Perturbation of the bacterial stress and starvation alarm

Lacassin, M.L.R.

DOI
10.4233/uuid:e7300328-8828-4c59-8a34-f004211b5a14

Publication date
2025

Document Version
Final published version

Citation (APA)
Lacassin, M. L. R. (2025). *Perturbation of the bacterial stress and starvation alarm*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:e7300328-8828-4c59-8a34-f004211b5a14

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Perturbation of the Bacterial Stress and Starvation Alarm

A PhD Thesis by Milan Lacassin

PERTURBATION OF THE BACTERIAL STRESS AND STARVATION ALARM

PERTURBATION OF THE BACTERIAL STRESS AND STARVATION ALARM

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen,
Chair of the Board of Doctorates,
to be publicly defended on Thursday July 3rd at 10.30am

by

Milan Lucien Roger LACASSIN

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. Y. M. Blanter,	Delft University of Technology, promotor
Dr. G. E. Bokinsky,	Delft University of Technology, promotor

Independent members:

Prof. dr. F.J. Bruggeman,	Vrije Universiteit Amsterdam
Dr. M. Consentino Lagomarsino,	University of Milan
Prof. dr. ir. P.A.S. Daran-Lapujade,	Delft University of Technology
Dr. A. Jakobi,	Delft University, of Technology
Dr. M. Pabst,	Delft University, of Technology
Prof. dr. G.H. Koenderink,	Delft University of Technology (reserve member)

Keywords: bacterial growth, proteome, chromatography, mass spectrometry, ppGpp, guanosine tetraphosphate, modeling, yeast, freeze-drying, anhydrobiosis.

Printed by: *Gildeprint, Enschede.*

Front & Back: *Another PhD Thesis in the Wall,*
digital painting by Konstantina D. Schoina.

ISBN 978-94-6384-812-1

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

Si vous allez trop loin, vous n'irez nulle part.

If you go too far, you will go nowhere.

Jacques Monod

CONTENTS

SUMMARY

From the identification of the pathogens causing major human diseases, through the understanding of the most basic forms of life, and in the struggle to develop new antibiotics to counter the increasing bacterial resistance to these widely used substances, research on microbial growth and survival has rightfully seen a booming interest across the last two centuries.

The guanosine tetraphosphate (ppGpp) signaling system is of particular significance for these fields of research. It is a bacterial response to stress and starvation which allows these organisms to activate the necessary genes to survive in these conditions. It also plays a key role in modulating the abundance of various machineries necessary for growth in order to maximize the rate at which bacteria are growing in different nutrient conditions. In poorer nutrient conditions requiring more enzymes to import and digest these nutrients into the same amount of essential building blocks for growth, production of ppGpp is triggered in response to the lack of these building blocks. Higher ppGpp concentrations lead to downregulation of the abundance of ribosomes, the machineries operating growth, allowing larger abundance of enzymes producing the building blocks from nutrients. Recent advances, by genetically modifying bacteria, allowed to finely tune ppGpp concentrations independently of growth conditions. Using this approach, we investigate the scope of the regulation operated by ppGpp: which proteins does it influence in changing growth conditions and to what extent?

In **Chapter 1**, we review the current knowledge of the ppGpp signaling system, with a focus on three points: results obtained in *E. coli*, the effect of this response on growth and what is known about the consequences of artificially perturbing ppGpp concentrations.

In **Chapter 2**, we present novel quantifications of different proteins and other relevant biomolecules in *E. coli* strains artificially modified to have higher or lower ppGpp concentration than naturally found. These results allow us to identify which proteins are regulated by ppGpp and why the right concentration of ppGpp is necessary for bacteria to achieve optimal growth. We confirm the role of ppGpp in upregulating the proteins related to the synthesis of new proteins: translation. We also show that, apart from translation-related proteins, ppGpp is not responsible for consistently regulating other groups of proteins, including some that do vary with growth rate variations caused by change in nutrient source, which were suspected to be under ppGpp's control.

In **Chapter 3**, we develop a simplified mathematical model of the ppGpp regulation and the way it is affected by artificial perturbation. We group compounds responsible for a common function, such as digesting nutrients into building blocks, assembling these building blocks into new biomass, or acheminating these building blocks to be assembled. With this mathematical model, we attempt to understand why slowing down bacterial growth by artificially increasing ppGpp requires a lot more ppGpp than naturally found. By doing so, we identify characteristics that a model seeking to explain ppGpp perturbations should have. These conditions strengthened our understanding of the

ppGpp system by refuting some of our intuitions and laying a foundation for future models elucidating the missing pieces.

In the last **Chapter 4**, we present another story regarding how yeast can survive one of the harshest stress. All known forms of life are based on interactions between biomolecules happening in water. For this reason, removing all water is deadly to most organisms. However, a few of them are able to survive while being completely desiccated. Their lives are halted and they stop growing, but when placed in water again, they are able to recover their usual functions. We studied the conditions that allow the yeast *Saccharomyces cerevisiae* to do so. We found that yeast cells face two major challenges in drying conditions: preservation of their cell membrane and of their ability to express genes from their DNA.

With these different studies, we extend our knowledge of two different systems relevant to growth and survival of microorganisms, which leads to potential new directions for those who seek to investigate such systems and answer some of the questions that our findings raise.

SAMENVATTING

Vanaf de identificatie van de ziekteverwekkers die belangrijke ziekten bij de mens veroorzaken, via het begrip van de meest fundamentele levensvormen, en de strijd om nieuwe antibiotica te ontwikkelen om de toenemende bacteriële resistentie tegen deze veelgebruikte stoffen tegen te gaan, heeft onderzoek naar microbiële groei en overleving terecht de afgelopen twee eeuwen een enorme belangstelling gekend.

Het guanosinetetrafosfaat (ppGpp) signaalsysteem is van bijzonder belang voor deze onderzoeksgebieden. Het is een bacteriële reactie op stress en honger, waardoor deze organismen de noodzakelijke genen kunnen activeren om onder deze omstandigheden te overleven. Het speelt ook een sleutelrol bij het moduleren van de overvloed aan verschillende machines die nodig zijn voor groei, om de snelheid waarmee bacteriën groeien onder verschillende voedingsomstandigheden te maximaliseren. In armere voedingsomstandigheden die meer enzymen vereisen om deze voedingsstoffen te importeren en te verteren tot dezelfde hoeveelheid essentiële bouwstenen voor groei, wordt de productie van ppGpp geactiveerd als reactie op het ontbreken van deze bouwstenen. Hogere ppGpp-concentraties leiden tot een neerwaartse regulatie van de overvloed aan ribosomen, de machines die de groei verzorgen, waardoor een grotere overvloed aan enzymen mogelijk wordt die de bouwstenen uit voedingsstoffen produceren. Recente ontwikkelingen, door bacteriën genetisch te modificeren, hebben het mogelijk gemaakt om de ppGpp-concentraties nauwkeurig af te stemmen, onafhankelijk van de groeiomstandigheden. Met deze aanpak onderzoeken we de reikwijdte van de regulering van ppGpp: welke eiwitten beïnvloedt het in veranderende groeiomstandigheden en in welke mate?

In **Hoofdstuk 1** bespreken we de huidige kennis van het ppGpp-signaleringssysteem, met de nadruk op drie punten: resultaten verkregen in *E. coli*, het effect van deze reactie op de groei en wat er bekend is over de gevolgen van het kunstmatig verstoren van ppGpp-concentraties.

In **Hoofdstuk 2** presenteren we nieuwe kwantificeringen van verschillende eiwitten en andere relevante biomoleculen in *E. coli*-stammen die kunstmatig zijn gemodificeerd om een hogere of lagere ppGpp-concentratie te hebben dan van nature voorkomt. Deze resultaten stellen ons in staat te identificeren welke eiwitten worden gereguleerd door ppGpp en waarom de juiste concentratie ppGpp nodig is voor bacteriën om optimale groei te bereiken. We bevestigen de rol van ppGpp bij het opreguleren van de eiwitten die verband houden met de synthese van nieuwe eiwitten: translatie. We laten ook zien dat ppGpp, afgezien van translatiegerelateerde eiwitten, niet verantwoordelijk is voor het consistent reguleren van andere groepen eiwitten, waaronder enkele die wel variëren met variaties in de groeisnelheid veroorzaakt door veranderingen in de voedingsbron, waarvan werd vermoed dat ze onder de controle van ppGpp stonden.

In **Hoofdstuk 3** ontwikkelen we een vereenvoudigd wiskundig model van de ppGpp-regulatie en de manier waarop deze wordt beïnvloed door kunstmatige verstoring. We groeperen verbindingen die verantwoordelijk zijn voor een gemeenschappelijke func-

tie, zoals het verteren van voedingsstoffen tot bouwstenen, het assembleren van deze bouwstenen tot nieuwe biomassa, of het achemineren van deze bouwstenen om te assembleren. Met dit wiskundige model proberen we te begrijpen waarom het vertragen van de bacteriegroei door het kunstmatig verhogen van ppGpp veel meer ppGpp vereist dan natuurlijk wordt aangetroffen. Door dit te doen identificeren we kenmerken die een model dat ppGpp-verstoringen probeert te verklaren zou moeten hebben. Deze omstandigheden versterkten ons begrip van het ppGpp-systeem door enkele van onze intuïties te weerleggen en een basis te leggen voor toekomstige modellen die de ontbrekende stukken ophelderen.

In het laatste **Hoofdstuk 4** presenteren we nog een verhaal over hoe gist een van de zwaarste stresssituaties kan overleven. Alle bekende levensvormen zijn gebaseerd op interacties tussen biomoleculen die in water plaatsvinden. Om deze reden is het verwijderen van al het water dodelijk voor de meeste organismen. Enkelen van hen kunnen echter overleven terwijl ze volledig uitgedroogd zijn. Hun leven wordt stopgezet en ze stoppen met groeien, maar wanneer ze weer in water worden geplaatst, kunnen ze hun gebruikelijke functies herstellen. We hebben de omstandigheden bestudeerd die de gist *Saccharomyces cerevisiae* daartoe in staat stellen. We ontdekten dat gistcellen bij het drogen met twee grote uitdagingen worden geconfronteerd: het behoud van hun celmembraan en van hun vermogen om genen uit hun DNA tot expressie te brengen.

Met deze verschillende onderzoeken breiden we onze kennis uit van twee verschillende systemen die relevant zijn voor de groei en overleving van micro-organismen, wat leidt tot potentiële nieuwe richtingen voor degenen die dergelijke systemen willen onderzoeken en enkele van de vragen willen beantwoorden die onze bevindingen oproepen.

1

Introduction: The bacterial stress and starvation alarm

Milan Lacassin, Yaroslav Blanter, Gregory Bokinsky

The dream of every cell is to become two cells.

François Jacob

1

The (p)ppGpp signaling system is a bacterial regulatory system involved in various important aspects of bacterial physiology, from maximizing growth rate to virulence of pathogenic bacteria. While (p)ppGpp synthesis triggers, the signal it responds to and its numerous targets have been extensively studied, the global physiological changes it operates remain poorly understood and the consequences of perturbing (p)ppGpp concentrations on growth cannot be predicted by most attempts to model this signaling response. In this chapter, we hope to provide the reader with relevant information about what is known about the (p)ppGpp response, the effects of perturbing (p)ppGpp, and the questions remaining.

1.1. AIM OF THIS THESIS

Since the discovery of bacteria in the late 17th century by Van Leeuwenhoek, and the proof of their involvement in major diseases by Louis Pasteur and Robert Koch in the 18th century, bacterial physiology has known a steadily growing interest. Understanding how these microorganisms grow, survive and evolve is interesting as a fundamental understanding of one of the earliest, simplest and most abundant forms of life. It often allows us to derive some principles which are true for more complex organisms. This knowledge is also key in the fight against bacterial diseases, including the most deadly infectious disease: tuberculosis. Since their discovery in 1928, antibiotics have drastically reduced casualties from bacterial infections. However, their use and abuse have pushed bacteria to evolve and resist them, calling for the development of new antibiotics, which is facilitated by bacterial physiology fundamental knowledge. Finally, bacteria are used for the biosynthesis of molecules of interest for medicinal, industrial, or scientific applications. The development of such processes requires an extensive understanding of bacterial metabolism, to develop new applications and ensure efficient processes.

In light of this brief description of non-exhaustive but major interests in bacterial physiology, any bacterial regulatory mechanism with implications in growth, antibiotic resistance, or having analogous systems in more complex organisms is particularly valuable to study. The (p)ppGpp signaling system gathers all these characteristics. This might explain, since its discovery in 1969 [1], the booming interest in understanding how the synthesis of the two nucleotides ppGpp (Figure 1.1) and pppGpp is triggered, what they regulate, and the mechanisms behind it. The mechanism through which the (p)ppGpp response is activated,

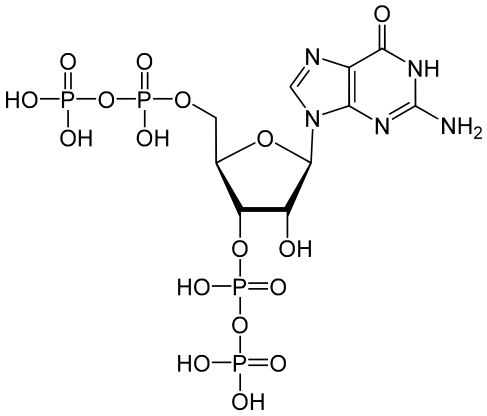


Figure 1.1: Chemical structure of guanosine tetraphosphate (ppGpp).

the triggers, and the signal have been has been extensively studied. While some targets repressed or activated by (p)ppGpp might still need to be discovered, numerous important ones have been identified, and the molecular mechanisms described. The system-level changes operated by (p)ppGpp, however, are not fully understood. Many models describing the (p)ppGpp signaling system cannot explain the effect on growth obtained when (p)ppGpp concentrations are perturbed, hinting towards an incomplete view of the effect (p)ppGpp on growth.

In the next section 1.2, we review (p)ppGpp's triggers, targets, and the mechanisms through which this system operates. In the following section 1.3 we discuss what is known and unknown about perturbing (p)ppGpp concentrations. Then, in the last section 1.4 of this chapter, we briefly summarize the relevant knowledge we reviewed and introduce the questions we attempt to answer in the coming chapters. In the next Chapter 2 we present novel results delineating the scope of (p)ppGpp regulation by titrating (p)ppGpp concentration. Then, in Chapter 3 we describe our attempts to model the (p)ppGpp regulatory system, its perturbations and the resulting effects on growth. Finally, in Chapter 4, we present a side story of how the budding yeast can survive desiccation.

1.2. THE (P)PPGPP SIGNALING SYSTEM: TRIGGERS, TARGETS AND MECHANISMS

When subject to various stresses or facing starvation, bacteria synthesize guanosine tetra- and pentaphosphate: (p)ppGpp. These signal molecules repress processes linked to fast growth and upregulate others that are required to face stress and increase metabolic flux to counter starvation. In this section, we review how ppGpp synthesis is triggered and the different aspects of bacterial physiology regulated by (p)ppGpp before, in the next section 1.3, discussing what is known about consequences of perturbing ppGpp concentrations and the questions that this current knowledge raises. Throughout this outline, we focus mainly on how the ppGpp[1] signaling system functions in *Escherichia coli* and the role it plays during exponential growth and carbon limitation.

1.2.1. HOW IS (P)PPGPP RESPONSE TRIGGERED?

RSH PROTEINS ARE SYNTHESIZING AND BREAKING DOWN (P)PPGPP

While it seems that ppGpp systems appear in almost every bacterial species, enzymes catalyzing synthesis and hydrolysis of (p)ppGpp vary across these species [2]. Apart from a few exceptions, all of them belong to the same group of proteins: the RelA-SpoT homolog (RSH) family (Figure 1.2). Though their properties may differ, these proteins have a similar structure comprising two domains. The N-terminal domain (NTD) contains the synthetase and the hydrolase domain. The synthetase domain catalyzes the addition of a pyrophosphate group to GDP or GTP, which respectively generates of guanosine tetra- (ppGpp) or pentaphosphate (pppGpp). The hydrolase domain catalyzes the opposite reaction from (p)ppGpp to GDP/GTP. The remaining of RSH proteins is composed of a C-terminal domain, whose function is not well established but speculated to regulate the NTD. The activity of the synthetase and hydrolase domains may vary in different

[1] In *E. coli* ppGpp is more abundant the pppGpp and has a higher regulatory activity

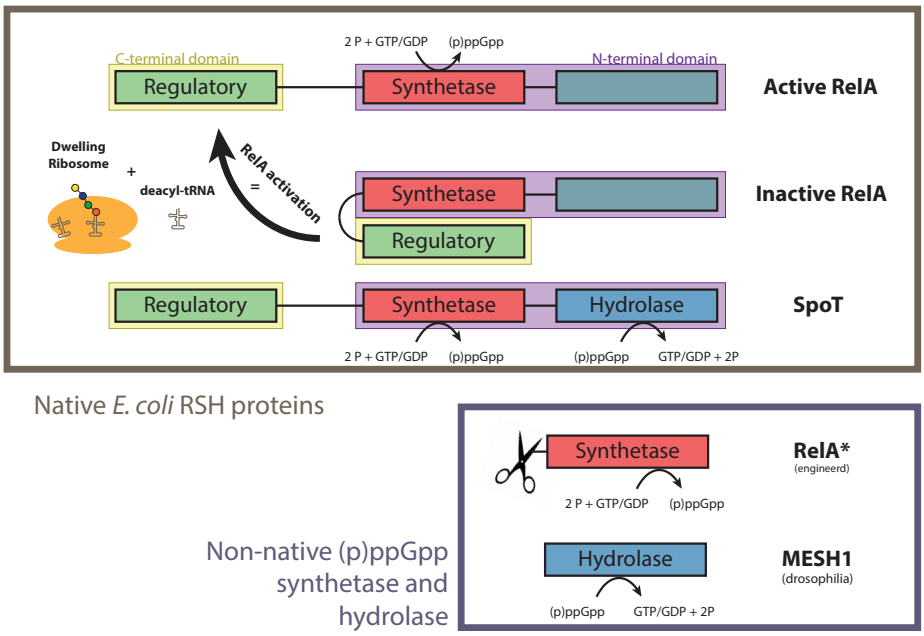


Figure 1.2: Non-native RSH proteins and truncated *E. coli* RSH can be used to perturb ppGpp concentrations. Simplified view the *E. coli* native RSH proteins as well as the non-native and synthetic ones used in Chapters 2 and 3, with the different domains and their identified fonctions; as well as the known mechanisms necessary to activate these proteins.

RSH proteins, some even losing one of the two. Many bacterial species have only one RSH protein responsible for hydrolysis and synthesis under different conditions. *E. coli* is endowed with two RSH proteins that gave its name to the family: RelA and SpoT. These two proteins have different synthesis and hydrolase activity and are responsible for (p)ppGpp response under different conditions. SpoT possesses both the ability to synthesize and hydrolyze (p)ppGpp, though its synthesis activity is decreased compared to other RSH proteins. It is responsible for all the (p)ppGpp hydrolysis in *E. coli*, as RelA lacks hydrolytic activity. SpoT responds to various nutrient stresses by reducing its hydrolytic activity [3]. The mechanisms through which SpoT responds to other nutritional stresses remain elusive. On the other hand, RelA responds specifically to amino acid starvation. Because of its relevance for carbon-limited exponential growth, we focus in the following paragraphs on RelA and go through what is known about the triggers for RelA-dependent (p)ppGpp synthesis, the various proposed models to explain how these triggers coordinate, and what (p)ppGpp is sensing through RelA at a system level.

DWELLING RIBOSOMES AND FREE TRNAS ARE RELA'S (P)PPGPP SYNTHESIS TRIGGERS

While RelA on its own can synthesize (p)ppGpp, it can be activated to reach a much higher activity. 50 years ago, and only four years after the discovery of (p)ppGpp, the two main triggers for this activation were identified. *In vitro* experiment then showed

that the following elements are necessary for full activation of RelA [4]: transfer RNAs (tRNAs) deprived of amino acids and dwelling ribosomes waiting for the right acyl-tRNA to resumes translation. Structural cryo-EM studies confirmed the interaction of RelA with these two triggers [5–7] and resolved the structure of the formed complex. The interpretation of this structure remains a debate with three competing views, which we review, along with arguments in their favor, in the next paragraph.

MECHANISTIC PICTURE OF (P)PPGPP SYNTHESIS BY RELA ACTIVATION REMAINS UNCLEAR

Following the identification of the two RelA triggers, the molecular mechanism for the activation of RelA remained unknown for almost three decades. Haseltine et al. [4] proposed that RelA might bind to stalled ribosomes that have a deacyl-tRNA in their acceptor site. Later, Wendrich et al. [8] confirmed the triggers and measured decreased affinity of RelA for the ribosomes after synthesis of (p)ppGpp thus suggesting a hopping mechanism where RelA binds to a ribosome, catalyzes (p)ppGpp synthesis, unbinds from the ribosome, and hops onto another stalled ribosome, providing it finds one. Due to another study [9] which claims that RelA enzymatic activity performs several cycles of (p)ppGpp synthesis off the ribosome, an activation model was proposed [10]. In this model, after being activated by the stalled ribosome, RelA is freed and stays in an active conformation long enough to synthesize several (p)ppGpp molecules before it needs to be re-activated. Recent cryo-EM studies [5–7] seem to indicate that RelA either binds first to dwelling ribosomes with a vacant acceptor site and is locked by deacyl-tRNA or forms a complex with deacyl-tRNA and then binds to a dwelling ribosome. The structure found in these studies shows that, when bound to the ribosome, RelA adopts an open conformation in which the C-terminal domain is separated from the enzymatically active N-terminal domain. A very recent study [11] seems to indicate that this conformation is stabilized by deacyl-tRNA and therefore proposes another model in which RelA is locked onto the ribosome by deacyl-tRNA and stays active until translation resumes. However, another recent study still defends the hopping mechanism, due to the link they find between translation elongation rate and ppGpp [12], which gives a system-level view of what ppGpp is sensing and is reviewed in the next paragraph.

