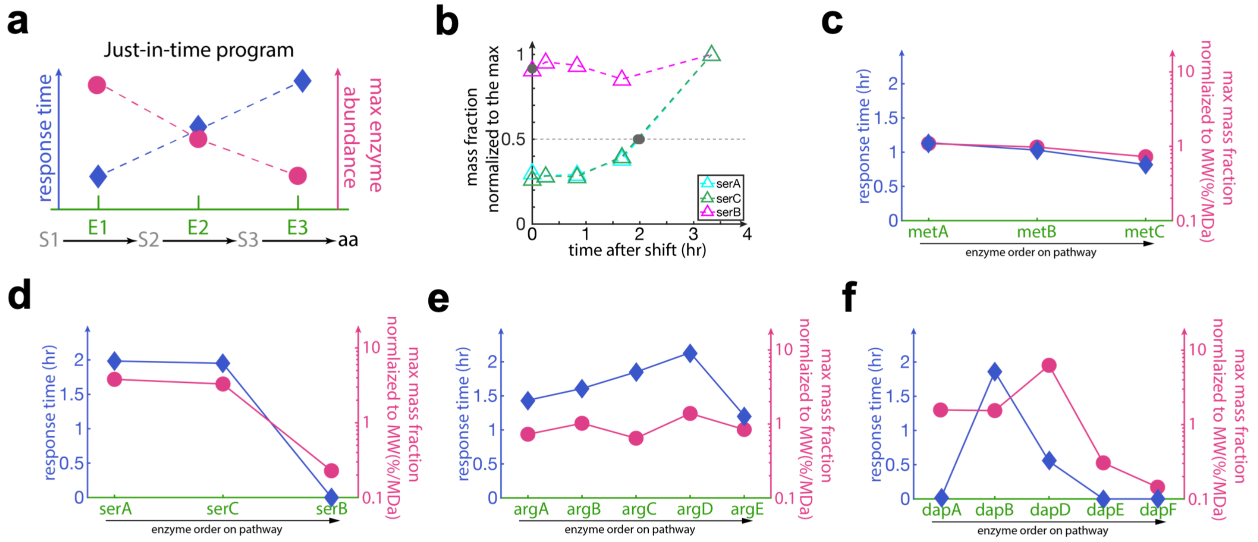


Figure 3.17: Applicability of the "just-in-time" program to enzyme recovery kinetics. (a) 'Just-in-time' program of enzyme recovery kinetics (51) states that during AA downshift, the response time of an enzyme in a linear pathway is shorter if it is located towards the beginning of the pathway (as indicated by the blue symbols), while the maximum abundance is higher towards the beginning of the pathway (as indicated by the red symbols). **(b)** Examples of response time calculated based on the proteomic measurements of SerA, SerB, SerC (**Fig. 3.10i**). Between the measured data points, we used linear interpolation (gray dash lines) to connect the data points. The x-coordinates of the solid dots represent the corresponding response times, defined as the time when the enzyme abundance reaches 50% of the maximum during the shift according to Zaslaver et al (51). For enzyme abundance exceeding 50% at the beginning of the shift (as in the case of SerB), we took the response time to be zero. **(c)-(g)** The observed response times and maximum enzyme abundances for 5 linear segments of the AAB pathways are plotted in the order of enzyme locations for met in panel **(c)**, ser in panel **(d)**, arg in panel **(e)**, lys in panel **(f)**. Neither the response time nor the maximum abundance exhibited the predicted trends shown in **(a)** for any of the 4 pathways**.** The derivation of the 'just-in-time' program shows that this strategy was formulated based on the assumption that the synthesis of an enzyme in a pathway depended only on the pool of its substrate. In reality during AA downshift, enzyme synthesis depended on the availability of all AAs, which are products located at the end of each AAB pathway.

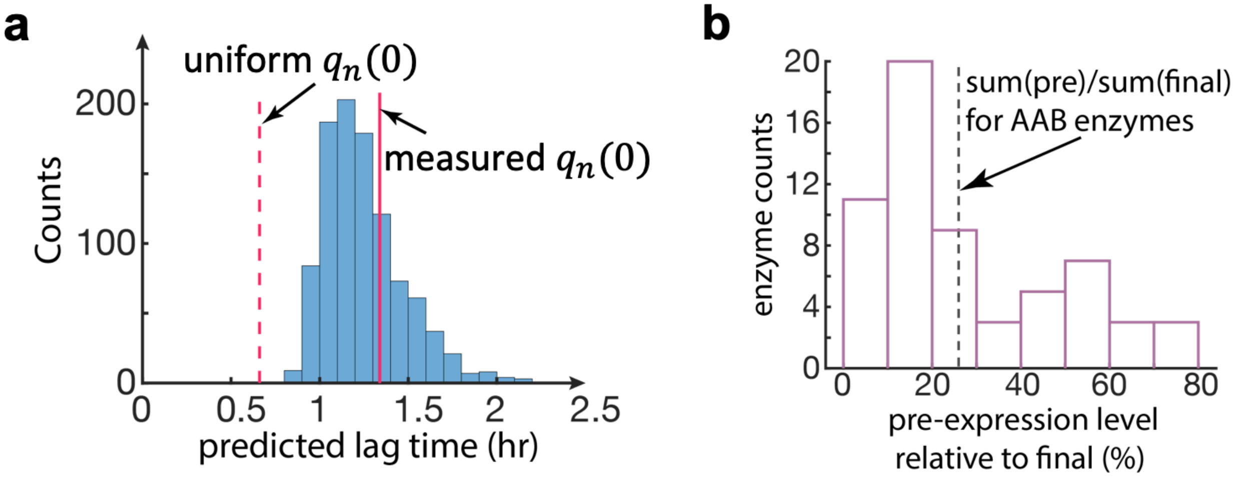


Figure 3.18: (a) Histogram of lag times predicted by our dynamic proteome allocation model, with randomly generated pre-expression levels $q_n(0)$ constrained to the total pre-shift AAB enzyme levels, $\sum_n \phi_{A,n}(0)$. The abundances in post-shift steady-state are kept the same. For the abundances in pre-shift steady-state, we generated 14 random numbers (uniformly distributed in interval (0,1)) and scaled them so that the sum equals the measured total pre-expression level of all AAB enzymes, which is 3.14% of the proteome. That gave us one set of initial condition $q_n(0)$ which we used to compute the lag time. This was repeated 1000 times and the histogram of the predicted lag time is shown as blue bars. The red solid line shows the predicted lag time using the measured values of $q_n(0)$. The red dash line shows the predicted lag time if every AAB groups adopts the same $q_n(0)$, ~26%. For fixed total abundance of AAB enzymes in the pre-shift and post-shift steady state, the model of Pavlov and Ehrenberg predicted that the shortest lag time would be obtained if all the AAB enzymes had the same $q(0)$ value (vertical dash line). **(b)** The histogram of the relative pre-expression levels ($q(0)$) observed for the 14 AAB enzyme groups. Since cysteine and tyrosine were not provided in the pre-shift media, enzymes on cysteine and tyrosine biosynthesis pathways are excluded here.

3.9 Tables

Table 3.1: List of strains employed in this work

Label	Genotype	description	reference
EQ353	Wide-type *E. coli* K12	MG1655 from CGSC	(56)
NCM3722	Wide-type *E. coli* K12	the parental strain for the strains below	(29, 30)
NQ1243	$zca-3633 :: \Phi(\text{FRT}: rrnBt: \text{P}_{\text{Ltet-O1}}: xylR)$ $\Delta ptsG468 :: \Phi(kan: rrnBt: \text{Pu})$	*ptsG* titratable strain	(10)
NQ1261	$\Delta ptsG468 :: \Phi(kan: rrnBt: \text{Pu})$	*ptsG* deletion strain	(17)
NQ399	$attB :: \Phi(amp: \text{P}_{\text{Llac-O1}}: xylR)$ $\Delta glpF224 :: \Phi(kan: rrnBt: \text{Pu})$	*glpFK* titratable strain	(12)

Table 3.2: Proteomic data sample conditions.

1. Sample P1,P8-P10 are the repeat of M1-M4, respectively. The average growth rates and protein abundances for these four conditions are plotted in **Fig. 1**.
2. O8-O10, P3 are used as reference for scaling the proteomic data in other conditions.
3. See strain description in Supplementary Tabel 1.
4. MOPS EZ rich defined medium kit M2105 from Teknova (glucose excluded).
5. Bacto^{TM} casamino acid.
6. Except LB condition, all other conditions are based on MOPS buffered media with 50mM NaCl, 1.32mM KH2PO4, 10mM NH4Cl and corresponding nutrients listed in the column. RDM condition and conditions for EQ353 are based on Neidhardt's MOPS (22) while the rest all based on Cayley's MOPS (31); see Methods for details.
7. 3-methylbenzyl alcohol.

proteomic data sample condition

Sample ID	symbol	growth rate (1/h)	Strain[3]	Nutrient	Medium[6]	Inducer	sample collection	Description
M1	○	0.96	NCM3722	0.2% glucose	MOPS		OD_{600}=0.4	Clim MOPS (WT)
M2	○	0.75	NQ1243	0.2% glucose	MOPS	200uM 3MBA[7]	OD_{600}=0.4	Clim MOPS (Pu-ptsG)
M3	○	0.54	NQ1243	0.2% glucose	MOPS	0 3MBA[7]	OD_{600}=0.4	Clim MOPS (Pu-ptsG)
M4	○	0.33	NQ1261	0.2% glucose	MOPS		OD_{600}=0.4	Clim MOPS (del ptsG)
N5	□	1.77	NCM3722	RDM[4]+0.2% Glucose	Neidhardt's MOPS		OD_{600}=0.4	WT-Rich
N6	■	1.42	NCM3722	RDM[4]	Neidhardt's MOPS		OD_{600}=0.4	WT-Rich
N7	○	1.58	NCM3722	0.2% CAA[5]+0.2% Glucose	MOPS		OD_{600}=0.4	WT-Rich
N8	●	1.32	NCM3722	0.2% CAA[5]	MOPS		OD_{600}=0.4	WT-Rich
P1[1]	○	0.95	NCM3722	0.2% glucose	MOPS		OD_{600}=0.4	Clim MOPS (WT)
P8[1]	○	0.73	NQ1243	0.2% glucose	MOPS	200uM 3MBA[7]	OD_{600}=0.4	Clim MOPS (Pu-ptsG)
P9[1]	○	0.53	NQ1243	0.2% glucose	MOPS	0 3MBA[7]	OD_{600}=0.4	Clim MOPS (Pu-ptsG)
P10[1]	○	0.35	NQ1261	0.2% glucose	MOPS		OD_{600}=0.4	Clim MOPS (del ptsG)
S1	△	1.45	NCM3722	0.4% glycerol+18AA	MOPS		OD_{600}=0.3	18AA (preshift steady state)
S2	△		NCM3722	0.4% glycerol (shift)	MOPS		time=15min	shift 18AA to 0
S3	△		NCM3722	0.4% glycerol (shift)	MOPS		time=50min	shift 18AA to 0
S4	△		NCM3722	0.4% glycerol (shift)	MOPS		time=100min	shift 18AA to 0
S5	△		NCM3722	0.4% glycerol (shift)	MOPS		time=200min	shift 18AA to 0
O8[2]		2.19	NCM3722	LB	LB		OD_{600}=0.24	WT-Rich
O9[2]		1.85	EQ353	RDM+0.2% Glucose	Neidhardt's MOPS		OD_{600}=0.4	MG1655 rich
O10[2]		0.66	EQ353	0.2% Glucose	Neidhardt's MOPS		OD_{600}=0.4	MG1655 glucose
P3[2]		0.64	EQ353	0.2% glucose	Neidhardt's MOPS		OD_{600}=0.4	MG1655 glucose

Table 3.3: Gene classification in Chapter 3.
*lpd was split half and half into TCA cycle and fermentation to acetate while calculating protein abundance of each group.

AA biosynthesis	arg group	argA	argB	argC	argD	argE	argF	argG	argH	argI	
	aro group	aroA	aroB	aroC	aroD	aroE	aroF	aroG	aroH	aroK	aroL
	cys group	cysC	cysD	cysE	cysH	cysI	cysJ	cysK	cysM	cysN	
	glt group	gdhA	gltB	gltD							
	his group	hisA	hisB	hisC	hisD	hisF	hisG	hisH	hisI		
	ilv group	ilvA	ilvB	ilvC	ilvD	ilvE	ilvG_1	ilvH	ilvI	ilvM	ilvN
	leu group	leuA	leuB	leuC	leuD						
	lys group	asd	dapA	dapB	dapD	dapE	dapF	lysA	lysC		
	met group	metA	metB	metC	metE	metH	metL				
	other group	alaA	alaC	asnA	asnB	aspC	glnA	glyA	proA	proB	proC
	phetyr group	pheA	tyrA	tyrB							
	ser group	serA	serB	serC							
	thr group	thrA	thrB	thrC							
	trp group	trpA	trpB	trpC	trpD	trpE					
Central carbon metabolism and energy	ATPase	atpA	atpB	atpC	atpD	atpE	atpF	atpG	atpH	atpI	
	fermentation to acetate	aceE	aceF	ackA	lpd*	pta					
	glycolysis &GNG	eno	fbaA	fbaB	fbp	gapA	glpX	gpmA	gpmM	maeA	maeB
		mdh	mqo	pck	pfkA	pfkB	pgi	pgk	ppsA	pykA	pykF
		tpiA	ybhA								
	TCA	acnA	acnB	fumA	fumB	fumC	fumD	fumE	gltA	icd	lpd*
		sdhA	sdhB	sdhC	sdhD	sucA	sucB	sucC	sucD		
Nucleotide biosynthesis		adk	cmk	deoB	dut	gmk	guaA	guaB	ndk	nrdA	nrdB
		nrdE	nrdF	prs	purA	purB	purC	purD	purE	purF	purH
		purK	purL	purM	purN	purT	pyrB	pyrC	pyrD	pyrE	pyrF
		pyrG	pyrH	pyrI	pyrL	thyA	tmk	trxA			
stationary phase proteins		dps	ecnB	elaB	katE	osmC	osmE	otsA	otsB	wrbA	
Translational proteins	affiliated translational apparatus	arfA	arfB	efp	frr	fusA	infA	infB	infC	lepA	prfA
		prfB	prfC	tsf	tufA	tufB					
	ribosomal proteins	rplA	rplB	rplC	rplD	rplE	rplF	rplI	rplJ	rplK	rplL
		rplM	rplN	rplO	rplP	rplQ	rplR	rplS	rplT	rplU	rplV
		rplW	rplX	rplY	rpmA	rpmB	rpmC	rpmD	rpmE	rpmF	rpmG
		rpmH	rpmI	rpmJ	rpsA	rpsB	rpsC	rpsD	rpsE	rpsF	rpsG
		rpsH	rpsI	rpsJ	rpsK	rpsL	rpsM	rpsN	rpsO	rpsP	rpsQ
		rpsR	rpsS	rpsT	rpsU	sra					
	tRNA synthase	alaS	argS	asnS	aspS	cysS	glnS	gltX	glyQ	glyS	hisS
		ileS	leuS	lysS	lysU	metG	pheS	pheT	proS	serS	thrS
		trpS	tyrS	valS							
Transporters	AA transport	abgT	alaE	ansP	argO	argT	aroP	artI	artJ	artM	artP
		artQ	brnQ	cadB	cstA	cycA	dppA	dppB	dppC	dppD	dppF
		dtpA	dtpB	dtpC	dtpD	eamA	eamB	frlA	gabP	gadC	glnH
		glnP	glnQ	gltI	gltJ	gltK	gltL	gltP	gltS	hisJ	leuE
		livF	livG	livH	livJ	livM	lysP	metQ	mmuP	mtr	oppA
		oppB	oppC	oppD	oppF	pheP	plaP	potA	potB	potC	potD
		potE	potF	potG	potH	potI	proV	proW	proX	proY	putP
		puuP	rhtA	rhtB	rhtC	sgrR	sstT	tcyJ	tcyL	tcyN	tcyP
		tdcC	tnaB	tyrP							
	carbon metabolism	aceA	aceB	acs	alsA	alsB	alsC	araF	araG	araH	crr
		fruA	fruB	fucK	fucP	galE	galF	galK	galM	galP	galR
		galS	galT	galU	gatA	gatB	gatC	lacY	lacZ	malE	malF
		malG	malK	malM	malP	malT	malX	manX	manY	manZ	melB

Table 3.3: Gene classification in Chapter 3, continued.

Transporters	carbon metabolism	mglA	mglB	mglC	mtlA	mtlD	nagE	ptsG	ptsH	ptsI	rbsA
		rbsB	rbsC	rbsD	rbsK	srlA	srlB	srlE	treB	ugpA	ugpB
		ugpC	ugpE	ulaA	ulaB	ulaC	xylA	xylB	xylE	xylF	xylG
		xylH									
	glycerol uptake	glpA	glpB	glpC	glpD	glpF	glpK				
	outer membrane porin	nmpC	ompA	ompC	ompF	ompG	ompN	phoE			

Table 3.4: Amino acid concentration in various medium.

1. AA concentration was calculated from AA composition reported by BD manufactory. Glutamine and asparagine is hydrolyzed during analysis. Here we split the percentile value of glutamate into glutamine and glutamate with 1:1 ratio. The same thing was done for aspartate. Reported cystine abundance was taken to be the abundance of cysteine.

compositions	Concentration (mM)						
	RDM	20AA	18AA	18AA-but-ser	18AA-but-met	18AA-but-ser/met	0.2% Casamino acid[1]
alanine	0.8	0.8	0.8	0.8	0.8	0.8	0.67
arginine	5.2	5.2	5.2	5.2	5.2	5.2	0.29
aspartate	0.4	0.4	0.4	0.4	0.4	0.4	0.18
asparagine	0.4	0.4	0.4	0.4	0.4	0.4	0.18
cysteine	0.1	0.1	0	0	0	0	0.02
glutamate	0.6	0.6	0.6	0.6	0.6	0.6	1.08
glutamine	0.6	0.6	0.6	0.6	0.6	0.6	1.09
glycine	0.8	0.8	0.8	0.8	0.8	0.8	0.37
histidine	0.2	0.2	0.2	0.2	0.2	0.2	0.10
isoleucine	0.4	0.4	0.4	0.4	0.4	0.4	0.61
leucine	0.8	0.8	0.8	0.8	0.8	0.8	0.76
lysine	0.4	0.4	0.4	0.4	0.4	0.4	0.71
methionine	0.2	0.2	0.2	0.2	0	0	0.19
phenylalanine	0.4	0.4	0.4	0.4	0.4	0.4	0.44
proline	0.4	0.4	0.4	0.4	0.4	0.4	1.39
serine	10	10	10	0	10	0	0.40
threonine	0.4	0.4	0.4	0.4	0.4	0.4	0.25
tryptophan	0.1	0.1	0.1	0.1	0.1	0.1	0
tyrosine	0.2	0.2	0	0	0	0	0.04
valine	0.6	0.6	0.6	0.6	0.6	0.6	0.96

Table 3.5: Carbon uptake mechanism

Carbons	Uptake mechanism/enzymes	reference
lactose	proton symporter LacY and lactose galactohydrolase lacY	(57)
glucose	Bacterial phosphotransferase system PtsG, Crr	(58, 59)
fructose	Bacterial phosphotransferase system FruAB	(59, 60)
xylose	ABC transporter XylFGH and proton symporter XylE	(61–63)
glycerol	Glycerol facilitator GlpF and glycerol kinase GlpK	(43)

3.10 Appendix

3.10.1 Kinetic equations for enzymes during shift

Here we provided the details on the deduction of the kinetic equations of AAB enzymes and ribosomes (Eq. (3.1) and (3.2)).

Through this work, we used mass fraction to represent protein abundance. For protein or protein groups x, its mass fraction ϕ_x can be written as

$$\phi_x = M_x/M_{p.} \tag{3.16}$$

M_x denotes the mass of protein (protein group) x in a standard culture volume, e.g., 1ml of culture. M_p denotes the total protein mass in the same culture volume. The time derivative of ϕ_x can thus be written as

$$\frac{d\phi_x}{dt} = \frac{1}{M_p}\left(\frac{dM_x}{dt} - \phi_x \frac{dM_p}{dt}\right). \tag{3.17}$$

(dM_p/dt) is the protein synthesis flux which is referred as J_R. In this work, we used χ_x to represent the fraction of protein synthesis flux that being directed to x. In other words,

$$\frac{dM_x}{dt} = \chi_x(t) J_R(t). \tag{3.18}$$

Then Eq. (3.17) can be rewritten as

$$\frac{d\phi_x}{dt} = \frac{J_R}{M_p}(\chi_x(t) - \phi_x(t)). \tag{3.19}$$

Given the definition of growth rate $\lambda(t) = \frac{d\ln\left(M_p(t)\right)}{dt}$, $J_R(t)$ can be written as the products of $\lambda(t)$ and $M_p(t)$. Eq. (3.19) then became

$$\frac{d\phi_x}{dt} = \lambda(t)(\chi_x(t) - \phi_x(t)). \tag{3.20}$$

Substituting AAB enzyme groups or ribosomal proteins for x, we then achieved the Eq. (3.1) and (3.2) in the main text for describing the kinetic changes of $\phi_{A,n}$ and ϕ_{Rb}.

Chapter 4

Kinetic response of E.coli to sub-lethal chloramphenicol

4.1 Introduction

Quantitative characterization of bacterial responses to antibiotics is essential for proper usage of antibiotics in treating bacterial infection. In this chapter, we will look into the kinetic responses of *E. coli* to sub-lethal chloramphenicol (Cm). Cm is an translational inhibitor useful for the treatment of number of bacterial infections, such as meningitis, plague. It is known that Cm inhibit translation elongation by binding to 50S ribosomal subunit. Previous work have well studied the physiological behavior of *E. coli* under sub-lethal Cm during exponential growth (referred as R-limitation). It showed that ribosomal abundance linearly increase with the decrease of growth rate with Cm addition (Fig. 1.2c, (3, 17)), while many catabolic and anabolic proteins behave the opposite (3, 7). Meanwhile, average translational elongation rate was also found increase under Cm (17), taken as a result of increased substrates (ternary complex with charged tRNAs). With the steady-state behavior under Cm R-limitation well established, it gave us a chance to quantify the kinetic response of *E. coli*. Taking the idea of the flux-controlled regulatory method used in nutrient shift (Chapter 2 & 3, (23)), we used translational elongation rate as a flux sensor and control the expression of ribosomal proteins and metabolic proteins. By extending the steady-state relationship to kinetic regime, our model successfully capture the adaptation kinetics.

The success of this model indicate a possible important rule of translational elongation rate in cellular global regulatory strategy, which is actually shown in Chapter 5.

4.2 Results

In order to obtain the kinetics of growth rate change with higher time resolution, we used turbidostat (Fig. 4.1) to monitor bacterial growth during chloramphenicol (Cm) downshift. We first let cells grow in glucose minimal medium. After they reach exponential growth, 4uM Cm was added to the cell culture to initiate Cm downshift. From the measured time course of OD_{600} (a proxy of biomass), we were able to calculate the time course of flux J and instantaneous growth λ rate as shown in Fig. 4.2. (Notice that the data points in Fig. 4.2 are the average of original data (Fig. 4.5) with certain window size for clear viewing.) The growth rate in Fig. 4.2c shows a special response with an immediate drop followed by a 1-hr slow decrease. After 1hr, cells reached the steady-state growth rate and kept the same growth rate during the time of measurements (Fig. 4.5). Similar two-stage growth rate drop were also observed when adding higher Cm concentration (adding 8uM Cm to glucose medium; Fig. 4.3), or with a different carbon source (adding 4uM Cm to fructose medium; Fig. 4.4). We also measured the kinetics of ribosomal content (reported by total RNA abundance; green diamonds in Fig. 4.2a) as well as catabolic protein content (reported by LacZ expression; red triangle in Fig. 4.2a). As expected, ribosomal content accumulated faster than biomass after adding Cm, while catabolic protein content rarely increase in the first 50mins after shift despite the increase in biomass. Comparing with the growth rate change in Fig. 4.2c, cells resumed the synthesis of catabolic proteins (or at least LacZ) roughly at the same time when growth rate relax to the post-shift steady state. It was also seen in other Cm downshifts shown in Fig. 4.3 and 4.4. Eventually, the proteome allocation also reached a new steady-state with higher fraction of ribosomal content and lower fraction of catabolic proteins.

In order to understand the kinetic response to Cm, let's first review the feature of Cm involved translational elongation inhibition. In brief, the function of Cm is equivalent to taking away active ribosomes but not affecting average translational elongation rate of complete proteins as studied in Dai et al (17). To be detailed, due to the low dissociation rate between Cm and ribosomes, the half-life of Cm bound ribosome is more than 8mins as shown in Fig. 4.6a. This value means that the percentage of ribosome stalled by Cm but can still finish translation process is very low due to mRNA degradation and ribosome abortion, estimated to be no more than 4% of total ribosomes (17). In that case, the effect of Cm can be simplified as the decrease in active ribosomal fraction (Fig. 4.6a). In detail, we divided total ribosomes into two sub group—inactive ribosomes and active ribosomes. Inactive ribosomes does not contribute to protein synthesis, whereas the elongation rate of active ribosomes does not affected by added Cm directly. f_{active} is used to represent fraction of active ribosomes among total ribosomes. Fig. 4.6b shows that steady-state f_{active} estimated from the dissociation constant of Cm is in a good agreement of measured f_{active} (data from Dai et al (17)). Fig. 4.6c shows that measured steady-state translational elongation rate of complete protein is either constant or increased under Cm treatments (data from Dai et al (17)). Both observations support the above simplification of Cm's effect on translation.

With f_{active} represents the effect of Cm, we established a mathematical model based on the flux-controlled regulatory model used in Erickson et al and Chapter 2 & 3 to describe the kinetics of transition as shown in Fig. 4.7a. It allows us to capture the proteome reallocation process without getting into the details of molecular interactions.

Qualitatively, according to the steady-state behavior under Cm, cells upregulate translational related proteins to compensate for the loss of active ribosomes during the transition. Meanwhile, other growth-rate dependent enzymes (non-translational related, such as catabolic and anabolic proteins) are down-regulated under the constrain of limited translational capacity. Based on that, we used the three-component partition introduced in Scotts et al to coarse-grained cellular proteome into three parts: translational proteins (ribosomal and co-regulated translation-affiliated proteins), metabolic proteins (other growth-rate dependent enzymes, mainly catabolic and anabolic proteins), and others (fixed core proteins, growth-rate independent under Cm R-limitation). Using mass fraction (defined as $\phi_x = M_x/M_P$ for group x where M_P is the mass of total proteins) to describe the protein concentration, the partition above give an intrinsic constrain:

$$\phi_R + \phi_E = 1 - \phi_{other} = \phi_{max}, \tag{4.1}$$

where ϕ_R is the mass fraction of translational proteins, ϕ_E is the mass fraction of metabolic proteins and ϕ_{other} is the mass fraction of others. Since in steady-state ϕ_{other} is invariant under R-limitation, the sum of ϕ_R and ϕ_E are fixed to be ϕ_{max} under R-limitation as shown in Eq. (4.1).