SYSTEMIC LEVEL VIEWS: (P)PPGPP SENSES ELONGATION RATE

While a consensus still has to be found on the details of how dwelling ribosomes and deacyl-tRNA articulate to activate RelA, a recent study elegantly identified the signal that ppGpp is responding to through these triggers: a slow down in translation rate [12]. The authors measured translation elongation rate and ppGpp levels in three different scenarios: carbon-limited exponential growth, carbon downshift, and sub-lethal translation inhibition. They found that in all three scenarios, ppGpp concentration follows the same linear relationship with the inverse of translation elongation rate. They further identify that this inverse relation is equivalent to the proportionality of ppGpp to the time spent by ribosomes dwelling over the one spent translating. They hypothesize that this can be explained through a hopping mechanism leading to ppGpp synthesis proportional to the amount of dwelling ribosomes and predict that ppGpp breakdown is scaling with the amount of translating ribosomes. There is, however, no biochemical or structural evidence yet for this prediction. While further studies are necessary to explain the link between these system-level measurements and the precise mechanisms of RelA activation,

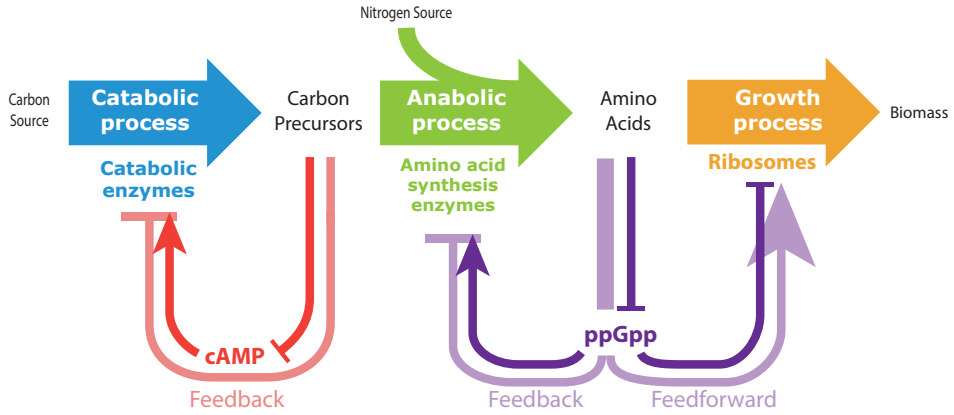


Figure 1.3: Guanosine tetraphospate (ppGpp) acts as a feedforward loop on ribosomes and a feedback one on amino acid synthesis enzymes while cyclic-AMP acts as a feedback loop on catabolic enzymes. Graphical summary of the role of signal molecules (p)ppGpp and cAMP in balancing growth-relevant fluxes. The horizontal arrows indicate the fluxes from carbon and nitrogen found in nutrients to new biomass. The red and purple arrows at the bottom indicate how the signalling system is thought to operate, with the darker arrows showing the actual interactions and the dimmer ones the overall effect on different fluxes.

it seems now clear that through interactions with dwelling ribosomes and deacyl-tRNA, ppGpp acts as a reporter of a slowdown of the elongation rate, the translation flux per ribosome.

1.2.2. WHAT DOES (P)PPGPP TARGET AND REGULATE?

Among *E. coli*'s regulators with the longest list of targets [13], (p)ppGpp plays a role in numerous key aspects of bacterial physiology from DNA replication to pathogenicity and resistance to antibiotics. We focus in this introduction on the role of (p)ppGpp in downregulating ribosome abundance, why this is relevant for steady state growth and nutritional transitions, the mechanisms through which this repression occurs, and the role of (p)ppGpp in activating amino acid biosynthetic genes. For other roles of (p)ppGpp in bacterial physiology, we refer the reader to the following insightful reviews: [3, 10, 14]. Finally, we briefly introduce the regulator responsible for balancing carbon input and amino acid biosynthesis, cyclic AMP, because of its implication in carbon limitation and our interest in investigating how it coordinates with (p)ppGpp.

(P)PPGPP DOWNREGULATES RIBOSOME ABUNDANCE

In the 1960s and early 1970s, it was shown that the fraction of bacterial mass constituted of RNA increases with growth rate when bacterial cells are grown in media with carbon sources of various qualities, supporting slower or faster growth rates. As ribosomes constitute most of bacterial RNA, it was deduced that the number of ribosomes increases with growth rate [15]. A decade later, it was shown that the amount of ribosomes follows a linear relationship with growth rate. Concentration of (p)ppGpp, on the other hand, anticorrelates with growth rate, and thus ribosome abundance [16]. When knocking

out the two (p)ppGpp synthetases RelA and SpoT [17], the ribosome content remains constant as a function of growth rate, establishing the causal link between (p)ppGpp and ribosomal downregulation at slower growth rate. These results were confirmed and given an interpretation later with the development of high-throughput proteomics and resource allocation theories [18–21]: as proteins constitute most of the bacterial biomass, a large number of the translation machineries producing them is required to support fast growth. When cells are provided with so-called poor nutrients, a greater fraction of the proteome must be allocated to converting the nutrients into precursors. The fraction of the proteome allocated to ribosomes must thus shrink and growth rate decreases. This change happens through (p)ppGpp signaling, which acts as a feedforward loop to equilibrate maximize metabolic and translational fluxes [12, 22, 23] to maximize steady-state growth rate. See Figure 1.3 for a graphical representation of this regulatory system. But how does (p)ppGpp respond when the nutritional conditions, and thus these fluxes, suddenly change? In the next paragraph, we review what is known about (p)ppGpp dynamics during nutrient transitions before explaining our understanding of these in the next one.

NUTRIENT TRANSITIONS LEAD TO FAST AND LARGE CHANGES IN (P)PPGPP LEVEL

When supplemented with higher quality nutrients or depleting a nutrient source in favor of a poorer one, the (p)ppGpp signaling system reacts by respectively decreasing and increasing (p)ppGpp concentration [25, 26]. These changes happen within seconds and are accompanied by an important overshoot before reaching the new (p)ppGpp concentration, as shown in Figure 1.4. When entering the starvation phase, (p)ppGpp also shows a drastic change: a large spike in concentration [26]. In Figure 1.4 we also show a schematic view of how growth rate varies during these transitions. During the transition to a richer carbon source, growth rate displays a fast increase followed by a slower one. During runout of a rich carbon source, growth rate sees a very drastic drop to almost no growth before being re-established to its final value. Growth dynamics are similar in the case of amino acid runout in a minimal media supplemented with glucose and knocking out the main (p)ppGpp synthetase RelA resulted in a much slower adaption [27], showing (p)ppGpp's role in ensuring rapid adaptation to this transition. (p)ppGpp is known to suppress ribosome expression but also to upregulate stress-response and ribosome hibernation operon, which might be relevant to these transitions. In the next paragraph, based on the known targets of (p)ppGpp we explain our understanding of the (p)ppGpp dynamics: the reasons behind such responses and what it might be doing.

(P)PPGPP LIKELY PLAYS A KEY ROLE IN NUTRIENT TRANSITIONS

In the case of a carbon upshift, the metabolic flux likely increases quickly due to the addition of a higher quality carbon source, which probably induces accumulation of amino acid and as a consequence, very few ribosomes dwelling. This would explain very low concentrations of (p)ppGpp after the shift, before they increase back to the value found in fast steady growth once the new proteome allocation with higher ribosome content has been found [28]. When running out of a good carbon source, (p)ppGpp spiking at high levels is likely due to a sudden drop in the metabolic flux causing amino acid starvation. Once more resources are allocated to metabolizing the poorer carbon source and the metabolic and translational fluxes re-equilibrated, the amino acid pool is likely

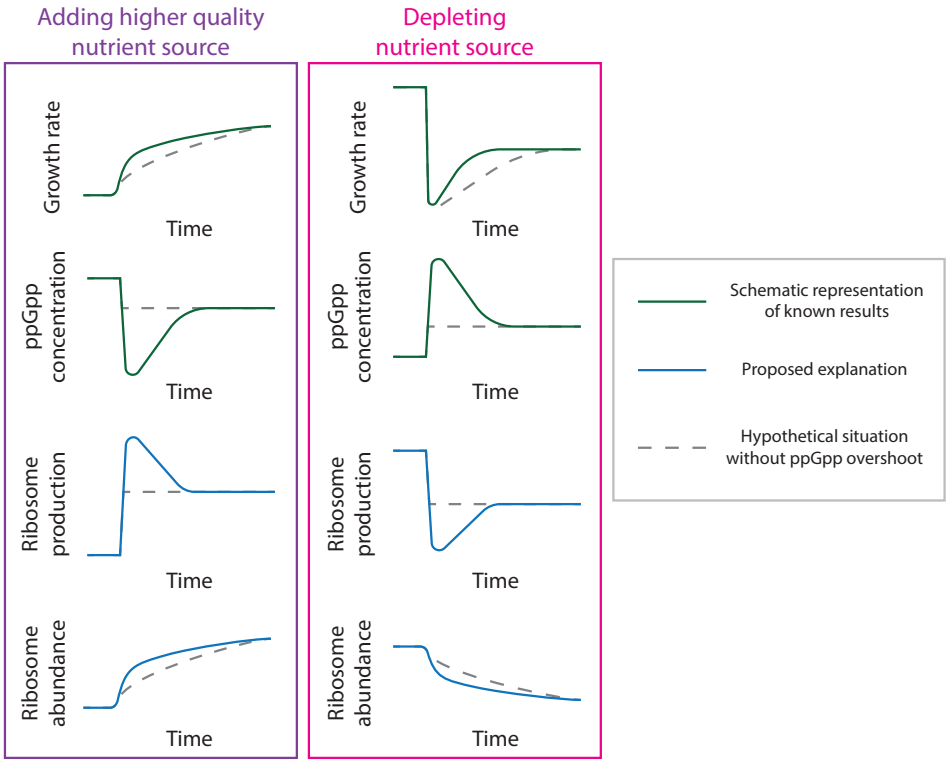


Figure 1.4: During nutrient transitions (p)ppGpp sees drastic changes and likely plays a key role because of its effect on ribosome abundance. Schematic view of the effect of nutrient upshift and runout on growth rate and corresponding ppGpp concentration as well as ribosome abundance and production. The timescale of the displayed behaviors displayed is of about 3 to 5 hours, while the first reaction to the transition can be observed within seconds. These graph show simplified interpreted behaviors and are built from data seen in various articles investigating these transitions [24–27] as well as our understanding of these phenomena. We refer the reader to the cited articles for precise data.

replenished. This is probably the cause of (p)ppGpp concentrations decreasing towards a slightly higher concentration than the initial one before the shift. It is probably due to this transient starvation following the nutrient downshift that growth is severely hampered before being restored with a new lower growth rate as shown in Figure 1.4 and measured by Erickson et al. [28]. It is thus likely because (p)ppGpp is sensing the amino acid pool rather than the quality of the nutrient source that (p)ppGpp concentrations overshoots during these nutrient transitions. From an evolutionary point of view, this overshoot is probably advantageous. Indeed, as depicted in Figure 1.4, it likely allows faster adaptation towards the new optimal proteome allocation. The spike in (p)ppGpp concentration appearing when facing nutritional challenges, by upregulating stress proteins and ribosome hibernation factors, might also promote long-term survival to allow bacteria to wait for better conditions. In the next paragraph, we quickly review the mechanism through

which (p)ppGpp is able to downregulate ribosomal operons and many others in *E. coli*.

(P)PPGPP TARGETS RNA POLYMERASE TO REPRESS TRANSCRIPTION

The mechanisms through which (p)ppGpp downregulates transcription of ribosomal operons vary across the bacterial kingdom. In *Bacillus subtilis* and many other bacteria, synthesis of (p)ppGpp sequesters GTP and downregulates the GTP biosynthesis pathway [29]. As GTP is required to initiate the transcription of ribosomal operons, the production of (p)ppGpp indirectly downregulates these operons. A similar mechanism was thus hypothesized for *E. coli*. It was later on found that rather than simply sequestering GTP, (p)ppGpp also has a direct effect on the activity of RNA polymerase, together with the protein DksA which strengthens the repressor effect of (p)ppGpp [30]. Two binding sites of (p)ppGpp were later identified by structural studies [31–33].

(P)PPGPP ACTIVATES AMINO ACID BIOSYNTHESIS GENES

Apart from its numerous repressive activities, (p)ppGpp is also involved in the activation of amino acid biosynthetic genes [13]. While the precise mechanism remains to be found, the latest studies seem to indicate a direct effect of (p)ppGpp on the transcription of these genes. It seems that the (p)ppGpp response attempts to replenish the amino acid pool by upregulating genes responsible for the synthesis of amino acids. Indeed, the lack of amino acids can be due to a insufficient enzymes synthesizing them. However, it can also be due to a insufficient carbon precursors, from which they are synthesized. When the quality of the carbon source is decreased, fewer carbon precursors are produced for the same amount of catabolic enzymes responsible for synthesizing these precursors. Biosynthesis of amino acid is then limited due to a lack of these precursors rather than insufficient biosynthetic enzymes. In this case, bacteria need to upregulate catabolic genes, not amino acid biosynthesis ones as (p)ppGpp is doing. Another regulator, cyclic AMP (cAMP), is involved in the regulation of the catabolic genes, responsible for carbon metabolism. We briefly discuss the cAMP signaling in the next paragraph.

ANOTHER REGULATOR, CAMP, BALANCES AMINO ACID SYNTHESIS AND CARBON INPUT

The metabolic flux leading to amino acid synthesis is in fact constituted of two fluxes: the catabolic flux, importing carbon and metabolizing it into precursors such as α-ketoacids, and the anabolic flux which imports nitrogen and synthesizes amino acids from these precursors. Nutrient conditions can impose different constraints on these fluxes. The regulator cAMP ensures that these fluxes are balanced by sensing the carbon precursor pool and repressing anabolic genes in favor of catabolic enzymes [20]. How cAMP is regulated and how it coordinates with (p)ppGpp is too poorly understood. By perturbing (p)ppGpp, we can force a sub-optimal translational flux, which will likely lead to the accumulation or depletion of some precursors and might lead to insights on these questions.

1.3. (P)PPGPP PERTURBATIONS: KNOWNS AND UNKNOWNS

Varying nutritional conditions, apart from impacting ribosome number, operates major changes in the bacterial proteome [37]. Which of these are operated by the global

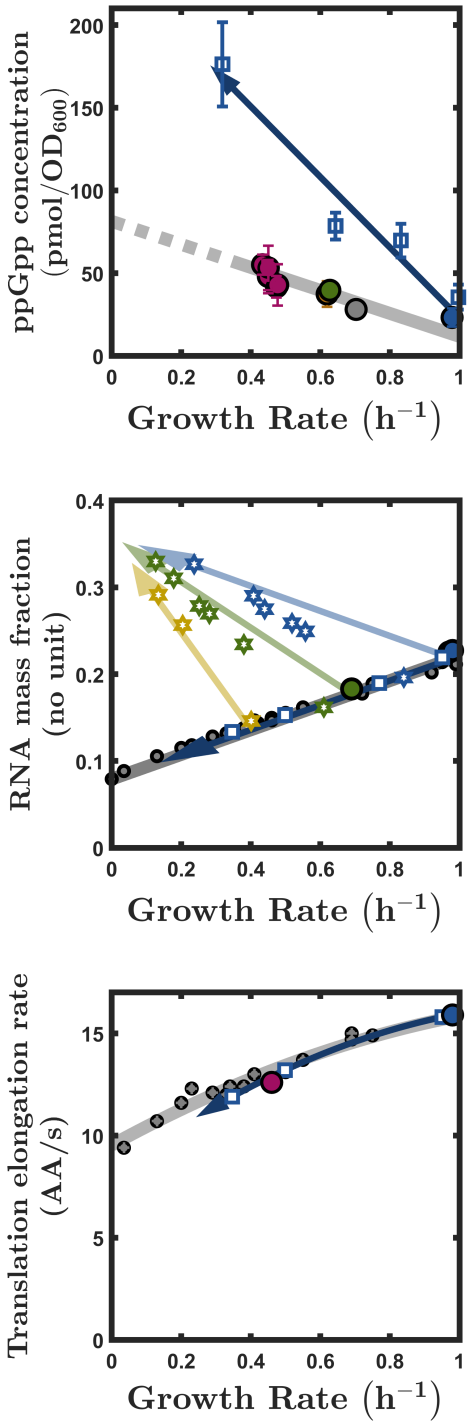


Figure 1.5: State of the art for ppGpp perturbation.
(top) ppGpp concentration results adapted from Noga et al. [34] showing divergence of ppGpp with regards to growth rate between carbon source quality limited and excess ppGpp conditions, inspiring this study. Circles show wild-type steady state ppGpp concentration vs. growth rate in minimal media supplemented with various carbon sources. Blue squares correspond to the same quantities measured in a strain engineered with RelA*, the catalytic domain of RelA, under control of a TetR-pTet system. This strain is grown in minimal media + glucose, which correspond in the wild-type to the fastest growth rate shown with blue circle. Each blue square corresponds to a different concentration of inducer, each yielding a different growth rate and ppGpp concentration. This result is reproduced and extended to more conditions in Chapter 2 Figure 2.1A.
(middle) RNA mass fraction results adapted and combined from [35, 36] displaying RNA mass fraction with regards to growth rate. Circles show wild-type bacteria growing in minimal media supplemented with carbon sources of various qualities and squares display excess ppGpp through various inductions of a IPTG-inducible RelA* strain growing in minimal media supplemented with glucose. Stars symbols display depleted ppGpp induced at different levels using a strain with drosophila ppGpp hydrolase MESH1 under the control of the IPTG-inducible Ptac promoter in three carbon sources, each represented with a different color.
(bottom) Translation elongation rate with regards to growth rate for wild-type bacteria growing in minimal media supplemented with various carbon sources (circles) and IPTG-inducible RelA* strain growin in minimal media with glucose (blue squares) at various inductions, results adapted and combined from from [35, 36].

regulator (p)ppGpp? Is (p)ppGpp fully or partially responsible for each of these changes? These questions, very hard to tackle by only changing nutrient conditions, can be better addressed through perturbation of the (p)ppGpp signaling system. One approach, discussed above in the case of growth rate control of ribosome abundance, is to knockout RelA or both RelA and SpoT. While this approach led to major advances in the understanding of the (p)ppGpp response, its limits are that the resulting strains exhibit heavily modified phenotype and it is often hard to determine if the differences with wild-type are due to the direct effect of (p)ppGpp or to an indirect effect resulting from adaptation to the perturbation. Another approach, which we focus on in this section, was used for the first time in 1995 by Schreiber et al. [38]. It consists in inserting genetic constructs to artificially increase or decrease (p)ppGpp concentrations. Concentrations of (p)ppGpp can be increased by the (p)ppGpp synthetic domain of RelA, denoted RelA* in this thesis, under the control of an inducible promoter. Because only the synthetic domain of RelA is expressed, the activity does not depend on activators such as dwelling ribosomes and transfer RNA, allowing to bring (p)ppGpp level above natural concentration. Oppositely, expression of drosophilian SpoT homolog MESH1, which only possesses a (p)ppGpp hydrolase domain [39], has been proven to successfully decrease (p)ppGpp concentrations [35, 40]. This approach revealed the effect of (p)ppGpp on cell size [40] and the way membrane synthesis coordinates with growth rate [34]. It has also been used to confirm the effect of (p)ppGpp on ribosome abundance [35], which we discuss in the next paragraph, before reviewing the questions that arise from these results, and the ones yet to be explored using such perturbations.

EXCESS AND DEPLETED (P)PPGPP REVEAL OPTIMALITY OF RIBOSOME ABUNDANCE

A recent study [35] showed that excess and depleted ppGpp respectively lead to higher and lower RNA mass fraction, thus ribosomal content, see Figure 1.5B for detailed results. In both cases, they found that the stronger they induced RelA* or MESH1, the further from wild-type ribosome abundance and the lower the growth rate. These results revealed the optimal behavior of wild-type ribosome content. In the case of depleted (p)ppGpp, more ribosomes are produced at the expense of metabolic enzymes, resulting in depletion of the amino acid pool and slower growth. In the case of excess (p)ppGpp through RelA* induction in minimal media supplemented with glucose, insufficient ribosomes limits growth by decreasing total translational flux. The RNA mass fraction follows the same linear relationship as under carbon limitation. The authors also measured the translation elongation rate, which is the translational flux through one ribosome. This flux slightly decreases following the same relationship both with carbon-limitation and excess (p)ppGpp. While a decrease of translation elongation in excess ppGpp might seem surprising at first as amino acids accumulate [41] due to slowed down growth, it can be explained with either insufficient transfer RNA caused by (p)ppGpp inhibition, direct repression of (p)ppGpp on translation, or a combination of both. This study confirmed and clarified the key role of (p)ppGpp in optimizing ribosome abundance for maximal steady-state growth rate.

EXTENT OF (P)PPGPP REGULATION IN STEADY STATE GROWTH IS UNCLEAR

(p)ppGpp perturbation has shown its potential to unravel new roles of the (p)ppGpp signalling system as well as precise quantitative laws arising from this system. However,

to this date, a broad study of the effects of perturbing (p)ppGpp concentrations on the bacterial proteome in different conditions is still lacking. Titrating (p)ppGpp and measuring proteins in an untargeted way in relations with (p)ppGpp concentrations would likely allow us to identify the role it plays in correlating with growth rate some key groups of proteins, such as the ones responsible for translation, catabolism and amino acid synthesis. It would better delineate the extent of ppGpp regulation. Indeed, some regulation have been attributed to (p)ppGpp from studies using nutritional differences or (p)ppGpp knockouts, which leaves the possibility that these proteins are responding to other signals affected by these.

SLOWING DOWN GROWTH WITH EXCESS PPGPP REQUIRES HIGH CONCENTRATIONS

Effect of artificially increasing ppGpp concentration [34] on growth was measured by the Bokinsky lab [34]. In comparison with the anticorrelation of ppGpp with growth rate under carbon limitation, slowing down growth with excess ppGpp in minimal media supplemented with glucose required much higher concentrations of ppGpp, as shown in Figure 1.5A. This result contrasts with results of RNA mass fraction and translation elongation Figure 1.5B-C for which induction of RelA* follows the same relationship with regards to growth rate, and is in that sense surprising. This preliminary result calls for a more extensive study of the effect of various artificially perturbed (p)ppGpp concentration on growth. In the next section, we summarize the state-of-the-art elements we have reviewed and introduce the questions we will try to answer throughout this thesis, which includes confirming and explaining the surprising result we just described.

1.4. CONCLUSION AND PERSPECTIVES

In the previous section, we reviewed the pre-existing knowledge relevant to this thesis. Briefly, we identified that the main triggers for (p)ppGpp synthesis are dwelling ribosomes and deacyl-tRNA, and that (p)ppGpp senses a slowdown in translation elongation rate through these triggers. We also reviewed the main effects of (p)ppGpp: downregulation of ribosome abundance and activation of genes responsible for amino acid biosynthesis. Finally, we reviewed the recent experiments using artificially tuned (p)ppGpp concentrations. We explained how using such perturbation of the (p)ppGpp signaling system revealed the optimality of ribosomal content in various conditions. Finally, we highlighted a preliminary result that seems to indicate that forcing down growth by artificially increasing ppGpp concentration requires much higher concentrations than the ones found in natural slow-growing conditions.

In the coming Chapter 2, we identify the scope of (p)ppGpp regulation during steady state growth. To do so, we titrate (p)ppGpp and measure the resulting nucleotide and protein concentrations. This allows us to distinguish: within the proteins and group of proteins that correlate with growth rate, which ones that are potentially regulated by (p)ppGpp? With this knowledge we also attempt to explain the reasons behind effects of perturbed (p)ppGpp concentrations on growth.

In Chapter 3, we investigate the reasons for the surprisingly high levels of ppGpp required to slow down growth, which we confirmed in Chapter 2 by extensively mapping the relationship between (p)ppGpp and growth rate when the ppGpp concentration is perturbed. We answer the following question: can higher ribosome saturation in

excess ppGpp explain this result by leading to faster growth for the same (p)ppGpp concentration? We then investigate what else could lead to a faster growth rate for the same (p)ppGpp concentration, using some of our other experimental results.