In steady-state growth, the proteomic allocation between ϕ_R and ϕ_E are set so that the flux they provided balance with each other as stated in Scott et al. So in order to understand the proteomic reallocation process while adding Cm, it's important to look into the flux change. Translational proteins, whose mass are labeled as M_R, provides protein synthesis flux J_R (Fig. 4.7a). The relationship between M_R and J_R were deduced in Dai et al (17) as

$$J_R = \alpha M_R \cdot f_{active} \cdot \epsilon, \tag{4.2}$$

where αM_R is the mass of ribosomal proteins (α ≅62.5% is an empirical constant possibly due to co-regulation of ribosomal and other translation-affiliated proteins, (17)) and ϵ represents elongation rate. Notice that ϵ has absorbed the factor of $(1/n_{Rb})$ where n_{Rb} is the number of amino acid residues in one ribosome, in order to make the unit consistent. For convenience, we let ϵ absorb the factor of α in the following text. With the definition of growth rate λ ($\lambda = \frac{dlnM_P}{dt} = \frac{J_R}{M_P}$), Eq. (4.2) can be rewritten in terms of λ and ϕ_R as

$$\lambda = \phi_R \cdot f_{active} \cdot \epsilon, \tag{4.3}$$

Metabolic proteins, on the other hand, provide nutrient influx j_{AA} as shown in Fig. 4.7a. Quantitatively, they are written as

$$j_{AA} = M_E \cdot k, \tag{4.4}$$

where k is the efficient catalytic rate of metabolic proteins. Under steady-state growth, j_{AA} balance with J_R—

$$J_R = c \cdot j_{AA}, \tag{4.5}$$

where c is conversion factor from amino acid to protein (taken as constant). Combining Eq. (4.4)-(4.5) with the growth rate definition, ϕ_E is also linked with growth rate λ under flux balance via

$$\lambda = \phi_E \cdot \nu, \tag{4.6}$$

where $\nu = k \cdot c$. Notice that Eq. (4.6) is only true under flux balance while Eq. (4.3) always holds. Under R-limitation, both f_{active} and ϵ change as stated above, while ν is irrelevant to Cm conc. but only reflects nutrient states (3). Combining Eq. (4.6) with the intrinsic constrain Eq. (4.1), we get the relationship between λ and ϕ_R under R-limitation

$$\phi_R = \phi_{max} - \frac{\lambda}{\nu}. \tag{4.7}$$

Fitting Eq. (4.7) to steady-state measurements as shown in Fig. 4.7b (solid lines), we can then get ν and ϕ_{max} for specific nutrient conditions as shown in Table 4.1. (Here ϕ_{max} is a little different in different nutrient conditions. For accurately capturing the steady states, we didn't force the same ϕ_{max} across various media as previous work (3).) With ν, we can then evaluate scaled nutrient influx $c \cdot j_{AA}$ based on the change of metabolic proteins during the shift. Notice here ν is assumed to be constant through the whole shift. (In reality, due to allosteric inhibitions, ν could be vary even with the same nutrient source. We will discuss about it more later.)

As we know, before Cm is added, f_{active} is constant and stays high. While Cm is added, it brings down f_{active}, leading to a drop in protein synthesis flux J_R. According to the on/off rate of Cm, the estimated time to achieve the equilibrium of Cm-Rb binding is within 10mins for various concentration of Cm (Fig. 4.6d). Our model does not intend to catch fast metabolic process, so we assumed f_{active} drop to post-shift steady-state value f_{active}^{ss} at time 0 when Cm is added. f_{active}^{ss} is determined by post-shift steady-state measurements via Eq. 4.3. With decreased consumption rate of amino acids (AAs), AAs accumulates, initiating the reallocation of protein synthesis flux towards translational proteins and metabolic proteins. Regulation function χ_i is used to represent the fraction of protein synthesis goes to group i. In other words,

$$\frac{dM_R}{dt} = \chi_R(t)J_R(t) \tag{4.8}$$

$$\frac{dM_E}{dt} = \chi_E(t)J_R(t) \tag{4.9}$$

Molecularly, χ_R and χ_E are directly or indirectly set by the AAs pools via gene regulations (e.x., ppGpp signal system as shown in Fig. 4.7a). With much molecular knowledge on the regulatory interactions, it's still impossible to quantitatively characterize it. Considering AAs pool also set translational elongation rate ϵ, we then used ϵ as a representative of AAs pool and set χ_R and χ_E (Eq. 4.10 and Eq. 4.11 in Fig. 4.7a).

The forms of $\hat{\chi}_R(\epsilon)$ and $\hat{\chi}_E(\epsilon)$ are obtained from their steady-state relationship. Under R-limitation, it was found the elongation rate ϵ and ϕ_R follows Michaelis-Menten relationship

$$\epsilon = \frac{\epsilon_{max}}{1 + K_m/\phi_R}, \tag{4.12}$$

shown as the solid line in Fig. 4.7c . Considering $\chi_R \equiv \phi_R$ in steady state, Eq. (4.12) also give a relationship between ϵ and χ_R under R-limitation. Now we extended this relation to kinetic region and got

$$\hat{\chi}_R(\epsilon) = \frac{K_m}{\frac{\epsilon_{max}}{\epsilon} - 1} \tag{4.13}$$

The extension was done assuming quasi-steady state during transition (fast reaction between metabolites comparing to protein synthesis). The value of K_m and ϵ_{max} are listed in Table 4.1. Similarly, for $\hat{\chi}_E(\epsilon)$, we extended the steady-state constraint (Eq. 4.1) to transient and got

$$\hat{\chi}_E(\epsilon) = \phi_{max} - \hat{\chi}_R(\epsilon) = \phi_{max} - \frac{K_m}{\frac{\epsilon_{max}}{\epsilon} - 1}. \tag{4.14}$$

The value of ϕ_{max} is taken to be the same as the steady-state value given by Fig. 4.7b as explained above. It's based on the assumption that the ϕ_{other} stays constant during transient.

The last remaining piece of the picture is the kinetic of translational elongation rate ϵ. As a central dynamic variable, the kinetics of ϵ determines the kinetics of gene expressions through Eq. (4.13)-(4.14) and biomass accumulations through Eq. (4.2)-(4.3). First, let's look at the initial value of ϵ right after shift. Different from protein level or biomass, translational elongation rate can jump at our coarse-grained time scale due to the fast change in metabolites. From the drop of growth rate at time 0, we can calculate the boost in elongation rate from Eq. 4.3 (Table 4.1). We can see that in different Cm downshift cases, translational elongation rate jumped up 2 to 3 fold. The increase in translational elongation rate can be understood as a result of sudden increased substrates. The actual boost is probably much smaller since f_{active} doesn't drop instantaneously. The fitted value of boost incorporate the artificial effect of simplified dynamics of f_{active}, so the value boosted elongation rate looks much higher than those in steady states, even ϵ_{max}.

Under the condition when flux is balanced, the dynamics of ϵ can be derived through Eq. (4.2) and Eq. (4.4) as below

$$\frac{d\epsilon}{dt} = \frac{d}{dt}\left(\frac{J_R(t)}{f_{active} M_R(t)}\right) = \frac{d}{dt}\left(\frac{c \cdot j_{AA}(t)}{f_{active} M_R(t)}\right) = \frac{d}{dt}\left(\frac{\nu M_E(t)}{f_{active} M_R(t)}\right). \tag{4.15}$$

Simplifying Eq. (4.15), we got

$$\frac{d\epsilon}{dt} = \epsilon\big(\nu\hat{\chi}_E(\epsilon) - \epsilon f_{active}\hat{\chi}_R(\epsilon)\big), \tag{4.16}$$

which can be solved with given initial condition of ϵ. However, from measured instantaneous growth rate, the obvious growth rate drop right after Cm addition clearly shows a drop in protein synthesis flux. Since M_E can't drop instantaneously, the nutrient influx proportional to M_E (assuming constant ν) surely surpass the protein synthesis flux during a finite time after Cm is added (the yellow shaded area in Fig. 4.8a). During the flux imbalance regime, we can't use Eq. 4.16 to solve the kinetics of ϵ. Instead, we assumed that ϵ stays constant at the boosted level until the flux balance is re-gained (Fig. 4.8b). During this regime, χ_R is set to be the maximal possible value ϕ_{max} and χ_E is set to be the minimal possible value 0. With the increase of translational proteins and the dilution of metabolic proteins, eventually the protein synthesis flux balance with the nutrient influx (Eq. 4.5). Then the kinetics of ϵ is calculated through Eq. 4.16, and χ_R and χ_E is set by $\hat{\chi}_R(\epsilon)$ and $\hat{\chi}_E(\epsilon)$ (the hard boundary of $\chi_R < \phi_{max}$ and $\chi_E < 0$ still applies). With the time course of ϵ being calculated as above, the time courses of $M_P(t)$, $M_R(t)$, $M_E(t)$ can then be solved by integrations. Furthermore, fluxes can be calculated by corresponding protein abundance. The model predictions are shown as solid lines in Fig. 4.2-4.4. They nicely catch the recovery dynamics under various Cm downshifts. In this model, all the parameters come from steady-state measurements except the boost in ϵ at time 0 is determined by the sudden drop of growth rate at time 0.

4.3 Discussion

Figure 4.8 marked a special flux imbalance region in our model when capturing Cm downshift. It doesn't exist for carbon downshift (8) or AA downshift described in Chapter 3-4. Here I want to discuss about the physical meaning of the flux imbalance region. It doesn't necessary mean that amino acids are constantly leaking out during that time period. The more accurate statement is that the protein synthesis flux is smaller compare to the 'maximal nutrient influx' ϕ_E can provide during that region. Here the 'maximal influx' means ϕ_E is working at the same rate as in the steady-state growth, which is exactly the assumption of constant ν (or k) we made in the model. In reality, while protein synthesis flux can't match the maximal nutrient influx, AAs will accumulate and inhibit the activity of ϕ_E and lead to a drop in ν (or k). In other words, nutrient influx may actually balance with protein synthesis flux in all or part of the marked 'flux imbalance region' in Fig. 4.8. But this effect on ν won't affect our model results. If ν is repressed in the flux imbalance region, it means AAs pool must have increased and the substrates for translation is more abundant. So it is still reasonable to assume the elongation rate stays high during that region.

Notice different from the model in carbon shift (8) or in AA downshifts, the central dynamic variable in this model is translational ER ϵ. Comparing to previously used σ (defined as J_R/M_{Rb}), ϵ has an clear physiological meaning. That lead to the question whether translational ER indeed contains important regulatory information, which is answered in Chapter 5.

4.4 Methods

4.3.1 Strain

E. coli K-12 NCM3722 with lacI deletion (NQ354 in (12)) was the strain used in this work. LacZ expression was unrepressed under NQ354 (lacI null) strain.

4.3.2 growth medium

All the growth medium used in this work were based on MOPS buffered media described in Cayley et al (31). The medium contains 40mM Mops and 4mM Tricine (adjust to pH 7.4 by NaOH), 0.1mM $FeSO_4$, 0.276mM Na_2SO_4, 0.5μM $CaCl_2$, 0.523mM $MgCl_2$, and also micronutrients mixtures used in Neidhardt et al (22). It was supplied with 50mM NaCl, 1.32mM KH_2PO_4, 10mM NH_4Cl and different carbon sources (0.2% glucose or 0.2% fructose). The chloramphenicol (Cm) is stocked at 0.1M dissolved in ethanol. When adding Cm, the volume of the stock added was maintained to be less than 0.1% of total cell culture.

4.3.3 cell growth and chloramphenicol addition

For growth curve measurements, we used turbidostat (Fig. 4.1) for batch culture growth and media changes. For β-galactosidase and total RNA quantifications, we grew cells either in test tubes or flasks in the shaker with 240 rpm. For both cases, the temperature was maintained at 37°C by either warm room or water bath. Both experiments followed a standard three-step cell growth round: seed culture, pre-culture and experimental culture. Except for the setting in experimental culture, the settings of seed culture and pre-culture were the same for turbidostat experiments and shaker experiments. For seed culture, cells were inoculated to LB broth from fresh LB plate and grew in the test tube in the shaker for about 7hrs. Afterwards cells were washed and diluted to the desired MOPS medium (pre-culture) at starting OD_{600} around 10^{-3}. For Cm addition experiments, the medium for pre-culture is the same as experimental culture but without Cm. After overnight pre-culture (grew in test tube in the shaker), cells were then transferred to the same MOPS medium and kept growing either in turbidostat or in the test tube/flasks in the shaker (experimental culture). Cm addition was executed after cells reached exponential growth in experimental culture.

The settings for turbidostat experiments were from Suckjoon Jun's lab. The construction details were described in previous papers (64). For this study, all the turbidostat settings were kept at the 37°C warm room. Cells from overnight pre-culture was diluted to the same medium in the turbidostat tube as shown in Fig. 4.1. The starting OD_{600} was smaller than 0.01. The volume of experimental culture was maintained at 20ml by the waste tubing. Oxygen level was maintained by the constantly pumping the air to the air tubing. Magnetic stirrer kept spinning during the whole experiments to avoid any heterogeneities in the culture. OD_{600} was read every 15s at the bottom of the tube, above the magnetic stirrer and below the bubbles came out of the air tubing. The dilution was set to happen around $OD_{600} \cong 0.4$. While the OD_{600} reached the threshold, the fresh medium (also at 37°C) came in through the medium tubing and the cell culture were pushed out through the waste tubing until the culture was diluted to $OD_{600} \cong 0.1$. For exponential growth measurements, the medium used for was the same as the medium in the tube. After the first dilution, OD_{600} was kept between 0.1 and 0.4 all the time. For Cm addition experiments, the medium in the tube didn't contain Cm. The medium used for dilution contained the desired amount of Cm (e.x., if shifting to 4uM Cm, then the medium contains 4uM Cm). Cm addition was done between OD_{600} 0.1 and 0.2 before the cell density first reaching $OD_{600} \cong 0.4$, that is, before the first dilution. The time it took for manually adding Cm to the tube is less than 2 mins.

For batch growth in the shaker, cells from overnight pre-culture were diluted to test tube or flasks. If not specific, the starting OD_{600} was around 0.01. For exponential growth at growth rate smaller than 0.3/hr (e.x., with Cm), we let the pre-culture grow to OD_{600} around 0.4 and then diluted to experimental culture with starting OD_{600} around 0.05. For Cm addition experiments, Cm was added between OD_{600} 0.1 and 0.2.

The growth rates measured in turbidostat was consistently higher than the growth rates measured in the shaker for the same condition. In order to match the growth measurements with the RNA and β-galactosidase measurements, the growth curve data in Fig. 4.2-4.2 was scaled so that the steady-state growth rates obtained from turbidostat agree with the steady-state growth rates obtained from traditional batch measurements. The scaling factor was 1.1.

4.3.4 Total RNA quantification

The RNA quantification method is based on the method described previously (12).

4.3.5 β-galactosidase quantification

β-galactosidase assay was based on the Miller assay with ONPG (ortho-Nitrophenyl-ß-galactoside) as described previously (12).

4.5 Acknowledgements

We are grateful to the technical support for turbidostat from Sarah Cox, Dongyang Li and Suckjoon Jun.

Chapter 4 is coauthored with Terence Hwa. The dissertation author was the primary author of this chapter.

4.6 Figures

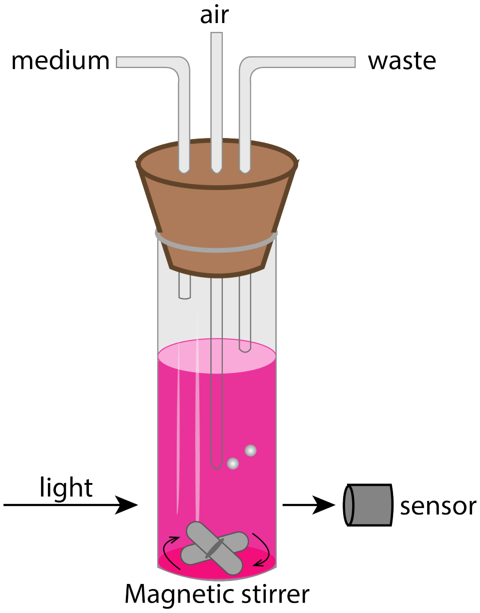


Figure 4.1. Turbidostat setting. The glass tube shown above was used for cell growth, cell density control and cell density reading in turbidostat. The cap was tightly sealed to the tube so that any medium or air exchange were only implemented by the tubings across the cap. From left to right, the three tubings shown above connect to fresh medium, air pump and waste bottles respectively. The medium tubing and air tubing each connected to a valve for flux control. The flux through waste tubing was passively set by the pressure in the tube. In other words, the length of waste tubing inside the tube controls the height (volume) of the cell culture. The length of medium tubing inside the tube was set to be much shorter than the waste tubing. The air tubing goes in the cell culture as deeper as possible but not interfere with cell density reading (absorbance). The position of OD_{600} sensor was set right above the magnetic stirrer.

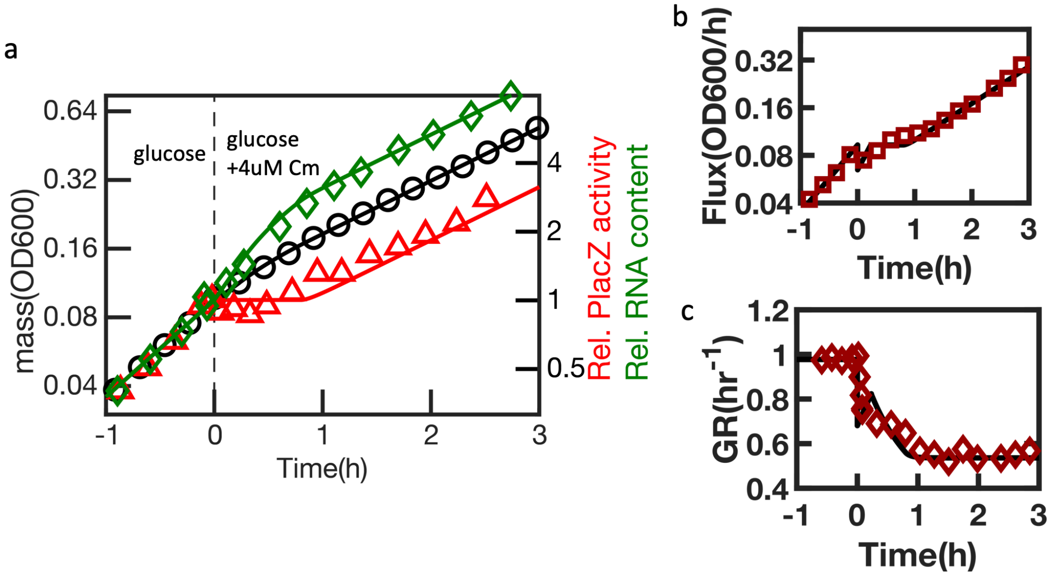


Figure 4.2. Cm downshift in glucose condition. NQ354 (E.coli K-12, lacI-) is used in OD_{600}, RNA and β-galactosidase (represent PlacZ activity) measurements during Cm downshift. (**a**) Before adding Cm, cells were growing exponentially in Mops glucose minimal medium. 4uM Cm (final concentration) was added at time 0 to trigger the shift. OD_{600} (black circle) was obtained by turbidostat. RNA amount per culture volume (green diamond) and PlacZ activity per culture volume (red triangle) were both normalized to the value at t=0. (**b**) shows biomass accumulation flux calculated from OD_{600} (time derivative of OD_{600}) verse time. (**c**) shows instantaneous growth rate verse time. Solid lines in all the plots are model predictions.

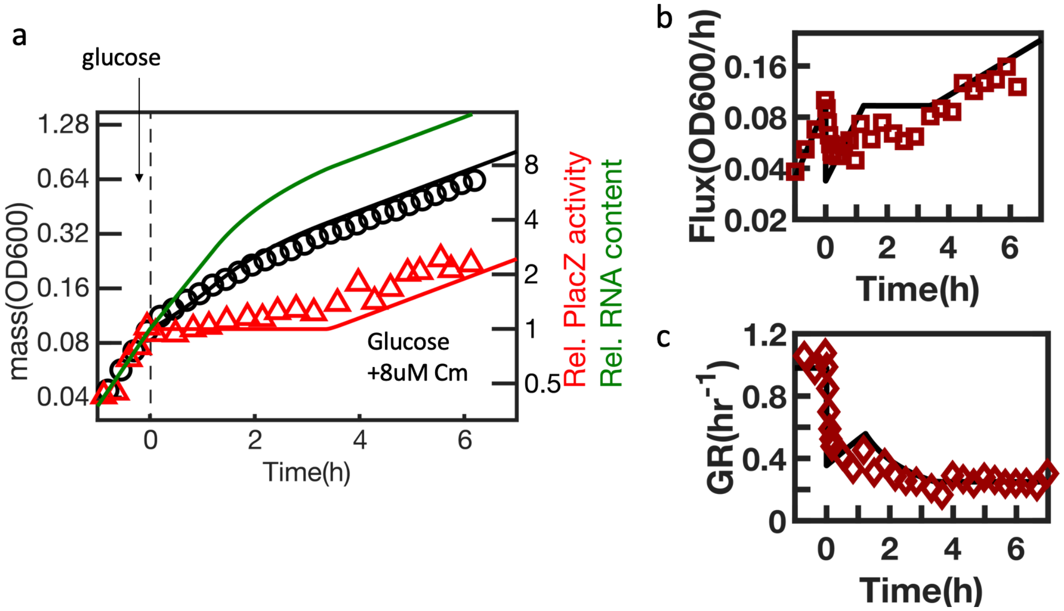


Figure 4.3. Downshift with 8uM Cm addition in glucose medium. 8uM Cm (final concentration) was added at time 0 to trigger the shift. OD_{600} (black circle) was obtained by turbidostat. PlacZ activity per culture volume (red triangle) was normalized to the value at t=0. (**b**) shows biomass flux calculated from OD_{600} (time derivative of OD_{600}) verse time. (**c**) shows instantaneous growth rate verse time. Solid lines in all the plots are model predictions. Same strain was used in this figure as in Fig. 4.2.

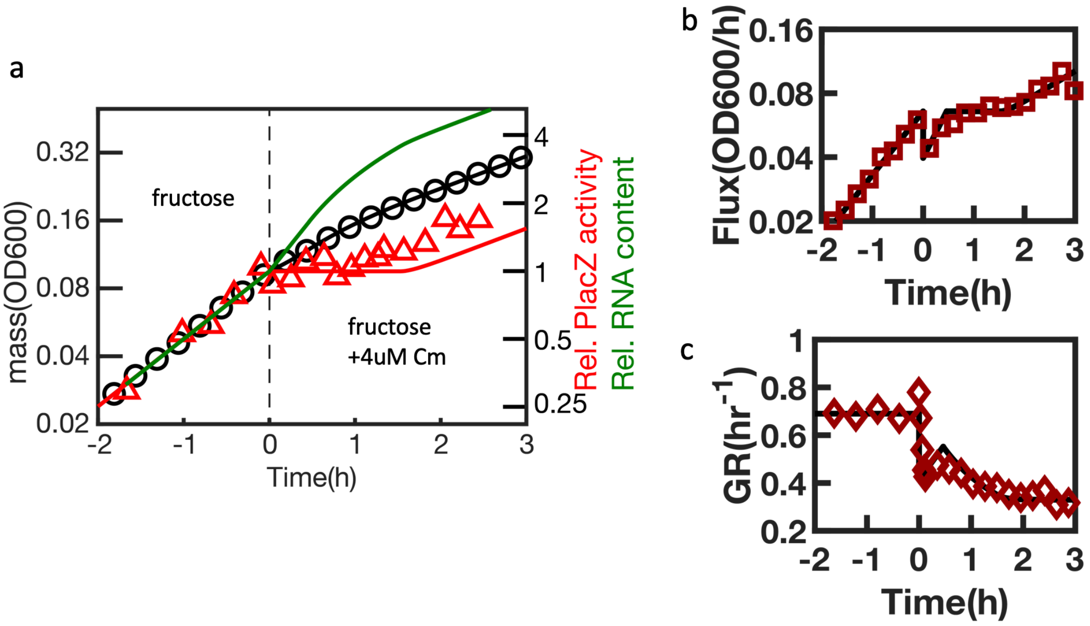


Figure 4.4. Downshift with 4uM Cm addition in fructose medium. 4uM Cm (final concentration) was added at time 0 to trigger the shift. OD_{600} (black circle) is obtained by turbidostat. PlacZ activity per culture volume (red triangle) was normalized to the value at t=0. (**b**) shows biomass flux calculated from OD_{600} (time derivative of OD_{600}) verse time. (**c**) shows instantaneous growth rate verse time. Solid lines in all the plots are model predictions. Same strain was used in this figure as in Fig. 4.2.

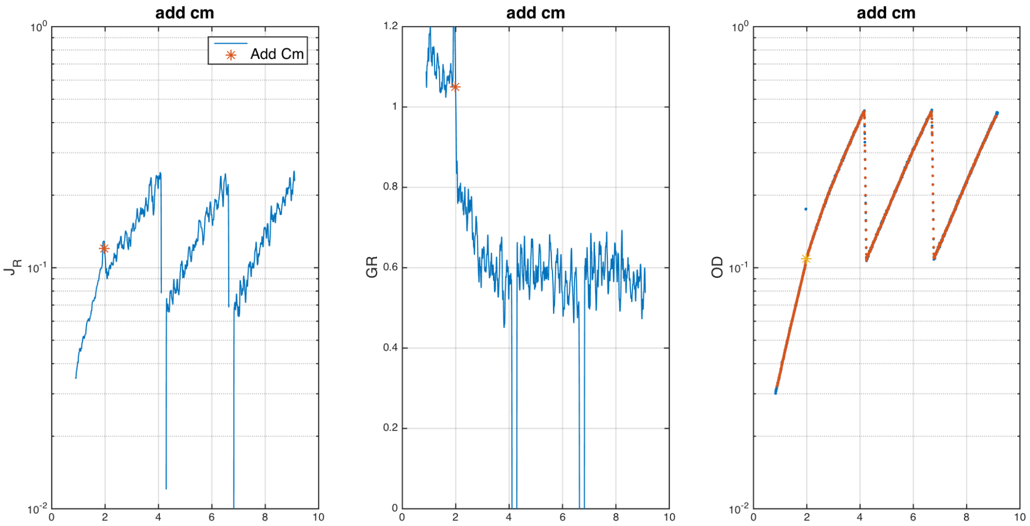


Figure 4.5. Raw growth data of Cm addition experiment shown in Figure 4.2. From left to right, it shows the time course of flux, instantaneous GR and OD_{600} while adding 4uM Cm to glucose minimal medium. Notice the growth rate in turbidostat is higher than that measured from batch experiments. The data reported in Figure 4.2 was scaled from the raw data to match the growth rate (Methods).