Finally, in Chapter 4 we investigate the necessary conditions for yeast to be able to survive in a desiccated state. We study ways to increase the fraction of cells surviving the process of freeze-drying. We also investigate why some desiccated cells are not able to resume their replicative life after being rehydrated.

REFERENCES

[1] M. Cashel and J. Gallant, *Two compounds implicated in the function of the rc gene of escherichia coli,* Nature **221**, 838 (1969).

[2] G. C. Atkinson, T. Tenson, and V. Hauryliuk, *The rela/spot homolog (rsh) superfamily: distribution and functional evolution of ppgpp synthetases and hydrolases across the tree of life,* PloS one **6**, e23479 (2011).

[3] K. Potrykus and M. Cashel, *(p) ppgpp: still magical?* Annu. Rev. Microbiol. **62**, 35 (2008).

[4] W. A. Haseltine and 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,* Proceedings of the National Academy of Sciences **70**, 1564 (1973).

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

[6] A. Brown, I. S. Fernández, Y. Gordiyenko, and V. Ramakrishnan, *Ribosome-dependent activation of stringent control,* Nature **534**, 277 (2016).

[7] S. Arenz, M. Abdelshahid, D. Sohmen, R. Payoe, A. L. Starosta, O. Berninghausen, V. Hauryliuk, R. Beckmann, and D. N. Wilson, *The stringent factor rela adopts an open conformation on the ribosome to stimulate ppgpp synthesis,* Nucleic acids research **44**, 6471 (2016).

[8] T. M. Wendrich, G. Blaha, D. N. Wilson, M. A. Marahiel, and K. H. Nierhaus, *Dissection of the mechanism for the stringent factor rela,* Molecular cell **10**, 779 (2002).

[9] B. P. English, V. Hauryliuk, A. Sanamrad, S. Tankov, N. H. Dekker, and J. Elf, *Single-molecule investigations of the stringent response machinery in living bacterial cells,* Proceedings of the National Academy of Sciences **108**, E365 (2011).

[10] V. Hauryliuk, G. C. Atkinson, K. S. Murakami, T. Tenson, and K. Gerdes, *Recent functional insights into the role of (p) ppgpp in bacterial physiology,* Nature Reviews Microbiology **13**, 298 (2015).

[11] H. Takada, M. Roghanian, J. Caballero-Montes, K. Van Nerom, S. Jimmy, P. Kudrin, F. Trebini, R. Murayama, G. Akanuma, A. Garcia-Pino, et al., *Ribosome association primes the stringent factor rel for trna-dependent locking in the a-site and activation of (p) ppgpp synthesis,* Nucleic acids research **49**, 444 (2021).

[12] C. Wu, R. Balakrishnan, N. Braniff, M. Mori, G. Manzanarez, Z. Zhang, and T. Hwa, *Cellular perception of growth rate and the mechanistic origin of bacterial growth law,* Proceedings of the National Academy of Sciences **119**, e2201585119 (2022).

[13] L. U. Magnusson, A. Farewell, and T. Nyström, *ppgpp: a global regulator in escherichia coli,* Trends in microbiology **13**, 236 (2005).

[14] S. E. Irving, N. R. Choudhury, and R. M. Corrigan, *The stringent response and physiological roles of (pp) pgpp in bacteria,* Nature Reviews Microbiology **19**, 256 (2021).

[15] P. P. Dennis, *Regulation of ribosomal and transfer ribonucleic acid synthesis in escherichia coli b/r,* Journal of Biological Chemistry **247**, 2842 (1972).

[16] J. Ryals, R. Little, and H. Bremer, *Control of rrna and trna syntheses in escherichia coli by guanosine tetraphosphate.* Journal of bacteriology **151**, 1261 (1982).

[17] K. Potrykus, H. Murphy, N. Philippe, and M. Cashel, *ppgpp is the major source of growth rate control in e. coli,* Environmental microbiology **13**, 563 (2011).

[18] D. Molenaar, R. Van Berlo, D. De Ridder, and B. Teusink, *Shifts in growth strategies reflect tradeoffs in cellular economics,* Molecular systems biology **5**, 323 (2009).

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

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

[21] S. Hui, J. M. Silverman, S. S. Chen, D. W. Erickson, M. Basan, J. Wang, T. Hwa, and J. R. Williamson, *Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria,* Molecular systems biology **11**, 784 (2015).

[22] M. Scott and T. Hwa, *Shaping bacterial gene expression by physiological and proteome allocation constraints,* Nature Reviews Microbiology , 1 (2022).

[23] G. Chure and J. Cremer, *An optimal regulation of fluxes dictates microbial growth in and out of steady-state,* bioRxiv (2022).

[24] Y. K. Kohanim, D. Levi, G. Jona, B. D. Towbin, A. Bren, and U. Alon, *A bacterial growth law out of steady state,* Cell reports **23**, 2891 (2018).

[25] J. D. Friesen, N. Fiil, and K. Von Meyenburg, *Synthesis and turnover of basal level guanosine tetraphosphate in escherichia coli.* Journal of Biological Chemistry **250**, 304 (1975).

[26] H. D. Murray, D. A. Schneider, and R. L. Gourse, *Control of rrna expression by small molecules is dynamic and nonredundant,* Molecular cell **12**, 125 (2003).

[27] M. Zhu and X. Dai, *Stringent response ensures the timely adaptation of bacterial growth to nutrient downshift,* Nature Communications **14**, 467 (2023).

[28] D. W. Erickson, S. J. Schink, V. Patsalo, J. R. Williamson, U. Gerland, and T. Hwa, *A global resource allocation strategy governs growth transition kinetics of escherichia coli,* Nature **551**, 119 (2017).

[29] L. Krásnỳ and R. L. Gourse, *An alternative strategy for bacterial ribosome synthesis: Bacillus subtilis rrna transcription regulation,* The EMBO journal **23**, 4473 (2004).

[30] B. J. Paul, M. M. Barker, W. Ross, D. A. Schneider, C. Webb, J. W. Foster, and R. L. Gourse, *Dksa: a critical component of the transcription initiation machinery that potentiates the regulation of rrna promoters by ppgpp and the initiating ntp,* Cell **118**, 311 (2004).

[31] J. J. Lemke, P. Sanchez-Vazquez, H. L. Burgos, G. Hedberg, W. Ross, and R. L. Gourse, *Direct regulation of escherichia coli ribosomal protein promoters by the transcription factors ppgpp and dksa,* Proceedings of the National Academy of Sciences **108**, 5712 (2011).

[32] Y. Zuo, Y. Wang, and T. A. Steitz, *The mechanism of e. coli rna polymerase regulation by ppgpp is suggested by the structure of their complex,* Molecular cell **50**, 430 (2013).

[33] W. Ross, C. E. Vrentas, P. Sanchez-Vazquez, T. Gaal, and R. L. Gourse, *The magic spot: a ppgpp binding site on e. coli rna polymerase responsible for regulation of transcription initiation,* Molecular cell **50**, 420 (2013).

[34] M. J. Noga, F. Büke, N. J. van den Broek, N. C. Imholz, N. Scherer, F. Yang, and G. Bokinsky, *Posttranslational control of plsb is sufficient to coordinate membrane synthesis with growth in escherichia coli,* MBio **11**, e02703 (2020).

[35] M. Zhu and X. Dai, *Growth suppression by altered (p) ppgpp levels results from non-optimal resource allocation in escherichia coli,* Nucleic acids research **47**, 4684 (2019).

[36] X. Dai, M. Zhu, M. Warren, R. Balakrishnan, V. Patsalo, H. Okano, J. R. Williamson, K. Fredrick, Y.-P. Wang, and T. Hwa, *Reduction of translating ribosomes enables escherichia coli to maintain elongation rates during slow growth,* Nature microbiology **2**, 1 (2016).

[37] 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, et al., *From coarse to fine: the absolute escherichia coli proteome under diverse growth conditions,* Molecular systems biology **17**, e9536 (2021).

[38] G. Schreiber, E. Z. Ron, and G. Glaser, *ppgpp-mediated regulation of dna replication and cell division in escherichia coli,* Current microbiology **30**, 27 (1995).

[39] D. Sun, G. Lee, J. H. Lee, H.-Y. Kim, H.-W. Rhee, S.-Y. Park, K.-J. Kim, Y. Kim, B. Y. Kim, J.-I. Hong, et al., *A metazoan ortholog of spot hydrolyzes ppgpp and functions in starvation responses,* Nature structural & molecular biology **17**, 1188 (2010).

[40] F. Büke, J. Grilli, M. C. Lagomarsino, G. Bokinsky, and S. J. Tans, *ppgpp is a bacterial cell size regulator,* Current Biology **32**, 870 (2022).

[41] G. Bokinsky, E. E. Baidoo, S. Akella, H. Burd, D. Weaver, J. Alonso-Gutierrez, H. García-Martín, T. S. Lee, and J. D. Keasling, *Hipa-triggered growth arrest and β-lactam tolerance in escherichia coli are mediated by rela-dependent ppgpp synthesis,* Journal of bacteriology **195**, 3173 (2013).

2

REGULATORY SCOPE OF THE MAGIC SPOT[1] DURING STEADY-STATE GROWTH

Milan LACASSIN, Maxime DEN RIDDER, Martin PABST, Yaroslav BLANTER, Gregory BOKINSKY

In the fields of observation chance favors only the prepared mind.

Louis Pasteur

[1] The nucleotide ppGpp was initially named "magic spot" due to its unidentified nature [1]; this name stayed in use due to its unusual number of targets and global regulatory nature [2].

2.1. ABSTRACT

The ppGpp regulatory system, through the proof of its involvement in downregulating ribosomal content and the discovery of the broad impact it has on the proteome, has seen growing interest since its discovery. However, the changes it causes on the proteome during steady-state growth have only been investigated through ppGpp synthase knockouts or nutritional changes, making it impossible to precisely delineate the scope of ppGpp regulation. By artificially titrating ppGpp and measuring nucleotide and protein concentrations, we confirm the key role of ppGpp is downregulating most of proteins responsible for translation. We also show that most other proteins are not regulated by ppGpp, even though some of them respond to growth rate. We identify the ones that do see ppGpp regulation. We find that, while ppGpp has been shown necessary to express some amino acid pathways, the concentration of these enzymes does not depend on ppGpp; and that the abundance of TCA cycle enzymes anticorrelates with growth rate independently of ppGpp. While setting the boundaries of ppGpp regulation, our results highlight the potential of titrating concentration of regulatory molecules, including the ones that regulate the proteins outside the reach of ppGpp.

2.2. INTRODUCTION

Bacterial growth is amazingly adaptable. As a phylogenetic kingdom, bacteria are able to grow using a wide variety of chemicals, from carbon dioxide to cellulosic biomass to plastic. This is due to the incredible versatility of microbial catabolic pathways that convert environmental nutrients into metabolic precursors to cellular biomass. Less appreciated is the ability of cells to precisely tune the rest of the cell to match conditions. The use of global approaches to study cellular physiology provides insight into bacterial adaptations and metabolic flexibility and their consequences on the cellular growth rate.

Cellular growth is simply the production of biomass, which includes proteins, RNA, DNA, and the cell envelope. As proteins contribute most to *Escherichia coli* biomass (~ 50%), cellular growth is directly proportional to the rate at which proteins can be produced from amino acids. Environments lacking amino acids require the cell to produce those from basic metabolites for growth. Cells do so by expressing amino acid biosynthesis pathways. However, some nutrients are converted less efficiently into essential metabolites. Cells respond to poor nutritional conditions by increasing the abundance of catabolic proteins, which increases the overall rate at which essential metabolites are produced. These adaptations enable many bacteria to grow on a wide variety of substrates, broadening their ecological versatility.

However, such adaptations come at a cost. Because cells have a limited bandwidth for protein synthesis, increasing the abundance of catabolic and anabolic proteins reduces the abundance of proteins dedicated to biomass synthesis. Conversely, nutrient-rich environments decrease the need for catabolic and anabolic pathways, which liberates protein synthesis bandwidth for biomass producing pathways. Hence, the overall rate of biomass production (and thus growth) is faster in richer media because they require fewer catabolic and anabolic proteins, and thus allow a higher abundance of ribosomes [3–5].

Understanding how *E. coli* allocates the fractions of the proteome among nutrient processing, building block synthesis, and biomass production is a topic of intense study. Much attention is focused upon guanosine tetraphosphate (ppGpp), an established regulator of ribosome abundance. ppGpp has long been understood as a signal indicating acute starvation and stress, during which it achieves high concentrations. However, ppGpp is also present during steady-state growth in smaller amounts, even without stress or starvation. So-called "basal" ppGpp concentrations vary inversely with steady-state growth rate [6–8]. Studies in which the basal ppGpp concentration is artificially raised or lowered using ppGpp synthesis and degradation enzymes demonstrated that these perturbations strongly influence ribosome abundance [5, 9]. These studies further found that any change in ribosome abundance decreases the growth rate. This result strongly indicates that the concentration of ribosomes set by basal ppGpp is optimal for *E. coli* growth.

Precisely how basal ppGpp is set by growth conditions via the enzymes RelA and SpoT is poorly understood. The main ppGpp synthesis enzyme RelA specifically detects tRNA that lacks an amino acid and ribosomes waiting for an acyl-tRNA to resume translation [10]. ppGpp thus likely varies in response to the intracellular amino acid supply. In doing so, ppGpp balances protein synthesis with the catabolic and anabolic pathways that ultimately supply the ribosomes with amino acids. By tuning ribosome abundance in response to nutrient availability [11], ppGpp establishes the first growth law [8] stipulating that steady-state growth rate scales with ribosome abundance [5, 6, 8, 12].

Given the importance of ppGpp for optimizing resource allocation, identifying the specific genes that respond to basal ppGpp is essential for understanding growth rate maximization. The tight correlation between basal ppGpp concentrations and steady-state growth rate complicates the identification of genes that are directly regulated by basal ppGpp. This challenge is compounded because ppGpp is known to control rRNA expression, and thus a major portion of the total cellular proteome. Previously, Traxler et al. compared the transcriptome of wild-type and ppGpp-deficient strains to identify ppGpp-induced genes [13]. This important study established the broad reach of ppGpp and how various subsets of regulatory targets are activated by basal ppGpp or starvation-level ppGpp: genes for amino acid biosynthesis, DNA and fatty acid synthesis, stress responses, cell cycle regulation, and stationary phase adaptation are influenced by ppGpp. Given its broad array of regulatory targets, it is possible that ppGpp plays additional roles beyond its effect on ribosome content during steady-state growth. However, these experiments used nutrient exhaustion to indirectly activate ppGpp synthesis, thus introducing confounding factors in that some genes may be triggered by nutrient starvation. Furthermore, as ribosomes comprise 25-50% of the cell biomass, ppGpp-driven variations in ribosome abundance could indirectly affect proteome sectors that are not ppGpp regulated.

Here, we determine the contributions of incremental ppGpp variations on proteome allocation by artificially perturbing ppGpp during steady-state growth. This approach disentangles the effects of ppGpp from other influences such as nutrient and metabolic conditions. We use untargeted proteomics to quantify the abundance of cellular proteins in each condition explored. We obtain a global view of the influence of ppGpp variations on proteome allocation in cellular growth or catabolic pathways. We find that while ribosome abundance is highly sensitive to basal ppGpp, most sectors are either

inconsistently affected by ppGpp, or appear surprisingly indifferent. We identify the specific consequences of ppGpp dysregulation on the balance between translation and catabolism. Our results better delineate the specific roles of ppGpp and other cellular signals in adapting physiology to the environment.

2.3. RESULTS

2.3.1. NATURAL PPGPP LEVELS ARE OPTIMIZED TO MAXIMIZE GROWTH RATE

We first measured concentrations of basal ppGpp within batch cultures of wild-type (WT) *Escherichia coli* NCM3722 during growth on minimal growth media. Nucleotide concentrations were quantified using tandem liquid chromatography mass spectrometry (LCMS). In four selected conditions, we synthetically titrated ppGpp by using an inducible promoter (Ptet) to vary expression of either the catalytic domain of the ppGpp synthesis enzyme RelA (RelA*) or the ppGpp hydrolase Mesh1. Growth was followed using regular measurements of culture turbidity by an automated absorbance measurement system. Samples were taken at 0.35 OD after growth rate stabilized (usually 3 to 5 hours after inducing for RelA* and 5 to 8 for Mesh1).

Consistent with many long-standing studies, basal ppGpp concentrations correlate inversely with growth rate (Figure 2.1A). As expected, expressing RelA* and Mesh1 respectively increases and decreases ppGpp. It also decreases the growth rate progressively with the magnitude of the ppGpp titration. This indicates that the naturally-occurring basal ppGpp concentrations are optimized to maximize growth rate in each condition examined.

PPGPP TITRATION EXERTS VARYING EFFECTS ON CAMP

We quantified the catabolism global regulator cAMP in parallel with ppGpp. In agreement with previous observations, basal cAMP is also inversely correlated with growth rate (Figure 2.1B), varying from a low concentration (~ 2pmol/OD) in glucose medium to ~ 22pmol/OD in glutamate medium. Synthetic ppGpp titrations exerted different effects on basal cAMP depending upon the carbon source. Elevating ppGpp in glucose and succinate medium did not significantly affect cAMP, while cAMP increased in glycerol and decreased in acetate ($p<0.05$)[2]. The few measurements obtained suggest that ppGpp depletion also decreases cAMP in glucose, succinate, and glycerol medium, while not significantly affecting cAMP in acetate medium ($p>0.05$)[3].

2.3.2. UNTARGETED PROTEOMICS IDENTIFIES PROTEIN SECTORS THAT COVARY WITH GROWTH RATE

We quantified *E. coli* proteins during steady-state growth in three carbon sources (glucose, succinate, and acetate) using the same untargeted LCMS method as den Ridder et al. [14]. Very briefly, proteins we extracted from collected culture samples and digested into peptides before being separated and analyzed with LCMS. Next, we quantified proteins after inducing ppGpp overproduction and depletion in each media, obtaining 9 condi-

[2] Here, significance testing was performed by applying an unpaired t-test on all wild-type samples against the two most extreme sets of technical replicates in terms of growth rate gathered.

[3] Significance is calculated in the same way as previously with the exception of succinate and glycerol for which only one set of technical replicates was measured and thus used here.

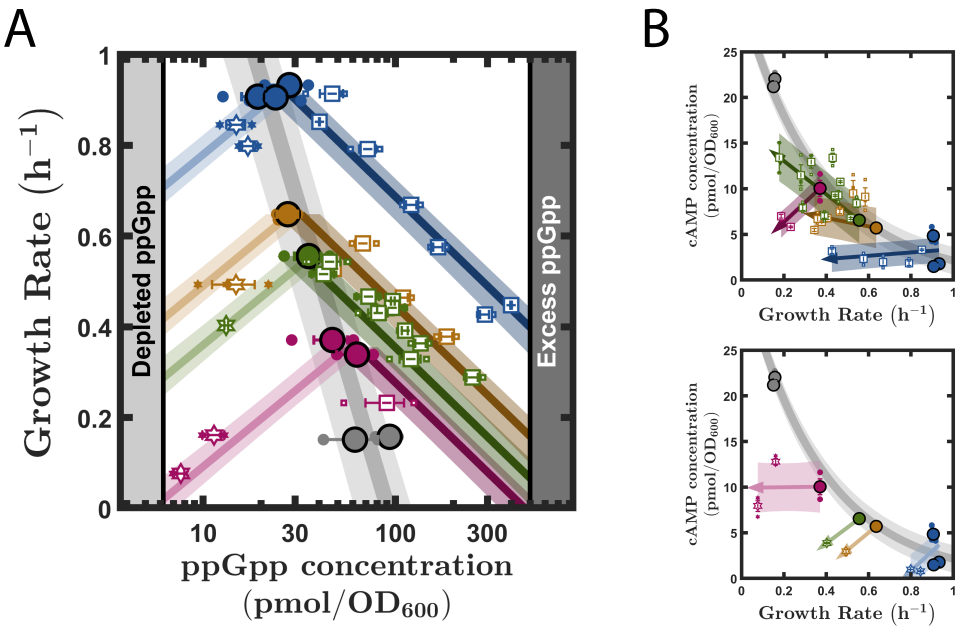


Figure 2.1: Artificially perturbing ppGpp concentrations reveals their optimality while resulting in various effects on cAMP concentrations. (A) Results of ppGpp concentrations and growth rates. (B) Results of cyclic-AMP concentrations as a function of growth rate in various carbon source conditions and under excess (top) and depleted (bottom) ppGpp. Circles represent WT samples, stars represent depleted ppGpp (Mesh1) and squares correspond to excess ppGpp (RelA*). Each color corresponds to a different carbon source: blue for glucose, ochre for succinate, green for glycerol, and magenta for acetate.

tions. To identify proteins that correlate or anti-correlate with growth in the absence of artificial ppGpp variation, we determined the concentration ratio of each protein between glucose and acetate medium. The ratios were logarithmically plotted against statistical significance p to obtain volcano plots. Among the 1185 proteins quantified, 211 show a positive and statistically significant correlation with growth rate ($p<0.05$) whereas 107 show a negative correlation with growth rate (Figure 2.2). Due to limited data and our focus on the global effects of ppGpp, we disregarded the significance criterion, unlike most studies that focus on individual proteins. Instead, we set a threshold for growth rate correlation (increasing or decreasing from acetate medium by at least 25%), which leads to similar results: 277 proteins increase by at least 25% from acetate to glucose medium (are positively correlated with growth rate), whereas 168 proteins decrease in abundance by at least 25% (are negatively correlated with growth rate). Remarkably, the majority of proteins (740, or 62% of all proteins quantified) vary by less than 25% across the conditions sampled (are growth rate-insensitive).