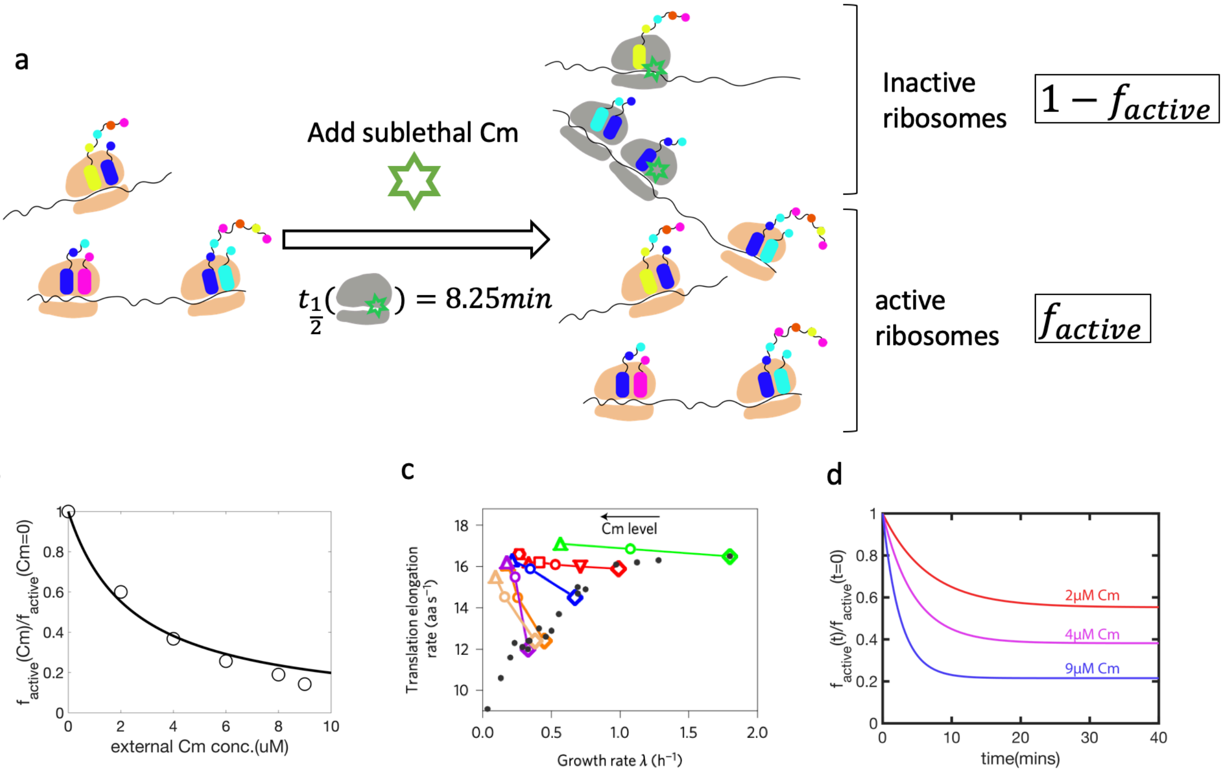


Figure 4.6. Chloramphenicol (Cm) inhibits protein synthesis via decreasing active ribosome fraction. (a) shows a cartoon of Cm effects. When a translating ribosome is hit by a Chloramphenicol molecule, it can't bind to its substrate and will get stalled on mRNA. The half life of Rb-Cm complex is 8.25min (17, 65). Due the the tight binding, most stalled ribosomes won't finish translation (mRNA degradation, ribosomal rescue, etc.). The stalled ribosomes are then considered as inactive ribosomes that won't contributed to protein synthesis in our model. f_{active} (active ribosomal fraction) is used to describe the percentage of active ribosomes over all the ribosomes. (**b**) Compares estimated steady-state f_{active} from Cm dissociate constant with calculated f_{active} from measured growth rate, RNA/protein ratio (represents total Rb mass fraction) and translation elongation rate (ER) in glucose condition through Eq. 4.3. Data from (17). (**c**) came from (17), showing that ER is either constant or increased under Cm treatments. Different color represent different carbon source. ER here only means ER of complete proteins. (**d**) shows simulated kinetics of f_{active} upon Cm addition from the on/off rate of Cm.

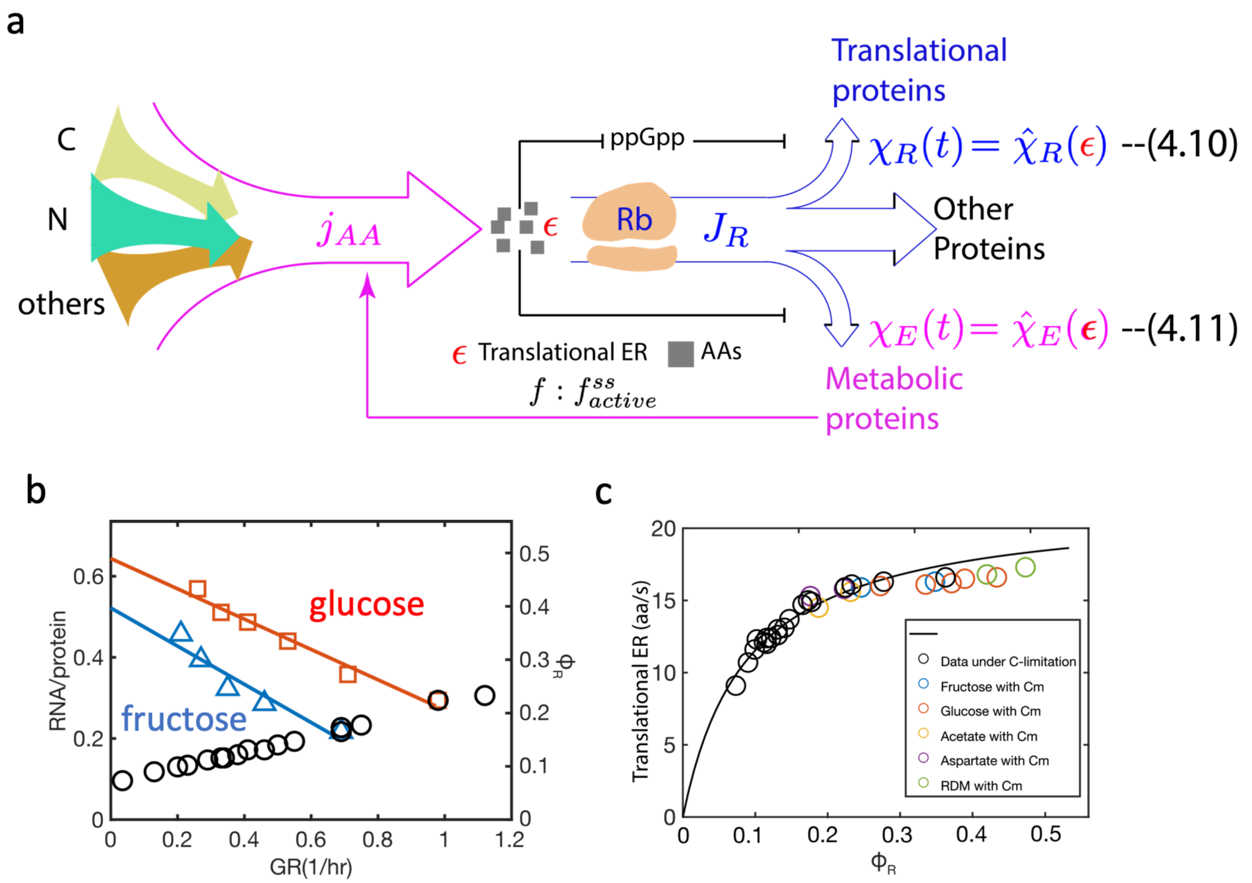


Figure 4.7. Cm downshift model. (**a**): Bacterial proteome are coarse-grained into 3 parts: metabolic proteins (catabolic and anabolic proteins), translational proteins (ribosomal and co-regulated translation-affiliated proteins) and others (fixed core proteins that are Cm-independent) based on Scotts 2010. Metabolic proteins convert external nutrient source to amino acids, which are used as substrates of translational proteins. Amino acid synthesis flux j_{AA} is proportional to the abundance of metabolic proteins M_E and efficient rate k that is only determined by nutrient conditions as shown in the main text. Protein synthesis flux J_R depends on abundance of translational proteins M_R, active ribosomal fraction f_{active} and translational elongation rate ϵ as shown in the main text, where f_{active} is assumed to drop to the post-shift steady-state value f_{active}^{ss} instantaneously when Cm is added. Drop of f_{active} lowers protein synthesis flux. AA pool increases as a results, initiating reallocation of protein synthesis flux towards metabolic proteins (χ_E) and translational proteins (χ_R) via gene regulation through ppGpp (14, 45, 66). Molecularly, AA pool also directly impact translational elongation rate ϵ. Taking the idea of dimension reduction from FCR model (8), ϵ is used as a representative of AA pool and set the global regulation function χ_R and χ_E as shown in Eq. (4.10) and (4.11). The forms of $\hat{\chi}_R(\epsilon)$ and $\hat{\chi}_E(\epsilon)$ are obtained from steady-state relationships. (**b**) Measured steady-state relationship between R/P (represent ϕ_R) and GR under R-limitation (blue triangle: fructose, red square: glucose). Data from (17). Black circles are data under C-lim for comparison. Solid lines are linear fit to Eq. 4.7. Parameters ϕ_{max} and ν are obtained via the fitting. (**c**) Measured steady-state relationship between mass fraction of translational proteins (ϕ_R) and translational ER (in unit of aa/s) under C-lim (black circles) and Cm treatment (colored circles). Data from (17). Solid line is a fit to Michaelis-Menten relation (Eq. 4.12). (ϵ_{max} used in our model later absorbed the factor of α/n_{Rb} to be consistent.)

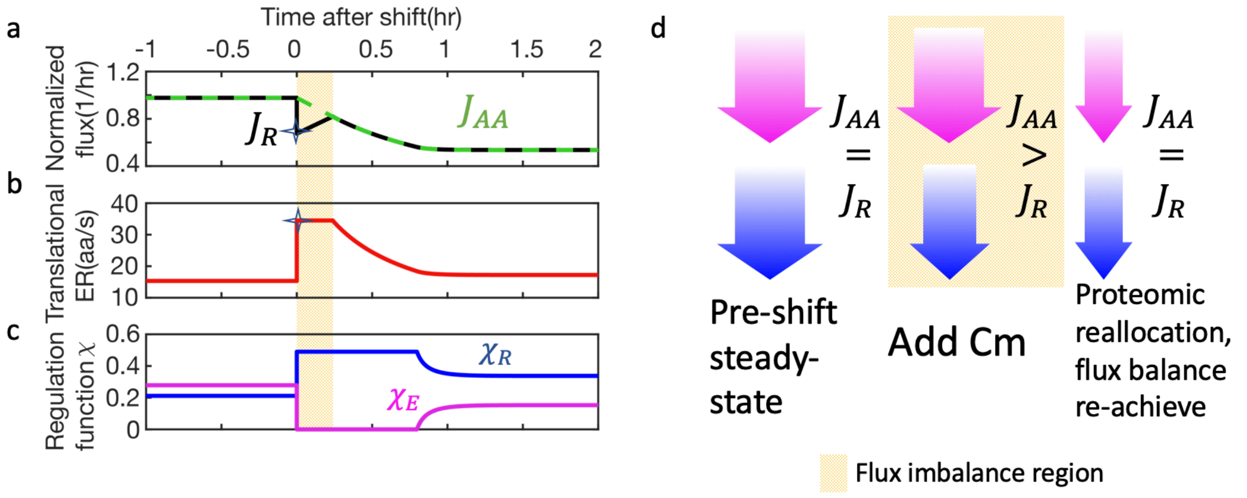


Figure 4.8. Model predictions of flux, translational ER and regulation functions changes with time for the shift in Figure 4.2. Flux imbalance region is yellow shaded in (**a**)-(**d**). (a) shows the AA synthesis flux (green dash line) and protein synthesis flux (black solid line), both normalized to biomass (OD_{600}), during Cm downshift. Here for easy comparison, $J_{AA} = c \cdot j_{AA}$. (**b-c**) shared the same x-axis as (**a**). Right after Cm added (t=0), J_R has an immediate drop (marked by a 4-point star). The decrease in J_R is determined by the drop in f_{active}(fixed by post-shift steady-state GR) and the jump in translational ER (marked by 4-point start in (**b**)). ER right after shift is chosen so that the predicted drop in J_R match the experimental data. J_{AA}, however, changes continuously after Cm addition (**a**), leading to a flux imbalance region marked with yellow background in (**a-d**). When flux is not balanced, Eq. 4.16 can't be applied to described the change in ϵ. ϵ is set to be constant as shown in (**b**). (**c**) shows the time course of regulation function χ_λ (solid purple line) and χ_R (solid blue line). They are functions of ϵ with their natural boundary $0 \le \chi_E, \chi_R \le \phi_{max}$. (**d**) gives an intuitive view of flux in different stages of Cm downshift with arrow size representing flux abundance (purple arrow: J_{AA}, blue arrow: J_R).

4.7 Tables

Table 4.1. Parameter table for Fig. 4.2-4.4.

	Glu+4uM Cm	Glu+8uM Cm	Fru+4uM Cm
λ_i	0.977	0.98	0.69
ϕ_{max}	0.49	0.49	0.3976
ν	3.512	3.512	2.7872
$\epsilon_{max}(\frac{1}{hr})$	10.7	10.7	10.7
K_M	0.088	0.088	0.088
λ_f	0.536	0.25	0.33
boost	2.27	3.3	2.8

Chapter 5

Cellular perception of growth rate and mechanistic origin of bacterial growth laws

5.1 Abstract

Bacteria organize many activities according to their grow rate via the ppGpp signaling system. Yet it is not clear how this signaling system "knows" how fast cells grow. Through quantitative experiments, we show that ppGpp responds inversely to the rate of translational elongation in E. coli. Together with its roles in inhibiting ribosome biogenesis and activity, ppGpp closes a key regulatory circuit that enables the cell to perceive and control the rate of its growth across conditions. The celebrated linear growth law relating the ribosome content and growth rate emerges as a consequence of keeping a supply of ribosome reserves while maintaining elongation rate in slow growth conditions. Further analysis suggests the elongation rate itself is detected by sensing the ratio of dwelling and translocating ribosomes, a strategy employed to collapse the complex, high-dimensional dynamics of the molecular processes underlying cell growth to perceive the physiological state of the whole.

5.2 Introduction

In the past decade, much efforts have been devoted towards characterizing and understanding the allocation of bacterial proteome across different growth conditions (2, 3, 7, 33, 67, 68). Central to the bacterial proteome allocation strategy is the approximate linear increase of the ribosome content with growth rate (69–71), when growth is varied by using different nutrients. This classic bacterial "growth law" is rationalized by Maaloe in term of the need of more ribosomes to synthesize proteins to achieve faster growth rate, when the rate of translational elongation by ribosomes is saturated[9]. This strategy of producing ribosomes "as needed" in different growth conditions forms the basis of the optimal resource allocation strategy, which posits that cells allocate its resources (the proteome in this case) in such a way to maximize its growth (26, 68). However, it is actually long known that the translational elongation rate itself varies across growth conditions (72), which poses a challenge to the rationalization by Maaloe. Moreover, it is known that in conditions where cells are hardly growing, a significant pool of ribosome (the "ribosome" reserves) is kept idle, presumably for rapid transition to fast growth when favorable growth conditions returns (1, 73, 74). Intriguingly, the ribosome reserve kept by cells is not limited to slow growth, but maintained at a constant amount above the minimum needed across growth rates (3, 74, 75). This behavior again challenges the notion of optimal resource allocation for the current growth condition.

One approach towards understanding the bacterial proteome allocation strategy is to follow its regulatory mechanisms, to see how the linear growth law is implemented mechanistically. This involves the sensing and control of the cell's growth rate, since proteome allocation strategy is strongly dependent on the growth rate. Towards this end, we note that Guanosine tetraphosphate (ppGpp) (76) is a key signaling molecule involved in bacterial response to environmental changes and in coordinating growth-rate dependent responses (27, 66, 77, 78). ppGpp signaling has been extensively studied (14, 27, 66, 78–81), both for mechanisms contributing to its synthesis and degradation, and for its downstream effects on hundreds of genes (25, 47, 82, 83), including the synthesis (84, 85) and activity (66, 80) of the translation machinery. Without a doubt, ppGpp-signaling plays a crucial role in responding to the cellular growth rate. Yet, despite the wealth of information at the molecular level, quantitative understanding of how bacteria perceive the state of cell growth is lacking. Here we reveal the underlying signaling strategies employed by E. coli to perceive and respond to growth, established through a series of experiments in which ppGpp and other key physiological variables are quantitatively measured. These strategies provide important insight on the initial question on bacterial proteome allocation strategy as we will discuss at the end.

5.3 Results

5.3.1 A simple, robust relation between ppGpp and translational elongation rate

During environmental changes such as diauxic shifts, *E. coli* responds by producing ppGpp (86). Fig. 5.1a shows a typical diauxic growth curve in minimal medium containing glycerol and a small amount of glucose as the only carbon sources: cells grow exponentially on glucose without utilizing glycerol until glucose is depleted (10), followed by a period of growth arrest (approximately 40-50 min in this case), before fully resuming growth on glycerol. We followed the kinetics of ppGpp accumulation by performing such growth transition experiments in the presence of ^{32}P-orthophosphate. Throughout the transition, labelled nucleotides were extracted and resolved by thin-layer chromatography (Fig. 5.1b). The ppGpp level relative to that of steady-state growth in glucose, denoted as $g(t)$, increased by over 8-fold within the first 10 min of glucose depletion before relaxing to a new steady-state level (Fig. 5.1c).

We also characterized changes in the translational elongation rate (ER, denoted by ε) during the growth recovery period by assaying for the delays in LacZ induction, as previous studies have established that ER determined from LacZ is representative of that of typical proteins (17, 87), and single-molecule study of translation kinetics *in vivo* suggested little variation in ER in the absence of antibiotics (88, 89). Using the induction time obtained at various time t (Fig. 5.6a-c) and taking into account of the initiation time which showed little variation (Fig. 5.7), the instantaneous ER, $\varepsilon(t)$, was deduced throughout the transition period (Fig. 1d, Fig. 5.6d). The data shows an abrupt drop in ER immediately following glucose depletion and a gradual recovery before growth resumed. During the period of growth arrest, the time course of $\varepsilon(t)$ strikingly mirrored that of the relative ppGpp level $g(t)$ (Fig. 5.1c). Scatter plot of the ppGpp level with the reciprocal of ER exhibits a striking linear relation (Fig. 5.6e). Defining the value of the extrapolated ER at $g = 0$ to be ε_{max}, the maximum elongation rate (to be justified below), the empirical relation between the relative ppGpp level and ER can be expressed as

$$g = c \cdot \left(\frac{\varepsilon_{max}}{\varepsilon} - 1\right) \tag{5.1}$$

where $\varepsilon_{max} \approx 19.4\ aa/s$ and $c \approx 4.0$; see Fig. 5.1e.

Since the ppGpp level and ER are both known to change with the cellular growth rate (13, 17), we further examined their mutual relationship during steady-state growth. We grew *E. coli* with different nutrient composition at growth rates ranging from 0.13 h^{-1} to 0.96 h^{-1} (Table 5.1) and measured the steady-state ppGpp levels relative to that in glucose, as well as the corresponding translational elongation rates; see Methods. As growth rate was reduced, ppGpp levels increased while ER decreased, (Fig. 5.2a, 5.2b, red squares), consistent with earlier reports (13, 17) Additionally, ER has recently been shown to increase in the presence of sub-lethal amounts of chloramphenicol (Cm) (17). Accordingly, we observed ppGpp levels to decrease and ER to increase during steady-state growth in the presence of increasing doses of Cm (Fig. 5.2a, 5.2b, green triangles). Owing to the difficulty of detecting low ppGpp levels, we used a Δ*ptsG* strain (NQ1261) (17) which has reduced glucose intake and thereby shows elevated ppGpp levels in the absence of Cm. This strain allowed us to quantify changes in ppGpp level and compare it to the changes in ER under Cm treatment. Scatter plot of the steady state ppGpp level with the reciprocal of ER under carbon limitation again exhibited a linear relation (red squares, Fig. 5.2c). Moreover, those from Cm-inhibited cells fell on the same linear relationship (green triangles). Strikingly, this is the same relationship as the one observed during the diauxic shift (compare with blue circles in Fig. 5.2d), i.e., Eq. (5.1) with the same intercept and proportionality constant.

5.3.2 Regulatory circuit mediated by translation rate links ppGpp quantitatively to growth rate

A steady state relationship between ER and ppGpp level allows the cell to link ppGpp uniquely to the steady state growth rate via a simple regulatory circuit (Fig. 5.3): Due to negligible rate of protein turnover (90, 91), the rate of protein synthesis is given by the product of ER and the total number of active (translating) ribosomes per cell, N_R^{act}. During exponential growth at rate λ, the total peptide synthesis rate is $\lambda \cdot M_P$, where M_P is the total protein mass per cell (in unit of the mass of an amino acid). The number of active ribosomes is the difference between the total number of ribosomes per cell (N_R) and the number of inactive ribosomes per cell (N_R^{inact}). Thus,

$$\lambda M_P = \varepsilon \cdot \left(N_R - N_R^{inact}\right). \tag{5.2}$$

The ratio N_R/M_P, which is proportional to the cellular ribosome concentration (since M_P is proportional to the cell volume (1)), is set by the ppGpp level through regulation of rRNA expression (85). We take this regulatory function to be

$$N_R/M_P \equiv R(g) = a/g \tag{5.3}$$

(with an unknown constant a) since the RNA/protein mass ratio, which is proportional to R, scales as $1/g$; see Fig. 5.4a. The inactive ribosome concentration, which is proportional to N_R^{inact}/M_P, is difficult to quantify directly but can arise due to two potential effects of ppGpp. (i) the binding of ribosomes to various ribosome hibernation factors, e.g., Rmf, Hpf, and RaiA(92–94), the abundances of the latter proteins all increase linearly with the ppGpp level as growth rate is reduced by limiting carbon uptake (Fig. 5.8) and (ii) the inhibition of the translation initiation factor IF2(95–97). Accounting for these roles leading to ribosome inactivation, we assume the form

$$N_R^{inact}/M_P \equiv H(g) = b \cdot g \tag{5.4}$$

for simplicity, with another unknown constant b.

Putting together the form of the regulatory factors in Eqs. (5.3) and (5.4) into Eq. (5.2) leads us to a relationship between the growth rate and the ppGpp level for exponentially growing cells:

$$\lambda = \varepsilon(g) \cdot \big(R(g) - H(g)\big) = \varepsilon(g) \cdot \left(\frac{a}{g} - b \cdot g\right). \tag{5.5}$$

where $\varepsilon(g) = \varepsilon_{max}/(1 + g/c)$ is obtained from the steady-state relation between ε and g (Fig. 5.2d), which is mathematically the same as inverting Eq. (5.1). The two constants a and b in Eq. (5.5) specify the magnitudes of the two regulatory interactions. With appropriate choices of these two constants, the simple model defined by Eq. (5.5) is able to quantitatively capture all the observed correlations among the growth rate λ, the ppGpp level (g), the ribosome content (R), and the elongation rate (ε) under nutrient limitation (Fig. 5.4b-e), with model predictions based on best-fitted values of a and b shown as solid lines. In particular, the model recapitulated the well-known inverse relation between the growth rate and ppGpp level (Fig. 5.4b) (13, 98). This illustrates the general principle that the cell can perceive its own growth rate by incorporating the sensing of ER (Eq. (5.1)) into a simple regulatory circuit (Fig. 5.3) that controls the active ribosome content by the sensor. Equally importantly, the simple correspondence between the growth rate and ppGpp level enables the cell to implement growth-rate dependent control of many cellular functions, ranging from metabolism to cell division control, by simply using ppGpp to control the expression of the relevant genes (64, 77, 82, 99, 100). In a previous study on bacterial growth control by Erickson et al (8), an *ansatz* was introduced in which the translational activity ($\sigma \equiv \lambda/R$) was used to control gene expression during diauxic shifts. The results here establish a one-to-one relation between σ and g (Fig. 5.9), thereby justifying the *ansatz* used in (8).

5.3.3 Linear bacterial growth law obtained with a special condition on elongation rate

At a quantitative level, our model captured the approximate linear relation between the ribosome content and the growth rate (Fig. 5.4c), the celebrated growth law discovered long ago (69, 71, 101). Additionally, the model captured the Michaelis-like relation between the ER and the ribosome content (Fig. 5.4d), substantiated with extensive data collected from many conditions as reported in Dai et al (17). Notably, a fit of the data to the Michaelis-Menten relation (Fig. 5.10) recovers a maximum elongation rate (20.0 ± 1.9 aa/s) that is indistinguishable from $\varepsilon_{max} = 19.4 \pm 1.4$ aa/s defined by taking $g \to 0$ in Eq. (5.1) (Fig. 5.6e). Finally, the model captures the weak dependence of the elongation rate at different growth rate (Fig. 5.4e), with the minimal ER in slow growth condition, denoted by ε_0, whose value is close to one-half of ε_{max}.

The emergence of a simple linear relation between the ribosome content R and the growth rate λ is surprising, given the nonlinear regulatory effects exerted by ppGpp (Eq. (5.5)). In fact, a variety of relationships among these quantities is possible for generic values of a and b (Fig. 5.11). However, if the regulatory parameters a and b are such that the ratio $\varepsilon_0 \colon \varepsilon_{max}$, is exactly one-half, then mathematically the model yields an exact linear relation between R and λ, with the slope given by $1/\varepsilon_{max}$, and an exact Michaelis-Menten relation between the ER and the ribosome content, with the maximal ER being ε_{max} (Appendix 1). Thus, to the extent that Eqs. (5.3) and (5.4) capture the forms of the regulatory functions, prescribing the appropriate regulatory parameters a and b to enforce ε_0 being approximately one-half of ε_{max} is necessary for the emergence of the approximate linear growth relation between the ribosome content and the growth rate. (We have separately shown that adding offsets to the simplest forms of the regulatory functions used in Eqs. (5.3) and (5.4) do not affect the quality of the fit; see Fig. 5.12.)

Maintenance of ER above a minimal level is clearly of physiological importance, as too low an ER would lead to problems in the processivity of protein synthesis (102, 103). Another physiological requirement is the maintenance of a sufficient ribosome reserve at slow growth, denoted by R_0, needed for rapid growth recovery when favorable nutrient conditions return (1). Both physiological requirements are satisfied by employing hibernation factors to inactivate ribosomes. By employing both positive and negative regulation through distinct promoters (Fig. 5.3), the cell can readily attain the required values of ε_0 and R_0 by simply prescribing the regulatory parameters a and b; see Fig. 5.13 and Appendix 1. To keep ε_0 high while also maintaining a ribosome reserve is possible in principle; see Fig. 5.14. With high ER (e.g., above 90% of ε_{max}, dashed line in Fig. 5.14a), the ribosome content would even be moderately reduced at fast growth rate (Fig. 5.14b), a fitness benefit from the proteome allocation perspective (104, 105). However, this strategy would also require significant improvement in detecting very short translational dwell time (right axis, Fig. 5.14d). Thus, the choice of using $\varepsilon_0 \approx \varepsilon_{max}/2$ may reflect a compromise between the physiological demand for keeping ε_{min} high and the molecular constraint for detecting small changes in ER in order to sense slow growth and enforce growth-rate dependent regulation.

5.3.4 Mechanism of sensing the translational elongation rate

It is important to point out that the steady-state results presented here (Eq. (5.5) and Fig. 5.3) depend only on the existence of the empirical relation given by Eq. (5.1). We now return to discuss the causal link and the mechanistic origin. Towards this end, the first question to address is whether ppGpp or ER is the primary driver of this response. One scenario is that ppGpp rises in response to some unknown "starvation signal" and the resulting increase in ppGpp then reduces the ER. However, we will show shortly below that a mutant in which ppGpp does not rise instantaneously still exhibits a strong immediate drop in ER. An alternative scenario is that the drop of ER occurs first, and this drop is itself the signal that drives up ppGpp. The latter scenario is supported by metabolomic study which found the amino acid pools (and particularly the glutamate pool) to drop sharply and immediately following glucose runout (106), thus imposing obligatory reduction in ER. Sensing the drop in ER could therefore be an effective strategy to sense the nutritional status of the cell.

We next examine the form of the response (5.1) in terms of the known mechanisms of ppGpp synthesis and degradation. It will be convenient to re-express ER and ε_{max} in Eq. (5.1) in terms of the elemental steps of the translation cycle (Fig. 5.5a): a time τ_{dwell} where the ribosome dwells on the A-site waiting for the cognate charged tRNA, and a time τ_{trans} for peptidyl transfer and translocation to the next codon. This changes Eq. (5.1) to $g = c \cdot \tau_{dwell}/\tau_{trans}$, with the maximal elongation rate $\varepsilon_{max} = \tau_{trans}^{-1}$ identified with the case where $\tau_{dwell} \to 0$. (Incidentally, the special limit of elongation rate at slow growth being one-half of ε_{max} corresponds simply to $\tau_{dwell} = \tau_{trans}$ at slow growth.)