PROTEINS INVOLVED IN TRANSLATION POSITIVELY CORRELATE WITH GROWTH RATE

To obtain a functional understanding of growth rate regulation, we classified proteins according to function using an adapted version of Kyoto Encyclopedia of Genes and Genomes (KEGG) [15]. By grouping proteins according to their functions into sectors,

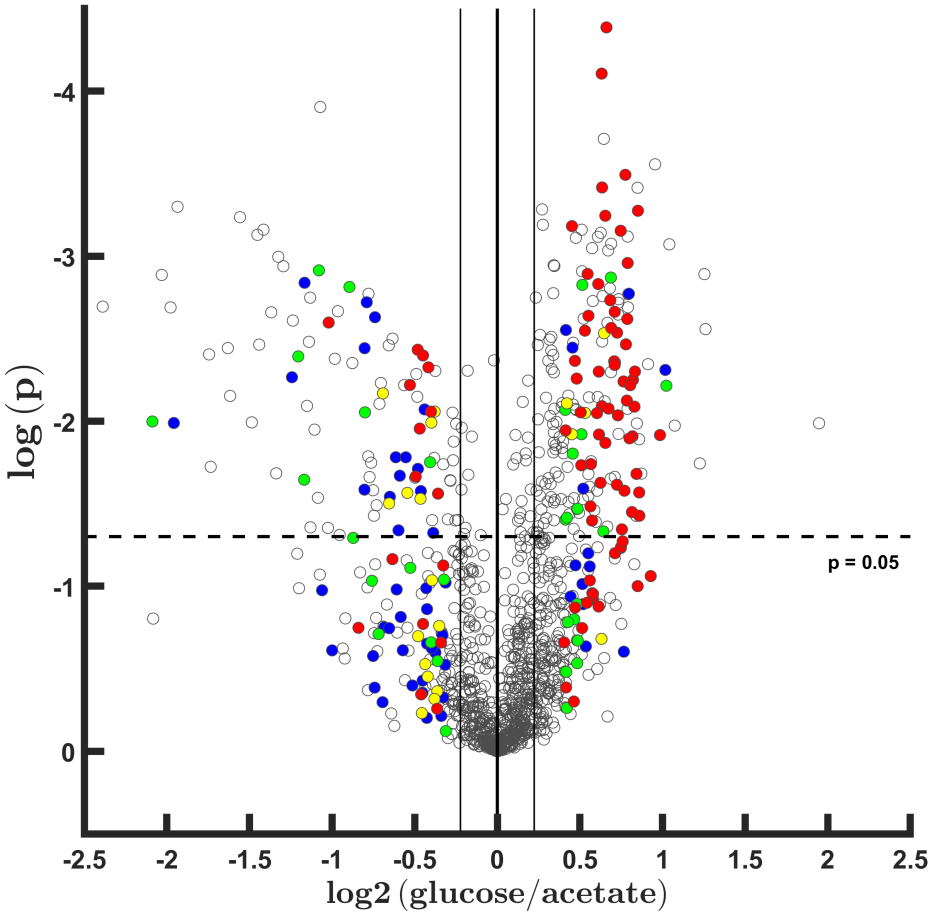


Figure 2.2: Only a minority of proteins see their level correlate with growth rate, of which only some are regulated by ppGpp. Volcano plot of the variation in concentration from glucose to acetate for all 1185 measured proteins: x-axis shows this variation in log scale while y-axis displays the statistical significance of this change. Colors are relative to the classification described in section 2.3.3 and Figure 2.8. White: no consistent influence of ppGpp. Blue: influenced by RelA* similarly as with decreased growth rate. Green: influenced by MESH1 similarly as with increased growth rate. Yellow: RelA* influence is consistent. Red: MESH1 influence is consistent. Thin vertical lines correspond to 20% decrease and 25% increase from glucose to acetate.

we estimate the proteome allocation into broad cellular functions such as translation, catabolism, and anabolism. Liebermeister et al. [16] adapted this classification such that each protein is assigned to a single sector for each of the three hierarchical levels. We use intensity-based absolute quantification (iBAQ) spectral counting [17, 18] to estimate the fraction of proteome mass represented by each protein. Summing spectral counts of all proteins comprising a sector leads to an estimate of the sector mass. The mass of these sectors in glucose medium is displayed per level in a graphical way in Figure 2.3.

For the two largest top level sectors of the adapted KEGG classification, we provide in Figure 2.4-2.5 plots showing quantification of total proteome mass. Figure 2.4 displays these results for the 3 sectors composing "Genetic Information Processing": “Folding, Sorting, and Degradation,” "Translation," and "Transcription". Of these 3 sectors, growth rate most strongly affects the Translation sector, which varies from 20% to 30% of the total proteome spectral counts from acetate to glucose medium. Of the Level 3 sectors within the Translation sector, the Ribosome sector (13%-21%) and Translation Cofactor sector (5.5-7.5%) are most sensitive to growth rate.

As sector-averaged trends obscure the responses of individual proteins, we use volcano plots to depict how the abundance of each protein varies between acetate and glucose. A volcano plot of the Translation sector reveals that the vast majority of proteins increase in parallel with growth rate. Of the 54 ribosome proteins, all except 2 (stationary phase-associated) increase with growth rate, by average of 55% from acetate to glucose medium. All 13 proteins classified as translation cofactors increase with growth rate. This is consistent with translation and growth rate increasing in parallel.

GLYCOLYSIS SECTOR ENZYMES EXHIBIT MIXED RESPONSES TO GROWTH RATE DEPENDING UPON GLYCOLYSIS/GLUCONEOGENESIS CONDITIONS

Next, we looked at the total mass fraction of metabolic proteins through the Level 1 proteome sector "Metabolism” (Figure 2.5-2.6). Of the two most abundant Level 2 sectors within the Metabolism sector (Central Carbon Metabolism and Biosynthesis), only Central Carbon Metabolism increases with carbon limitation (Figure 2.5). To discern trends within the Central Carbon Metabolism sector, we generated volcano plots for the four Level 3 sectors that represent most of the Central Carbon Metabolism sector: "Glycolysis"/Gluconeogenesis, “Pentose phosphate pathway” and “Other central metabolic enzymes”. The proteome fractions of these sectors only slightly vary. However, of the 40 proteins within the Glycolysis sector many are highly sensitive to growth rate. 9 proteins highly expressed in acetate media perform roles in gluconeogenesis (e.g. PckA, PpsA, FbaB). Conversely, the 11 proteins that increase in glucose medium specialize in glycolysis (e.g. FbaA and PfkA). These observations are consistent with gluconeogenesis occurring in acetate and succinate media, while growth on glucose requires glycolysis enzymes. Thus, proteome allocation within the Glycolysis sector reflects a switch between glycolysis and gluconeogenesis, while the total proteome allocation to the overall sector remains stable.

TCA CYCLE SECTOR PROTEINS NEGATIVELY CORRELATE WITH GROWTH RATE

The Level 3 sector “TCA and anapleurotic enzymes” is the most growth rate-sensitive of the sectors within Central Carbon Metabolism: it decreases from 14% of total proteome in acetate to 6% in glucose medium (Figure 2.5). 14 out of 19 of its proteins decrease by more

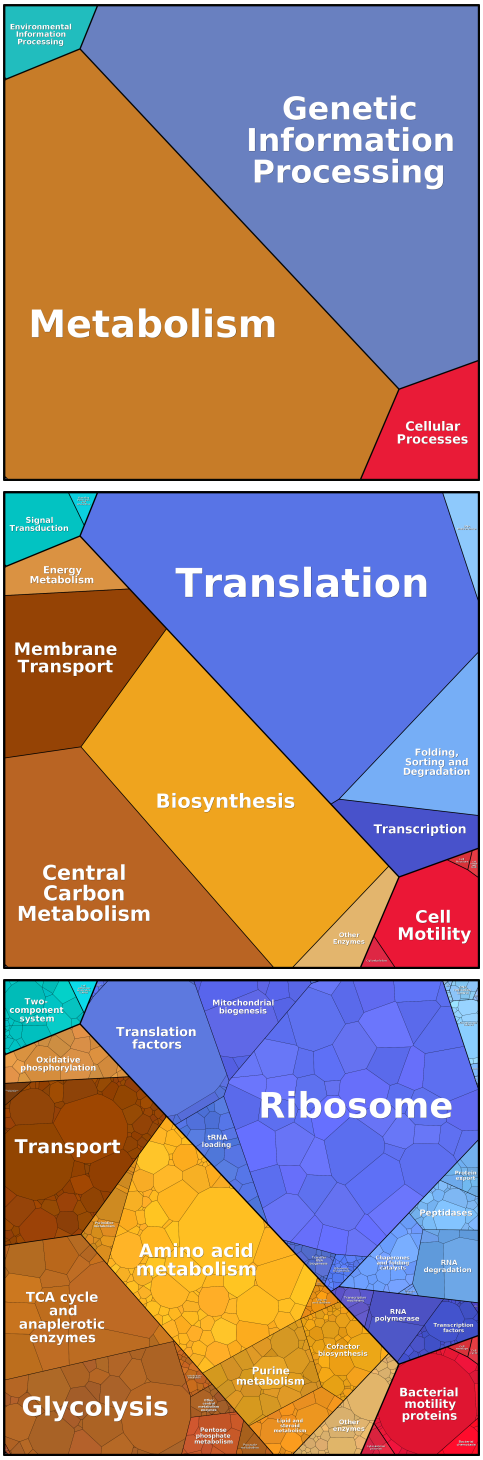


Figure 2.3: Proteins can be gathered in a 3-level hierarchical classification. These proteomaps show relative mass of proteome sectors (represented with polygonal area size) at the three hierarchical levels of adapted KEGG classification proposed by by Liebmeister et al. [16]. Axes do not have any meaning. These maps were generated from average spectral count of our wild-type samples growing in glucose medium using the platform offered by Liebmeister et al. [16].

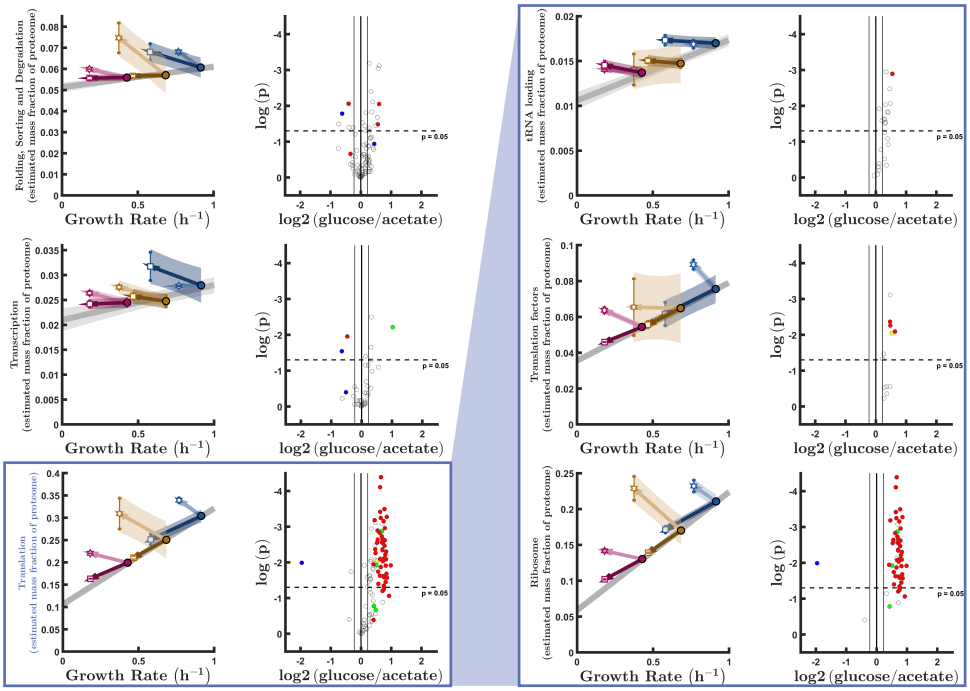


Figure 2.4: Ribosomes and translation cofactors are in large parts regulated by ppGpp.
Total proteome mass fraction and single proteins volcano plots for the three largest sectors constituting the "Genetic Information Process" sector (left) as well the three largest ones composing its sub-sector "Translation" (right). Colored frame corresponds to proteomap displayed in Figure 2.3 and depicts the hierarchical relation between these sectors. Proteome mass plots follow the same colors and symbols as Figure 2.1: Circles represent WT samples, stars represent depleted ppGpp (Mesh1) and squares correspond to excess ppGpp (RelA*). Each color corresponds to a different carbon source: blue for glucose, ochre for succinate, and magenta for acetate. Volcano plots follow the same conventions as Figure 2.2: each color corresponds to a different level of our classification identifying basal ppGpp-regulated proteins.

than 20% from acetate to glucose medium. In particular, enzymes of the glyoxylate shunt (AcnA, AcnB, GltA, Mdh) accumulate to 1% of the total proteome in acetate medium. Neglecting the anapleurotic enzyme Ppc, which replenishes the TCA cycle during growth on minimal medium, abundance of individual proteins within this sector decreases an average of 40% from acetate to glucose media.

AMINO ACID BIOSYNTHESIS PATHWAYS POSITIVELY CORRELATE WITH GROWTH RATE

The “Biosynthesis” sector includes anabolic pathways for synthesis of amino acids, nucleotides, cofactors, and membrane components. Oppositely to central carbon metabolism, this sector increases as a fraction of proteome from 14 to 18% from acetate to glucose media (Figure 2.6). 78 out of its 267 proteins increase by more than 20% with growth rate, while only 12 decrease with growth rate, 3 of which are misclassified fatty acid degradation enzymes. Of the 78 proteins that increase from acetate to glucose media, 47 are from amino acid biosynthesis pathways. They represent approximately 10% of

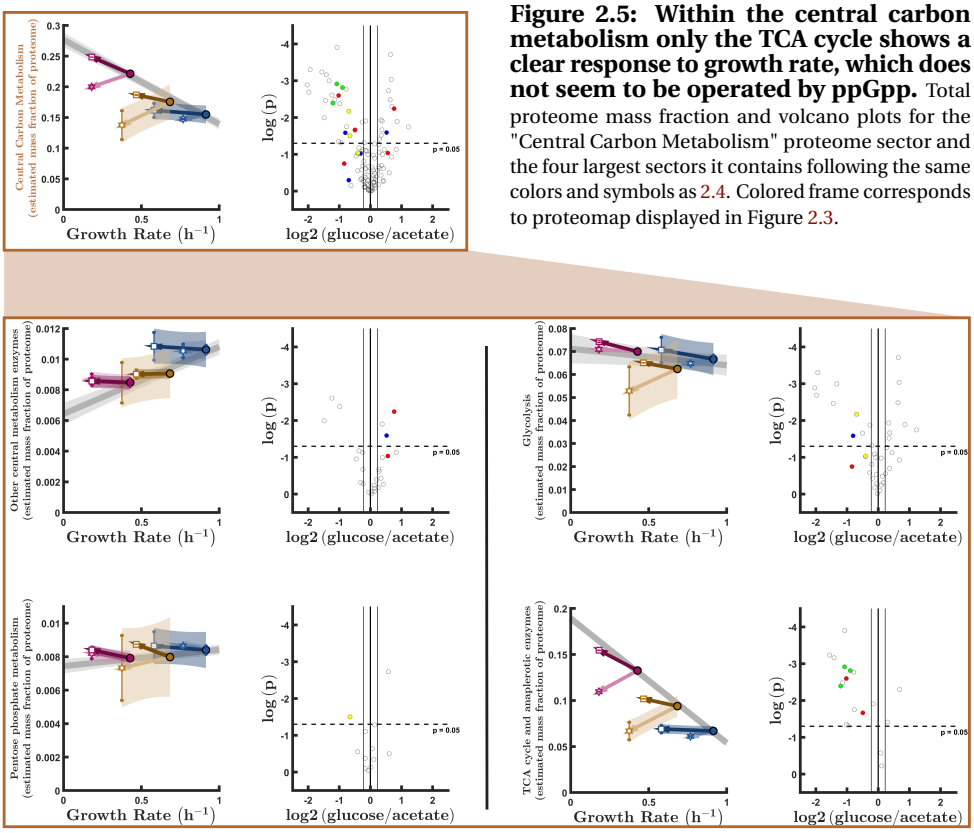


Figure 2.5: Within the central carbon metabolism only the TCA cycle shows a clear response to growth rate, which does not seem to be operated by ppGpp. Total proteome mass fraction and volcano plots for the "Central Carbon Metabolism" proteome sector and the four largest sectors it contains following the same colors and symbols as 2.4. Colored frame corresponds to proteomap displayed in Figure 2.3.

total proteome. A volcano plot (Figure 2.6, bottom right) of proteins from the Level 3 "Amino acid metabolism" sector indicates that most enzymes within this sector increase in parallel with growth rate, with the interesting exception of branched-chain amino acid biosynthesis pathway enzymes. We speculate that correlated expression of biosynthesis genes with growth rate reflects increasing demand for de novo amino acid biosynthesis to supply translation.

ABUNDANT OUTER MEMBRANE CHANNEL PROTEINS VARY WITH GROWTH RATE

While most of the remaining Level 2 sectors did not vary substantially with growth rate, several individual proteins that individually contribute more than 0.5% and comprise a significant fraction of the proteome (total ~ 18%) do vary with growth rate. Three outer membrane proteins (NmpC, OmpA, and OmpF) show contrasting trends (Figure 2.7): the Omp family of proteins increase in parallel with growth rate, while NmpC decreases with growth rate. Interestingly, the total proteome fraction of outer membrane proteins (Omp proteins plus NmpC) vary only from 7.5% to 9.1%, suggesting a shift in outer channel composition while maintaining abundance. As the specific roles of individual outer membrane channels are unknown, the significance of these trends is unclear.

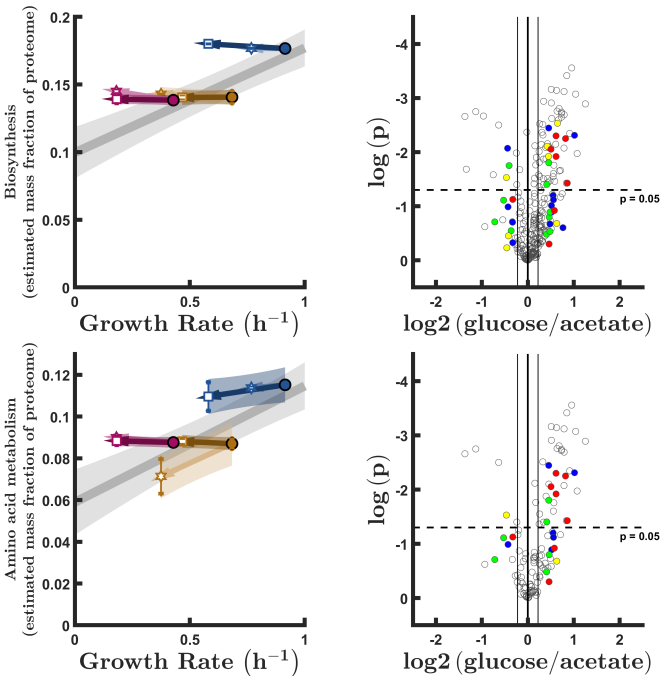


Figure 2.6: Total proteome mass of amino acid synthesis pathways correlates with growth rate independently of basal ppGpp. Total proteome mass fraction for the "Biosynthesis" proteome sector its largest component "Amino Acid Metabolism" following the same colors and symbols as Figure 2.4.

2

2.3.3. SYNTHETIC PPGPP TITRATION REVEALS SCOPE OF TRANSCRIPTIONAL REGULATION BY BASAL PPGPP

As basal ppGpp concentrations consistently vary inversely with growth rate, proteins whose abundance varies with growth rate are potentially regulated by basal ppGpp, or by other signals that correlate or anti-correlate with growth rate such as cAMP. To identify proteins regulated specifically by basal ppGpp, we quantified proteins following synthetic ppGpp titrations. A survey of Level 2 sectors indicates that synthetic ppGpp titrations exert varying effects. For instance, the Translation sector as a whole is consistently repressed by increased ppGpp and induced by depleted ppGpp (Figure 2.4), while the Biosynthesis sector is not significantly changed by ppGpp titrations (Figure 2.6). Interestingly, the effect of ppGpp titrations on the Central Carbon Metabolism sector changes between growth media (Figure 2.5). The sector-average response to ppGpp titration may not reflect the responses of individual proteins within the sector. Therefore, to identify proteins whose expression is likely regulated by basal ppGpp, we applied several criteria (Figure 2.8):

- 1) Protein abundance correlates (or anti-correlates) with growth rate. Proteins whose expression remains constant between growth conditions are apparently insensitive to varying basal ppGpp (although this does not exclude all forms of ppGpp regulation).

- 2) Increasing ppGpp with RelA must vary expression in a manner similar to basal ppGpp. In other words, synthetically increasing ppGpp must vary protein abun-

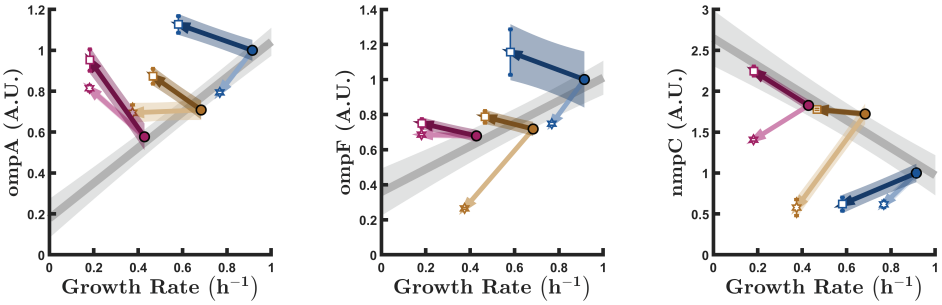


Figure 2.7: Outer membrane proteins show strong and contrasting responses to nutritional conditions as well as ppGpp titration. Relative concentration of OmpA, OmpF and NmpC. Rescaled by concentration obtained from wild-type sample in glucose minimal media. Colors and symbols as in Figure 2.1

dance in the same manner as lowering growth rate through poorer nutritional conditions. Applying this criteria excludes proteins whose abundance varies with carbon source but are apparently insensitive to ppGpp such as many central carbon metabolism enzymes or tRNA-aminoacyl synthetases.

- 3) Decreasing ppGpp with Mesh1 must vary protein abundance in a manner opposite of increasing ppGpp. Applying this criteria excludes proteins that are sensitive to growth rate, but apparently not sensitive to ppGpp. Proteins in this group include methionine synthase MetE, which contributes between 1-2% of the total proteome.

- 4) Synthetic ppGpp titration causes consistent variation. Trends driven by RelA or Mesh1 induction are clearly observed in all three conditions: these inductions lead to similar changes in all three carbon sources. Applying this constraint excludes proteins that become insensitive to ppGpp in specific carbon sources.

A total of 86 proteins meet all 4 criteria and are thus considered to be consistently regulated by ppGpp. They represent 28.7% of total proteome in glucose to 19.2% in acetate. 50 of these proteins are classified as "Translation" and represent both ribosomal proteins and translation factors (e.g. TufA, FusA). A closer inspection reveals 5 additional proteins with clear roles in translation but are listed as "Other Enzymes" or unclassified: Tig, Tgt, TrmB, RimO, YhbY, for a total of 55 translation-related proteins. Interestingly, 10 ribosomal proteins do not meet this strict criteria. 7 of these are non-essential for translation in some conditions (RplP [19], RpmC [20, 21], RpmF [22], RpmG, RpmH [23–26], RpsT [27] and RpsU [28, 29]) and two are associated with stationary phase (Sra and YkgM), which might explain their weaker dependance to basal ppGpp levels. The last missed ribosomal protein, RpsR, shows a weaker depency to growth rate, thus not making it through our filter for no evident biological reason. With this classification, we show that ppGpp control alone determines the size of the "Translation" sector through its repressive effect of ribosomes and translation cofactors. Apart from these two sub-sectors, proteins within this sector are in majority not basal ppGpp-regulated: tRNA-aminoacyl

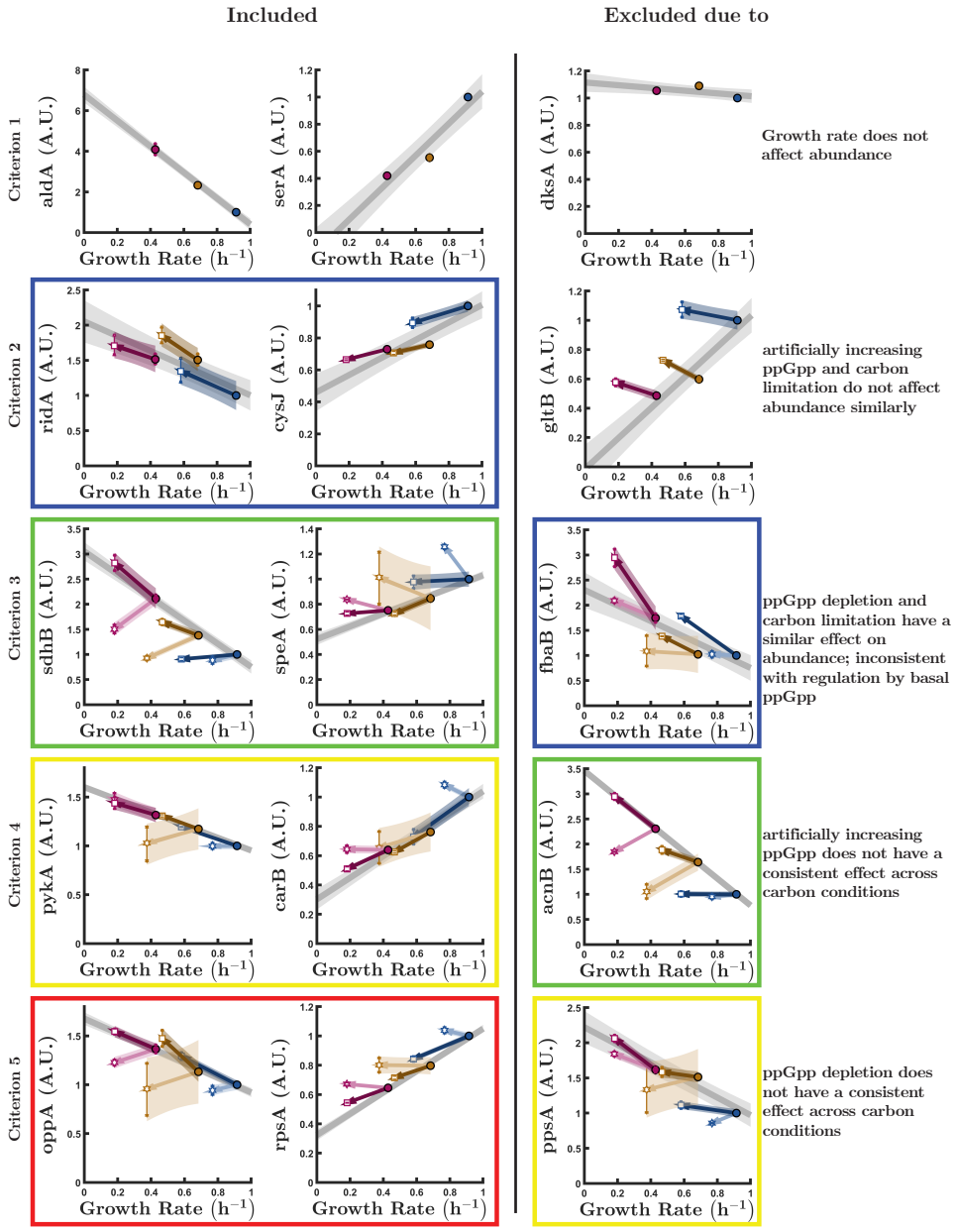


Figure 2.8: A 5-filter classification allows identification of proteins regulated by basal ppGpp. Representation, through examples of protein relative concentration, of the five filters used in our classification and describe in section 2.3.3. Each line corresponds to one of the filters used to classify with on the left two examples of proteins passing this filter and on the right an example of a protein passing all previous filters but not the considered one. Frame colors correspond to the color of each of these proteins used in various volcano plots found in this chapter. Plots show concentration rescaled by concentration in glucose minimal media and with the same colors and symbols as Figure 2.1.

synthetases, RNA helicases involved in ribosome biogenesis, and methyltransferases that modify rRNA.

The remaining 31 proteins classified as basal ppGpp-regulated originate from several sectors and collectively contribute less than 0.1% of the total proteome. 8 proteins classified as "biosynthesis" are from amino acid pathways and include 6 proteins from the arginine biosynthesis pathway. Only 5 out of 119 proteins from "Central Carbon Metabolism" are classified as basal ppGpp-regulated. Two of these are from the TCA cycle (citrate synthase GltA and stress-resistant aconitase AcnA) while another one (glucose-1-phosphatase scavenging enzyme Agp) is misclassified as a Glycolysis enzyme. Four proteins exhibit clear links to stress and detoxification: catalase enzymes KatG and KatE, glutathione transferase GstB, and aldehyde reductase YahK. Of transcriptional regulation proteins, only Crp and anti-sigma factor Rsd are basal ppGpp-regulated.

EXPRESSION OF AMINO ACID BIOSYNTHESIS ENZYMES PARADOXICALLY INCREASES WITH BOTH GROWTH RATE AND SYNTHETIC PPGPP

E. coli strains unable to produce ppGpp (ΔrelA ΔspoT) are also auxotrophic for specific amino acids (aspartic acid, glutamic acid, phenylalanine, histidine, isoleucine, leucine, threonine, valine, and serine) [30], indicating multiple amino acid biosynthesis pathways require activation by ppGpp for expression. The increased expression of amino acid synthesis operons in parallel with growth rate (and thus lower basal ppGpp) is therefore unexpected. Our results are consistent the role of ppGpp in inducing several amino acid biosynthesis genes with the increased expression of at least 15 amino acid enzymes by synthetically increasing ppGpp (Figure 2.6). Pathways that are paradoxically both ppGpp-induced and growth rate-correlated synthesize the amino acids histidine, threonine, and serine, which are also among the biosynthesis operons that require ppGpp for induction (Figure 2.9). Interestingly, 12 proteins from the amino acid biosynthesis sector are repressed by ppGpp, including the arginine biosynthesis pathway. Remarkably, the methionine biosynthesis protein MetE, which is positively correlated with growth rate, is repressed by both ppGpp accumulation and depletion. This indicates that MetE is synchronized with growth rate regardless of ppGpp concentrations (Figure 2.9). This result suggests that although a minimum concentration of basal ppGpp is required for expression of several amino acid biosynthesis pathways, further variations above this minimum level do not change expression. Transcription of amino acid biosynthesis operons are also feedback-regulated by their corresponding product: amino acids. Therefore, the steady-state concentrations of amino acids might synchronize expression of biosynthesis operons with translation demand, with some minimal concentration of ppGpp being required to induce expression.

TCA CYCLE ENZYMES ARE NOT CONSISTENTLY AFFECTED BY SYNTHETIC PPGPP VARIATION

The inverse correlation between the Level 3 sector "TCA cycle and anapleurotic enzymes" and growth rate might suggest that this sector is incrementally induced by basal ppGpp. However, increasing ppGpp by RelA expression failed to increase expression in most cases, with the notable exception of acetate medium (Figure 2.5). This indicates that expression is unlikely to be directly adjusted by basal ppGpp. Many of the proteins are known to be induced by the regulator cAMP, which also correlates negatively with growth rate (Figure 2.1). Therefore, TCA cycle enzymes are likely coordinated with growth by basal cAMP, not

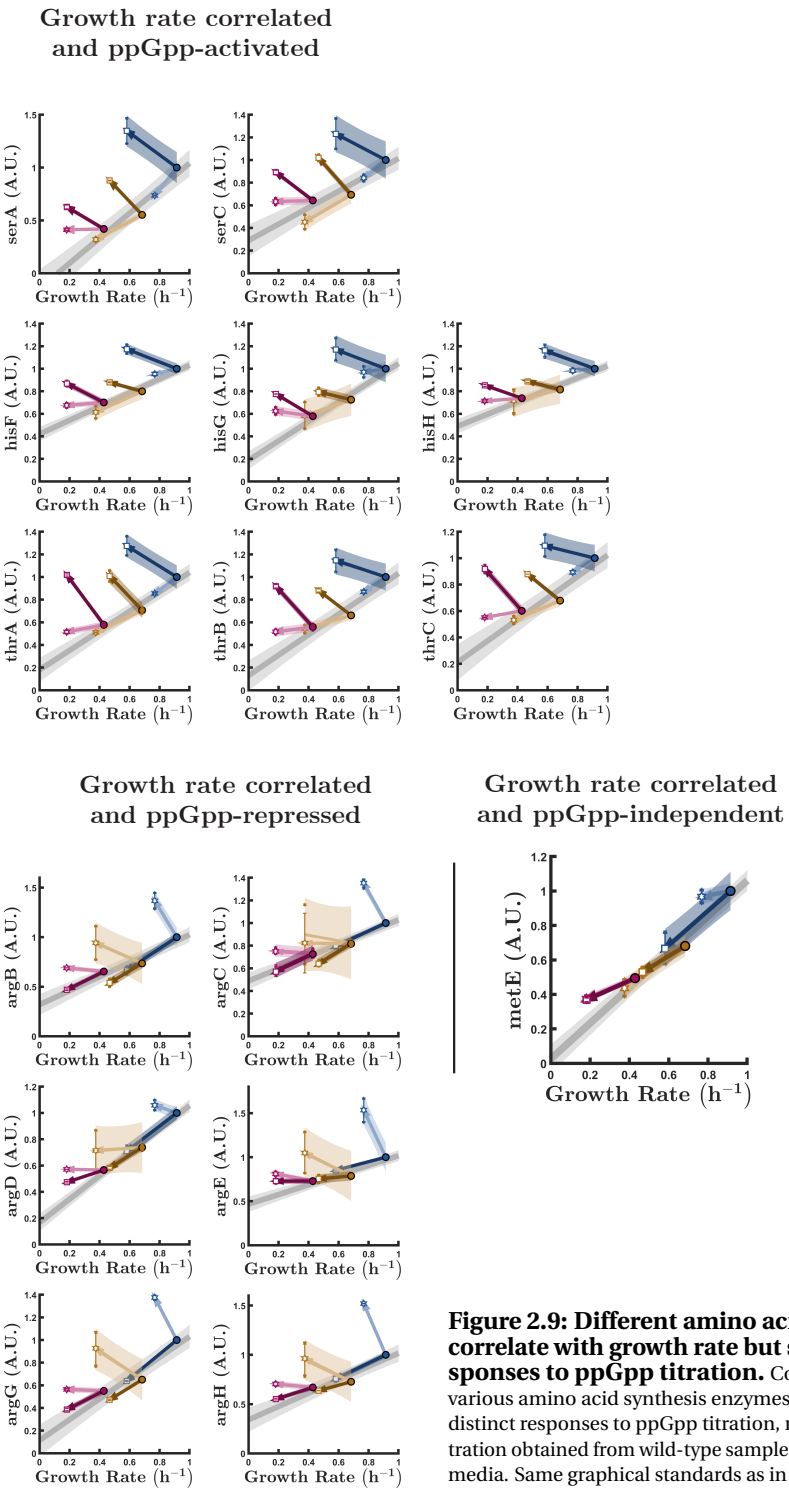


Figure 2.9: Different amino acid pathways all correlate with growth rate but see different responses to ppGpp titration. Concentration of the various amino acid synthesis enzymes showing clear and distinct responses to ppGpp titration, rescaled by concentration obtained from wild-type sample in glucose minimal media. Same graphical standards as in Figure 2.1

basal ppGpp.

However, Mesh1 expression consistently decreases expression of TCA cycle enzymes (Figure 4). This suggests that increasing ribosome expression by depleting ppGpp may indirectly decrease expression of TCA cycle enzymes. As growth rate in acetate medium is highly sensitive to abundance of TCA cycle enzymes [31, 32], further reducing their abundance by ppGpp depletion is likely decreasing growth rate. Disrupting the natural balance between ribosomes and TCA cycle enzymes is likely responsible for reducing growth rate in acetate medium.

OUTER MEMBRANE PORINS RESPOND INCONSISTENTLY TO PPGPP TITRATION

Outer membrane porins mediate permeability of the outer membrane. While the roles of individual porins are unclear, they represent a considerable fraction of the proteome mass. We consider four porins, NmpC, OmpA, OmpC, and OmpF, that occupy nearly 10% of the total cellular proteome. OmpC and OmpF are classified under the Level 3 sector "Transport" (within Metabolism), while NmpC and OmpA are not classified. NmpC abundance decreases from acetate to glucose (from 7% to 4%), and is consistently repressed by ppGpp depletion, but is not consistently induced by ppGpp overproduction (Figure 2.6). The Omp proteins considered here consistently increase with growth rate (OmpA increases from 1% to 2% of total proteome), but show inconsistent responses to ppGpp variation. For instance, OmpA is consistently induced by ppGpp overproduction, while OmpC is consistently induced by ppGpp depletion. As the specific roles of these porins are not entirely clear (and regulation is highly complex), it is difficult to understand the role of ppGpp in varying the abundance of these porins. However, none of these proteins exhibit behaviours that are consistent with regulation by basal ppGpp alone.

2.3.4. IN CONTRAST WITH EXCESS PPGPP, CONSEQUENCES OF PPGPP DEPLETION ARE CONDITION DEPENDENT

Any change in basal ppGpp decreases the growth rate. The proteome response to ppGpp perturbations provides insight into why basal ppGpp concentrations are optimal. ppGpp concentrations above basal exert clear effects on ribosomal protein abundance, consistent with well-established inhibition of ribosomal RNA transcription. This inhibition decreases ribosome abundance, thus decreasing the translation rate. In contrast, the specific effects of depleting ppGpp below basal concentrations are less clear (aside from increasing Translation sector proteins). While increasing ribosome abundance over the optimum level is likely to divert resources away from other cellular functions, it is not clear which specific resource becomes depleted by the expression of excess ribosomes. We therefore analysed ppGpp depletion conditions to identify the specific cellular functions that become growth-limiting when ppGpp is artificially depleted.

PPGPP DEPLETION LIMITS GROWTH BY REDUCING EXPRESSION OF CONDITION-DEPENDENT CATABOLIC ENZYMES

The three conditions examined use different pathways for the initial steps of carbon catabolism. We therefore examined the effects of artificial ppGpp titration on those specific pathways. Acetate uptake pathways (AckA, ActP, Acs) are not affected by ppGpp depletion (Figure 2.10). Instead, ppGpp depletion reduces expression of TCA and glyoxylate cycle enzymes. Cells in acetate medium rely upon the glyoxylate shunt in particular:

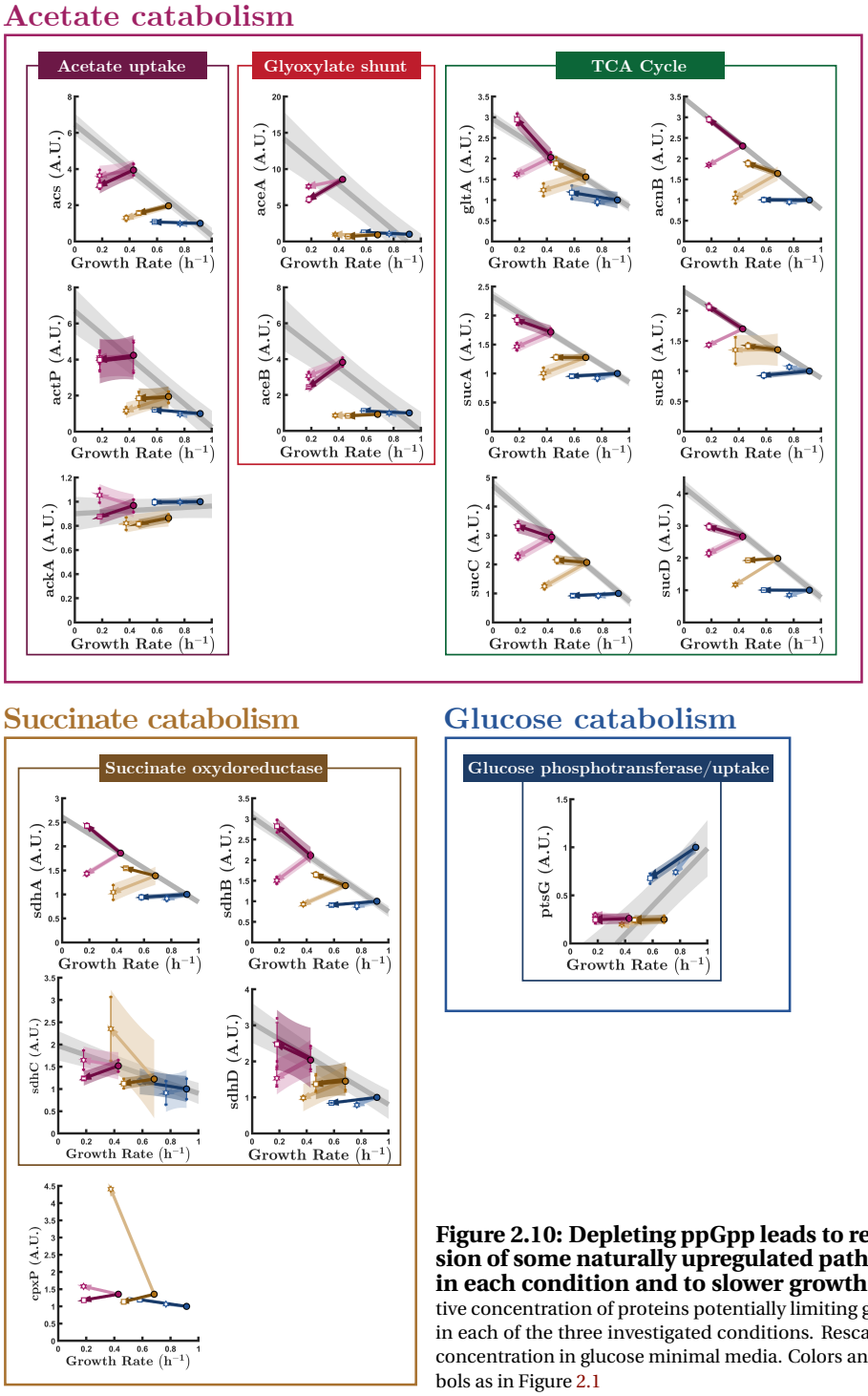


Figure 2.10: Depleting ppGpp leads to repression of some naturally upregulated pathways in each condition and to slower growth. Relative concentration of proteins potentially limiting growth in each of the three investigated conditions. Rescaled by concentration in glucose minimal media. Colors and symbols as in Figure 2.1

expression of enzymes AceA and AceB increase over 8 and 3-fold respectively compared to glucose medium. Depleting ppGpp with Mesh1 decreases AceB expression by 25%. Additional TCA cycle enzymes upregulated in acetate medium are also decreased by ppGpp depletion, including aconitase AcnB, oxyglutarate dehydrogenase (SucAB), succinyl-CoA synthase SucCD, and succinate dehydrogenase complex proteins SdhABCD. This suggests decreased growth due to ppGpp depletion in acetate medium might come from insufficient glyoxylate shunt and/or TCA cycle proteins rather than less proteins responsible for acetate uptake. Interestingly, in glucose medium ppGpp depletion does not affect abundance of these enzymes.

The first step of succinate catabolism is mediated by the succinate dehydrogenase complex (SdhABCD), which oxidizes succinate to malate and produces NADH which is oxidized by cytochrome b. ppGpp depletion reduces expression of SdhABCD by approximately 30% (Figure 2.10). This repression of SdhABCD expression provides a clear suggestion as to how ppGpp depletion can slow growth in succinate medium. Interestingly, CpxP, a transcriptional regulator associated with cell envelope stress, increases by 3-fold during ppGpp depletion, but only in succinate medium. As ppGpp is not known to regulate CpxP, membrane stress may be a downstream consequence of ppGpp depletion in succinate medium.

Interestingly, enzymes repressed by ppGpp depletion in acetate and succinate medium are only slightly repressed in glucose medium. Thus, ppGpp depletion in glucose medium is unlikely to reduce growth rate by reducing expression of TCA cycle enzymes. ppGpp depletion also minimally affects expression of enzymes within the Glycolysis sector. However, the glucose transporter PtsG, which is minimally expressed in acetate and succinate medium, increases by 4-fold in glucose medium (Figure 2.10). ppGpp depletion does not affect PtsG expression in acetate or succinate medium, but reduces expression in glucose medium by 25%. PtsG titration has been used to vary growth rate in glucose medium [33]. Thus, ppGpp depletion likely reduce growth rate in glucose medium by reducing expression of PtsG.