Next, detailed analysis based on flux balance (Appendix 2) establishes a simple relation between two pools of actively translating ribosomes, those in the dwelling state (of concentration R_{dwell}) and those in the process translocation (of concentration R_{trans}):

$$R_{dwell} \cdot \tau_{dwell}^{-1} = R_{trans} \cdot \tau_{trans}^{-1}, \tag{5.6}$$

Eq. (5.6) is simply a condition of *detailed balance* between the flux of ribosomes transitioning from the dwelling state to the translocation state, and the flux transition from the translocation state back to the dwelling state, with $R_{dwell} + R_{trans} = R_{act}$ being the total concentration of actively translating ribosomes. In terms of these ribosome pools, Eq. (5.1) then becomes

$$g = c \cdot R_{dwell}/R_{trans}, \tag{5.7}$$

i.e., the ratio of the two pools of ribosomes.

In a simple model of ppGpp comprised of rapid equilibration between synthesis and degradation such that $\frac{d}{dt}g = \alpha - \beta \cdot g$, the ppGpp level is given by the ratio of the synthesis rate (α) and the specific degradation rates (β). A simple scenario giving rise to the empirical relation (5.1) or (5.7) is therefore to have the synthesis rate $\alpha \propto R_{dwell}$ and the degradation rate $\beta \propto R_{trans}$. The effect of dwelling ribosomes on ppGpp synthesis is well-supported molecularly based on the known structure of the ppGpp synthetase RelA in complex with the ribosome (107, 108) as well as earlier biochemical studies (109, 110), which find that RelA is activated (i.e., synthesizing ppGpp) only when it is complexed with an uncharged tRNA together at the A-site, i.e., when the ribosome is in the dwelling state (Fig. 5.5b left). Less is known about ppGpp degradation, which is solely catalyzed by SpoT in *E. coli* (78, 111). The empirical relation (5.1) excludes a model with constitutive hydrolysis of ppGpp by SpoT (Appendix 2) and instead predicts SpoT hydrolysis activity to be stimulated by the translocating ribosomes (Fig. 5.5b right). Consistent with these ideas, we find deletion of *relA* to disrupt the linear relationship between ppGpp and ER during the diauxic transition (Fig. 5.16), with the remaining nontrivial ppGpp dynamics attributed to the response of SpoT, which is also the only other enzyme capable of ppGpp synthesis in *E. coli* (78, 81). [In this strain, a much slower accumulation of ppGpp occurred following glucose runout compared to the wild type (Fig. 5.16b); yet ER dropped more and for a longer period, thus negating the afore-mentioned scenario that the drop of ER could be a result of ppGpp accumulation.]

Intriguingly, *relA* deletion has no effect on either the ppGpp level or the elongation rate in steady state growth (Fig. 5.17), consistent with the knowledge that RelA is not essential during steady-state growth (13). The maintenance of the relationship (5.1) by the Δ*relA* strain can only be attributed to SpoT. Given Eq. (5.7), our data thus suggest that in steady state, SpoT has acquired the ability to sense both the dwelling and translocating ribosomes, in ways that it is not capable of during transient shift. How SpoT may sense different states of the ribosome is however not known. While there is some evidence of SpoT associated with the ribosome (112), the functional significance of this association is unclear. It is likely that this process is aided by some unknown mediators which interact with the ribosomes in ways analogous to RelA (Fig. 5.5b) and convey that information to SpoT.

5.4 Discussion

Coordination of bacterial behaviors with growth rate is widely observed (7, 64, 113, 114). While the ppGpp signaling system is known to be central to the growth-rate dependent responses (27, 77), how growth rate is perceived and used for regulation is not known at the quantitative level. Our work establishes a missing central element in E. coli's strategy to perceive its state of growth and respond to it: Changes in nutrient environment is immediately reflected in the rate of translational elongation, and this in turn sets the ppGpp level within a time scale of 5-10 mins (Fig. 5.1). ppGpp's well-established regulation of the ribosome content and activity, together with its robust relation with the elongation rate (Fig. 5.3), then forges a unique relation between ppGpp and growth rate since the latter is given quantitatively by the product of active ribosomes and their elongation rate.

As translational elongation is affected by many metabolic processes in the cell, monitoring the elongation rate is an effective strategy to diagnose the state of cell growth independently of specific nutrient bottlenecks. This is the origin of the generality of the phenomenological growth laws, i.e., why the same quantitative relation between the ribosome content and the growth rate is sustained regardless of whether cell growth is limited by carbon supply (3, 12, 69), nitrogen supply (115), partial auxotrophy (7, 12), or drugs which inhibits tRNA charging (116). This mechanism also predicts generally that for perturbations not captured by a repartition of the two ribosomal states, including antibiotics (17, 117), phosphate limitation (115) and lipid stress (118), growth rate perception is distorted and the linear relation between ribosome content and growth rate is altered.

One surprising finding of our analysis is the important role played by factors that inactivate the ribosome. These factors control the amount of ribosome reserve kept by cells at slow growth (119), and the ribosome reserve is important because it affects the rapidity of growth recovery when good growth condition returns (1, 73, 74, 119). However, the necessity of employing ribosome inactivating factors can only be appreciated in light of cell physiology at slow growth: If inactivating factors are not used, then keeping a finite ribosome reserve would require the translational elongation rate to drop to low levels in slow growth conditions, leading to problems in protein synthesis68. Thus, these inactivating factors serve as an effective tool to maintain elongation rate while setting aside ribosome reserve at slow growth.

Nevertheless, the deployment of these factors complicates the overall regulation of active ribosomes, making it difficult to understand how the well-known linear relation between ribosome content and growth rate arises. Our analysis shows that the linear relation emerges for a special choice of regulatory parameters such that the elongation rate at very slow growth is one-half of the maximal elongation rate, or alternatively, the time a ribosome spends in the dwelling state of the translation cycle is not longer than the time it spends in the translocating state, i.e., $\tau_{dwell} < \tau_{trans}$. It should be noted that this implementation of the linear ribosome-growth rate relation is very different from the existing ideas, based e.g., on a Michaelis-relation between translational factors and the elongation rate (75), or on optimizing the steady-state growth rate (26, 68). Regarding the former, this work shows that the appearance of a Michaelis relation between the translational factors and the elongation rate (17) is actually a consequence of regulation involving the ribosome inactivating factors. Regarding the latter, the existence of ribosome reserve which is detrimental to steady state growth but needed for rapid transition from slow to fast growth, is typically glossed over in optimization theories. However, the specific mechanism used to set the reserve, i.e., the use of ribosome inactivating factors which are needed for maintaining a finite elongation rate in the presence of inactive ribosomes, and which produces a constant "offset" (R_0) in the linear ribosome-growth rate relation across the entire growth rate range, calls into question the popular notion that the linear ribosome-growth rate relation is predominantly driven by the optimization of steady-state growth. Instead, it suggests that setting aside a pool of ribosomes as a reserve is something specific that the cell intends to accomplish in its proteome allocation strategy, despite having a cost to the steady-state growth rate.

Turning to the mechanisms for sensing the translational elongation rate, our data and analysis show that it is based on the ratio of translating ribosomes in their two alternating states (Fig. 5.5). Sensing of the dwelling ribosomes fits well with the elaborate molecular design known for RelA (81, 110). However, our analysis suggests that this RelA-based mechanism is insufficient by itself: As RelA is not required in steady state, ppGpp synthesis activity by SpoT must somehow also be able to respond to dwelling ribosomes. The synthesis activity of SpoT was recently shown to be correlated with the levels of acetyl phosphate, a glycolytic intermediate (120). But glycolytic flux is not necessarily a proxy for dwelling ribosomes. While SpoT has been shown to interact with ribosome associated proteins[70], currently it remains unclear what interactions can enable SpoT to sense dwelling ribosomes. Moreover, even in the presence of RelA, the empirical relation observed between ppGpp and elongation rate during growth transition requires additional regulation by SpoT (Appendix 2). Finer details of the ppGpp signaling system may be revealed by quantifying how the ppGpp-ER relation is modified for various combinations RelA and SpoT mutants in future studies.

At a broader level, this study provides a rare, trackable example of how cells perform dimensional reduction at the molecular level to attain crucial physiological information at the cellular level (121). The key "trick" E. coli uses to collapse the high-dimensional complexity of the metabolic state of a cell, e.g., involving 20 amino acid synthesis pathways and the charging of over 60 tRNAs (Fig. 5.5c), is to take advantage of detailed balance between the two alternating states of the elongation ribosome, so that the ratio of the ribosome dwelling and translocation time, which reflects a weighted average of the tRNA charging ratios, can be deduced from the ratio of the dwelling and translocation pools of ribosomes regardless of molecular details (Fig. 5.5a, b and Appendix 2). Identifying and elucidating further instances of such strategies of dimensional reduction employed by cells will be instrumental in fundamentally advancing our understanding of the connection between molecular interaction and cellular physiology for prokaryotes as well as for eukaryotes.

5.5 Methods and Materials

5.5.1 Growth media composition and culture conditions

Steady-state and growth transitioning cultures were grown in MOPS based minimal media (122) supplemented with various carbon sources and chloramphenicol as indicated in **Table 5.1**. All cultures were grown at 37°C with shaking at 250 rpm. For every experiment, culturing was carried out through sequential propagation of seed cultures in LB, pre-cultures in the experimental medium, and the experimental cultures. Single colonies from fresh LB agar plates were first grown in LB broth for 6 hrs as the "seeding culture". In the pre-culturing step, depending on the experiment, cells from seeding cultures were diluted into appropriate media such that the pre-cultures would still be in exponential growth phase after overnight growth. Media used for pre-culturing in steady-state experiments were same as the experimental media (**Table 5.1**). For the glucose to glycerol transitions, optical density was monitored at 600 nm (OD_{600}) to follow the growth transition kinetics. Pre-cultures were grown in medium supplemented with 20mM glucose and 20mM glycerol, to avoid glucose run-out during the pre-culturing step. Exponentially growing pre-cultures were then diluted in the appropriate experimental medium (pre-warmed) at an initial OD_{600} of ~0.005 and various measurements were carried in the OD_{600} range of 0.1-0.4.

5.5.2 Strain construction

Wild type *E. coli* K-12 NCM3722 (29, 30) and its derivatives were used in this work. HE838 *(ΔrelA*) was constructed using the λ-Red recombinase method (123) as follows. The km resistance gene was amplified from pKD13 using chimeric oligos relA1-P1 and relA2-P2 (**Table 5.2**). The PCR products were electroporated into NCM3722 cells expressing Lambada-Red proteins encoded by pKD46. The Km resistant colonies were confirmed by PCR and sequencing for the replacement of the region harboring *relA* by the km gene.

5.5.3 Translation elongation rate (ER) measurement

ER was measured using LacZ as a reporter as described in Dai et al (17) with modifications. Depending on the experiment, 10 ml cultures were either grown in different steady state conditions or as undergoing glucose to glycerol growth-transition. When cultures reached OD_{600}=0.4 (for steady-state growth) or at specific time-points during the growth transition, 5 mM isopropyl-β-D-thiogalactoside (IPTG) was added to induce the *lac* operon. Immediately after induction, 500 ml samples were taken at 10 s or 15 s intervals to pre-cooled (-20°C) tubes with 20 ml of 0.1M chloramphenicol and then rapidly frozen by dry ice. Samples were stored at -80°C before beta-galactosidase assay. 4-methylumbelliferyl-D-galactopyranoside (MUG, a sensitive fluorescence substrate) was used to measure LacZ activities in this work. Briefly, each sample was diluted by Z-buffer by 5-fold and added to 96-well plate to a volume of 200 ml. Plate was warmed at 37°C for 10 mins before adding MUG. Tecan (SPARK) plate reader was used for MUG injection and fluorescence readings. 20 ml of 2mg/ml MUG was injected to each well and fluorescence intensity (365nm excitation filter, 450nm emission filter) was measured every 4 mins for 2 hours. In the linear range of fluorescence intensity vs time plot, a linear fit was applied to obtain the slope as the relative LacZ activity for each sample. By plotting the square root of the relative LacZ activity above basal level against time (124) (Schleif plot), the lag time for the synthesis of the first LacZ molecule (T_{first}) was obtained for each sample; see **Fig. 5.6a-c**). Similar measurements using the α-complement of LacZ (strain NQ1468, (17)) allows us to estimate the translational initiation time; see Fig. 5.7. The ER measured in this work was found to be slightly higher than that reported by Dai et al (17), likely due to the higher sensitivity of the substrate MUG compared to ONPG as used previously (17).

5.5.4 ppGpp measurement

ppGpp measurements were carried out as described by Cashel (125) with minor modifications. Typically, experimental cultures were grown in 3ml volumes. Labelling was carried out when the experimental cultures grew to $OD_{600} = 0.02$ by adding 0.1mCi ^{32}P-orthophosphate (Perkin Elmer) per ml culture. For steady-state growth, 20 µl aliquots were drawn at various OD_{600} values between the range 0.1-0.4 (see **Fig. 5.18a**), and added to an equal volume of pre-chilled 10% formic acid. For cultures undergoing diauxic shift, 20 µl aliquots were drawn at various time points during the transition, and added to an equal volume of pre-chilled 10% formic acid. Formic acid-extracts were spun down at 13k rpm for 10 minutes and a total of 2 µl supernatant was spotted 0.5 µl at a time near the base of a PEI-Cellulose F thin layer chromatography plate (Millipore). The spots were dried and nucleotides were resolved using freshly prepared 1.5M KH_2PO_4 (pH 3.4). The TLC plates were dried and exposed to a phosphorimaging screen for 24-36 hours. Chromatograms were imaged using a Typhoon FLA 9500 scanner (GE) and analyzed using Fiji software. For steady state growth conditions, the slope of ppGpp signal intensities versus OD600 were compared among different cultures to obtain the relative ppGpp levels, as shown in **Fig. 5.18c**. For any batch of measurements, the ppGpp level from a sample of NCM3722 growing steadily in MOPS glucose was always included as a reference. All measurements in that batch were normalized to the glucose-grown reference in the same batch.

5.5.5 Total RNA and Protein measurement

Total RNA was measured using the method of Benthin et al (126), and protein was measured using the Biuret method (127), with minor modifications as described by You et al (12).

5.6 Acknowledgements

We are grateful to technical advices from Xiongfeng Dai and Manlu Zhu in early stages of this work and helpful discussions on the general issue of biological dimension reduction with Boris Shraiman. This work was supported by the NSF through Grant MCB-1818384 and the NIH through Grant GM109069.

Chapter 5, in full, has been approved for publication by PNAS, 2022. The authors are Chenhao Wu, Rohan Balakrishnan, Nathan Braniff, Matteo Mori, Gabriel Manzanarez, Zhongge Zhang, Terence Hwa. The dissertation author and Rohan Balakrishnan were the primary investigators and authors of this material.

5.7 Figures

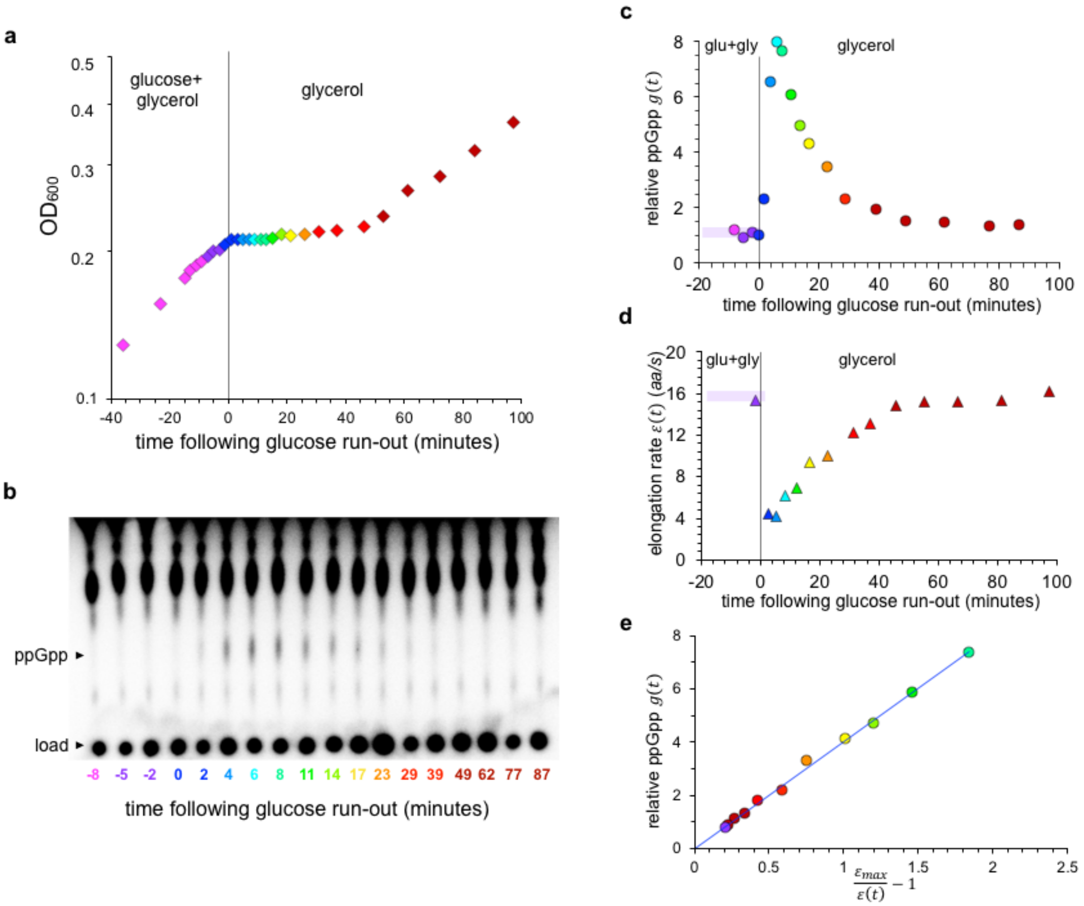


Figure 5.1: Relation between ppGpp and the translational elongation rate during growth transition. **a,** Growth kinetics of *E. coli* K-12 NCM3722 monitored by measuring the optical density at 600 nm (OD_{600}) during the diauxic transition from glucose to glycerol. The same color-scheme is used across the panels to match different measured quantities to samples taken at different time during the growth transition. **b,** Resolution of ppGpp in cells sampled at different time during the growth transition by thin layer chromatography (TLC). The spots at bottome correspond to sample loading and the migrated ppGpp are indicated alongside the chromatogram. Signal intensity from the chromatograph was used as the measure of the ppGpp level. **c,** Signal intensity for ppGpp obtained from panel b is normalized by that before the growth shift, denoted by $g(t)$, is plotted against the time t at which the sample was taken. **d,** The translational elongation rate $\varepsilon(t)$ was obtained as described in Fig. 5.6, and plotted against the time t at which sample was taken. **e,** ppGpp levels $g(t)$ are plotted against the reciprocal of the corresponding elongation rates normalized by the maximum elongation rate (ε_{max}), defined as the value of the elongation rate extrapolated to $g = 0$ (see Fig. 5.6e). The line is the best-fit of the data to Eq. (5.1), with $c \approx 4.0$ and $\varepsilon_{max} = 19.4 \pm 1.4$ aa/s.

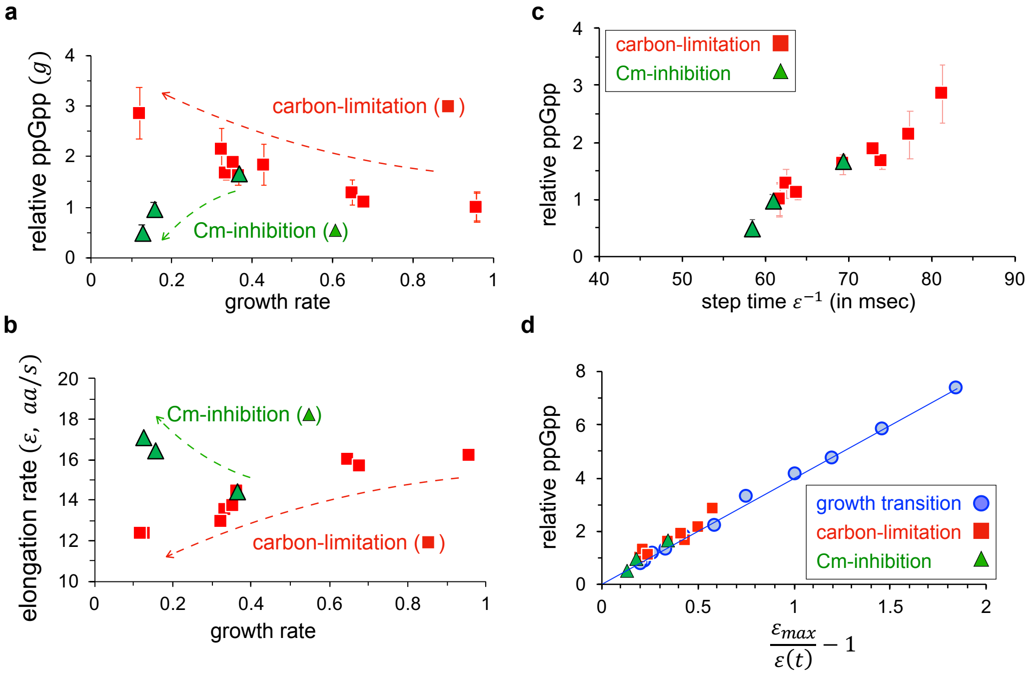


Figure 5.2: Relation between ppGpp and translational elongation rate during steady-state growth. a, ppGpp levels relative to that in the glucose minimal medium, g, are plotted against the corresponding growth rates for cells under steady state carbon-limited growth (red) and translation-limited growth (green) for cells treated with sub-lethal doses of chloramphenicol (Cm); see Table 5.1. For each growth medium, ppGpp level was obtained by measuring 4 samples taken from exponentially growing cells at different ODs and using linear regression (Methods and Fig. 5.18). Error bars represent the uncertainty in the linear fit. **b,** Translation elongation rates are plotted against the steady state growth rates for carbon-limited (red) and Cm treated (green) cells. **c,** Scatter plot of reciprocal elongation rates (or the step time for ribosome advancement) in milliseconds and the relative ppGpp levels (g) measured during steady-state growth for wild-type *E. coli* under carbon limitation (red) and translation inhibition (green). **d,** The same measurements from steady-state growth (panel c) are replotted by normalizing the elongation rate to ε_{max} together with the data collected under growth transition from Fig. 5.1e (blue symbols) for comparison. The line is the same as that in Fig. 5.1e.

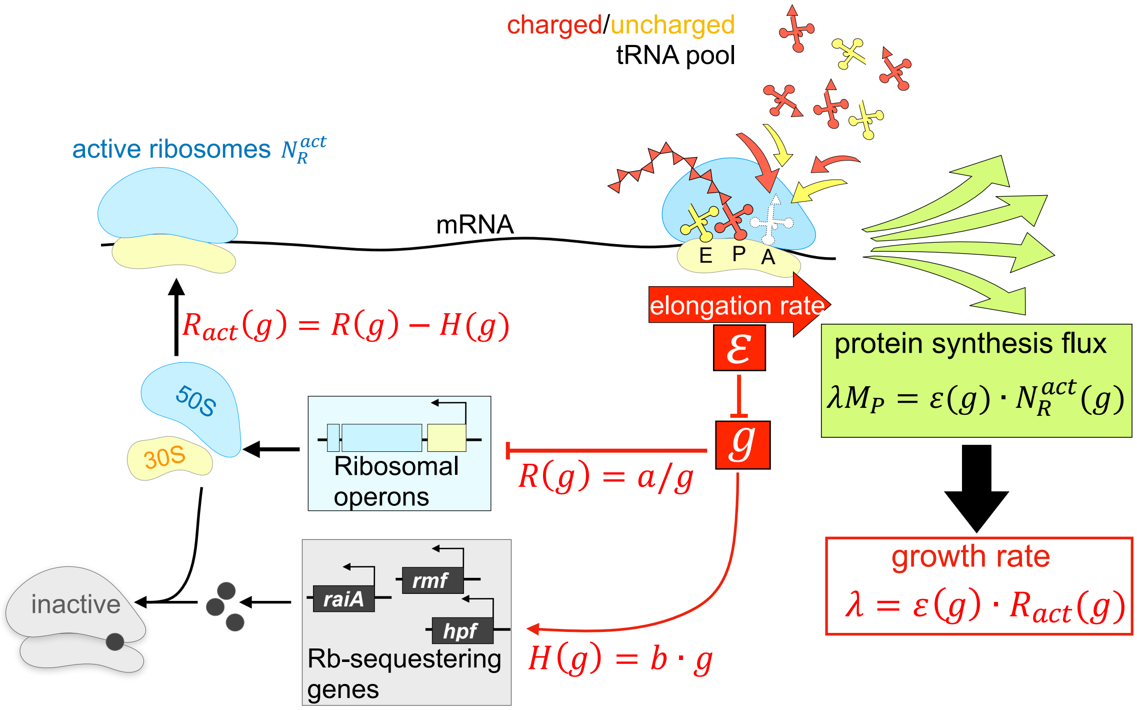


Figure 5.3: Regulatory circuit connecting ppGpp to growth rate. The steady state protein synthesis flux $\lambda \cdot M_P$ is given by the product of the elongation rate ε and the number of active ribosomes, N_R^{act}. Because ε is simply connected to the ppGpp level g (Fig. 5.2d as summarized by Eq. (5.1)), and the active ribosome content $R_{act} \equiv N_R^{act}/M_P$ is given by the difference between the total ribosome content $R(g)$ (Fig. 5.4a and Eq. (5.3)) and the content of the ribosome-sequestering elements $H(g)$ (Fig. 5.8 and Eq. (5.4)), each of which is a function of g due to ppGpp-mediated regulation, it follows that $\lambda = \varepsilon(g) \cdot R_{act}(g)$ is a function of g. This gives rise to the correlation between the ppGpp level and growth rate (13, 98) (Fig. 5.4b). By using ppGpp to regulate a spectrum of cellular processes (14, 66, 81), the cell thus manages to link the regulation of these processes to the growth rate, leading to the appearance of "growth-rate dependent" control (77, 99, 113)

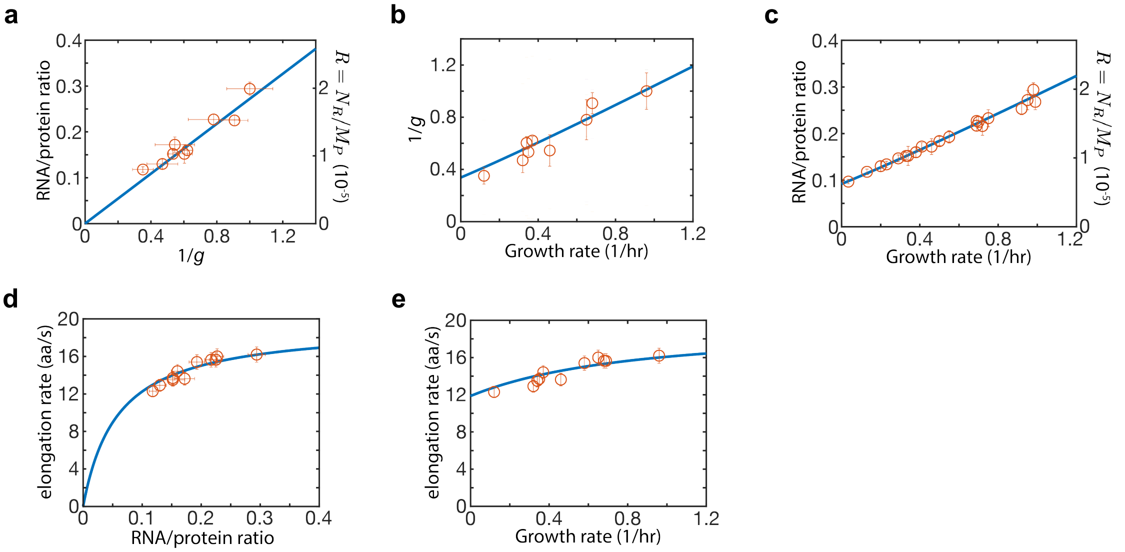


Figure 5.4: Model data comparison. a, RNA-protein ratio (circles, left vertical axis) is proportional to the reciprocal of ppGpp level. This ratio is taken to be proportional to the ribosome content, N_R/M_P (right vertical axis), and is used as an input to the model; see Eq. (5.3). The proportionality constant between RNA-protein ratio and N_R/M_P, denoted by η, is one of the 3 fitting parameters of the model. The best-fit is shown by the line. **b,** The approximate linearity between the RNA-protein ratio and the growth rate (circles, left vertical axis) is well accounted for by the model (line), as is the approximate linear relation between the growth rate and the reciprocal of ppGpp level (panel **c**), and the weak relation between the ER and growth rate (panel **d**), and the Michaelis relation between ER and the RNA-protein ratio (panel **e**). The model is described by Eqs. (5.1) and (5.5). The values of the best-fit parameters were $a \cong 1.85 \times 10^{-5}$, $b \cong 2.11 \times 10^{-6}, \eta \approx 6.8 \times 10^{-5}$.