PPGPP DEPLETION ACTIVATES INDIVIDUAL AMINO ACID SYNTHESIS PATHWAYS IN A CONDITION-DEPENDENT MANNER

The clear regulation of translational proteins highlights the link between translation and ppGpp. The best-understood enzyme responsible for maintaining ppGpp concentrations is the ppGpp synthase RelA, which becomes active in response to binding of deacyl-tRNA to ribosomes. While the processes that balance ppGpp synthesis and degradation to arrive at optimal basal ppGpp levels are poorly understood (largely because the ppGpp hydrolase enzyme, SpoT, is difficult to study), the abundance of amino acids likely plays a role in determining RelA activity and thus basal ppGpp. Depleting basal ppGpp and inducing excess ribosome expression might be expected to deplete amino acid pools, which should in turn increase expression of amino acid pathways as a response. Therefore to further identify consequences of metabolic limitation by ppGpp depletion, we looked for specific amino acid pathways that become upregulated following mild ppGpp depletion by Mesh1 expression. Interestingly, the overall abundance of Level 3 sector “Amino acid metabolism” is not affected by ppGpp depletion. However, individual proteins within the sector are induced by Mesh1 expression, in a carbon source dependent manner.

ppGpp depletion in acetate medium increases expression of valine/isoleucine syn-

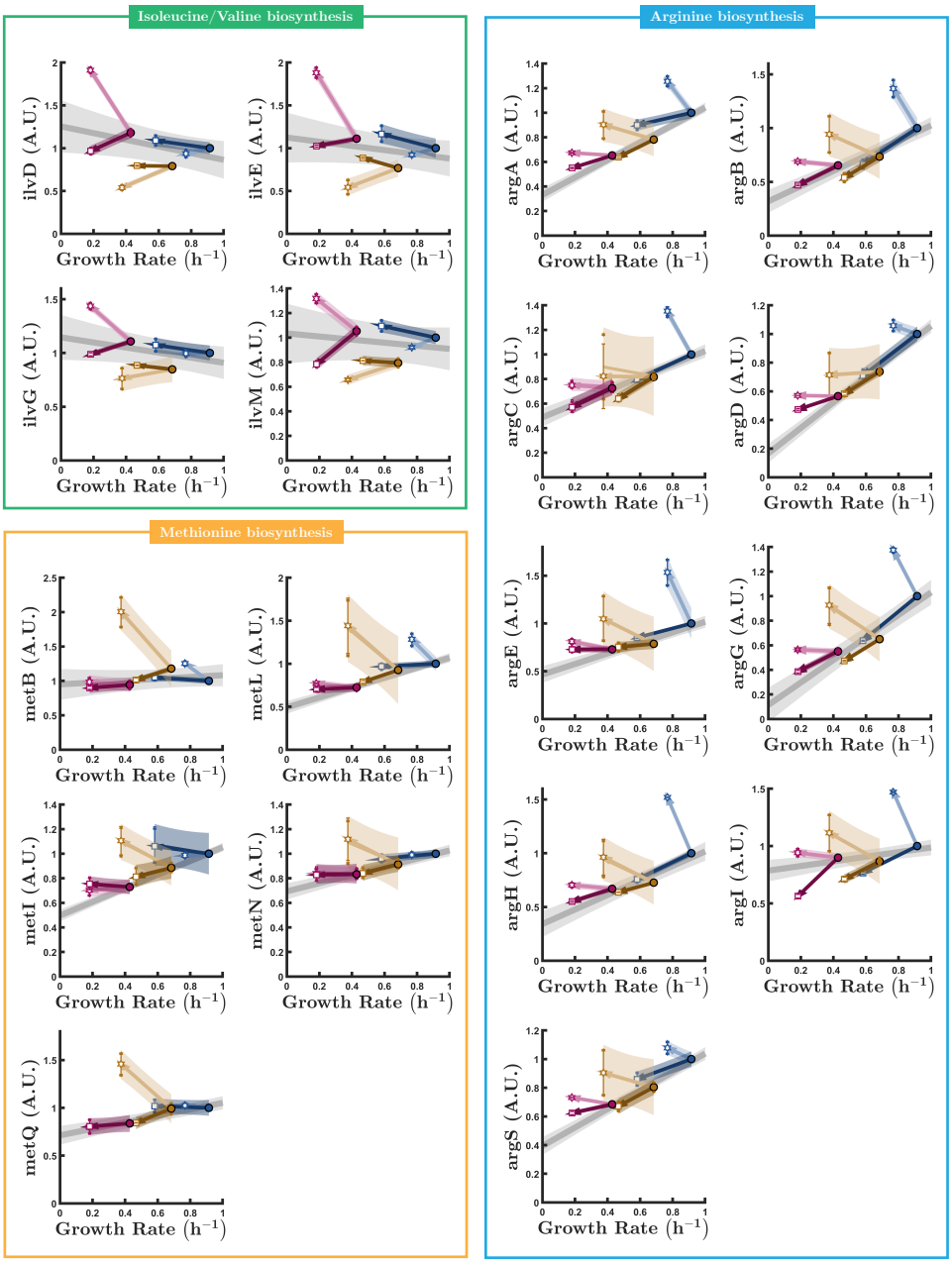


Figure 2.11: A single and different amino acid synthesis pathway is increased by depleted in each of the conditions investigated. Relative concentrations of amino acid synthesis enzymes belonging to pathways that respond the ppGpp depletion in a specific condition. Concentrations are rescaled by the concentration in glucose medium. Same graphical conventions as Figure 2.1

thesis genes IlvD, IlvE, IlvG, and IlvM (Figure 2.11), which are co-transcribed from a single operon. In both succinate and glucose media, ppGpp depletion slightly decreases expression of these genes, indicating that the increased expression is specific to ppGpp depletion in acetate medium. Thus, we predict that isoleucine or valine limit translation (and thus growth) when ppGpp is depleted in acetate medium. Similarly, in succinate medium ppGpp depletion increases expression of methionine synthesis genes MetB and MetL, as well as the methionine ABC transport proteins MetQ, MetI, and MetN. However the methionine synthesis gene MetE moves against this trend by decreasing. Finally, ppGpp deletion in glucose medium considerably increases expression of the arginine biosynthesis operon by 40% on average. Whether ppGpp depletion triggers starvation of specific amino acids can be verified using ribosomal pausing assays to determine translation pauses at corresponding codons in ppGpp-depleted conditions [34].

PPGPP TITRATION PERTURBS NUCLEOTIDE AND ACETYL-COA CONCENTRATIONS

To determine the metabolic consequences of synthetic ppGpp titration, we quantified purine nucleoside phosphates and acetyl-CoA. As previously reported, concentrations of ATP and GTP changed very little between the three media tested. However, incrementally increasing ppGpp increased ATP and GTP at first, while further increasing ppGpp sharply decreased ATP and GTP (Figure 2.12). In acetate increasing ppGpp seems to only decrease nucleoside triphosphate (NTP) concentration. Increasing ppGpp does not perturb the ATP charging ratio, which remains close to its normal value of 0.95. In contrast, the effects of depleting ppGpp with Mesh1 on nucleotide pools appears to depend upon the carbon source. With glucose and glycerol, depleted ppGpp increases GTP and ATP, while in acetate, depleted ppGpp decreases ATP and GTP and slightly decreases the ATP charging ratio (0.95 to 0.9).

The effects of synthetic ppGpp titration on acetyl-CoA are striking and consistent. Synthetically increasing ppGpp causes acetyl-CoA to accumulate in all conditions tested, whereas synthetic ppGpp depletion consistently depletes acetyl-CoA. As acetyl-CoA is a precursor of cellular building blocks such as amino acids and fatty acids, its depletion likely has profound consequences on downstream metabolism and cellular physiology.

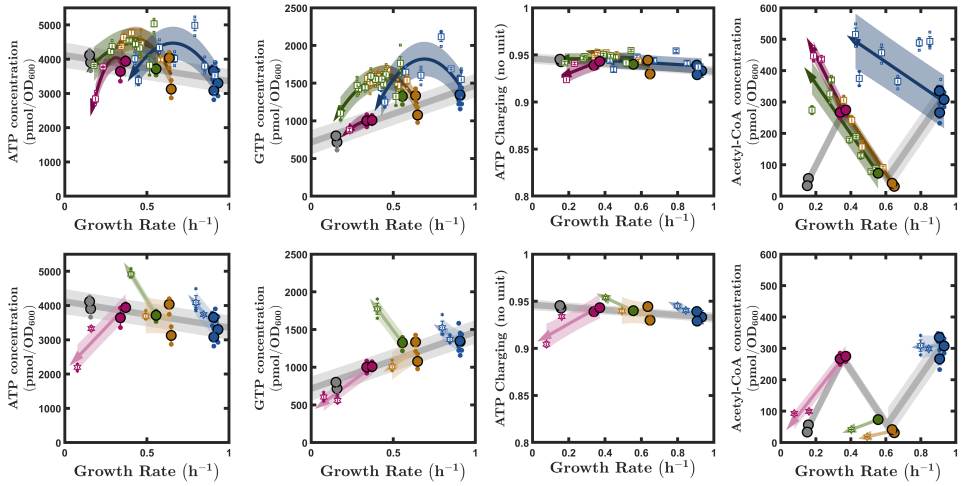


Figure 2.12: Excess and depleted ppGpp respectively accumulates and depletes acetyl-CoA, while effect on ATP and GTP depends on the titration magnitude and on the condition respectively. Concentrations of ATP, GTP and acetyl-CoA in various nutrient conditions for wild-type as well as RelA* (top) and Mesh1 (bottom). Colors and symbols are the same as in Figure 2.1.

2.4. DISCUSSION

In this Chapter, we go beyond Ref. [13], which established the broad reach of ppGpp using nutrient exhaustion to indirectly activate ppGpp synthesis. Manipulating ppGpp directly and setting criteria to identify genes that respond to basal-level ppGpp, we distinguish direct ppGpp regulation from indirect effects that follow from redistributing limited proteome resources or perturbing other global regulators such as cAMP with a set of criteria.

We find that proteins most directly related to translation (ribosomal proteins and translation cofactors) are highly sensitive to incremental variations in ppGpp. Surprisingly, many other proteins are insensitive to both ppGpp titration and growth rate. Furthermore, only a minority of proteins whose abundances are highly sensitive to growth rate meet our criteria, responding to both basal ppGpp and synthetically-titrated ppGpp in a consistent manner. The rest of these proteins are likely coordinated with growth in response to other signals such as cAMP. This work thus advances our understanding of basal ppGpp regulation by establishing boundaries to its reach.

What about the genes that are known to be directly activated and repressed by ppGpp? Our approach identified some genes thought to be regulated by ppGpp that do not match the expected trends of a ppGpp-controlled target. Many amino acid biosynthesis pathways require ppGpp for expression; however, their expression increases while basal ppGpp decreases with increasing growth rate. Expression of these pathways clearly cannot be titrated between growth conditions by ppGpp. It might be activated by a very low threshold of ppGpp concentration below the ppGpp concentration measured in glucose minimal medium. Increased expression of amino acid synthesis pathways with

higher growth rate may arise from stronger feedback regulation due to higher demand. This observation highlights that, in steady-state growth, most genes are not regulated by ppGpp alone. Instead, ppGpp acts in combination with additional signals such as starvation of a particular amino acid.

Our work highlights specific protein sectors that are balanced against the translation sector in the conditions we examined. Broadly speaking, decreased expression of the translation sector in succinate and acetate medium is a tradeoff required for expression of TCA cycle enzymes to increase. Disrupting this tradeoff by artificially decreasing ppGpp increases the translation sector, depleting the TCA cycle enzymes. This likely starves the cell of metabolites, as indicated by reduced acetyl-CoA, which in turn may trigger starvation of individual amino acids. This chain of events could be further tested with ribosomal pausing assays.

Our work also reveals some of the metabolic consequences of non-optimal ppGpp regulation leading to slower growth. Excess ppGpp is straightforward to understand at a detailed level: inhibition of ribosome synthesis by ppGpp reduces the abundance of ribosomes decreasing the overall rate of protein synthesis. As a consequence, the cell possesses a higher catabolic capacity than normal. Cells cannot channel this catabolic output into protein synthesis due to limited ribosome abundance, as indicated by the accumulation of acetyl-CoA in cells with excess ppGpp.

In contrast with excess ppGpp, the consequences of depleted ppGpp on the catabolic sector are not immediately clear. ppGpp depletion effects are obscured on the sector level because ppGpp depletion does not consistently deplete specific catabolic enzymes in every growth media. Instead, ppGpp depletion reduces expression of catabolic enzymes that are upregulated in response to the specific carbon source. For acetate and succinate media, expression of TCA and glyoxylate cycle enzymes are reduced by ppGpp depletion. Expression of these enzymes is barely affected in glucose medium while expression of the glucose importer PtsG, which is expressed in glucose but not in acetate or succinate, is specifically reduced during ppGpp depletion in glucose medium.

How does reducing ppGpp lower TCA cycle enzymes expression while those are not directly regulated by ppGpp? We note that ppGpp depletion reduces cAMP levels in three out of four investigated media (glucose, succinate and glycerol; $p<0.05$), which is a plausible explanation for why TCA cycle enzymes (and expression of PtsG in glucose medium) are expressed at lower levels. The reduction in cAMP concentrations is very surprising, as one would expect that excess ribosomes should deplete essential metabolites, as we showed for acetyl-CoA. Unfortunately, regulation of cAMP by AMP cyclase and cAMP hydrolase is poorly understood. Depleting ppGpp may cause specific metabolites monitored by the cAMP system to accumulate, "fooling" the system into thinking that conditions are better than they are.

Alternatively, ppGpp may affect expression via indirect mechanisms arising from its direct influence over a large fraction of the proteome. This so-called "passive" regulation has been speculated about [35, 36] and observed when a useless protein is strongly overexpressed [3]. It is a consequence of limited resources, such as RNA polymerases or ribosomes, available to transcribe genes and translate proteins: when more of them are busy synthesizing some proteins (here, ribosomal ones) become limiting to the production of others.

2.5. METHODS

STRAINS

All strains are created using WT strain NCM3722, referred as WT. The RelA* inducible strain, is the same strain described in Noga et al. 2020 [37]: pSC101-kanR-TetR-RelA*-mVenus. The Mesh1 inducible strain was constructed for this study and has the following genome p15A-kanR-TetR-Mesh1-YFP. To check whether the effect on growth of inducing this gene construct comes from the production of Mesh1 and its effect of ppGpp we also constructed p15A-kanR-TetR-YFP, which showed very minor to no effect on growth at the highest induction level used in this study and in the most sensitive condition to ppGpp depletion (acetate). Half of this induction concentration with the Mesh1 inducible strain in the same conditions induced lower growth rate than any condition investigated in our study.

GROWTH MEDIA AND CONDITIONS

Bacterial cultures are grown in MOPS minimal media using 15N isotope NH4Cl as nitrogen source and 2% carbon source: glucose, succinate, glycerol, aspartate or glutamate. These last two carbon sources, containing nitrogen, are also 15N isotopes. Cultures were inoculated directly from a single colony of the desired strain from a LB agar plate in a 250mL flask containing 25mL of the desired medium, which was brought to 20mL when setting up the growth monitoring setup.

GROWTH RATE MONITORING

Optical density at 600nm is monitored every 5sec to 1min using an automated growth measurement setup. Briefly, liquid culture is constantly pumped through a flow-through cuvette measured for optical density at frequent intervals and subsequently pumped back into the flask. Growth rate is extracted by smoothing OD measurements with a 5 minute wide gaussian filter on OD measurements and taking median of the time derivative of the neperian logarithm on a one hour window.

PPGPP TITRATION

To measure ppGpp concentrations at steady state, bacteria are directly induced in a 250mL flask containing 25mL of minimal media with the desired carbon source. For natural ppGpp levels, WT bacteria is simply grown up to 0.35OD and three technical replicate samples are taken as described in the paragraph below. For titrated ppGpp concentrations, the strains described in this sections are grown. These strains see a very strong evolutive disadvantage due to their low growth rate which can lead to the apparition of titration-resistant mutants. To ensure the selection of the desired strains, a concentration of 25µg/mL kanamycin is added to the media. We induce ppGpp titration when reaching 0.05OD to 0.08OD, depending on the strength of the induction. Steady state is considered reached when growth rate, stays within 10% variation for 30 minutes, which happened in around 3-5 hours for RelA* induction and 5-8 hours for Mesh1. Samples are then taken as described below at OD_{600} as close as possible to 0.35, always between 0.3 and 0.4. Growth was subsequently monitored for 3 hours after sampling. When a decrease or an increase in growth rate superior to 10% was measured after sampling, samples were discarded as the first one means the culture did not reach steady state and the second one is likely caused by selected mutants.

LCMS NUCLEOTIDE AND METABOLITE MEASUREMENTS

To measure nucleotide concentrations, three technical replicates were sampled by pipetting 1mL or 2.5mL (for lower ppGpp concentrations) of 0.35OD culture onto a wet filter under vacuum. The filter is immediately quenched upside down in 1mL ice-cold 2M formic acid. Immediately after filtering, a known concentration of the nucleotides of interest is added to the quenching solution. After 30 minutes in this solution, the filter is thoroughly washed and discarded. 25μL ammonium hydroxyde is added to the collected quenching solution in a 1mL tube which is frozen at $-80^{\circ}C$, which is kept at least overnight and up to one month. To measure nucleotides and metabolites, a sample is first thawed on ice the sonicated in ice-cold water for 10 minutes. Solid phase extraction was performed to purify nucleotides and get rid of other compounds that might cause matrix effects on the compounds of interest by altering the chromatographic separation or suppressing their ionization. To do so, during sonication, an Oasis Wax SPE cartridge 30mg 30µm is prepared by, first, flowing through three times 1mL methanol, then, three times 1mL 4.5 pH 50mM ammonium acetate. Sample is then applied on cartridge and slowly flowed through. Cartridge was washed with first 1mL 4.5 pH 50mM ammonium acetate then 1mL methanol and left to dry under vacuum for 5 minutes. Sample was eluted very slowly from cartridge into eppendorf tube with 200µL 5:3:1:1 MeOH:ACN:H2O:NH4OH. The obtained sampled was concentrate using a vacuum centrifuge after addition of 10µL of 5% trehalose for stabilization during drying. Once dried, the sample pellet was rediluted in 20µL 5:3:2 MeOH:ACN:H2O. The obtained sample was analyzed with Agilent triple quad mass spectrometer and ZIC-cHILIC column. For the first minute we maintained 90% mobile phase B 11.25mM ammonium acetate, 3.75mM acetic acid and 2mM acetylacetone in 80% ACN and 10% mobile phase A 3.75mM ammonium acetate, 1.25mM acetic acid and 2mM acetylacetone in water. Then we applied a gradient towards reaching 80% mobile phase A at 15 minutes. We maintained this concentration for one minute before applying a new gradient towards 100% mobile phase B between 16 and 18 minutes. We maintained this concentration until 22.5 minutes to flush out remaining compounds. Measuring low concentrations of ppGpp required to optimize peak height. In that purpose, the method described here above and used in this study was slightly adapted from the one developed and successfully used to measure nucleotides in Noga et al. 2020 [37].

NUCLEOTIDE AND METABOLITES DATA ANALYSIS

For each compound, 14N and 15N peaks were obtained, corresponding to standard concentration added and biological sample respectively. Concentrations in pmol/OD_{600} were obtained by multiplying ratio of the 15N over 14N peak areas area by the added concentration of compound and dividing by OD_{600} at sampling. The latter being obtained by average between the last OD_{600} point measured before sampling and the first one after sampling. From this calculation was obtained the concentration for each sample. Samples for which no peak above a certain threshold depending on the compound was measured were not used for this compound. When at least two technical replicates passed the threshold, the mean and standard deviation was extracted from the replicates estimated concentrations.

UNTARGETED PROTEOMICS MEASUREMENTS

To measure relative protein concentration, we performed untargeted proteomics method similarly as den Ridder et al. [14]. Briefly, 20mL of 20% of ice-cold Trichloroacetic acid was added to 20mL of 0.35OD culture. Content of the culture flask was transferred to a 50mL falcon tube and centrifuged 10 minutes at 1000g at $4^{\circ}C$. Supernatent was discarded and cell pellets were frozen at $-80^{\circ}C$. Cells were lyzed with bead beating, proteins were precipitated in TCA before being reduced in DTT, alkylated in IAA and digested with trypsin. Samples were labelled using TMT10plex reagents and analyzed by a QE plus Orbitrap mass spectrometer. For more details on this procedure, we refer the reader to den Ridder et al. [14].

SPECTRAL COUNTING UNTARGETED PROTEOMICS

To obtain absolute quantification of proteome mass fraction for each measured protein in each condition from our untargeted proteomics data, we used iBAQ spectral counting [17, 18]. To obtain absolute protein abundance estimations, we divided for each identified protein in each condition the number of counts by the number of theoretically observable peptides, which we obtained using the Matlab bioinformatics toolbox function cleave.m, selecting the ones with a mass comprised between 800 and 2400 Da. We thus ignored missed cleavages. We multiplied iBAQ absolute protein abundances by the mass of the considered protein and divided by the sum of these products over all proteins to obtain proteome mass fraction estimates.

REFERENCES

[1] M. Cashel and J. Gallant, *Two compounds implicated in the function of the rc gene of escherichia coli,* Nature **221**, 838 (1969).

[2] W. Ross, C. E. Vrentas, P. Sanchez-Vazquez, T. Gaal, and R. L. Gourse, *The magic spot: a ppgpp binding site on e. coli rna polymerase responsible for regulation of transcription initiation,* Molecular cell **50**, 420 (2013).

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

[4] S. Hui, J. M. Silverman, S. S. Chen, D. W. Erickson, M. Basan, J. Wang, T. Hwa, and J. R. Williamson, *Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria,* Molecular systems biology **11**, 784 (2015).

[5] M. Zhu and X. Dai, *Growth suppression by altered (p) ppgpp levels results from non-optimal resource allocation in escherichia coli,* Nucleic acids research **47**, 4684 (2019).

[6] J. Ryals, R. Little, and H. Bremer, *Control of rrna and trna syntheses in escherichia coli by guanosine tetraphosphate.* Journal of bacteriology **151**, 1261 (1982).

[7] F. Büke, J. Grilli, M. C. Lagomarsino, G. Bokinsky, and S. J. Tans, *ppgpp is a bacterial cell size regulator,* Current Biology **32**, 870 (2022).

[8] C. Wu, R. Balakrishnan, N. Braniff, M. Mori, G. Manzanarez, Z. Zhang, and T. Hwa, *Cellular perception of growth rate and the mechanistic origin of bacterial growth law,* Proceedings of the National Academy of Sciences **119**, e2201585119 (2022).

[9] M. Zhu and X. Dai, *Stringent response ensures the timely adaptation of bacterial growth to nutrient downshift,* Nature Communications **14**, 467 (2023).

[10] W. A. Haseltine and 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,* Proceedings of the National Academy of Sciences **70**, 1564 (1973).

[11] B. J. Paul, M. M. Barker, W. Ross, D. A. Schneider, C. Webb, J. W. Foster, and R. L. Gourse, *Dksa: a critical component of the transcription initiation machinery that potentiates the regulation of rrna promoters by ppgpp and the initiating ntp,* Cell **118**, 311 (2004).

[12] X. Dai, M. Zhu, M. Warren, R. Balakrishnan, V. Patsalo, H. Okano, J. R. Williamson, K. Fredrick, Y.-P. Wang, and T. Hwa, *Reduction of translating ribosomes enables escherichia coli to maintain elongation rates during slow growth,* Nature microbiology **2**, 1 (2016).

[13] M. F. Traxler, S. M. Summers, H.-T. Nguyen, V. M. Zacharia, G. A. Hightower, J. T. Smith, and T. Conway, *The global, ppgpp-mediated stringent response to amino acid starvation in escherichia coli,* Molecular microbiology **68**, 1128 (2008).

[14] M. den Ridder, P. Daran-Lapujade, and M. Pabst, *Proteome dynamics during transition from exponential to stationary phase under aerobic and anaerobic conditions in yeast,* bioRxiv , 2022 (2022).

[15] M. Kanehisa, S. Goto, S. Kawashima, Y. Okuno, and M. Hattori, *The kegg resource for deciphering the genome,* Nucleic acids research **32**, D277 (2004).

[16] W. Liebermeister, E. Noor, A. Flamholz, D. Davidi, J. Bernhardt, and R. Milo, *Visual account of protein investment in cellular functions,* Proceedings of the National Academy of Sciences **111**, 8488 (2014).

[17] B. Schwanhäusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, and M. Selbach, *Global quantification of mammalian gene expression control,* Nature **473**, 337 (2011).

[18] L. Arike, K. Valgepea, L. Peil, R. Nahku, K. Adamberg, and R. Vilu, *Comparison and applications of label-free absolute proteome quantification methods on escherichia coli,* Journal of proteomics **75**, 5437 (2012).

[19] F. J. Franceschi and K. H. Nierhaus, *Ribosomal proteins l15 and l16 are mere late assembly proteins of the large ribosomal subunit. analysis of an escherichia coli mutant lacking l15.* Journal of Biological Chemistry **265**, 16676 (1990).

[20] S. Shoji, C. M. Dambacher, Z. Shajani, J. R. Williamson, and P. G. Schultz, *Systematic chromosomal deletion of bacterial ribosomal protein genes,* Journal of molecular biology **413**, 751 (2011).

[21] G. Kramer, T. Rauch, W. Rist, S. Vorderwülbecke, H. Patzelt, A. Schulze-Specking, N. Ban, E. Deuerling, and B. Bukau, *L23 protein functions as a chaperone docking site on the ribosome,* Nature **419**, 171 (2002).

[22] T. Nakayashiki and H. Mori, *Genome-wide screening with hydroxyurea reveals a link between nonessential ribosomal proteins and reactive oxygen species production,* Journal of bacteriology **195**, 1226 (2013).