Figure 5.5: Sensing of the translational elongation rate by ppGpp. a, A cycle of translation elongation constitutes the loading of a cognate charged tRNA to the A site (taking time τ_{dwell}), followed by peptidyl transfer and the translocation of mRNA/tRNA (taking time τ_{trans}). The total time for one cycle, given by the reciprocal of the translational elongation rate ε, is thus given by $\varepsilon^{-1} = \tau_{dwell} + \tau_{trans}$. τ_{trans} depends on the molecular properties of the translation machinery and τ_{dwell} depends on the concentration of uncharged tRNAs. Hence, long dwell times would lead to slow elongation speeds after a nutrient downshift owing to the increased uncharged tRNA levels, while the fastest elongation speed is obtained when $\tau_{dwell} \to 0$, in which case $\varepsilon \to \tau_{trans}^{-1} \equiv \varepsilon_{max}$. The empirical relation observed between ppGpp and the elongation rate ε, Eq. (5.1), can thus be alternatively written as $g = c\tau_{dwell}/\tau_{trans}$. According to the analysis of Appendix 2, the ratio of dwelling and translocation time is given by the ratio of the dwelling to translocating ribosomes, whose concentrations are R_{dwell} and R_{trans} as indicated in the figure. As the transition from the dwelling state to the translocation state occurs with rate τ_{dwell}^{-1} while the transition from the translocation to the dwelling state occurs with the rate τ_{trans}^{-1}, the condition of detail balance imposes that $\tau_{dwell}/\tau_{trans} = R_{dwell}/R_{trans}$ regardless of the values of hundreds of molecular parameters underlying the translation process (Appendix 2). It then follows that the empirical relation (5.1) can be obtained if, e.g., the synthesis of ppGpp is proportional to R_{dwell} and the hydrolysis of ppGpp is proportional to R_{trans}. **b,** According to biochemical and structural studies (81, 107, 108, 110, 128), ppGpp synthesis is activated when the A-site of the ribosome is loaded by a RelA-bound uncharged tRNA. This provides a mechanistic model for the control of ppGpp synthesis rate by R_{dwell}, the concentration of dwelling ribosomes. Although this RelA-mediated synthesis activity would provide elevated levels of ppGpp in poor nutrient conditions, it is insufficient to generate the empirical relation described in Eq. (5.1); see Figure 5.15 with details in Appendix 2. Instead, the involvement of state-dependent ppGpp hydrolysis is predicted for a full account of the empirical relation. **c,** As charged tRNAs must be delivered to the ribosomal A-site to complete each step of translation, the distribution of dwelling and translocating ribosomes is dependent on the metabolic fluxes directed towards all the biosynthetic precursors needed for protein synthesis, represented here by the orange cloud: triangles, amino acids; clover leaves, tRNA; colored arrows, fluxes. The translational elongation rate is therefore a single quantity that reflects the combined status of the hundreds of diverse metabolic reactions underlying protein synthesis and cell growth (Appendix 2). In this sense, the mechanism of ER-sensing described in panel b is a dimensional reduction technique employed by *E. coli* to convey the nutrient status by a single quantity, the level of ppGpp. The latter is further connected to the growth rate via the regulatory circuit of Figure 5.3.

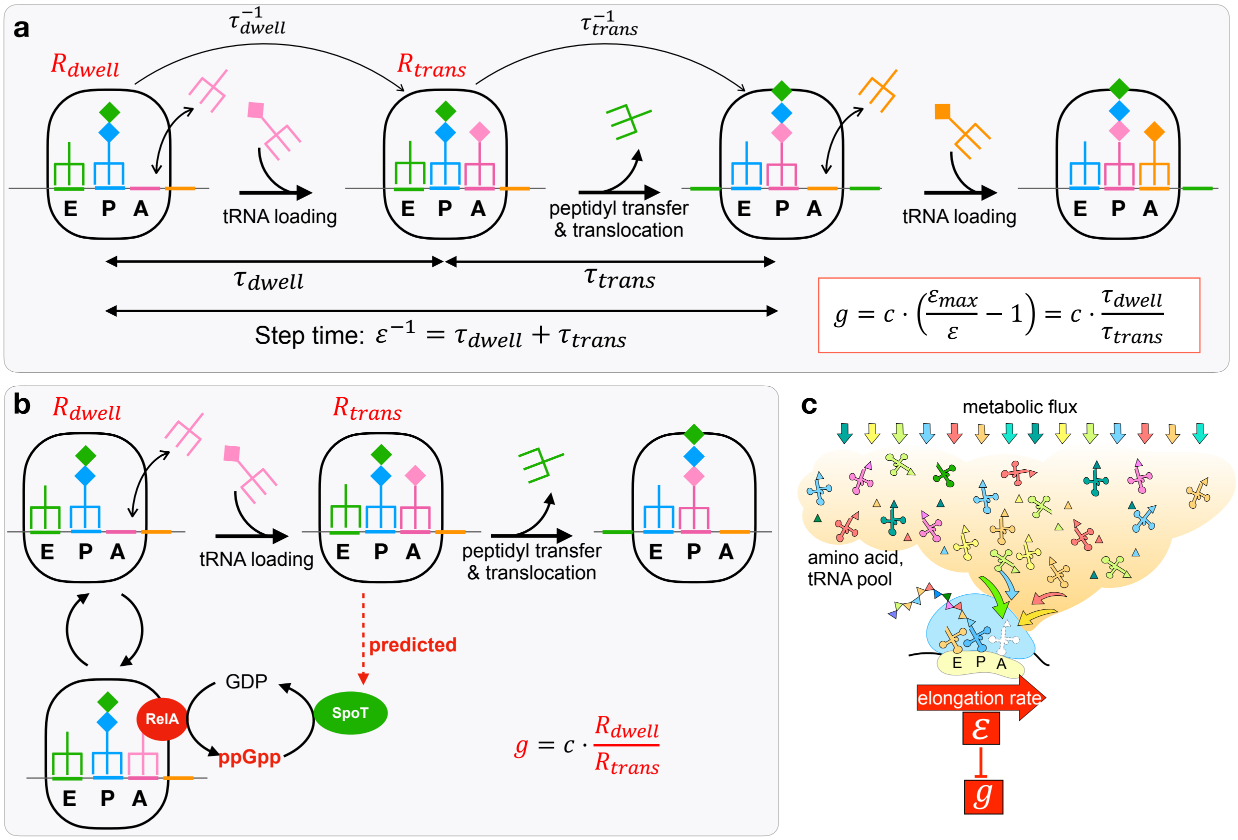

a
τ_{dwell}^{-1}
τ_{trans}^{-1}
R_{dwell}
R_{trans}
E P A
tRNA loading
peptidyl transfer & translocation
tRNA loading
τ_{dwell}
τ_{trans}
Step time: $\varepsilon^{-1} = \tau_{dwell} + \tau_{trans}$
$g = c \cdot \left(\frac{\varepsilon_{max}}{\varepsilon} - 1\right) = c \cdot \frac{\tau_{dwell}}{\tau_{trans}}$
b
R_{dwell}
R_{trans}
tRNA loading
peptidyl transfer & translocation
predicted
GDP
RelA
SpoT
ppGpp
$g = c \cdot \frac{R_{dwell}}{R_{trans}}$
c
metabolic flux
amino acid, tRNA pool
E P A
elongation rate
ε
g

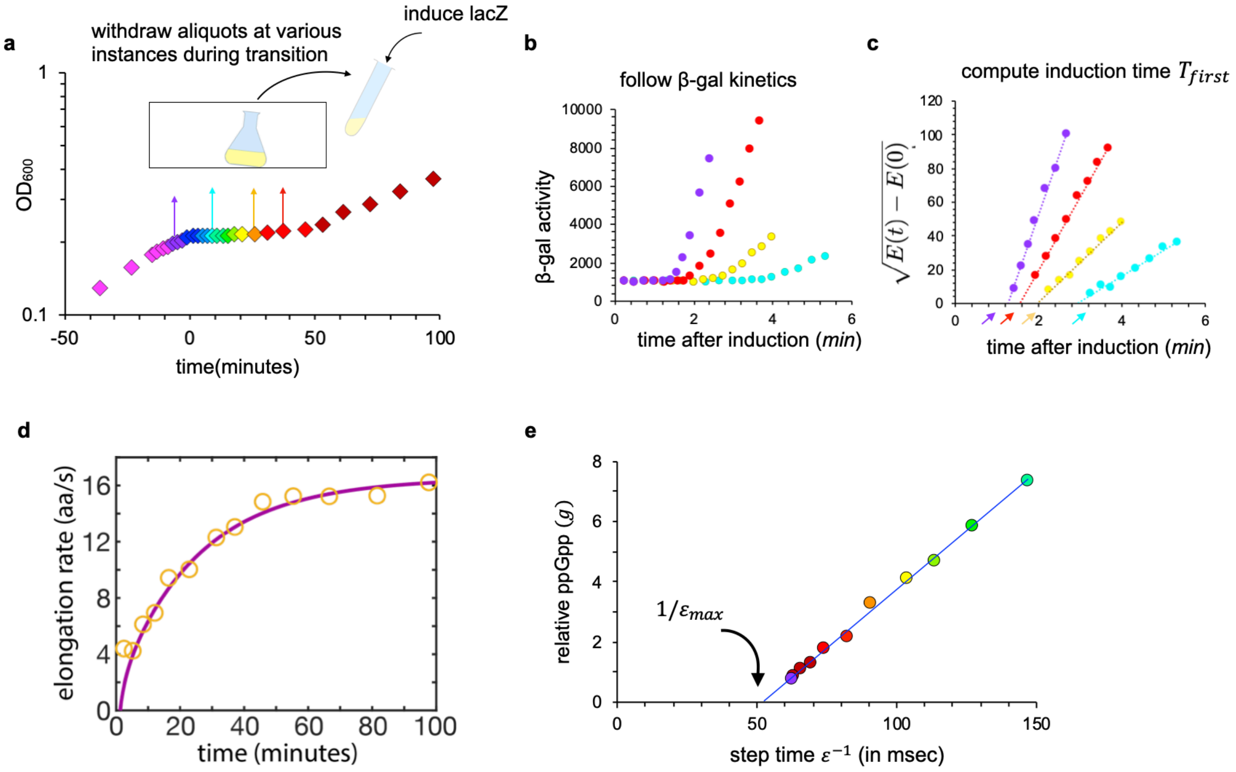
Figure 5.6: Elongation rate measurements during diauxic shift. a, Scheme for measuring the instantaneous translation elongation rates $\varepsilon(t)$ during growth transition. At different time (t) during the diauxic transition, 10 ml aliquots were removed into a fresh tube and the synthesis of the reporter protein LacZ was induced immediately by the addition of IPTG. **b,** LacZ induction curves for the four samples taken (indicated by the arrows in panel **a**). LacZ induction kinetics of each sample was followed by monitoring the β-galactosidase activity using 4-methylumbelliferyl-D-galactopyranoside (MUG) assay (see Methods and Materials for details). **c,** The square root of lacZ activity above basal level were plotted against induction time to obtain the lag time for the synthesis of the first LacZ molecule (T_{first}, shown by the arrows (124)). The translational elongation rates shown in Fig. 5.1 were obtained as $\varepsilon = L_{\mathrm{LacZ}}/(T_{first} - T_{init})$ with L_{LacZ} being the length of LacZ monomer (1024aa) and T_{init} being the time taken for the initiation steps (10s across various nutrient conditions (17) and during transition; see **Fig. 5.7**). **d,** Comparison of the elongation rate obtained from the naïve approach described above with a more detailed calculation using all of the LacZ induction data together. The yellow circle represents the naïve calculation of ER at the various time where samples were taken. The purple line represents $\varepsilon(t)$ calculated from solving the following integral equation

$$\int_{t_i+T_{init}}^{t_i+T_{first}(t_i)} \varepsilon(t)dt = \int_0^{t_i+T_{first}(t_i)} \varepsilon(t)dt - \int_0^{t_i+T_{init}} \varepsilon(t)dt = L_{\mathrm{LacZ}}, \tag{5.8}$$

where t_i is the time when IPTG is added. $T_{first}(t_i)$ denotes the lag time obtained from the LacZ induction curve when IPTG was added at time t_i. T_{init} was taken to be 10s as explained in **Fig. 5.7**. Taking the derivative of t_i on both sides of Eq. (5.8), we obtain

$$\varepsilon\left(t_i + T_{first}(t_i)\right) \times \left(1 + \frac{dT_{first}(t_i)}{dt_i}\right) = \varepsilon(t_i). \tag{5.9}$$

By fitting $T_{first}(t_i)$ to an exponential function $y = a_1 \exp(a_2 x) + a_3$, we estimated $\frac{d}{dt_i}T_{first}(t_i)$ at each time point during the shift. For $t_i > 40$ min, since the change in T_{first} is relatively small (circles in (d) represent the inverse of T_{first}), we Taylor-expanded $\varepsilon\left(t_i + T_{first}(t_i)\right)$ at t_i and solved Eq. (5.9) analytically with the boundary condition $\varepsilon(100\text{ min}) = \varepsilon_{glycerol} = 16aa/s$. For $t_i < 40$ min, we directly used Eq. (5.9) to calculate $\varepsilon(t_i)$ from $\varepsilon\left(t_i + T_{first}(t_i)\right)$ at a later time numerically. Notice that the difference between the two ways of calculating the instantaneous ER is negligible. This is because the time scale of T_{first}, a few minutes as seen in panel (c), is overall much smaller than the time scale of the shift (~40 min). So we simply reported the result of the naive ER calculation through this work. **e,** Relative ppGpp level $g(t)$ obtained in Fig. 5.1c are plotted against the reciprocal of the elongation rate, or the step-time for ribosome advancement. $\varepsilon(t)$ values for the time points at which ppGpp levels were measured were estimated from fitting the naïve calculation of ER vs time in panel (d) to an exponential function. A linear fit of the data in panel (e) yields the x-intercept $1/\varepsilon_{max}$, with $\varepsilon_{max} = 19.4 \pm 1.4\ aa/s$.


a
withdraw aliquots at various instances during transition
induce lacZ
OD600
1
0.1
-50
0
50
100
time(minutes)
b
follow β-gal kinetics
β-gal activity
10000
8000
6000
4000
2000
0
0
2
4
6
time after induction (min)
c
compute induction time T_{first}
$\sqrt{E(t) - E(0)}$
120
100
80
60
40
20
0
0
2
4
6
time after induction (min)
d
elongation rate (aa/s)
16
12
8
4
0
0
20
40
60
80
100
time (minutes)
e
relative ppGpp (g)
8
6
4
2
0
$1/\varepsilon_{max}$
0
50
100
150
step time ε^{-1} (in msec)

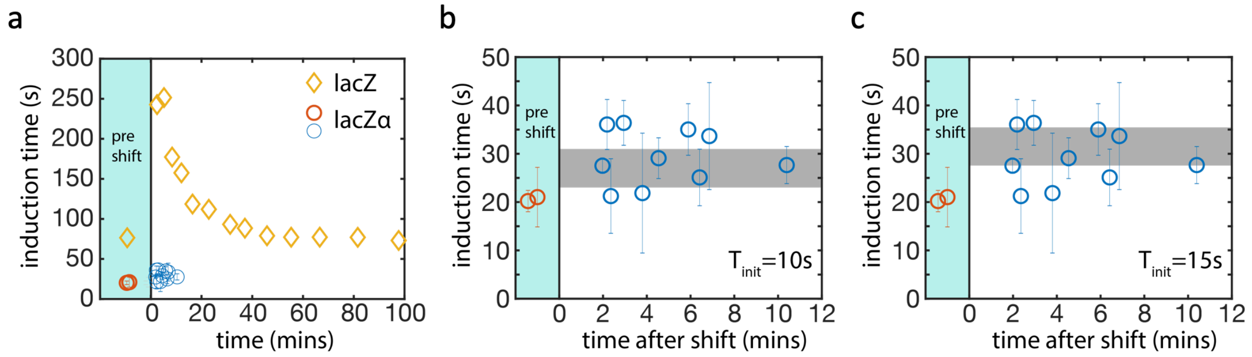


Figure 5.7. Determination of the translational initiation time. The induction time of lacZ monomer (T_{first} shown in **Fig. 5.7c**) during the shift were plotted as yellow diamonds in panel (**a**). To obtain the contributions of translational initiation towards this induction time, we also determined the induction kinetics of a short LacZ alpha fragment ($LacZ\alpha$, the N-terminal 1-90aa of lacZ) before and right after the shift. The induction time obtained, $T_{first,LacZ\alpha}$, were plotted as red and blue circles respectively. The error bars indicated the 95% confidence interval from fitting and technical repeats. The induction time obtained can in principle be used to estimate the translational initiation time, $T_{init} = T_{first,LacZ\alpha} - L_{LacZ\alpha}/\tilde{\varepsilon}$, where as a first estimate of the elongation rate, we used $\tilde{\varepsilon} = (L_{lacZ} - L_{lacZ\alpha})/(T_{first,LacZ} - T_{first,LacZ\alpha})$ assuming the same initiation time for LacZ and LacZ alpha, with $T_{first,LacZ\alpha}$ used for T_{init} of LacZα. This was done for the pre-shift condition, with the result $T_{init} = 15s \pm 6s$ (error from uncertainty in $LacZ\alpha$ induction time), consistent with previous measurements (17). To estimate the initiation time immediately following the shift, we looked more closely into the induction time data collected, shown as the blue circles in panels (**b**) and (**c**). Due to the uncertainties in $LacZ\alpha$ induction time determination, together with the uncertainty in determining the precise shift time (i.e., time "0") between the strains expressing LacZ and LacZα, it is difficult to calculate the initiation time based on the difference $T_{first,LacZ} - T_{first,LacZ\alpha}$ while $T_{first,LacZ}$ varied by strongly in the first 10 min. Instead, we assumed a constant initiation time (T_{init}) and predicted the range of $LacZ\alpha$ induction time given measured $LacZ$ induction time in this 10min-window (160s-250s). The grey bar in **b** and **c** shows the predicted range for $T_{first,LacZ\alpha}$ assuming $T_{init} = 10s$ and $T_{init} = 15s$, respectively. Comparing with our measurements, it suggest the initiation time in transition is within the range of 10s-15s, similar to the initiation time in pre-shift condition. For simplicity, we used $T_{init} = 10s$ for both pre- and post-shift throughout this work. This does not mean that the initiation time is not affected during the transition, but that our method is too coarse to resolve differences in the initiation time.

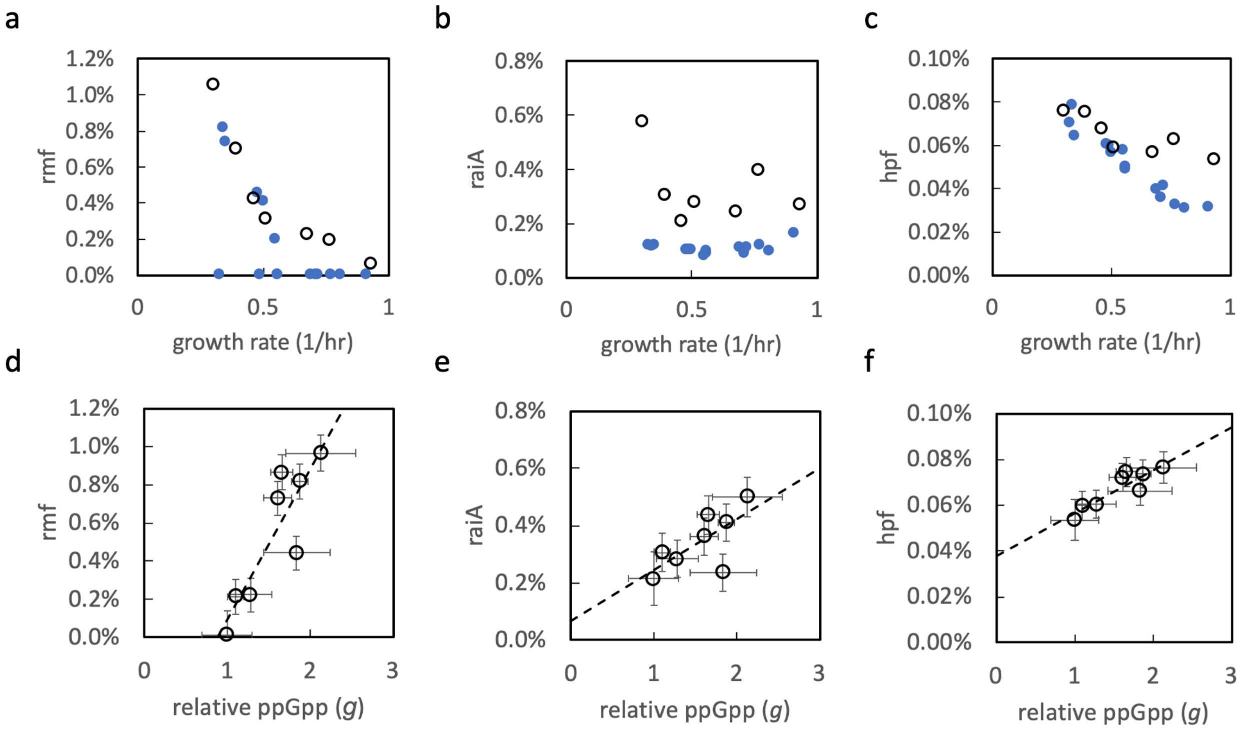


Figure 5.8: Relation between the ribosome-sequestering proteins and the ppGpp levels. a, The number of mRNA (black circles) and protein (blue circles) as a percentage of the total number of mRNA and protein, respectively, for the ribosome remodeling factor Rmf obtained during steady state growth are plotted against the growth rate for a number of cultures under varying degrees of carbon limitation. Data for mRNA abundance is obtained from Balakrishnan et al (129) and for protein abundance is obtained from Mori et al (2). **b,** Same as panel a, but for the mRNA and protein of the ribosome associated inhibitor gene *raiA* (130). **c,** Same as panel a, but for the mRNA and protein of the gene encoding the hibernation promoting factor *hpf* (93). For each case shown in panels a-c, the protein and mRNA abundances approximately match each other within 2-fold, as is typical of many genes expressed during exponential growth. For Rmf, the protein level showed zeero for a number of conditions due to lack of detected peptides. **d-f,** The fractional number abundances of the mRNA (black circles) of *rmf*, *raiA* and *hpf* are plotted against the relative ppGpp of the corresponding conditions. The black dash lines show the linear fit of the data. To directly compare the expression of *rmf, raiA* and *hpf* with ppGpp levels, we fitted the mRNA vs GR data in panel a-c with 3rd-degree polynomials and estimated their mRNA levels at the same growth rates for which ppGpp was measured. The horizonal error bars in panels d-f comes from ppGpp measurements, while the vertical error bars come from the polynomial fitting (95% prediction interval).

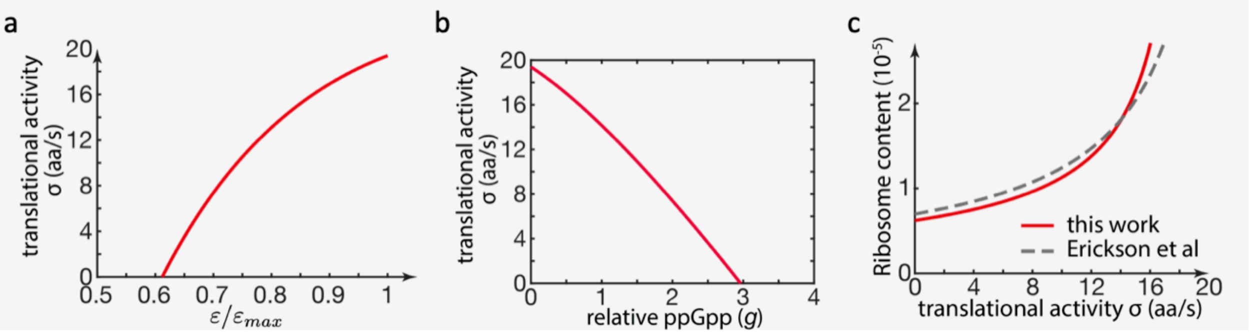


Figure 5.9. a, Translational activity σ, defined as λ/R, and thus given by $\varepsilon \cdot (R-H)/R$, is plotted against normalized elongation rate for the best-fit model parameters used in **Fig. 5.5**. **b,** Translational activity σ plotted against the relative ppGpp level g. The one-to-one relationship between the translational activity and the ppGpp level justifies the use of the translational activity as the dynamic variable in the kinetic model of diauxic transition developed by Erickson et al (8). **c,** The plot of ribosomal content vs σ used by Erickson et al (8) as the regulatory function for ribosome biogenesis is shown as the grey dashed line. The same relation according to the model in **Fig. 5.5** is shown as the solid red line. The two lines are very similar, indicating that Erickson et al (8) correctly inferred the regulatory function even though the relation between σ and g (panel **c**) was not known at the time.

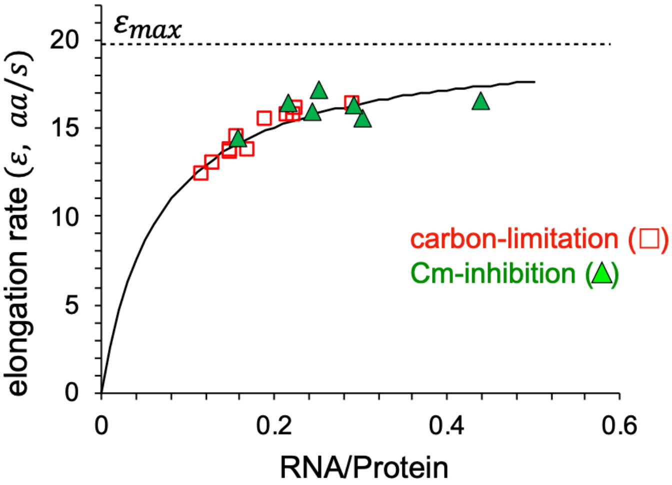
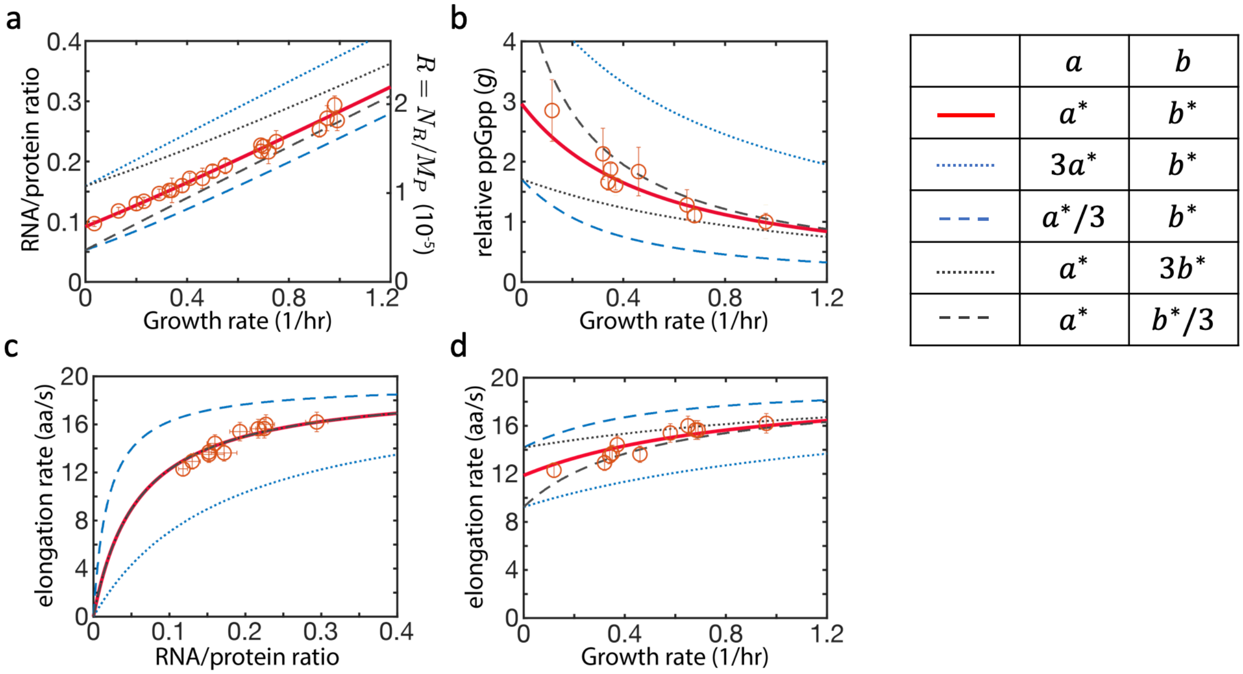
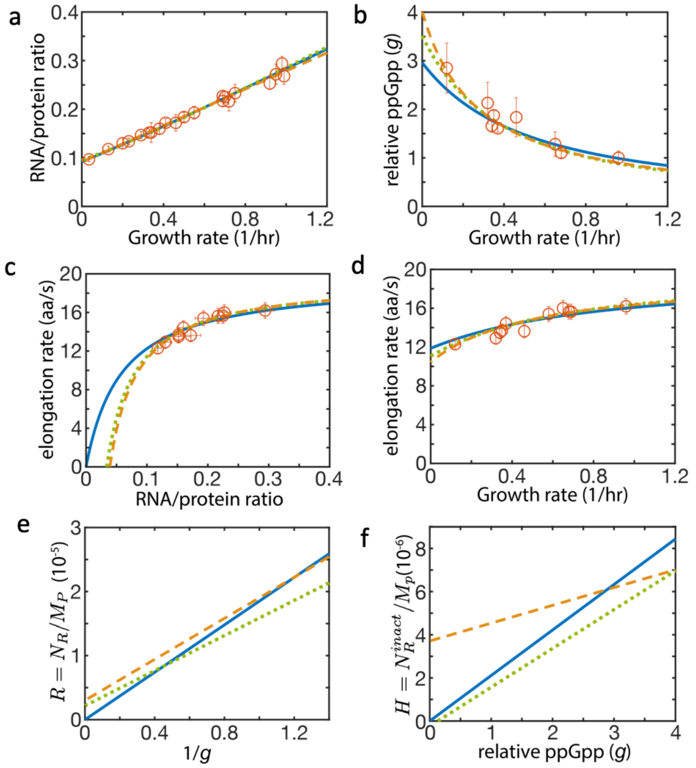


Figure 5.10: Estimating the maximum translation elongation rate in steady-state growth. Translation elongation rates measured in the carbon-limited and chloramphenicol-treated cultures are plotted against the ratio of RNA and protein measured in the same conditions. The data are collectively fitted to the Michaelis-Menten relation $\varepsilon = \varepsilon_{max} \cdot r/(K_m + r)$ where r is the RNA-protein ratio. The best-fit gives $\varepsilon_{max} = 20.0 \pm 1.9 aa/s$ and $k_m = 0.066$. The maximum elongation rate ε_{max} estimated this way is indistinguishable from that estimated from the relation between ppGpp and the elongation rate during growth transition (see **Fig. 5.6e**).



	a	b
—	a^*	b^*
⋯	$3a^*$	b^*
- - -	$a^*/3$	b^*
⋯	a^*	$3b^*$
- - -	a^*	$b^*/3$

Figure 5.11: The effect of parameter choices on model output. The measured data points (red open circles) and the best-fit model outputs (solid red lines) in panels a-d are the same as those shown in **Fig. 5.5b-e**. When keeping parameter a the same as the best-fit value a^* while varying b to be 3-fold larger or smaller, the model gave predictions shown as the dotted and dash black lines, respectively. When keeping parameter b the same as the best-fit value b^* and varying a to be 3-fold larger or smaller, the model gave predictions shown as the dotted and dash blue lines, respectively; See the legends table. The results show that model outputs are comparatively more sensitive to the choice of the parameter a than b. Fortunately, the parameter a involved in the relation between ppGpp and the ribosome content (see Eq. (5.3)) is well established by **Fig. 5.5a**.



	$g_H = 0$, $g_R = \infty$, $\frac{rRNA}{RNA}$ fitted	g_H fitted g_R fitted $\frac{rRNA}{RNA}$=0.86	g_H fitted g_R fitted $\frac{rRNA}{RNA}$ fitted
	———	·········	- - - -
a	1.85×10^{-5}	1.37×10^{-5}	1.6×10^{-5}
b	2.11×10^{-6}	1.8×10^{-6}	8.25×10^{-7}
g_R		6.3	5.3
g_H		−0.13	4.5
$\frac{rRNA}{RNA}$	0.92		1.03

$$R(g) = a \cdot (1/g + 1/g_R) \quad \text{---} \quad (5.10)$$

$$H(g) = b \cdot (g + g_H) \quad \text{---} \quad (5.11)$$

Figure 5.12: The effect of constant offsets in the regulatory functions on model output. Here we investigate the effect of altered forms of regulatory functions from the simplest forms assumed in Eqs. (5.3) and (5.4). The altered forms are shown as Eqs. (5.10) and (5.11) below the legend table on the right, with constant offsets parameterized by g_R and g_H. **a-d** show the same four outputs as in **Fig. 5.11** with different modeling results. The table on the right compares the best-fit parameter values under different model settings. The solid blue line is the same as that shown in **Fig. 5.5** (where offsets in $R(g)$ or $H(g)$ are not considered, left column of the legend). The orange dash lines are the best-fit of the model with g_R and g_H treated as fitting parameters also. Results of the fits are listed in the right column of the legend table. While the fits may be marginally better, one aspect of the fitting parameter becomes quite unreasonable: the relation between the ribosome content (N_R/M_P) and the RNA-protein ratio is determined by the fraction of ribosome RNA within all RNA (the last row in legend table). In the case where all parameters are fitted (right column), the rRNA-RNA ratio is ~1, which is impossible given that tRNA comprises ~10% of the total RNA. In comparison, the simplest model not considering offsets gave an rRNA-RNA ratio of 92% (left column) which is closer to the accepted range. If we fix the rRNA-RNA ratio to 86%, estimated previously for a few growth conditions (4, 65) , the best-fit model outputs are shown as the green dotted lines, with the best-fit parameters shown in the middle column. For the ease of assessing the effect of these offsets on the form of the regulatory functions $R(g)$ and $H(g)$, we display these forms for the 3 cases considered in panel **e** and **f**, using the same line styles as those indicated in the legend table. We see that little difference is made to the model outputs despite substantial changes to the forms of the regulatory functions, thus indicating that these results are robust to the exact forms of the regulatory functions assumed.

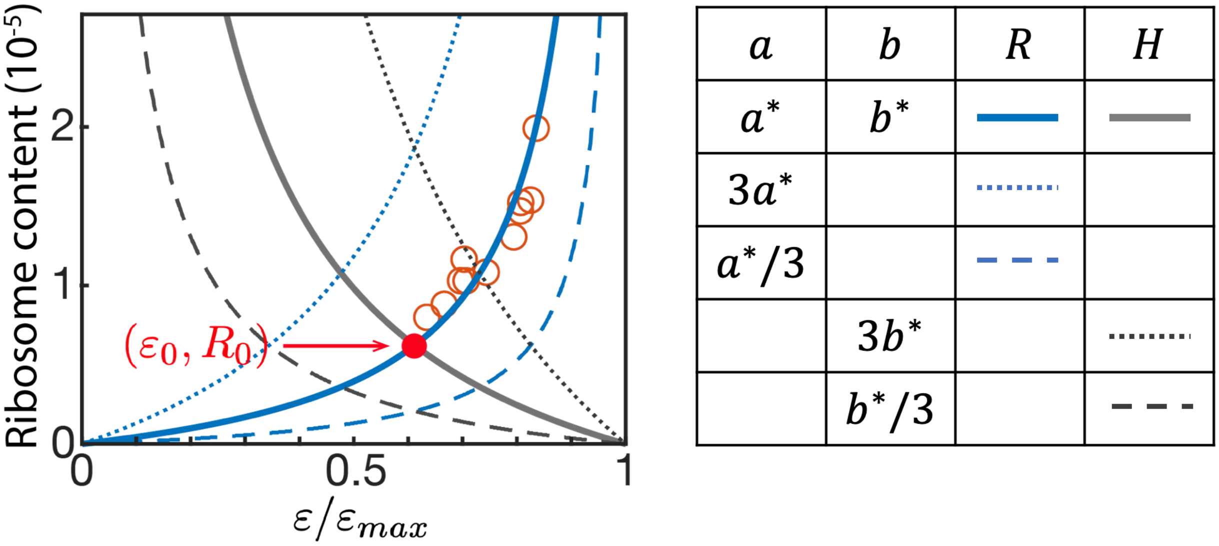


Figure 5.13. Ribosome content are plotted against normalized elongation rate for various model parameters. Using the best-fit parameters in **Fig. 5.5** (labeled as a* and b* in legend table), total ribosome content (solid blue line) increases with elongation rate while inactive ribosome content (solid grey line) decrease with elongation rate. Their intersection (marked by the filled red circle) represents the elongation rate and ribosome content while growth rate approaches zero, i.e., ε_{min} and R_0, respectively. When the parameter a is changed to 3-fold smaller or larger, the relation between total ribosome content and elongation rate is changed to the blue dashed or dotted lines, respectively. When the parameter b is changed 3-fold smaller or larger, the relation between total ribosome content and elongation rate is changed to the black dashed and dotted lines, respectively. As shown in the figure, the different parameters result in different values of the intersection point and hence different values of (ε_0, R_0).

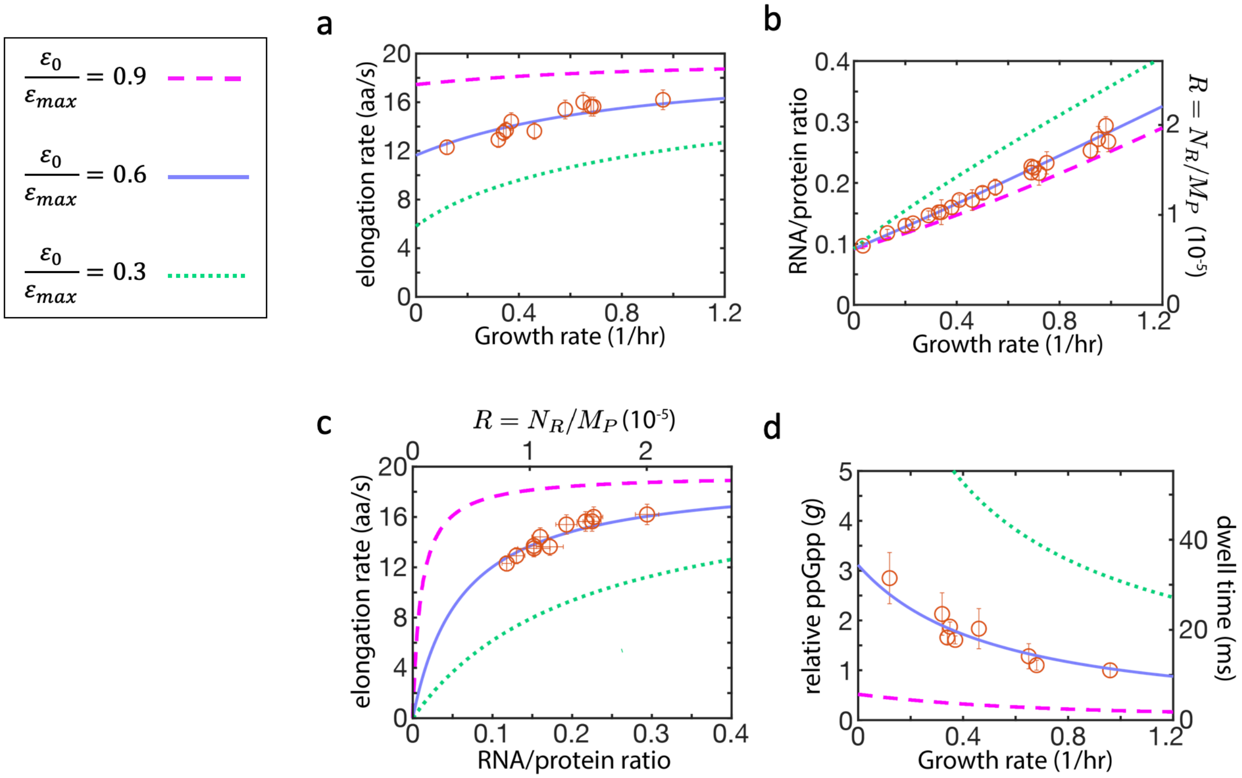


Figure 5.14: The effect of the value of ε_0 on model output. As shown in Supplementary Note, the model defined by Eqs. (5.1) and (5.5) has a particularly simple solution for $\varepsilon_0 = \varepsilon_{max}/2$, where ε_0 is the elongation rate as the growth rate $\lambda \to 0$. The best-fit output of the model has $\varepsilon_0 : \varepsilon_{max} \approx 0.6$, which is close to the simple linear solution, thus rationalizing the approximate linear correlations observed in **Fig. 5.5**. In this figure, we show the model output for choices of the parameters a and b such that $\varepsilon_0 : \varepsilon_{max} = 0.3$ and 0.9 (dotted and dashed lines, respectively). The red open circles are the same data as those shown in **Fig. 5.5**. Deviations from the linear growth law are clearly seen for the expected RNA/protein ratio in panel b, with the high ratio of $\varepsilon_0 : \varepsilon_{max}$ (dashed purple lines) exhibiting reduced ribosome content. In principle, large ER and reduced RNA/protein ratio are advantageous. However, this would lead to very low dwell time for ribosome on A-site as seen in panel d (right vertical axis): 10-30 ms for the best-fit parameters, and < 5ms for the purple dashed line with ER maintained above 90% of ε_{max}.

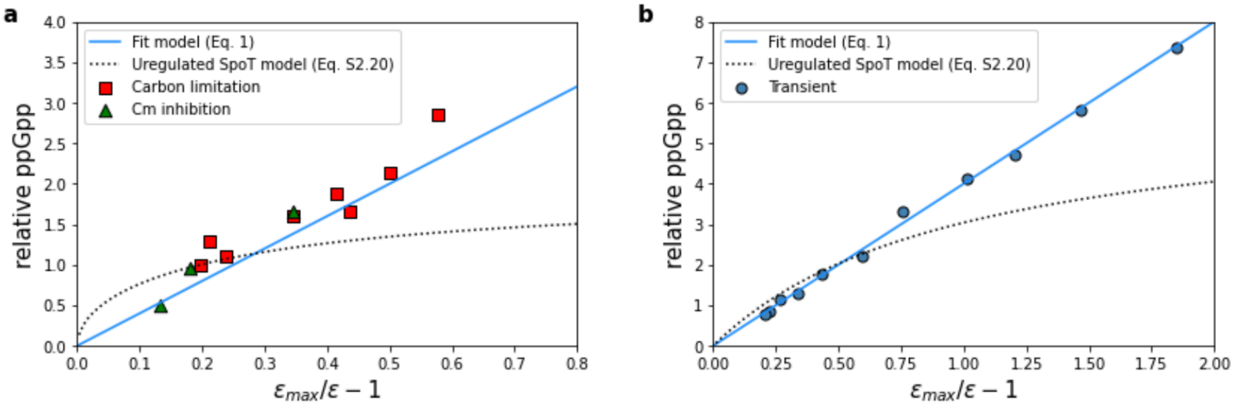


Figure 5.15: Comparing predictions for unregulated hydrolysis to the observed relation between elongation and ppGpp. a, Observed ppGpp levels for steady state growth under carbon limitation and Cm inhibition (Fig. 5.2d) are compared to the predicted ppGpp levels for unregulated SpoT hydrolysis derived in Eq. (5.40). Using Eq. (5.40), $\tau_{dwell}/(\tau_{dwell} + \tau_{trans})$ is computed from the elongation rate as $1 - \varepsilon/\varepsilon_{max}$. The ppGpp level, g, is then solved assuming R_{act} is the difference between $R(g)$, main text Eq. (5.3), and H(g), main text Eq. (5.4). The proportionality constant for Eq. (5.40) has been set so that the ppGpp level in the glucose reference condition is 1. The fit phenomenological model, main text Eq. (5.1), is shown for comparison with $\varepsilon_{max} \approx 19.4 aa/s$ and $c \approx 4.0$. **b**, Observed ppGpp levels during the growth transition (Fig. 5.1e) are compared to predictions for unregulated SpoT hydrolysis computed from Eq. (5.40). Here R_{act} is assumed to remain at pre-shift levels, and the same proportionality constant is used from panel (a). The fit phenomenological model, main text Eq. (5.1), is again shown for comparison.

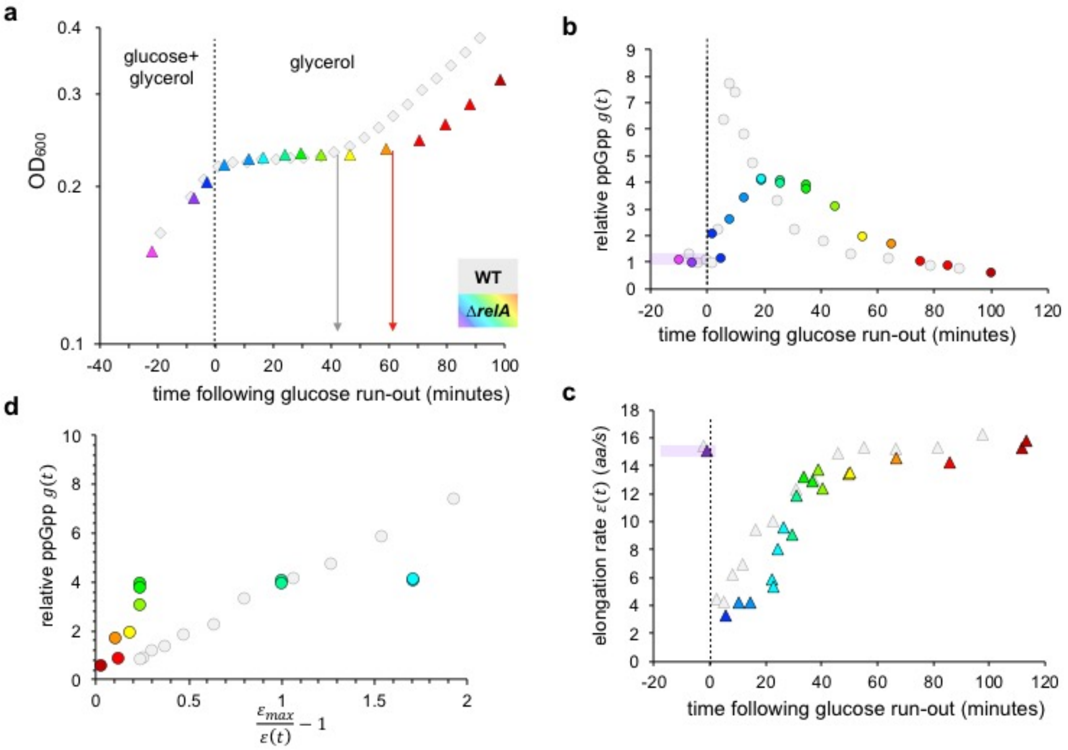


Figure 5.16: Relation between the ppGpp level and elongation rate for the Δ*relA* strain during diauxic shift. **a,** Growth kinetics of the Δ*relA* strain (colored triangles) and the wild type (grey circles, same as in **Fig. 5.1a**) during the glucose-to-glycerol diauxic transition. Vertical arrows indicate the duration of the lag period following glucose depletion before growth resumes on glycerol. Color-scheme corresponding to individual samples for Δ*relA* is shared across panels in this figure. **b,** ppGpp levels measured for the Δ*relA* strain (colored) and the wild type (grey, same as those in **Fig. 5.1c**), relative to the WT ppGpp level in glucose minimal medium, during the glucose-to-glycerol transition. Note that pre-shift ppGpp levels of the two strains are indistinguishable. **c,** Instantaneous translation elongation rate measured for the Δ*relA* strain (colored) and the wild type (grey, same as those in **Fig. 5.1d**) during the glucose-to-glycerol transition. **d,** The relation between ppGpp level and the instantaneous elongation rate for the Δ*relA* strain (colored) and the wild type (grey, same as those in **Fig. 5.1e**) during the glucose-to-glycerol transition.

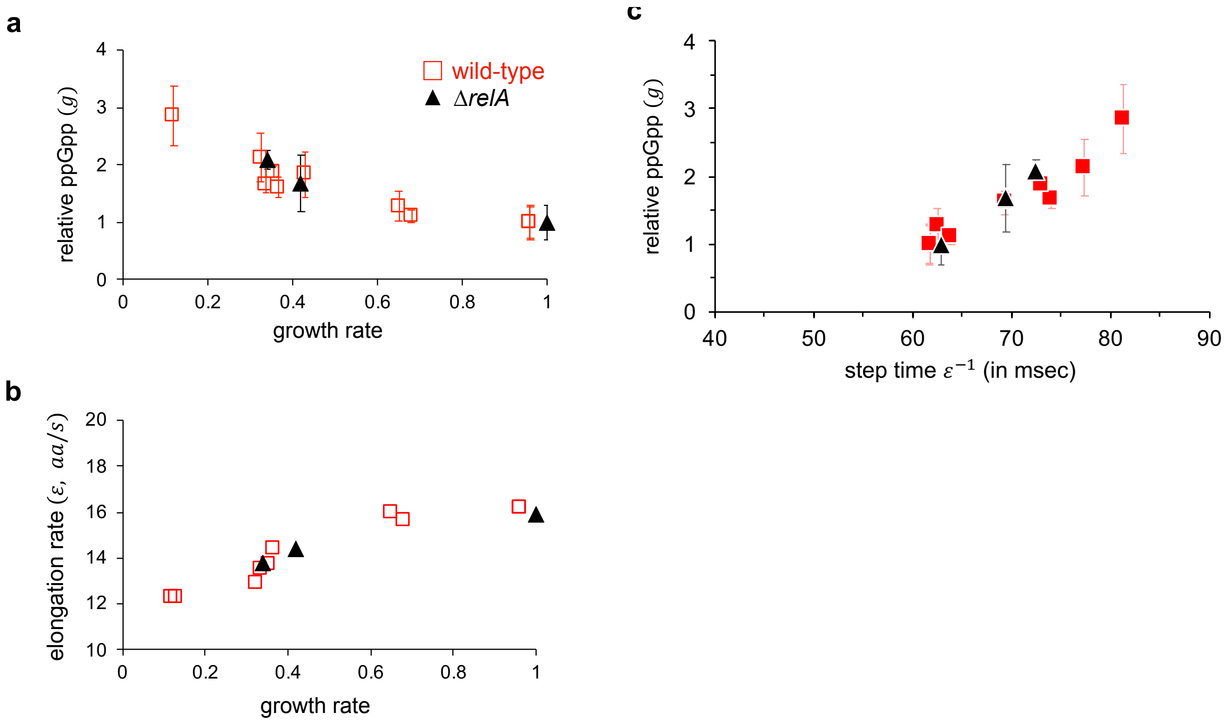


Figure 5.17: Relation between the ppGpp level and elongation rate for the Δ*relA* strain in steady-state growth. a, ppGpp levels relative to that in the reference growth condition (wild type grown in minimal glucose medium) are plotted against the steady state growth rates for wild-type (red) and Δ*relA* (black) grown in various carbon sources (See **Table 5.1**). Error bar represents the uncertainty in the linear fit over four measurements; see **Fig. 5.18**. **b,** Translation elongation rates are plotted against the steady state growth rates for wild-type (red) and Δ*relA* (black) grown in various carbon sources (See **Table 5.1**). **c,** Steady-state relative ppGpp levels are plotted against the reciprocal of the translational elongation rate for wild-type (red) and Δ*relA* (black) grown in the same conditions. Both the ppGpp levels and the elongation rates in steady state growth are not affected by *relA* deletion, unlike the case during diauxic transition shown in **Fig. 5.16**.

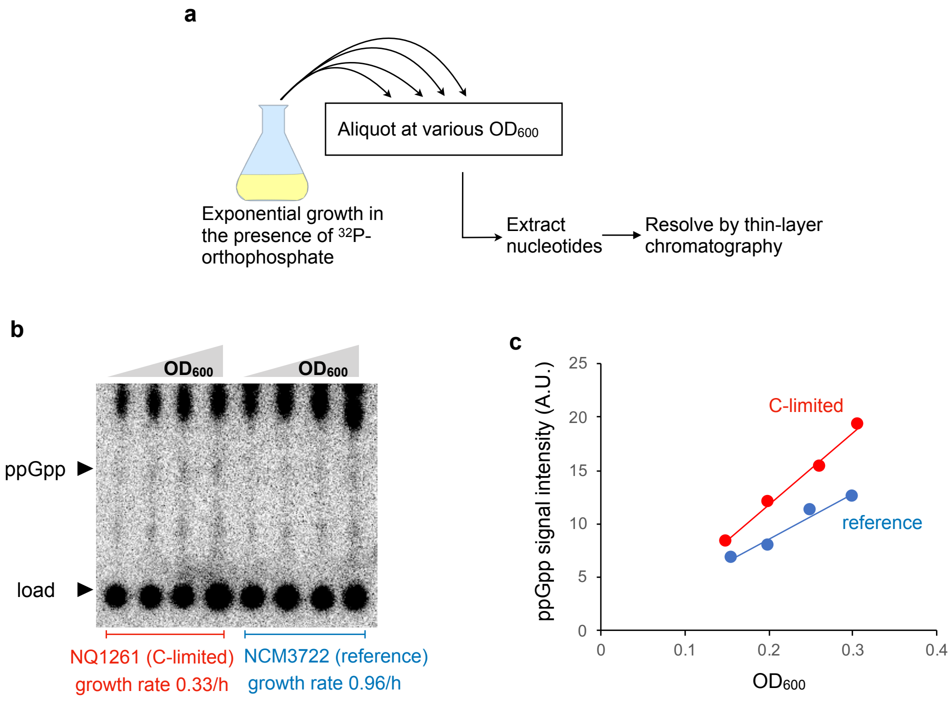


Figure 5.18: Measurement of ppGpp levels in steady-state growth conditions.
a, Scheme for ppGpp measurement under balanced growth. *E. coli* strains were grown in either glucose minimal media (reference condition, growth rate =0.96/h) or various poor carbons at different steady-state growth rates. At various OD_{600} values during the exponential growth phase of each culture, aliquots were withdrawn, nucleotides were extracted and ppGpp was resolved using TLC. **b,** An example of ppGpp measurement for the wild type strain (reference condition, blue) and NQ1261, a Δ*ptsG* strain defective in glucose uptake, both grown in glucose minimal medium (red). NQ1261 strain is used to create reduced growth on glucose (see Methods). Nucleotides extracted at four different OD_{600} were spotted on the bottom of TLC plates and ppGpp was resolved. **c,** ppGpp from the wild type and NQ1261 strains are plotted against OD_{600}. The slope of each plot gives the ppGpp level (per OD_{600}) for that strain and condition. In **Fig. 5.3** of the main text, we report the ppGpp level relative to that in glucose steady state. In this case, the relative ppGpp level of NQ1261 strain grown in glucose is just the ratio of the slope of the red line to that of the blue line. Error bar in the estimate of ppGpp was taken to be the uncertainty in the slopes of the linear fit.