[23] S. H. Coleman, B. A. Maguire, and D. G. Wild, *Ribosome assembly in three strains of escherichia coli with mutations in the rpmb, g operon,* Microbiology **139**, 707 (1993).

[24] B. A. Maguire and D. G. Wild, *The roles of proteins l28 and l33 in the assembly and function of escherichia coli ribosomes in vivo,* Molecular microbiology **23**, 237 (1997).

[25] B. A. Maguire and D. G. Wild, *Mutations in the rpmbg operon of escherichia coli that affect ribosome assembly,* Journal of bacteriology **179**, 2486 (1997).

[26] B. A. Maguire and D. G. Wild, *The effects of mutations in the rpmb, g operon of escherichia coli on ribosome assembly and ribosomal protein synthesis,* Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression **1353**, 137 (1997).

[27] M. Bubunenko, T. Baker, and D. L. Court, *Essentiality of ribosomal and transcription antitermination proteins analyzed by systematic gene replacement in escherichia coli,* Journal of Bacteriology **189**, 2844 (2007).

[28] W. A. Held, M. Nomura, and J. W. Hershey, *Ribosomal protein s21 is required for full activity in the initiation of protein synthesis,* Molecular and General Genetics MGG **128**, 11 (1974).

[29] J. Van Duin and R. Wijnands, *The function of ribosomal protein s21 in protein synthesis,* European Journal of Biochemistry **118**, 615 (1981).

[30] K. Potrykus, H. Murphy, N. Philippe, and M. Cashel, *ppgpp is the major source of growth rate control in e. coli,* Environmental microbiology **13**, 563 (2011).

[31] K. Walsh and D. E. Koshland Jr, *Characterization of rate-controlling steps in vivo by use of an adjustable expression vector.* Proceedings of the National Academy of Sciences **82**, 3577 (1985).

[32] M. J. Gruer, A. J. Bradbury, and J. R. Guest, *Construction and properties of aconitase mutants of escherichia coli,* Microbiology **143**, 1837 (1997).

[33] M. Basan, S. Hui, H. Okano, Z. Zhang, Y. Shen, J. R. Williamson, and T. Hwa, *Overflow metabolism in escherichia coli results from efficient proteome allocation,* Nature **528**, 99 (2015).

[34] A. R. Subramaniam, B. M. Zid, and E. K. O'Shea, *An integrated approach reveals regulatory controls on bacterial translation elongation,* Cell **159**, 1200 (2014).

[35] M. Scott and T. Hwa, *Shaping bacterial gene expression by physiological and proteome allocation constraints,* Nature Reviews Microbiology **21**, 327 (2023).

[36] K. Kochanowski, H. Okano, V. Patsalo, J. Williamson, U. Sauer, and T. Hwa, *Global coordination of metabolic pathways in escherichia coli by active and passive regulation,* Molecular systems biology **17**, e10064 (2021).

[37] M. J. Noga, F. Büke, N. J. van den Broek, N. C. Imholz, N. Scherer, F. Yang, and G. Bokinsky, *Posttranslational control of plsb is sufficient to coordinate membrane synthesis with growth in escherichia coli,* MBio **11**, e02703 (2020).

3

TITRATING PPGPP REVEALS ITS LIMITED CONTROL OVER RESOURCE ALLOCATION

Milan LACASSIN, Yaroslav BLANTER, Gregory BOKINSKY

With four parameters I can fit an elephant,
and with five I can make him wiggle his trunk.

John von Neumann to Enrico Fermi, reported by Freeman Dyson[1]

3.1. ABSTRACT

Since the discovery of the ppGpp signaling system, many attempts have been made to design a model explaining its role in maximizing bacterial growth rate in various nutrient conditions. However, these models often focus on subsets of the available relevant data or rely on empiric optimality of growth rate rather than the mechanisms through which ppGpp operates. None of these models have been applied to recent results of the impact of artificial ppGpp concentrations on ribosome content and growth rate. In this chapter, we start with the simplest bacterial growth model and add the necessary components upon this basic architecture to explain more results. In particular, we investigate the reasons for the observed divergence of artificially increased ppGpp concentration with the ones found in carbon limitation. After showing that this cannot be explained by a higher acyl-tRNA saturation of ribosomes, we demonstrate mathematically that a model of the ppGpp signaling system can not explain the effect of excess ppGpp on growth and ribosome content if it considers that ribosome abundance depends on ppGpp concentration only. We investigate whether expressing ribosome production as proportional to the RNA polymerase concentration, activated by nucleoside tri-phosphates and competitively inhibited by ppGpp could explain such result. We find that defining ribosome production in such a way does not give rise to the expected ribosome abundance, even in the natural case, highlighting the importance of a better understanding of the factors contributing to ribosome production. Finally, we provide results concerning other regulators and processes known or suspected to impact ribosome production. We found that most of these factors show no evidence that they could counteract the effect of excess ppGpp on ribosomes, with the exception of RNA polymerase availability and passive regulation which might arise from the lack of upregulation of catabolic proteins in excess ppGpp, in comparison with carbon limitation.

3.2. INTRODUCTION

PPGPP: A KEY REGULATOR IN BALANCING METABOLISM AND RIBOSOMES

The ppGpp regulatory system is a bacterial signalling system known to respond to various nutrient starvations and bacterial stresses. Apart from its role during starvation and stress, ppGpp also has a key role during steady-state growth: it downregulates some processes linked to growth, such as the production of ribosomes [2, 3], and upregulates processes involved in nutrient uptake and amino acid synthesis. By doing so in response to insufficient amino acid pool, ppGpp is thought to find the balance between growth (operated mostly by ribosomes) and the supply in amino acids [5], necessary to sustain the first process. For so-called high quality carbon sources, low amount of ppGpp is present, leading to large ribosome abundance to sustain fast growth. In poorer quality carbon, which requires more enzymes to lead to the same production of amino acids, more ppGpp is present, which downregulates ribosome content in favor of more enzymes responsible for import and digestion of carbon. This feedback maximizes the growth rate in various nutrient conditions. For a more detailed review of the way ppGpp is triggered and its effects, we refer the reader to Chapter 1.

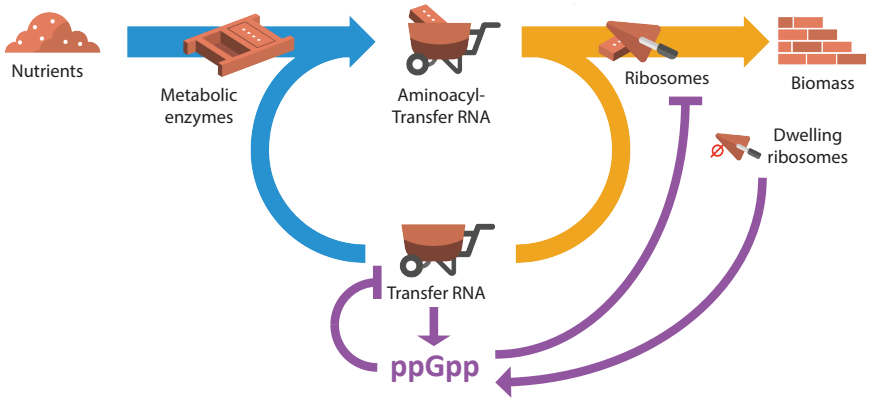


Figure 3.1: The ppGpp regulatory system acts on key actors of the growth process. Schematic view of the actors of bacterial growth and the role of ppGpp. A series of metabolic enzymes import and digest nutrients, in the form of a carbon and nitrogen source, into amino acid and bind them to the cognate transfer RNA (tRNA). The aminoacyl-tRNA are utilized by ribosomes to produce new proteins, the main component of bacterial biomass, thereby making these tRNAs available to be bound to amino acids again. In *E. coli*, ppGpp production is activated by deacyl-tRNA and dwelling ribosomes. ppGpp downregulates the production of new tRNAs and ribosomes.

PPGPP PERTURBATIONS: FROM KNOCKOUTS TO ALTERING CONCENTRATION

The first identifications of ppGpp's functions, like its control of ribosomal content, have been made through knocking out the ppGpp synthetases RelA and/or SpoT. Recent studies used a different approach by expressing various levels of either the catalytic domain of the ppGpp synthetase RelA or the ppGpp hydrolase MESH1. This approach allows to finely increase or decrease the ppGpp concentration to identify quantitative relationships between ppGpp concentration, growth rate and other relevant compounds independently from nutritional effects. This enabled investigations of the effect of ppGpp on cell size [4], the relationship between ribosomal content and growth rate in non-optimal ppGpp conditions [5], or the role ppGpp plays in ensuring timely adaptation to sudden amino acid starvation [6]. Finally, this approach allowed us to obtain results that better delineate the scope of ppGpp regulation on the proteome (Chapter 2).

LARGE AMOUNTS OF PPGPP SEEM TO BE NECESSARY TO FORCE DOWN GROWTH

Recent work investigating the coordination of phospholipid flux with growth [7], by measuring growth rate and ppGpp concentration, obtained a result which indicates that forcing down growth rate by inducing excess of ppGpp requires much higher concentrations of ppGpp than the ones found in the natural case. This is surprising as it clashes with the idea that ppGpp sets the ribosome abundance which is the main growth rate contributor. Indeed, following this idea ppGpp and growth rate should follow a single relationship when ribosomes are sufficiently supplied (which should be the case in excess ppGpp). This called for a mapping of ppGpp concentrations in relation with growth rate to better understand the effect of altering ppGpp levels and test whether this phenomenon can be observed in other conditions.

In Chapter 2, we used the same method as Noga et al. [7] and mapped the relationship between artificially titrated ppGpp concentrations and growth rate in various minimal media supplemented with different carbon sources. By doing so, we confirmed that, regardless of the condition, much larger concentrations are required to force down growth by artificially increasing ppGpp than the ones found in wild type bacteria growing at the same rate (Figure 3.2A). Throughout the current chapter, we propose a model attempting to explain this puzzling feature of the experimental data to which we refer to as "divergence of excess ppGpp with the carbon line". Can higher ribosome saturation explain divergence of excess ppGpp and the carbon line? Can NTPs activation of ribosomes counteract the effect of ppGpp on ribosomes to yield faster growth in excess ppGpp? Can other regulators or processes influence ribosome content differently in excess ppGpp and carbon limitation? What are the quantitative laws that dictate ppGpp production and its effects on growth? Before investigating these questions with our model, we briefly review existing ppGpp-signaling models focusing on the impact on growth rate, as well as relevant experimental data we based our model on.

EXISTING PPGPP SIGNALING MODELS OFTEN RELY ON GROWTH OPTIMALITY OR PREDICTED PARAMETERS AND HAVE NOT YET BEEN APPLIED TO PPGPP PERTURBATIONS

Based on their experiments that beautifully show that ppGpp responds to a slowdown in translation elongation rate, Wu et al. [8] built a model using this empirical observation for ppGpp synthesis. In turn, ppGpp not only downregulates the production of ribosomes but also inactivates ribosomes. Growth rate is then obtained as a product of the elongation rate and the active ribosome abundance. While this model is one of the most comprehensive models of ppGpp regulation and explain many results (ribosome abundance, ppGpp concentration, elongation rate and growth rate), it relies on the direct inactivation of ribosomes by basal ppGpp (which lacks sufficient evidence). Attempting to bridge the obtained relation between elongation rate and ppGpp production using known triggers, the authors had to introduce a predicted term, for which there is no experimental confirmation to date. Finally, this model has not been applied to ppGpp perturbations.

Erickson et al. [9] proposed another interesting model that considers the tradeoff between metabolism and ribosome content as governed by ppGpp. In turn, ppGpp influences the abundance of ribosomes and metabolic proteins able to metabolize a specific substrate. However, this model primarily focuses on explaining nutrient transition results and relies on optimal growth rather than explicating the way ppGpp is produced and acts on ribosome and metabolic enzyme abundance. This limits its applicability to explaining the effects of perturbing ppGpp concentrations on steady growth.

Another recent model by Chure et al. [10] also investigates the effect ppGpp has on the metabolic-ribosomal tradeoff. It considers ppGpp as proportional to the ratio of deacyl- to acyl-tRNA. Using this framework it considers three scenarios: fixed ribosomal production, optimized translation rate, optimized growth rate. By deriving the ribosome production as a function of ppGpp that corresponds to each of these scenarios it shows that the latter (optimized growth) is consistent with many experimental observations. However, rather than expliciting the mechanisms of ppGpp production and its action on ribosomes, it relies on mathematical solutions leading to a specific behavior, once again limiting its application to understand the divergence of excess ppGpp with the carbon line.

Finally, the model proposed by Bosdriesz et al. [11] is, to our knowledge, one of the only models to explicitly model ppGpp production and its effect on ribosome abundance based on the current understanding of the mechanisms involved: it considers ppGpp production as a constant production by SpoT plus a RelA production activated by ribosomes bound to uncharged tRNA. Effect of ppGpp on ribosome production is modelled as a competitive inhibition. This model successfully predicts key experimental results of ribosome abundance and nutritional shifts. Unfortunately, this model has not been applied to relevant results with artificially perturbed ppGpp concentrations, which came later. While starting with a simpler model, we inspired ourselves from these models and re-used some of their elements.

THE METABOLIC-RIBOSOMAL TRADEOFF MODEL: A GOOD STARTING POINT

Kohanim et al. [12] developed a simple model to explain a sudden increase in growth rate when switching to a more abundant carbon source. This model considers the presence of sub-saturated ribosomes that can quickly utilize the more abundant nutrient. This model does not include the ppGpp signaling system. Instead, it is built on an empirical optimal allocation that maximizes the growth rate. Due to the simplicity of the model, this might allow us to add the necessary features to explain how this optimal allocation is found through ppGpp, and to do it in a way that facilitates making predictions for ppGpp perturbations. Additionally, this model is designed to describe the saturation of ribosomes by substrates and predicted the presence of subsaturated ribosomes in poor nutrient conditions. This prediction sparked our interest and we thus set out to explore whether having subsaturated ribosomes could explain why higher levels of ppGpp are required to artificially hamper growth than the ones found in the natural case. Indeed, one could speculate that, oppositely to limiting the quality of the carbon supply, excess ppGpp keeps a higher ribosomal saturation than carbon limitation. This could lead to the higher growth rates observed in excess ppGpp for the same level of ppGpp (Figure 3.2 for details of this data). Based on these reasons, we start in the next section 3.3 with an adapted version of this model and include the necessary elements to explicitly incorporate the ppGpp regulatory system and its perturbations. Before doing so, we briefly describe experimental data we seek to reproduce and based our model on.

PERTURBING PPGPP LEVELS AWAY FROM THEIR NATURAL OPTIMUM INHIBITS GROWTH

The first result we hope to reproduce with our model is ppGpp concentration results obtained in Chapter 2. As mentioned earlier we want to explain perturbations of the ppGpp concentrations and how they impact growth. We found (Figure 3.2A), in accordance with previous results [13], that limiting carbon source quality increases ppGpp concentration, along with slower growth (gray dashed line in Figure 3.2A). We also induced ppGpp perturbations: both excess and depleted ppGpp had a negative impact on growth (dark and light colored dashed line respectively in Figure 3.2A), revealing the optimal behavior of WT ppGpp concentrations. We hope, by explaining with our model these concentrations, to gain understanding on how optimal growth is achieved by ppGpp.

PPGPP PERTURBATION IMPACTS RIBOSOME CONTENT

The second and key result we seek to reproduce is the so called first growth law which stipulates that ribosome content scales with bacterial steady-state growth rate for wild-type *E. coli* in minimal media supplemented with carbon sources of increasing quality

(supporting increasingly fast growth). This result has been for a long time attributed to down-regulation of ribosomes by ppGpp. Adding to this, Zhu et al. [5] more recently showed that perturbing ppGpp concentration affects ribosomal allocation in the following way: excess ppGpp leads to insufficient ribosomes and, oppositely, depleted ppGpp leads to excess ribosomes, both hampering growth. The ribosome content can be estimated through RNA:protein ratio as most of bacterial RNA is ribosomal and most of cellular mass is made of proteins. Relevant data of such ratio, obtained by Zhu et al. and Dai et al. [5, 14] is compiled in Figure 3.2B.

TRNA CONTENT SHOWS A ROUGHLY CONSTANT RATIO TO TOTAL RNA

Amino acids are bound to specific transfer RNA (tRNA) which are used by ribosomes to recognize amino acids and assemble them into proteins. The tRNAs are freed in the process and therefore act as a carrier cycle delivering amino acids to ribosomes, as depicted in Figure 3.1. Deacyl-tRNA have been reported to activate ppGpp synthesis, together with dwelling ribosomes waiting for the right acyl-tRNA to resume translation [15]. Implementation of the tRNA carrier cycle in a model might thus allow a better description of the way amino acids contribute to translation and growth and enable a mechanistic description of the way ppGpp synthesis is triggered. To do so in a way that respects experimental evidence, we are interested in how the tRNA pool varies with growth rate. Measurements of the fraction of RNA which is tRNA in carbon-limited cultures have been long available and we provide such results, obtained by Rosset et al. [16] in Figure 3.2C. These measurements show that, except at low growth rates, WT *E. coli* displays a rather constant tRNA to total RNA ratio; unfortunately such measurements of tRNA acylation are difficult to implement, and have not been performed during ppGpp titration.

TRANSLATION RATE DECREASES BOTH WITH CARBON LIMITATION AND EXCESS PPGPP

Finally, because we are interested in ribosomal saturation we also include in our benchmark data regarding the rate at which individual ribosomes are translating proteins: the translation elongation rate. Dai et al. [14] measured this rate for carbon limitation and excess ppGpp and kindly made their data available. These measurements show that both when inducing carbon limitation and excess ppGpp, the translation elongation rate decreases towards about two third of its maximum value at null growth rate (Figure 3.2D). In section 3.3.4, we compare these results to the one obtained through our model.

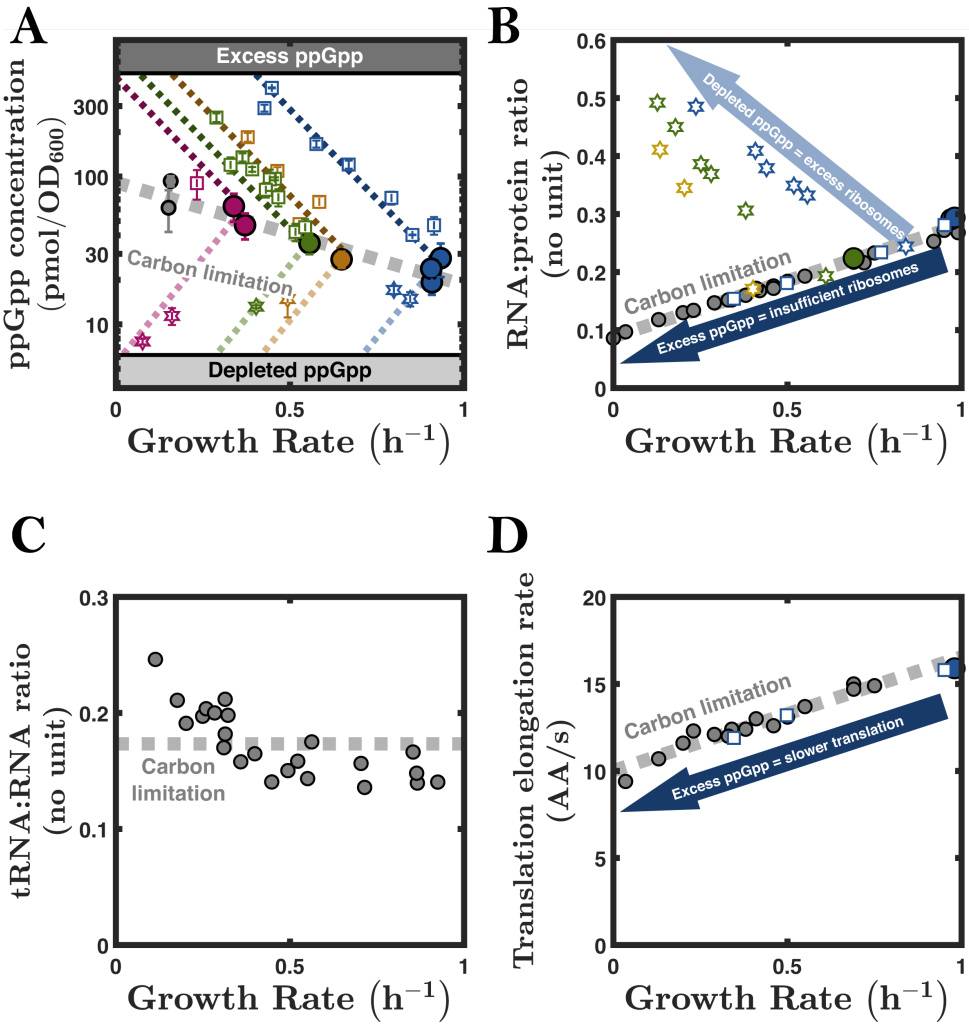


Figure 3.2: Benchmark of experimental results to validate ppGpp regulation model including ppGpp perturbations. (A) ppGpp concentration versus growth rate results. Circles represent wild-type *E. coli* in minimal media with various carbon sources, squares represent excess ppGpp through induction of a gene construct expressing the catalytic domain of the ppGpp synthetase RelA and stars represent depleted ppGpp through induction of the ppGpp hydrolase MESH1. Each color represents a carbon source: acetate in magenta, glycerol in green, succinate in ochre and glucose in blue. For more details on how this data was obtained, we refer the reader to Chapter 2. (B) RNA:protein ratio depending on growth rate. Colors and markers similarly as previous plot. Adapted from Zhu et al. and Dai et al. [5, 14]. (C) tRNA:RNA ratio in carbon limited wild-type *E. coli* cultures. Digitalized and adapted from Rosset et al. [16]. (D) Translation elongation rate measurements in carbon-limited cultures and under excess ppGpp. Adapted from Dai et al [14].

3.3. RESULTS

In this section 3.3 we start with a simple model inspired by the one proposed by Kohanim et al. [12] and describing the tradeoff between metabolic enzymes and ribosome abundance. After solving this simple model and comparing its prediction to experimental RNA:protein ratio data in the first part of this section 3.3.1, we incrementally add model features to attempt to explain more experimental evidence on growth in various carbon conditions and when ppGpp concentrations are perturbed. Firstly, in the second part of this section 3.3.2, in order to better describe the way amino acid are delivered to ribosomes and to facilitate the implementation of the ppGpp signaling system later on, we add the entity reponsible for providing amino acids to ribosomes: tRNA. We then explore how this addition changes the previously obtained predictions. Next, we implement the ppGpp regulatory system in the third part of this section 3.3.3 and attempt to explain both the existing RNA:protein data and the ppGpp concentrations we obtained in Chapter 2. In the next part of this section, 3.3.4, we explore the conditions a model needs to satisfy to explain both of these results. Then, in part 3.3.5, we develop a model of ribosome production including the activation by NTPs. Finally in the last part 3.3.6, we investigate through our experimental results obtained in Chapter 3, other regulators or processes that might influence ribosome production.