5.8 Tables

Table 5.1: Strains and growth conditions used in this study.

strain name	description	growth condition	growth rate (h^{-1})
NCM3722 (29, 30)	Wild type *E. coli* K-12	MOPS+ 20mM glucose	0.96
		MOPS+20mM succinate	0.68
		MOPS+40mM glycerol	0.65
		MOPS+30mM acetate	0.43
		MOPS+3mM mannose	0.33
		MOPS+20mM aspartate + 10mM NH_4Cl	0.35
		MOPS 0.4% glycerol+10mM arginine	0.32
		MOPS+20mM glutamate + 10mM NH_4Cl	0.13
NQ1261 (17)	*ΔptsG*	MOPS+20mM glucose	0.36
		MOPS+20mM glucose + 2μM Cm	0.16
		MOPS+20mM glucose + 3μM Cm	0.12
HE838 (this work)	*ΔrelA*	MOPS 20mM glucose	0.99
		MOPS+4mM mannose	0.42
		MOPS+20 mM aspartate+10mM NH_4Cl	0.34
NQ1468 (17)	Inducible $LacZ\alpha$ with constitutively expressed $LacZ\omega$	MOPS 20mM glucose	0.90
NCM 3722 and HE838 under diauxic transition		MOPS+2mM glucose+40mM glycerol	N.A.

Table 5.2: Primers used in this study.

name	sequence	use
relA-P1	cgatttcggcaggtctggtccctaaaggagaggacgatggttgcggtaagatgtgtaggctggagctgcttc	Chromosomal deletion of *relA* gene
relA-P2	atatcaatctacattgtagatacgagcaaatttcggcctaactcccgtgcaattccggggatccgtcgacctg	Chromosomal deletion of *relA* gene
relA-ver-R	tacgctactgtggatcataaccctttc	Verification of deletion of *relA* gene

5.9 Appendix

5.9.1 Dependence of the steady-state model on ε_0 and R_0

In the main text, Eq. (5.5) described the relationship[1] between growth rate λ and ppGpp level (or elongation rate via Eq. (5.1)) using parameter a and b. Here we show that a and b can be expressed by ε_0 and R_0, defined as the elongation rate (ER) and ribosomal abundance, respectively, when the growth rate approaches zero. Also, we show that when $\varepsilon_0 : \varepsilon_{max} = 1:2$, the model yields simple dependences of R and g on λ.

In order for ER to be nonzero (with value ε_0) when the growth rate approaches zero, we must have $R = H$ (with value R_0) according to Eq. (5.5). Using Eqs. (5.1) and (5.3), we can express a in terms of ε_{min} and R_0:

$$a = c \cdot R_0 \cdot (h - 1), \tag{5.12}$$

where we introduced $h \equiv \varepsilon_{max}/\varepsilon_0$ for convenience. Similarly, from Eqs. (5.1) and (5.4), the parameter b can be expressed as

$$b = \frac{R_0}{c \cdot (h - 1)}. \tag{5.13}$$

Conversely, we can express the two physiological parameters R_0 and h in terms of the model parameters a and b, with $R_0 = \sqrt{a \cdot b}$ and $h = 1 + c\sqrt{a/b}$.

Substitute Eqs. (5.3), (5.12) and (5.13) into Eq. (5.5), we obtain a relationship between λ and R with h and R_0 being the parameters:

$$\lambda = \varepsilon_{max} \frac{R^2 - R_0^2}{R + R_0 \cdot (h - 1)}. \tag{5.14}$$

---

[1] In writing down the relationship in Eq. (5.2) and Eq. (5.5), we assumed proteins are stable. At very slow growth, the effect of protein degradation should in principle be included. However, based on the measured protein degradation rates (91, 133) this effect is negligible even at the smallest growth rate we studied here.

Notice that when $h = 2$ (or $\varepsilon_0 : \varepsilon_{max} = 1:2$), Eq. (5.14) is reduced to $\lambda = \varepsilon_{max} \cdot (R - R_0)$, or

$$R = R_0 + \lambda/\varepsilon_{max}, \tag{5.15}$$

i.e., an exact linear relationship between R and λ with $1/\varepsilon_{max}$ being the slope of R-λ plot.

We can similarly work out the relation between g and λ. From Eq. (5.12) above, we can rewrite Eq. (5.3) of the main text as $R = R_0 c \cdot (h-1)/g$. Defining the value of g as the growth rate $\lambda \to 0$ as g_0, we have $c \cdot (h-1) = g_0$, or

$$R = R_0 \cdot g_0/g. \tag{5.16}$$

Substituting Eq. (5.16) into Eq. (5.14), we obtain

$$\lambda = \varepsilon_{max} R_0 \frac{(g_0/g)^2 - 1}{g_0/g + (h-1)}. \tag{5.17}$$

For $h = 2$ (or $\varepsilon_0 : \varepsilon_{max} = 1:2$), Eq. (5.17) is reduced to $\lambda = \varepsilon_{max} R_0 \cdot (g_0/g - 1)$, or

$$g = \frac{g_0}{1 + \lambda/(\varepsilon_{max} R_0)}, \tag{5.18}$$

i.e., a simple hyperbolic dependence of the ppGpp level on the growth rate λ. Alternatively, g^{-1} has a simple linear dependence on λ.

Lastly, we examine the relation between R and ε. Inverting Eq. (5.1) of the main text, we have

$$\varepsilon = \frac{\varepsilon_{max}}{1 + g/c}. \tag{5.19}$$

Further using Eq. (5.16) and $g_0 = c \cdot (h-1)$, we obtain

$$\varepsilon = \frac{\varepsilon_{max}}{1 + (h-1)R_0/R}, \tag{5.20}$$

which is generally of the Michaelis form, ε_{max} being the maximal ER. For $h = 2$, the "Michaelis constant" becomes simply R_0.

5.9.2 Molecular-level models linking Elongation Rate and ppGpp

The observed relation between ppGpp levels and the elongation rate (Eq. 5.1) provides a new constraint on the molecular mechanisms underlying ppGpp synthesis and degradation. Given the mechanistic view of translation detailed in Figure 5.5 of the main text, the observed relation in Eq. 5.1 can be expressed as $g = c\tau_{dwell}/\tau_{trans}$, where τ_{trans} is the translocation time of charged ribosomes and τ_{dwell} is the dwell time of ribosomes waiting for charged tRNA. However, it remains to be established how the cell can sense these two timescales biochemically, and how the timescales in turn reflect the various tRNA charging levels that ultimately determine the elongation rate. In this note, we show that the cell can feasibly sense the ratio τ_{dwell}:τ_{trans} through the ratio of ribosomal states, R_{dwell}: R_{trans} where R_{dwell} is the concentration of ribosomes in the dwelling state and R_{trans} is the concentration in the translocation state. We show that this relation τ_{dwell}:$\tau_{trans} = R_{dwell}$: R_{trans} depends only on an assumption of flux balance for each tRNA species during the translation process. We then show how the ratio of ribosomal states is proportional to a weighted average of tRNA charging ratios, such that the cell can sense the imitation of any charged tRNA species via sensing the ratio R_{dwell}: R_{trans}. Finally, we show how the known mechanism of ppGpp synthesis via RelA cannot alone recapitulate the observed relation between ppGpp and the elongation. We propose that regulation of SpoT, either directly or indirectly, by translocating ribosomes is needed to reproduce the observed relation. The proposed regulation of SpoT also provides a feasible mechanism for the cell to sense RR_{dwell}: R_{trans}.

<u>*Elongation timescales and translocating/dwelling ribosomes*</u>

In order to produce the empirical relation between $\frac{\varepsilon_{max}}{\varepsilon} - 1$ and the ppGpp level, g, the cell must be able to sense the τ_{trans} and τ_{dwell} timescales detailed in Figure 5.5. In correspondence with these two timescales, we assume actively translating ribosomes can be in one of two states: translocating, with concentration R_{trans}, and dwelling, with concentration R_{dwell}, such that $R_{act} = R_{dwell} + R_{trans}$. From global flux balance, it follows that the growth rate, λ, must balance both 1) average elongation rate, ε, multiplied by the active ribosomes, R_{act}, as well as 2) the translocation rate, ε_{max}, multiplied by the translocating ribosomes, R_{trans}, such that

$$\lambda = \varepsilon \cdot R_{act} = \varepsilon_{max} R_{trans}. \tag{5.21}$$

Rearranging Eq. (5.21) we see that the ratio of the elongation rate to the maximum rate is equal to the fraction of active ribosomes that are translocating

$$\frac{\varepsilon}{\varepsilon_{max}} = \frac{R_{trans}}{R_{act}}. \tag{5.22}$$

As discussed in Fig. 5.5, the elongation rate decomposes into two timescales such that, $\varepsilon^{-1} = \tau_{trans} + \tau_{dwell}$, and the maximum elongation rate can be interpreted as and ${\varepsilon_{max}}^{-1} = \tau_{trans}$. By substituting these expressions, along with $R_{act} = R_{dwell} + R_{trans}$, into Eq. (5.22), we have

$$\frac{\tau_{trans}}{\tau_{trans} + \tau_{dwell}} = \frac{R_{trans}}{R_{dwell} + R_{trans}}, \tag{5.23}$$

Simplifying this expression then yields Eq. (5.6) in the main text; $R_{dwell} \cdot \tau_{dwell}^{-1} = R_{trans} \cdot \tau_{trans}^{-1}$. As mentioned, Eq. (65.) must hold to ensure detailed balance of fluxes between the two ribosomal states.

Eq. (5.22) can also be re-arranged to solve for elongation rate in terms of R_{trans} and R_{dwell}, such that

$$\varepsilon = \frac{1}{\tau_{trans}} \frac{R_{trans}}{R_{dwell} + R_{trans}}. \tag{5.24}$$

Using Eq. (5.24) we can then see that the empirical quantity of interest, $\frac{\varepsilon_{max}}{\varepsilon} - 1$, can be expressed in terms of translocating and dwelling ribosomes as

$$\frac{\varepsilon_{max}}{\varepsilon} - 1 = \frac{\tau_{dwell}}{\tau_{trans}} = \frac{R_{dwell}}{R_{trans}}. \tag{5.25}$$

Thus, $\frac{\varepsilon_{max}}{\varepsilon} - 1$ -- which corresponds to the ratio of the dwell and translocation times -- can be obtained by the cell simply as the ratio of the dwelling ribosomes to the translocating ribosomes. Note that this relation is valid *instantaneously*. In particular, Eq. (5.25) depends only on the ratio of ribosomes in different states, not on the overall abundance of the translating ribosomes.

tRNA charging and elongation rate

Past models (131, 132) have derived the dependence of the translation elongation rate on the individual tRNA abundances and charging levels. Here we use these models to connect the elongation timescales, τ_{trans} and τ_{dwell}, to the individual tRNA charging levels. In the past models, each active (i.e., translating) ribosome is assumed to be located at a specific codon associated with a given tRNA species, i. At a given moment, the active ribosomes are partitioned into *i* subspecies, with $N_{R,i}^{act}$ representing the number of ribosomes whose A-site resides at a codon associated with the *i*th tRNA species. The total number of active ribosomes is $N_R^{act} = \sum_i N_{R,i}^{act}$. The corresponding concentrations of active ribosomes are $R_{act,i} \equiv N_{R,i}^{act}/M_P$ and $R_{act} \equiv N_R^{act}/M_P$. The concentration of each tRNA species i is denoted as $t_{tot,i}$ and can be further partitioned as either charged, of concentration ta_i, or uncharged, of concentration t_i, such that $t_{tot,i} = t_i + ta_i$. Active ribosomes are assumed to undergo reversible binding with both charged and uncharged tRNA species, having codon-specific dissociation constants κ_i^{ta} and κ_i^{t} respectively. Binding to non-cognate tRNAs is neglected. Under the reversible binding model, active ribosomes associated with the ith tRNA can be further partitioned into three possible states based on the occupancy of their A-site: 1) bound with charged tRNA, of concentration $R_{ta,i}$ and referred to as "charged ribosomes", 2) bound with uncharged tRNA, of concentration $R_{t,i}$ and referred to as "uncharged ribosomes”, or 3) not bound to either, of concentration $R_{o,i}$ and referred to as “open ribosomes”. It then holds that $R_{act,i} = R_{ta,i} + R_{t,i} + R_{o,i}$. Using the reversible binding model, the fractional abundance of the charged, uncharged and open ribosome species can be written as

$$\rho_i^{ta} \equiv \frac{R_{ta,i}}{R_{act,i}} = \left(\frac{ta_i}{\kappa_i^{ta}}\right) \Big/ \left(1 + \frac{ta_i}{\kappa_i^{ta}} + \frac{t_i}{\kappa_i^t}\right), \tag{5.26}$$

$$\rho_i^{t} \equiv \frac{R_{t,i}}{R_{act,i}} = \left(\frac{t_i}{\kappa_i^{t}}\right) \Big/ \left(1 + \frac{ta_i}{\kappa_i^{ta}} + \frac{t_i}{\kappa_i^t}\right), \tag{5.27}$$

$$\rho_i^{o} \equiv \frac{R_{o,i}}{R_{act,i}} = 1 \Big/ \left(1 + \frac{ta_i}{\kappa_i^{ta}} + \frac{t_i}{\kappa_i^t}\right). \tag{5.28}$$

To relate these ribosome fractions to the translational rate of the ribosomes, we start with the incorporation flux of the amino acid associated with the i^{th} tRNA, denoted as J_i. The flux J_i is proportional to the number of ribosomes associated with the i^{th} charged tRNA such that $J_i = \varepsilon_{max}\rho_i^{ta} N_{R,i}^{act}$. The proportionality constant is the maximum specific elongation rate, ε_{max}, which is just the rate for ribosome translocation given that it is loaded with the cognate charged tRNA. Due to flux balance of individual amino acids in protein synthesis, the consumption of amino acids from each tRNA species must be balance by the overall protein synthesis flux, J_R, i.e., $J_i = f_i \cdot J_R$ where f_i is the fraction of all actively translated codons corresponding to the i^{th} tRNA species. It will be convenient to express the fluxes in intensive quantities. The protein synthesis flux is generally defined as $J_R \equiv \frac{d}{dt} M_P = \lambda(t) \cdot M_P$ where $\lambda(t) = \frac{d}{dt} \ln M_P$ is the instantaneous growth rate. This leads to

$$f_i \lambda = \varepsilon_{max} R_{act,i} \rho_i^{ta}, \tag{5.29}$$

which is valid not only in steady state but also during transient where λ and ρ may be strongly time-dependent.

To derive an expression for the elongation rate, we solve for $R_{act,i}$ from Eq. (5.29)

$$R_{act,i} = \frac{\lambda}{\varepsilon_{max}} \frac{f_i}{\rho_i^{ta}}, \tag{5.30}$$

and sum it over all tRNA species to obtain

$$R_{act} = \sum_i R_{act,i} = \frac{\lambda}{\varepsilon_{max}} \sum_i f_i / \rho_i^{ta} \, . \tag{5.31}$$

The translation elongation rate is then obtained from its definition as

$$\varepsilon \equiv \frac{\lambda}{R_{act}} = \varepsilon_{max} \Big/ \sum_i \frac{f_i}{\rho_i^{ta}} = \varepsilon_{max} \Big/ \left[1 + \sum_i f_i \frac{\kappa_i^{ta}}{ta_i} \left(1 + \frac{t_i}{\kappa_i^t} \right) \right] , \tag{5.32}$$

where the 2nd equality follows from the definition of ρ_i^{ta} in Eq. (5.26), so that the elongation rate can be expressed in terms of the charged and uncharged tRNA species. Importantly, Eq. (5.32) can be rearranged such that

$$\frac{\varepsilon_{max}}{\varepsilon} - 1 = \sum_i f_i \frac{\kappa_i^{ta}}{ta_i} \left(1 + \frac{t_i}{\kappa_i^t} \right) , \tag{5.33}$$

where the left-hand side is the quantity found empirically to be proportional to the ppGpp level (Eq. 5.1 of the main text) both in steady state and during transient shifts (Fig. 5.2d).

It is also useful to express Eq. (5.33) in terms of the ribosomes in their different states, to make contact with the results derived above based on global flux balance. From Eq. (5.26) and (5.30), we have $R_{ta,i} = R_{act,i} \cdot \rho_i^{ta} = \lambda f_i / \varepsilon_{max}$, and $R_{ta} \equiv \sum_i R_{ta,i} = \lambda / \varepsilon_{max}$ such that $R_{ta,i} = f_i \cdot R_{ta}$. Further, from Eqs. (5.26)-(5.28), we have

$$R_{t,i} = R_{ta,i} \frac{\rho_i^t}{\rho_i^{ta}} = f_i R_{ta} \cdot \left(\frac{k_i^{ta}}{ta_i} \cdot \frac{t_i}{k_i^t} \right), \tag{5.34}$$

$$R_{o,i} = R_{ta,i} \frac{\rho_i^o}{\rho_i^{ta}} = f_i R_{ta} \cdot \left(\frac{k_i^{ta}}{ta_i} \right), \tag{5.35}$$

Summing up each of the above expression over i and using the shorthand $R_t \equiv \sum_i R_{t,i}$, $R_o \equiv \sum_i R_{o,i}$, we obtain

$$R_t + R_o = R_{ta} \cdot \sum_i f_i \cdot \frac{k_i^{ta}}{ta_i}\left(1 + \frac{t_i}{k_i^t}\right). \tag{5.36}$$

Comparing the above relation with Eq. (5.33), we obtain

$$\frac{\varepsilon_{max}}{\varepsilon} - 1 = \frac{R_t + R_o}{R_{ta}}, \tag{5.37}$$

i.e., our central quantity of interest, $\frac{\varepsilon_{max}}{\varepsilon} - 1$, is given by the ratio of the total concentrations of ribosomes not bound to the charged tRNAs ($R_t + R_o$) and those bound to charged tRNAs (R_{ta}). We can thus identify the former group as "dwelling ribosomes" of concentration R_{dwell} referred to in the main text and earlier in this note, and the latter group as "translocating ribosomes" of concentration R_{trans}, with Eq. (5.37) being a mathematical derivation of Eq. (5.25).

Implications for ppGpp synthesis and degradation

It is known experimentally that ppGpp synthesis by RelA requires the binding of RelA-bound uncharged tRNA to the ribosome3. Thus, the pool of ribosomes that RelA samples is R_t in our classification. On the other hand, Eq. (5.37) requires that the observed ppGpp pool to be proportional to $\frac{R_t+R_o}{R_{ta}}$,. This implies that $R_o \ll R_t$, which occurs if $t_i \gg k_i^t$, i.e., it is much favorably for the A-site to be occupied by tRNA, even if uncharged, compared to not occupied at all. In this limit, we have $R_{dwell} \equiv R_t + R_o \approx R_t$, and Eq. (5.37) becomes

$$\frac{R^{dwell}}{R^{trans}} \approx \sum_i f_i \frac{\kappa_i^{ta}}{\kappa_i^t} \frac{t_i}{ta_i}, \tag{5.38}$$

which is a key relation of this Note. It shows that ppGpp levels are directly proportional to the ratio of dwelling to translocating ribosomes, senses a weighted average of the inverse of the tRNA charging ratios, ta_i: t_i, each of which is in turn dictated by the availability of the corresponding amino acid pool. Thus by sensing the ratio of the dwelling and translocating ribosomes, ppGpp is able to combine the charging levels of the many tRNA species and hence the availability of each amino acid into a single signal.

The ratio R_{dwell}/R_{trans} also provides the cell with a convenient biochemical ‘handle’ which can be used to regulate ppGpp synthesis and degradation. Considering a simple model for ppGpp dynamics,

$$\frac{dg}{dt} = \alpha - \beta g, \tag{5.39}$$

which yields $g = \alpha/\beta$. Given Eq. (5.38), our empirical observation $g \propto \frac{\varepsilon_{max}}{\varepsilon} - 1$ can be most conveniently explained if ppGpp is synthesized at a rate proportional to R_{dwell}, such that $\alpha \propto R_{dwell}$, and hydrolyzed at a rate proportional to $g \cdot R_{trans}$, such that $\beta \propto R_{trans}$. The hypothesized form of the synthesis rate is well-justified by the known mechanism of RelA activity as explained in Fig. 5.5 of the main text. The hypothesized form of the hydrolysis rate could arise if SpoT’s hydrolysis activity is stimulated by ribosomes bound to charged tRNA, or during the translocation of ribosomes. Alternatively, SpoT hydrolysis activity could be auto-regulated by ppGpp levels, e.g., $\beta \propto R_{act}/(1+g)$.

Regardless of the form, our analysis suggests that the regulation of SpoT hydrolysis activity is practically a requirement in order to produce the empirical relation between elongation rate and ppGpp level. (An exception is if ppGpp synthesis is controlled by the ratio of R_{dwell} and R_{trans}, which would be very difficult to implement molecularly and is not considered here.) If ppGpp synthesis via RelA was the only point of regulation, with $\alpha \propto R_{dwell}$ and β constant as proposed in (132), then ppGpp levels would be proportional to R_{dwell}. Combining Eq. (5.6) from the main text, $R_{dwell} \cdot \tau_{dwell}^{-1} = R_{trans} \cdot \tau_{trans}^{-1}$, and the constraint $R_{act} = R_{dwell} + R_{trans}$, we can show that $R_{dwell} = R_{act} \cdot \tau_{dwell}/(\tau_{trans} + \tau_{dwell})$. It then follows that, if RelA is the only point of regulation,

$$g \propto R_{act} \frac{\tau_{dwell}}{\tau_{trans} + \tau_{dwell}}. \tag{5.40}$$

In this form it is clear that RelA-exclusive regulation would pose a problem in slow growth conditions where g is observed to be in high abundance, because at slow growth $R_{act} \rightarrow 0$ while $\tau_{dwell}/(\tau_{trans} + \tau_{dwell})$ is limited by saturation. This is shown in Fig. 5.15 where we have plotted observations in steady state growth from the main text against predictions for unregulated SpoT hydrolysis computed from Eq. (5.40). Without regulation of SpoT, ppGpp would be much lower than observed. Likewise, there would be problems during the transient condition shown in Fig. 5.1c, where ppGpp reaches ~8 times their initial level. In such transitions the number of translating ribosomes R_{act}may not change significantly right after the shift. However, the factor $\tau_{dwell}/(\tau_{dwell} + \tau_{trans}) = 1 - \varepsilon/\varepsilon_{max}$ changes only ~3x, from $\varepsilon/\varepsilon_{max} \approx 0.75$ before the shift (Fig. 5.6e) to $\varepsilon/\varepsilon_{max} \approx 0.25$ shortly after the shift. This is clearly shown in Fig. 5.15b where we have plotted transition data from the main text against predictions from Eq. (5.40), assuming R_{act} does not change from pre-shift conditions. Without SpoT regulation, ppGpp is predicted to undershoot observed levels. These problems would only be exacerbated if RelA-mediated activation of ppGpp synthesis saturates according to a Michaelis-Menten relation (i.e. $g \propto R^t/(R^t + K)$), as is proposed in (132). In such a case, ppGpp levels could only be more limited in both steady state and transient conditions. Combined, the slow growth and transient ppGpp levels suggest that regulation of RelA by uncharged tRNA at the ribosomal A-site is by itself insufficient, and some regulation by SpoT is needed. It is therefore not surprising to see the ppGpp level responding in a different way in the $\Delta relA$ strain. However, quantitatively understanding the dynamics of the latter would require also knowing how ppGpp level is set in the steady state in the $\Delta relA$ strain.

Chapter 6

Summary

The work of my dissertation started from extending the model of carbon shift in Erickson et al (23) to multiple growth-transition scenarios, and ended with closing the circuit of the ppGpp based growth rate perception network. In the work, I studied the dynamics of proteomic reallocation during chloramphenicol addition and various amino acid depletion, which all displayed interesting phenotypes. In chloramphenicol downshift, I observed a two-phase growth rate change—a fast drop followed by a slow decrease, indicating an instantaneous translational elongation rate boost upon chloramphenicol addition. In all AA downshifts, we observed a positive linear relationship between the onset time and fractional pre-shift reserves across amino acid biosynthesis (AAB) pathways, indicating an as-needed gene expression pattern in AAB enzymes. Despite more complexities in growth conditions, we showed the power of the flux-controlled regulatory method in predicting kinetics of growth transitions with only qualitative knowledge on gene regulations and limited proteomic measurements at steady state. Although we only focused on three different growth transitions in my dissertation, it opens up the opportunity to quantitatively understanding bacterial behavior under various nutrient shifts (especially amino acid shifts) and other environmental changes in nature.

In my dissertation we also first revealed the rate-sensing strategy of ppGpp signaling system, answering the decades old puzzle of how ppGpp senses the metabolic flux involving 20 different AAs, and as a result leads to different growth rates (along with different proteomic allocations) in different conditions. Along with the downstream effects of ppGpp on ribosomes, we linked growth rate, translational elongation rate, ribosomal abundance and ppGpp level all together. In the past, the studies on bacterial physiological behavior and the studies on detailed molecular gene expression regulations are mostly two parallel lines. Therefore, it is very exciting to see that my work can connect the ribosomal growth law under nutrient limitations with this magical signaling molecule ppGpp. More inspiring, it opens a gate for future bacterial studies. For short term, the rate-sensing strategy allows us link ppGpp level with defined nutrient conditions. One can thus explore the ppGpp responses under other limitations with fixed nutrient condition, including the conditions directly affecting the translational process, which could lead to thoroughly understanding the mechanism of ppGpp synthesis and degradation. Also, it's interesting to extend the study to biological systems other than *E.coli*, involving the ones with different enzymes synthesizing/degrading ppGpp, to see whether the response of ppGpp to translational ER is universal. For long term, with ppGpp being the link between internal metabolites and growth rates, it is possible to look into the effects of internal metabolites on growth rate changes, especially in transient considering the fast responding time of metabolites to the environments. Previous study and my dissertation work on growth transitions based on proteomic allocation only allow us to view bacterial behavior at coarse-grained time scales (time resolution>10mins). But with the link to metabolites, we can quantitatively look into bacterial behavior with the time resolution around one minute. Furthermore, since the rate-sensing mechanism is very helpful for bacteria to obtained the integrated information from variable inputs, it is possible that the bacteria also used it in some other signaling systems. I hope it can inspire some new findings on the regulatory systems of bacteria and even other organisms.

Bibliography

1. M. Mori, S. Schink, D. W. Erickson, U. Gerland, T. Hwa, Quantifying the benefit of a proteome reserve in fluctuating environments. *Nat. Commun.* **8**, 1–8 (2017).

2. M. Mori, Z. Zhang, A. Banaei-Esfahani, J.-B. Lalanne, H. Okano, B. C. Collins, A. Schmidt, O. T. Schubert, D.-S. Lee, G.-W. Li, R. Aebersold, T. Hwa, C. Ludwig, From coarse to fine: the absolute Escherichia coli proteome under diverse growth conditions. *Mol. Syst. Biol.* **17**, e9536 (2021).

3. M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, T. Hwa, Interdependence of cell growth and gene expression: origins and consequences. *Science* **330**, 1099–1102 (2010).

4. H. Bremer, P. P. Dennis, Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential Growth Rates. *EcoSal Plus* **3** (2008).

5. J. Forchhammer, L. Lindahl, Growth rate of polypeptide chains as a function of the cell growth rate in a mutant of Escherichia coli 15. *J. Mol. Biol.* **55**, 563–568 (1971).

6. O. Maaløe, R. F. Goldberger, Biological regulation and development. *Biol. Regul. Dev.* **1**, 487–542 (1979).

7. S. Hui, J. M. Silverman, S. S. Chen, D. W. Erickson, M. Basan, J. Wang, T. Hwa, J. R. Williamson, Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. *Mol. Syst. Biol.* **11**, 784 (2015).

8. D. W. Erickson, S. J. Schink, V. Patsalo, J. R. Williamson, U. Gerland, T. Hwa, A global resource allocation strategy governs growth transition kinetics of Escherichia coli. *Nature* **551**, 119–123 (2017).

9. M. Basan, T. Honda, D. Christodoulou, M. Hörl, Y.-F. Chang, E. Leoncini, A. Mukherjee, H. Okano, B. R. Taylor, J. M. Silverman, C. Sanchez, J. R. Williamson, J. Paulsson, T. Hwa, U. Sauer, A universal trade-off between growth and lag in fluctuating environments. *Nature* **584**, 470–474 (2020).

10. H. Okano, R. Hermsen, K. Kochanowski, T. Hwa, Regulation underlying hierarchical and simultaneous utilization of carbon substrates by flux sensors in Escherichia coli. *Nat. Microbiol.* **5**, 206–215 (2020).

11. A. Kolb, S. Busby, H. Buc, S. Garges, S. Adhya, Transcriptional regulation by cAMP and its receptor protein. *Annu. Rev. Biochem.* **62**, 749–795 (1993).

12. C. You, H. Okano, S. Hui, Z. Zhang, M. Kim, C. W. Gunderson, Y.-P. Wang, P. Lenz, D. Yan, T. Hwa, Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature* **500**, 301–306 (2013).

13. J. Ryals, R. Little, H. Bremer, Control of rRNA and tRNA syntheses in Escherichia coli by guanosine tetraphosphate. *J. Bacteriol.* **151**, 1261–1268 (1982).

14. K. Potrykus, M. Cashel, (p)ppGpp: Still Magical? *Annu. Rev. Microbiol.* **62**, 35–51 (2008).

15. J. J. Lemke, P. Sanchez-Vazquez, H. L. Burgos, G. Hedberg, W. Ross, R. L. Gourse, J. W. Roberts, Direct regulation of Escherichia coli ribosomal protein promoters by the transcription factors ppGpp and DksA. **108** (2011).

16. M. Y. Pavlov, M. Ehrenberg, Optimal control of gene expression for fast proteome adaptation to environmental change. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 20527–20532 (2013).

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

18. M. Basan, S. Hui, H. Okano, Z. Zhang, Y. Shen, J. R. Williamson, T. Hwa, Overflow metabolism in Escherichia coli results from efficient proteome allocation. *Nature* **528**, 99–104 (2015).

19. I. Schomburg, A. Chang, D. Schomburg, BRENDA, enzyme data and metabolic information. *Nucleic Acids Res.* **30**, 47–49 (2002).

20. G.-W. Li, D. Burkhardt, C. Gross, J. S. Weissman, Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* **157**, 624–635 (2014).

21. M. A. Savageau, Escherichia coli Habitats, Cell Types, and Molecular Mechanisms of Gene Control. *Am. Nat.* **122**, 732–744 (1983).

22. F. C. Neidhardt, P. L. Bloch, D. F. Smith, Culture Medium for Enterobacteria. *J. Bacteriol.* **119**, 736–747 (1974).

23. D. W. Erickson, S. J. Schink, V. Patsalo, J. R. Williamson, U. Gerland, T. Hwa, A global resource allocation strategy governs growth transition kinetics of Escherichia coli. *Nature* **551**, 119–123 (2017).

24. H. Weissbach, N. Brot, Regulation of methionine synthesis in Escherichia coli. *Mol. Microbiol.* **5**, 1593–1597 (1991).

25. J. J. Lemke, P. Sanchez-Vazquez, H. L. Burgos, G. Hedberg, W. Ross, R. L. Gourse, Direct regulation of Escherichia coli ribosomal protein promoters by the transcription factors ppGpp and DksA. *Proc. Natl. Acad. Sci.* **108**, 5712–5717 (2011).

26. M. Scott, S. Klumpp, E. M. Mateescu, T. Hwa, Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Mol. Syst. Biol.* **10**, 747 (2014).

27. V. Hauryliuk, G. C. Atkinson, K. S. Murakami, T. Tenson, K. Gerdes, Recent functional insights into the role of (p)ppGpp in bacterial physiology. *Nat. Rev. Microbiol.* **13**, 298–309 (2015).

28. D. A. Smith, "S-Amino Acid Metabolism and its Regulation in. Escherichia coli and Salmonella typhimurium" in Advances in Genetics., E. W. Caspari, Ed. (Academic Press, 1971), pp. 141–165.

29. S. D. Brown, S. Jun, Complete Genome Sequence of Escherichia coli NCM3722. *Genome Announc.* **3**, e00879-15 (2015).

30. E. Soupene, W. C. van Heeswijk, J. Plumbridge, V. Stewart, D. Bertenthal, H. Lee, G. Prasad, O. Paliy, P. Charernnoppakul, S. Kustu, Physiological studies of Escherichia coli strain MG1655: growth defects and apparent cross-regulation of gene expression. *J. Bacteriol.* **185**, 5611–5626 (2003).

31. S. Cayley, B. A. Lewis, H. J. Guttman, M. T. Record, 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).

32. J. Steinier, Y. Termonia, J. Deltour, Smoothing and differentiation of data by simplified least square procedure. *Anal. Chem.* **44**, 1906–1909 (1972).

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

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

35. M. U. Caglar, J. R. Houser, C. S. Barnhart, D. R. Boutz, S. M. Carroll, A. Dasgupta, W. F. Lenoir, B. L. Smith, V. Sridhara, D. K. Sydykova, D. Vander Wood, C. J. Marx, E. M. Marcotte, J. E. Barrick, C. O. Wilke, The E. coli molecular phenotype under different growth conditions. *Sci. Rep.* **7**, 45303 (2017).

36. X. Wang, K. Xia, X. Yang, C. Tang, Growth strategy of microbes on mixed carbon sources. *Nat. Commun.* **10**, 1279 (2019).

37. L. V Holdeman, 1929-, W. E. C. Moore, E. P. Cato, Anaerobe laboratory manual. : ill.; 29 cm. (1977).

38. A. Chacko, J. H. Cummings, Nitrogen losses from the human small bowel: obligatory losses and the effect of physical form of food. *Gut* **29**, 809–815 (1988).

39. L. C. Gillet, P. Navarro, S. Tate, H. Röst, N. Selevsek, L. Reiter, R. Bonner, R. Aebersold, Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell.*

Proteomics **11**, O111.016717 (2012).

40. C. Ludwig, L. Gillet, G. Rosenberger, S. Amon, B. C. Collins, R. Aebersold, Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol. Syst. Biol.* **14**, e8126 (2018).

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

42. P. D. Karp, W. K. Ong, S. Paley, R. Billington, C. Fulcher, A. Kothari, M. Krummenacker, M. Latendresse, P. E. Midford, P. Subhraveti, S. Gama-Castro, L. Muniz-Rascado, C. Bonavides-Martinez, A. Santos-Zavaleta, A. Mackie, J. Collado-Vides, I. M. Keseler, I. Paulsen, “The EcoCyc Database” in *EcoSal Plus: Cellular and Molecular Biology of E. Coli, Salmonella, and the Enterobacteriaceae*, (ASM Press, 2018), pp. 200–400.

43. E. C. Lin, Glycerol dissimilation and its regulation in bacteria. *Annu. Rev. Microbiol.* **30**, 535–578 (1976).

44. C. A. Smith, Physiology of the bacterial cell. A molecular approach. By F C Neidhardt, J L Ingraham and M Schaechter. pp 507. Sinauer associates, Sunderland, MA. 1990. $43.95 ISBN 0–87893–608–4. *Biochem. Educ.* **20**, 124–125 (1992).

45. C. Wu, R. Balakrishnan, M. Mori, G. Manzanarez, Z. Zhang, T. Hwa, Cellular perception of growth rate and the mechanistic origin of bacterial growth laws. *bioRxiv*, 2021.10.16.464649 (2021).

46. T. Hwa, *The physics of living matter space, time and information : the proceedings of the 27th Solvay Conference On Physics, Brussels, Belgium, 19-21 October 2017*, D. Gross, A. Sevrin, B. Shraiman, Eds. (World Scientific, 2020).

47. P. Sanchez-Vazquez, C. N. Dewey, N. Kitten, W. Ross, R. L. Gourse, Genome-wide effects on Escherichia coli transcription from ppGpp binding to its two sites on RNA polymerase. *Proc. Natl. Acad. Sci.* **116**, 8310–8319 (2019).

48. Y. Yang, A. M. Pollard, C. Höfler, G. Poschet, M. Wirtz, R. Hell, V. Sourjik, Relation between chemotaxis and consumption of amino acids in bacteria. *Mol. Microbiol.* **96**, 1272–1282 (2015).

49. M. Zampieri, M. Hörl, F. Hotz, N. F. Müller, U. Sauer, Regulatory mechanisms underlying coordination of amino acid and glucose catabolism in Escherichia coli. *Nat. Commun.* **10**, 3354 (2019).

50. W. G. Gordon, W. F. Semmett, R. S. Cable, M. Morris, Amino Acid Composition of α-

Casein and β-Casein2. *J. Am. Chem. Soc.* **71**, 3293–3297 (1949).

51. A. Zaslaver, A. E. Mayo, R. Rosenberg, P. Bashkin, H. Sberro, M. Tsalyuk, M. G. Surette, U. Alon, Just-in-time transcription program in metabolic pathways. *Nat. Genet.* **36**, 486–491 (2004).

52. D. Gilis, S. Massar, N. J. Cerf, M. Rooman, Optimality of the genetic code with respect to protein stability and amino-acid frequencies. *Genome Biol.* **2**, research0049.1 (2001).

53. H. Hama, Y. Sumita, Y. Kakutani, M. Tsuda, T. Tsuchiya, Target of serine inhibition in Escherichiacoli. *Biochem. Biophys. Res. Commun.* **168**, 1211–1216 (1990).

54. H. Hama, T. Kayahara, M. Tsuda, T. Tsuchiya, Inhibition of homoserine dehydrogenase I by L-serine in Escherichia coli. *J. Biochem.* **109**, 604–608 (1991).

55. G. V Stauffer, Regulation of Serine, Glycine, and One-Carbon Biosynthesis. *EcoSal Plus* **1** (2004).

56. F. R. Blattner, G. 3rd Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, Y. Shao, The complete genome sequence of Escherichia coli K-12. *Science* **277**, 1453–1462 (1997).

57. H. R. Kaback, The lac carrier protein in Escherichia coli. *J. Membr. Biol.* **76**, 95–112 (1983).

58. J. B. Stock, E. B. Waygood, N. D. Meadow, P. W. Postma, S. Roseman, Sugar transport by the bacterial phosphotransferase system. The glucose receptors of the Salmonella typhimurium phosphotransferase system. *J. Biol. Chem.* **257**, 14543–14552 (1982).

59. O. Misset, M. Blaauw, P. W. Postma, G. T. Robillard, Bacterial phosphoenolpyruvate-dependent phosphotransferase system. Mechanism of the transmembrane sugar translocation and phosphorylation. *Biochemistry* **22**, 6163–6170 (1983).

60. P. W. Postma, J. W. Lengeler, G. R. Jacobson, Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol. Rev.* **57**, 543–594 (1993).

61. C. Ahlem, W. Huisman, G. Neslund, A. S. Dahms, Purification and properties of a periplasmic D-xylose-binding protein from Escherichia coli K-12. *J. Biol. Chem.* **257**, 2926–2931 (1982).

62. M. Sumiya, E. O. Davis, L. C. Packman, T. P. McDonald, P. J. Henderson, Molecular genetics of a receptor protein for D-xylose, encoded by the gene xylF, in Escherichia coli. *Receptors Channels* **3**, 117–128 (1995).

63. K. J. Linton, C. F. Higgins, The Escherichia coli ATP-binding cassette (ABC) proteins. *Mol. Microbiol.* **28**, 5–13 (1998).

64. F. Si, D. Li, S. E. Cox, J. T. Sauls, O. Azizi, C. Sou, A. B. Schwartz, M. J. Erickstad, Y. Jun, X. Li, S. Jun, Invariance of Initiation Mass and Predictability of Cell Size in Escherichia coli. *Curr. Biol.* **27**, 1278–1287 (2017).

65. R. J. Harvey, A. L. Koch, How partially inhibitory concentrations of chloramphenicol affect the growth of Escherichia coli. *Antimicrob. Agents Chemother.* **18**, 323–337 (1980).

66. S. E. Irving, N. R. Choudhury, R. M. Corrigan, The stringent response and physiological roles of (pp)pGpp in bacteria. *Nat. Rev. Microbiol.* **19**, 256–271 (2021).

67. D. Molenaar, R. van Berlo, D. de Ridder, B. Teusink, Shifts in growth strategies reflect tradeoffs in cellular economics. *Mol. Syst. Biol.* **5**, 323 (2009).

68. F. J. Bruggeman, R. Planqué, D. Molenaar, B. Teusink, Searching for principles of microbial physiology. *FEMS Microbiol. Rev.* **44**, 821–844 (2020).

69. M. Schaechter, O. MaalØe, N. O. Kjeldgaard, Dependency on Medium and Temperature of Cell Size and Chemical Composition during Balanced Growth of Salmonella typhimurium. *Microbiology* **19**, 592–606 (1958).

70. F. C. Neidhardt, B. Magasanik, Studies on the role of ribonucleic acid in the growth of bacteria. *Biochim. Biophys. Acta* **42**, 99–116 (1960).

71. O. Maaløe, "Regulation of the Protein-Synthesizing Machinery---Ribosomes, tRNA, Factors, and So On" in *Biological Regulation and Development: Gene Expression*, R. F. Goldberger, Ed. (Springer US, 1979), pp. 487–542.

72. R. Young, H. Bremer, Polypeptide-chain-elongation rate in Escherichia coli B/r as a function of growth rate. *Biochem. J.* **160**, 185–194 (1976).

73. H. Brunschede, T. L. Dove, H. Bremer, Establishment of exponential growth after a nutritional shift-up in Escherichia coli B/r: accumulation of deoxyribonucleic acid, ribonucleic acid, and protein. *J. Bacteriol.* **129**, 1020–1033 (1977).

74. A. L. Koch, The inefficiency of ribosomes functioning in Escherichia coli growing at moderate rates. *J. Gen. Microbiol.* **116**, 165–171 (1980).

75. S. Klumpp, M. Scott, S. Pedersen, T. Hwa, Molecular crowding limits translation and cell growth. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 16754–16759 (2013).

76. M. Cashel, J. Gallant, Two compounds implicated in the function of RC gene. *Nature* **221**, 838–841 (1969).

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

78. L. Fernández-Coll, M. Cashel, Possible Roles for Basal Levels of (p)ppGpp: Growth Efficiency Vs. Surviving Stress. *Front. Microbiol.* **11** (2020).

79. L. U. Magnusson, A. Farewell, T. Nyström, ppGpp: A global regulator in Escherichia coli. *Trends Microbiol.* **13**, 236–242 (2005).

80. A. Srivatsan, J. D. Wang, Control of bacterial transcription, translation and replication by (p)ppGpp. *Curr. Opin. Microbiol.* **11**, 100–5 (2008).

81. S. Ronneau, R. Hallez, Make and break the alarmone: Regulation of (p)ppGpp synthetase/hydrolase enzymes in bacteria. *FEMS Microbiol. Rev.* **43**, 389–400 (2019).

82. H. Bremer, M. Ehrenberg, Guanosine tetraphosphate as a global regulator of bacterial RNA synthesis: a model involving RNA polymerase pausing and queuing. *BBA - Gene Struct. Expr.* **1262**, 15–36 (1995).

83. M. F. Traxler, S. M. Summers, † Huyen-Tran Nguyen, V. M. Zacharia, G. A. Hightower, J. T. Smith, T. Conway, The global, ppGpp-mediated stringent response to amino acid starvation in Escherichia coli (2008) https:/doi.org/10.1111/j.1365-2958.2008.06229.x.

84. E. Baracchini, H. Bremer, Stringent and growth control of rRNA synthesis in Escherichia coli are both mediated by ppGpp. *J. Biol. Chem.* **263**, 2597–2602 (1988).

85. B. J. Paul, W. Ross, T. Gaal, R. L. Gourse, rRNA transcription in Escherichia coli. *Annu. Rev. Genet.* **38**, 749–770 (2004).

86. R. B. Harshman, H. Yamazaki, Formation of ppGpp in a Relaxed and Stringent Strain of Escherichia coli during Diauxie Lag. *Biochemistry* **10**, 3980–3982 (1971).

87. X. Dai, M. Zhu, M. Warren, R. Balakrishnan, H. Okano, J. R. Williamson, K. Fredrick, T. Hwa, Slowdown of translational elongation in Escherichia coli under hyperosmotic stress. *MBio* **9** (2018).

88. I. L. Volkov, M. Lindén, J. Aguirre Rivera, K. W. Ieong, M. Metelev, J. Elf, M. Johansson, TRNA tracking for direct measurements of protein synthesis kinetics in live cells. *Nat. Chem. Biol.* **14**, 618–626 (2018).

89. I. L. Volkov, A. C. Seefeldt, M. Johansson, Tracking of single tRNAs for translation kinetics measurements in chloramphenicol treated bacteria. *Methods* **162**–**163**, 23–30 (2019).

90. A. L. Goldberg, A. C. St John, Intracellular protein degradation in mammalian and bacterial cells: Part 2. *Annu. Rev. Biochem.* **45**, 747–803 (1976).

91. K. Nath, A. L. Koch, Protein degradation in Escherichia coli. II. Strain differences in the degradation of protein and nucleic acid resulting from starvation. *J. Biol. Chem.* **246**, 6956–6967 (1971).

92. K. Izutsu, A. Wada, C. Wada, Expression of ribosome modulation factor (RMF) in Escherichia coli requires ppGpp. *Genes to Cells* **6**, 665–676 (2001).

93. M. Ueta, R. L. Ohniwa, H. Yoshida, Y. Maki, C. Wada, A. Wada, Role of HPF (hibernation promoting factor) in translational activity in Escherichia coli. *J. Biochem.* **143**, 425–433 (2008).

94. Y. S. Polikanov, G. M. Blaha, T. A. Steitz, How hibernation factors RMF, HPF, and YfiA turn off protein synthesis. *Science (80-.).* **336**, 915–918 (2012).

95. P. Milon, E. Tischenko, J. Tomsic, E. Caserta, G. Folkers, A. La Teana, M. V Rodnina, C. L. Pon, R. Boelens, C. O. Gualerzi, The nucleotide-binding site of bacterial translation initiation factor 2 (IF2) as a metabolic sensor. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 13962–13967 (2006).

96. S. Diez, J. Ryu, K. Caban, R. L. J. Gonzalez, J. Dworkin, The alarmones (p)ppGpp directly regulate translation initiation during entry into quiescence. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 15565–15572 (2020).

97. D. S. Vinogradova, V. Zegarra, E. Maksimova, J. A. Nakamoto, P. Kasatsky, A. Paleskava, A. L. Konevega, P. Milón, How the initiating ribosome copes with ppGpp to translate mRNAs. *PLOS Biol.* **18**, 1–25 (2020).

98. E. Sarubbi, K. E. Rudd, M. Cashel, Basal ppGpp level adjustment shown by new spoT mutants affect steady state growth rates and rrnA ribosomal promoter regulation in Escherichia coli. *MGG Mol. Gen. Genet.* **213**, 214–222 (1988).

99. R. L. Gourse, T. Gaal, M. S. Bartlett, J. A. Appleman, W. Ross, rRNA Transcription and growth rate-dependent regulation of ribosome synthesis in Escherichia coli. *Annu. Rev. Microbiol.* **50**, 645–677 (1996).

100. P. Sanchez-Vazquez, C. N. Dewey, N. Kitten, W. Ross, R. L. Gourse, Genome-wide effects on Escherichia coli transcription from ppGpp binding to its two sites on RNA polymerase https:/doi.org/10.1073/pnas.1819682116.

101. F. Neidhardt, B. Magasanik, Studies on the role of ribonucleic acid in the growth of bacteria. *Biochem. Biophys. Acta* **42**, 99–116 (1959).

102. E. Samatova, J. Daberger, M. Liutkute, M. V. Rodnina, Translational Control by Ribosome Pausing in Bacteria: How a Non-uniform Pace of Translation Affects Protein Production and Folding. *Front. Microbiol.* **11** (2021).

103. Y. Liu, A code within the genetic code: Codon usage regulates co-translational protein folding. *Cell Commun. Signal.* **18**, 1–9 (2020).

104. M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, T. Hwa, Interdependence of Cell Growth. *Science (80-.).* **330**, 1099–1102 (2010).

105. M. Mori, E. Marinari, A. De Martino, A yield-cost tradeoff governs Escherichia coli's decision between fermentation and respiration in carbon-limited growth. *npj Syst. Biol. Appl.* **5** (2019).

106. H. Link, T. Fuhrer, L. Gerosa, N. Zamboni, U. Sauer, Real-time metabolome profiling of the metabolic switch between starvation and growth. *Nat. Methods* **12**, 1091–1097 (2015).

107. A. B. Loveland, E. Bah, R. Madireddy, Y. Zhang, A. F. Brilot, N. Grigorieff, A. A. Korostelev, Ribosome•RelA structures reveal the mechanism of stringent response activation. *Elife* **5**, e17029 (2016).

108. S. Arenz, M. Abdelshahid, D. Sohmen, R. Payoe, A. L. Starosta, O. Berninghausen, V. Hauryliuk, R. Beckmann, D. N. Wilson, The stringent factor RelA adopts an open conformation on the ribosome to stimulate ppGpp synthesis. *Nucleic Acids Res.* **44**, 6471–6481 (2016).

109. W. A. Haseltine, R. Block, Synthesis of guanosine tetra- and pentaphosphate requires the presence of a codon-specific, uncharged transfer ribonucleic acid in the acceptor site of ribosomes. *Proc. Natl. Acad. Sci. U. S. A.* **70**, 1564–1568 (1973).

110. T. M. Wendrich, G. Blaha, D. N. Wilson, M. A. Marahiel, K. H. Nierhaus, Dissection of the mechanism for the stringent factor RelA. *Mol. Cell* **10**, 779–788 (2002).

111. H. Xiao, M. Kalman, K. Ikehara, S. Zemel, G. Glaser, M. Cashel, Residual guanosine 3',5'-bispyrophosphate synthetic activity of relA null mutants can be eliminated by spoT null mutations. *J. Biol. Chem.* **266**, 5980–5990 (1991).

112. M. Jiang, S. M. Sullivan, P. K. Wout, J. R. Maddock, G-protein control of the ribosome-associated stress response protein SpoT. *J. Bacteriol.* **189**, 6140–7 (2007).

113. F. C. Neidhardt, GUEST COMMENTARY Bacterial Growth: Constant Obsession with dN/dt. *J. Bacteriol.* **181**, 7405–7408 (1999).

114. N. M. Belliveau, G. Chure, C. L. Hueschen, H. G. Garcia, J. Kondev, D. S. Fisher, J. A. Theriot, R. Phillips, Fundamental limits on the rate of bacterial growth. *bioRxiv* (2020) https:/doi.org/10.1101/2020.10.18.344382.

115. S. H. J. Li, Z. Li, J. O. Park, C. G. King, J. D. Rabinowitz, N. S. Wingreen, Z. Gitai, Escherichia coli translation strategies differ across carbon, nitrogen and phosphorus limitation conditions. *Nat. Microbiol.* **3**, 939–947 (2018).

116. J. Hughes, G. Mellows, Inhibition of isoleucyl-transfer ribonucleic acid synthetase in Escherichia coli by pseudomonic acid. *Biochem. J.* **176**, 305–318 (1978).

117. P. M. Bennett, O. Maaløe, The effects of fusidic acid on growth, ribosome synthesis and RNA metabolism in Escherichia coli. *J. Mol. Biol.* **90**, 541–561 (1974).

118. S. Vadia, J. L. Tse, R. Lucena, Z. Yang, D. R. Kellogg, J. D. Wang, P. A. Levin, Fatty Acid Availability Sets Cell Envelope Capacity and Dictates Microbial Cell Size. *Curr. Biol.* **27**, 1757-1767.e5 (2017).

119. T. Prossliner, K. Skovbo Winther, M. A. Sørensen, K. Gerdes, Ribosome Hibernation. *Annu. Rev. Genet.* **52**, 321–348 (2018).

120. L. Fernández-Coll, M. Cashel, Contributions of SpoT hydrolase, SpoT synthetase, and RelA synthetase to carbon source diauxic growth transitions in Escherichia coli. *Front. Microbiol.* **9**, 1–13 (2018).

121. T. Hwa, Cell Behavior and Control in *The Physics of Living Matter: Space, Time and Information*, (WORLD SCIENTIFIC, 2020), pp. 85–134.

122. S. Cayley, B. A. Lewis, H. J. Guttman, M. T. Record, 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).

123. K. A. Datsenko, B. L. Wanner, One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 6640–5 (2000).

124. R. Schleif, W. Hess, S. Finkelstein, D. Ellis, Induction kinetics of the L arabinose operon of Escherichia coli. *J. Bacteriol.* **115**, 9–14 (1973).

125. M. Cashel, The control of ribonucleic acid synthesis in Escherichia coli. IV. Relevance of unusual phosphorylated compounds from amino acid-starved stringent strains. *J. Biol. Chem.* **244**, 3133–3141 (1969).

126. S. Benthin, J. Nielsen, J. Villadsen, A simple and reliable method for the determination of cellular RNA content. *Biotechnol. Tech.* **5**, 39–42 (1991).

127. D. Herbert, P. J. Phipps, R. E. Strange, Chemical Analysis of Microbial Cells. *Methods Microbiol.* **5**, 209–344 (1971).

128. A. Brown, I. S. Fernández, Y. Gordiyenko, V. Ramakrishnan, Ribosome-dependent activation of stringent control. *Nature* **534**, 277–280 (2016).

129. R. Balakrishnan, M. Mori, I. Segota, Z. Zhang, R. Aebersold, C. Ludwig, T. Hwa, Principles of gene regulation quantitatively connect DNA to RNA and proteins in bacteria. *bioRxiv*, 2021.05.24.445329 (2021).

130. D. E. Agafonov, V. A. Kolb, A. S. Spirin, Ribosome-associated protein that inhibits translation at the aminoacyl-tRNA binding stage. *EMBO Rep.* **2**, 399–402 (2001).

131. J. Elf, M. Ehrenberg, Near-critical behavior of aminoacyl-tRNA pools in E. coli at rate-limiting supply of amino acids. *Biophys. J.* **88**, 132–146 (2005).

132. E. Bosdriesz, D. Molenaar, B. Teusink, F. J. Bruggeman, How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization. *FEBS J.* **282**, 2029–2044 (2015).

133. M. J. Pine, Steady-state measurement of the turnover of amino acid in the cellular proteins of growing Escherichia coli: existence of two kinetically distinct reactions. *J. Bacteriol.* **103**, 207–215 (1970).