3.3.1. TRADEOFF MODEL EXPLAINS RIBOSOME CONTENT SCALING WITH GROWTH RATE

ADAPTING TRADEOFF MODEL TO QUANTITATIVELY APPROACH EXPERIMENTAL RESULTS

As we seek to build a model that quantitatively reproduces the measured data, we slightly adapted the qualitative model used by Kohanim et al. by adding scaling parameters allowing the ribosomal mass fraction to vary in the range $[0; R_{max}]$ instead of $[0; 1]$. We present here the adapted version and refer the reader to Ref. [12] for the version of the model they developed. A schematic view of our version of the model is presented in FigRCtradeoffA In this model, bacterial growth is envisioned in a highly simplified manner that consists of two steps: nutrients are metabolized into a single precursor entity defined by its mass fraction[1] x, which is consumed by the growth process occurring at rate μ. The time derivative of this precursor therefore has a production term γ and a consumption term μ. This precursor, like other entities in this model, is also diluting through growth, adding a $-\mu x$ term. The precursor dynamics are thus described as follows:

$$\frac{dx}{dt} = \gamma - \mu - \mu x, \tag{3.1}$$

The metabolic process is driven by catabolic enzymes grouped into a metabolic sector represented by its mass fraction M. These enzymes metabolize nutrients at the rate γ proportional to M and the nutrient quality n. The nutrient quality describes the resource efficiency of each nutrient source, namely how much of this nutrient can be metabolized per unit of time and per metabolic enzyme. Additionally, for stability of the model and in order to better describe the metabolic process, the activity of metabolic enzymes is feedback inhibited by the precursor x, bringing the Michaelis-Menten multiplicative term

[1] For complete description of the definition through mass fraction see section 3.7.1.

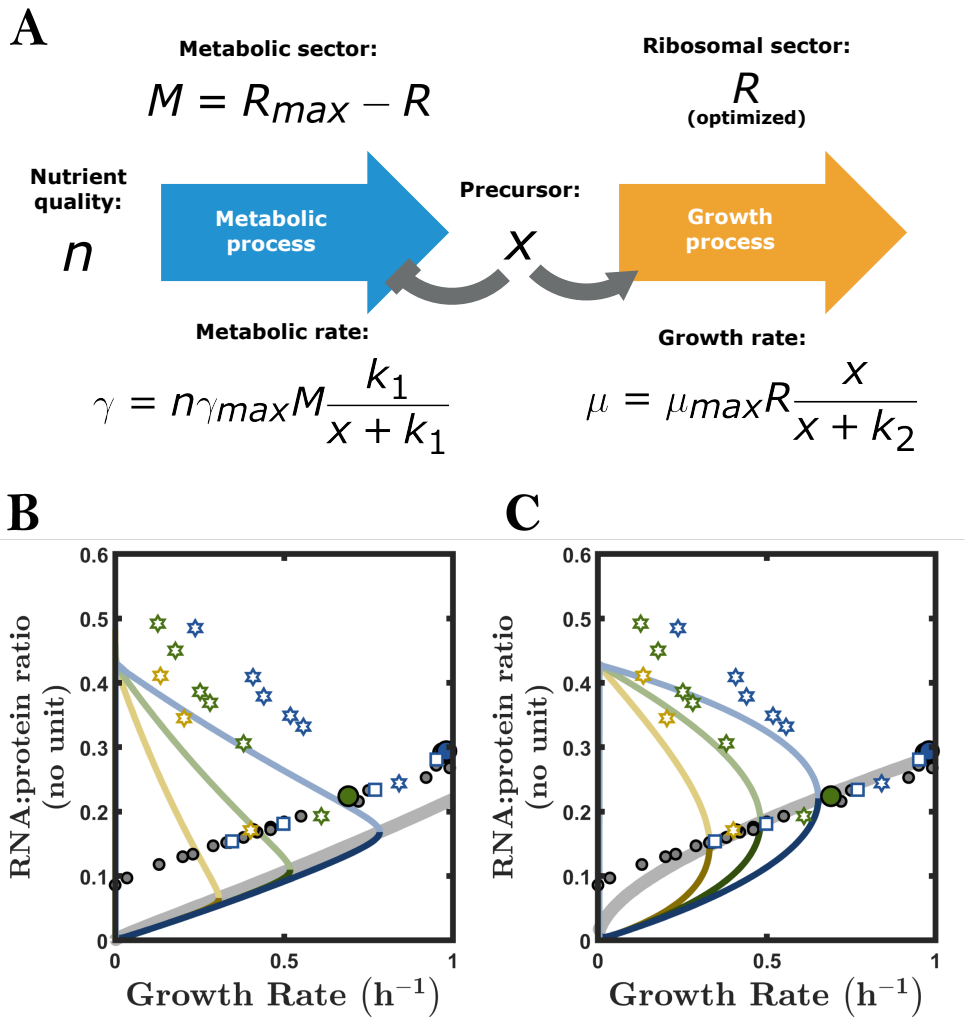


Figure 3.3: Minimal tradeoff model displays ribosomal allocation optimum scaling with growth rate in increasing nutrient quality (A) Schematic view of minimal model inspired by Kohanim et al. [12], adapted and described in section 3.3.1. (B-C) Predictions of this model for RNA:protein ratio in comparison corresponding experimental data. Model predicitions are obtained using two different parameter sets: (C) parameters equivalent to the ones found by Kohanim et al. with subsaturation of ribosomes $k_2/k_1 = 1$, (B) parameters with high ribosomal saturation $k_2/k_1 = 0.01$. Grey line represents model prediction for optimal RNA:protein ratio following a nutrient quality n scan. Optimal RNA:protein ratio is obtained from optimal ribosome mass fraction $R^\star$ through $R/P = R^\star/(1-R^\star)$. Dark colored lines represent insufficient ribosomes $R < R^\star$ and light colored lines represent excess ribosomes $R > R^\star$. Each color correspond to one carbon as in previous FigBenchmark to which is attributed a nutrient quality n. Experimental data points are displayed in the same way as in FigBenchmark.

$\frac{k_1}{x+k_1}$. This leads to the following metabolic rate:

$$\gamma = n\gamma_{max} M \frac{k_1}{x+k_1}. \tag{3.2}$$

Based on observation that mass fractions of metabolic and ribosomal proteome sectors sum a to a constant value [17], this model assumes that the metabolic sector mass fraction M and the ribosome mass fraction R follow the relationship:

$$M = R_{max} - R, \tag{3.3}$$

where the parameter R_{max} is the maximal ribosome content when no resources are allocated to metabolism or, in other words, the total mass fraction of metabolic and ribosomal sectors[2]. This equation describes the tradeoff between allocating newly made resources to metabolism to produce precursors or to ribosomes for growth. While growth is composed of various processes such as transcription, translation and membrane synthesis, a large fraction of the biomass is made of proteins [18] and production of other components strongly rely on proteins. Therefore, growth is here considered as the translational flux only and other processes are neglected. Because ribosomes are responsible for translation, the growth rate scales with the mass fraction of ribosomes R. The precusor x is considered as a substrate of the ribosomes necessary for them to drive growth, which is described by the Michaelis-Menten type term $\frac{x}{x+k_2}$. This term describes how the translation rate per ribosomes decreases when precursor supply is insufficient. This leads to the following equation for growth rate:

$$\mu = \mu_{max} R \frac{x}{x+k_2}. \tag{3.4}$$

SOLVING TRADEOFF MODEL TO FIND OPTIMAL RIBOSOME ABUNDANCE

Kohanim et al. used an analytical solution of the model but for consistency with the following presented models we use numerical solving to obtain μ depending on R. At various nutrient qualities n, we scanned values of R and computed the corresponding growth rate by numerically solving the time derivatives of $(x;R;M)$ equal to zero. For each nutrient quality, we found one optimal ribosomal mass fraction $R^\star$ yielding the maximum growth rate for the value of n. To this value corresponds the optimal metabolic allocation $M^\star = R_{max} - R^\star$. To describe the way the mass fractions of these two sectors evolve towards a new optimum, we introduce similarly as Kohanim et al. the dynamic equations $\frac{dM}{dt} = \mu M^\star - \mu M$ and $\frac{dR}{dt} = \mu R^\star - \mu R$. These dynamic equations, by using a production term proportional to growth and the optimal allocation in competition with dilution through growth, consistently find the optimal allocation at steady state. To summarize, the obtained model can be described with the following equations:

$$\begin{cases} \frac{dx}{dt} = \gamma - \mu - \mu x \\ \frac{dR}{dt} = \mu R^\star - \mu R \\ \frac{dM}{dt} = \mu\left(R_{max} - R^\star\right) - \mu M \end{cases} \quad \text{With:} \quad \begin{cases} \gamma = n\gamma_{max} M \frac{k_1}{x+k_1} \\ \mu = \mu_{max} R \frac{x}{x+k_2} \\ M = R_{max} - R \end{cases}, \tag{3.5}$$

[2] Kohanim et al. used the equation $M = 1 - R$, which is qualitatively similar but does not allow to quantitatively reproduce experimental data.

which leads, by taking all time derivatives equal to zero, to the following steady-state equations describing steady growth:

$$\begin{cases} x = \frac{\gamma}{\mu} - 1 \\ R = R^\star \\ M = R_{max} - R^\star \\ \gamma = n\gamma_{max}\left(R_{max} - R^\star\right)\frac{k_1}{x+k_1} \\ \mu = \mu_{max}R^\star\frac{x}{x+k_2} \end{cases} . \tag{3.6}$$

ADAPTED TRADEOFF MODEL DISPLAYS INCREASING OPTIMAL RIBOSOME CONTENT

With this model the authors of Kohanim et al. were able to reproduce the sudden jump in growth rate after a shift in nutrient abundance at the conditions that $k_1 \ll k_2$, which they interpreted as ribosome subsaturation as $\mu < \mu_{max}R^\star$ in most conditions. They also predicted an increasing optimal ribosomal allocation $R^\star$ with increasing nutrient quality. We are interested in repoducing ribosome content measurements as we seek to explain how ppGpp is able find the optimal ribosome abundance. Ribosome abundance has been inferred through RNA:protein ratio [5, 14], in wild-type *E. coli* and in strains with artificially perturbed ppGpp levels. We assumed the vast majority of the biomass to be made of proteins and ribosomal RNA, as has been shown to be the case [9]. In other words, writing P the protein mass fraction, we have: $R + P = 1$. Therefore we can obtain the RNA:protein with:

$$\frac{R}{P} = \frac{R}{(1-R)}. \tag{3.7}$$

For each value of nutrient quality n, we get one curve of $^{R}/_{P}$ with a maximum with regards to μ corresponding to $R = R^\star$. We split these curves between excess ribosomes $R > R^\star$, depicted with light colored lines, and insufficient ribosomes $R < R^\star$, depicted with dark colored lines, and display three of these curves in FigRCtradeoffB-C. For a large number of these curves scanning values of n, we also identify the evolution of the optimal RNA:protein ratio yielding the highest growth rate. We find that our model predicts, accordingly with experimental results, scaling of the optimal RNA:protein ratio with growth rate.

RIBOSOME SUBSATURATION IMPLIES SPLIT BETWEEN OPTIMAL AND INSUFFICIENT RIBOSOMES

Results in 3.3B correspond to high ribosome saturation $k_1 \ll k_2$. We compare these RNA:protein ratio results with results for ribosome subsaturation as identified by Kohanim et al.: $k_1 = k_2$ (3.3C). Because excess ppGpp induces insufficient ribosomes, leading to a lower growth rate, we are interested in the model predictions for insufficient ribosomes $R < R^\star$. In contrast with experimental evidence of insufficient ribosomes through excess ppGpp (Figure 3.3B), the model with $k_1 = k_2$ as proposed by Kohanim et al. because it explains best their shift results, predicts that insufficient ribosomes lowers RNA:protein ratio following a different relationship than the optimal RNA:protein ratio with regards to growth rate, as shown in Figure 3.3C. More precisely, for a same RNA:protein, insufficient ribosomes in a high nutrient quality yields a higher growth rate than a poor nutrient quality. Interestingly, we found that with $k_1 \ll k_2$, which corresponds to higher saturation

of ribosomes, predicted an overlap of the relationship between optimal RNA:protein ratio with the RNA:protein ratio arising from insufficient ribosomes, which resembles experimental data better.

3.3.2. ADDING THE TRNA CARRIER CYCLE RETAINS OPTIMALITY OF RIBOSOME CONTENT

Tranfer RNA (tRNA) delivers amino acids to the ribosomes to have them assembled into proteins. They are also, together with dwelling ribosomes, thought to activate the main ppGpp synthetase RelA. For these reasons, including tRNA in our model rather than a single precursor might improve our description of ribosomal saturation while allowing us to explicitly write the ppGpp regulation. In the next paragraph, we introduce tRNA and how they act as a cycle acheminating amino acids to ribosomes before investigating the consequences of adding this cycle on the model predictions. By implementing this cycle in the resource-allocation model, we hope to bridge the notion of ribosomal saturation with ppGpp regulation by tRNA pools.

IMPLEMENTING THE TRANSFER RNA CARRIER CYCLE

We add the tRNA carrier cycle to our model by grouping all tRNAs together into two entities: deacyl-tRNAs described by their mass fraction θ and acyl-tRNAs with mass fraction T. We consider the whole metabolic pathway as one process converting deacyl-tRNA into acyl-tRNA. A schematic view of the resulting model is provided in Figure 3.4A. In the previous model, the metabolic process, driven by the metabolic sector M is only regulated by feedback inhibition by the precursor. While we can include feedback inhibition by acyl-tRNA in this model, including only this regulation leads to metabolic rate greater than zero when no deacyl-tRNAs are present: $\gamma(\theta = 0, T \neq 0) > 0$. This would be absurd as deacyl-tRNAs are necessary substrates of the metabolic process. To prevent this and include the need for deacyl-tRNA to be able to produce acyl-tRNA we introduce in the metabolic rate a Michaelis-Menten type term on θ: $\frac{\theta}{\theta+k_\theta}$. As this term has a non-identical but similar effect as inhibition by T, namely that higher charging ratio of tRNA leads to slower metabolic process, we used only the term describing the need for uncharged for the sake of simplicity. This leads to to the metabolic rate:

$$\gamma = n\gamma_{max}(R_{max} - R)\frac{\theta}{\theta + k_\theta}. \tag{3.8}$$

For the growth rate, we keep a similar expression, replacing the precursor x by acyl-tRNA described by its mass fraction T as follows:

$$\mu = \mu_{max} R \frac{T}{T + k_T}. \tag{3.9}$$

TRNA ABUNDANCE SCALING WITH RIBOSOME CONTENT

In order to successfully add the tRNA cycle to our model and describe its dynamics, we must assess how the tRNA pool scales with growth rate and ribosomes in different conditions. Rosset et al. [16] have measured the fraction of RNA that is transfer RNA. They found that this fraction stays rather constant in different carbon sources yielding various growth rates, except for slow growth with $\mu < 0.3h^{-1}$ for which it is slightly higher. As a first approximation, we assume this fraction to be constant. As tRNA and ribosomal RNA

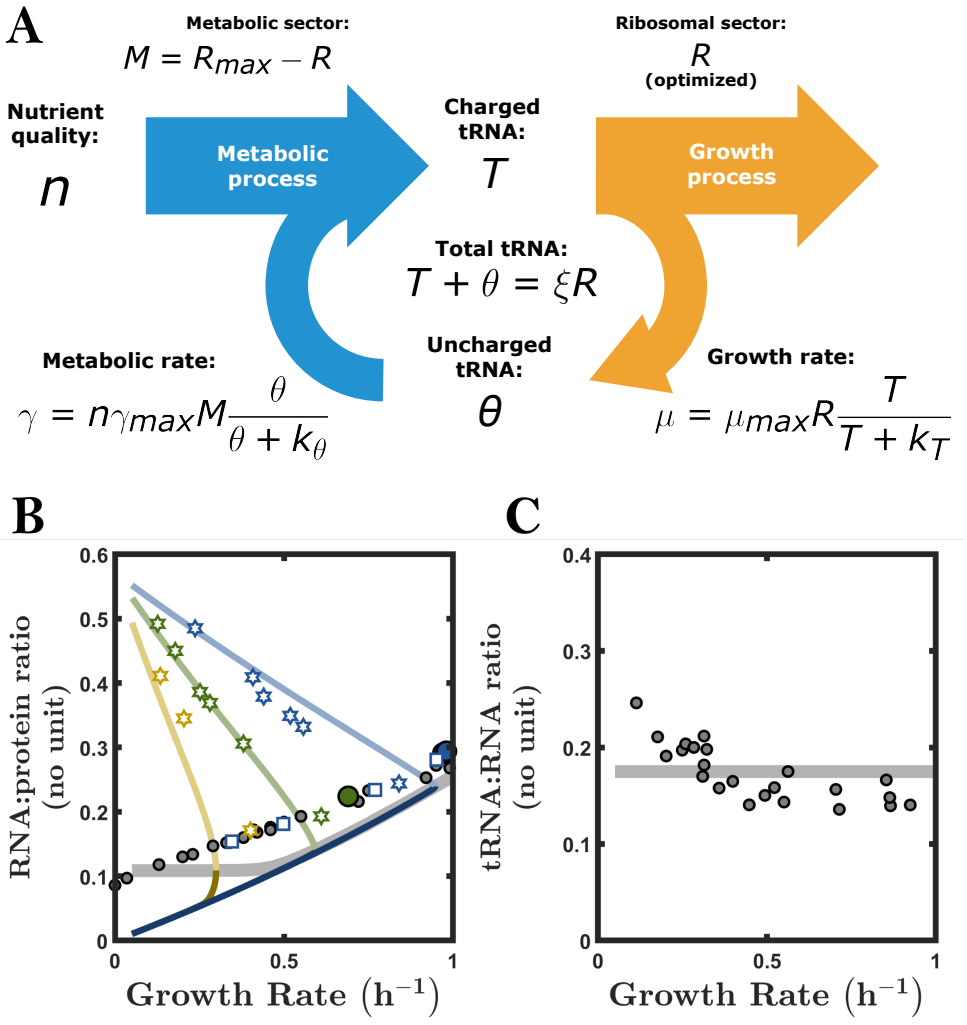


Figure 3.4: Adding transfer RNA carrier cycle to minimal model preserves most model features and allows to explain non-zero RNA:protein ratio at null growth rate. (A) Schematic view of the model with tRNA carrier cycle. (B-C) Model predictions and data displayed similarly as previous Figures 3.2 & 3.3: grey line represents optimal behavior, dark colored lines describe insufficient ribosomes $R < R^\star$ and light colored lines, excess of ribosomes $R > R^\star$. (B) RNA:protein ratio versus growth rate predictions of this model in comparison with experimental data. RNA:protein ratio is obtained through $R/P = R/(1-R)$ (C) Fraction of RNA that is transfer RNA as a function of growth rate in carbon limitation, comparison of model prediction and experimental data. In this graph, colored lines corresponding to lack and excess of ribosomes fully overlap with optimal behavior and are thus hidden by the grey line describing this behavior.

constitute the vast majority of bacterial RNA [18], we neglect the mass fraction of other types of RNA. With this assumption, the measurements of tRNA:RNA ratio correspond to $(\theta + T) / (R + \theta + T)$. This quantity being constant leads to $\theta + T \propto R$. In other words: total tRNA content is proportional to ribosome content. We define a new parameter $\xi = (T + \theta) / R$, which describes how the production of tRNAs scales with the production of ribosomes.

SOLVING THE TRNA CARRIER CYCLE MODEL

Similarly as in the previous model the ribosomal sector is optimized to $R = R^\star$ achieving the highest possible growth rate for a given nutrient quality. To be able to describe kinetics of this model and solve it numerically through finding steady state, we describe the dynamics by writing time derivative as follows:

$$\begin{cases} \frac{dT}{dt} = \gamma - \mu - \mu T \\ \frac{d\theta}{dt} = \mu - \gamma + \xi R^\star - \mu\theta \\ \frac{dR}{dt} = \mu R^\star - \mu R \\ \frac{dM}{dt} = \mu\left(R_{max} - R^\star\right) - \mu M \end{cases} \text{, with: } \begin{cases} \gamma = n\gamma_{max}(R_{max} - R^\star)\frac{\theta}{\theta + k_\theta} \\ \mu = \mu_{max} R^\star \frac{T}{T + k_T} \end{cases} . \quad (3.10)$$

These equations correspond to the steady-state values:

$$\begin{cases} T = \frac{\gamma}{\mu} - 1 \\ \theta = \xi R - \frac{\gamma}{\mu} + 1 \\ R = R^\star \\ M = R_{max} - R^\star \end{cases} . \quad (3.11)$$

We analyzed this model to investigate how adding this cycle affects the predictions on RNA:protein ratio. This model can be solved analytically to find $\mu(R, n)$. We provide such solution in supplementary materials 3.7.2 for any reader who wishes to use it. For consistency with following models, after checking that it matches the analytical solution, we analyze this model by minimizing derivatives for each set of (R, n) inputs considered. We obtained, for each set of input, a solution (T, θ, R, M) from which we can deduce the growth rate $\mu(R, T)$ and the RNA:protein ratio. For the latter we use, as previously, the assumption that proteins and RNA constitute the vast majority of bacterial mass and now include tRNA as follows:

$$\frac{R}{P} = \frac{R + T + \theta}{1 - R - T - \theta}. \quad (3.12)$$

ADDING TRNA CYCLE PRESERVES SCALING OF RIBOSOME CONTENT WITH GROWTH RATE

Similarly as with the previous model, we identify the optimal ribosomal allocation yielding the highest growth rate for a range of nutrient quality values. We find that, as expected due to the way this model is constructed, we obtain a tRNA:RNA ratio that is constant both in optimal and non-optimal ribosomal allocation, see Figure 3.4C. We also found that RNA:protein ratio prediction was similar though not identical to the simpler model investigated in the previous section 3.3.1 (Figure 3.4B).

THE TRNA CYCLE ALLOWS NON-ZERO RIBOSOME ABUNDANCE AT NULL GROWTH RATE

One of the main differences in the predictions of this model in comparison with the previous model is that the split between the carbon limitation and the excess ppGpp curve happens more drastically only under a specific growth rate, which is around $0.4h^{-1}$. Secondly, at very low nutrient quality, this model is able to achieve null growth rate $\mu = 0$ with RNA:protein ratio $R/P > 0$ as is observed in various studies investigating this ratio [5, 13, 14]. Therefore, at least in that sense, adding the tRNA carrier cycle improved the ability of the model to reproduce experimental results. However, the RNA:protein ratio corresponding to optimal behavior displays a kink which we do not see in the experimental data.

3.3.3. PPGPP SCALING WITH DWELLING RIBOSOMES REPRODUCES NATURAL PPGPP LEVELS AND FINDS OPTIMAL RIBOSOME CONTENT

In this section, we present an extension of the model analyzed in the previous section 3.3.2 by replacing the optimized ribosomal content that maximizes growth rate with a regulation through the ppGpp signaling system determining ribosome production. By doing so, we try to find a simple rule that explains how ppGpp finds the optimal ribosome content and make predictions on the effect of perturbing ppGpp concentrations on RNA:protein ratio and translation elongation rate.

IMPLEMENTING PPGPP PRODUCTION PROPORTIONAL TO DWELLING RIBOSOMES

As ppGpp is not a major component of bacterial mass and our experimental results assessed ppGpp concentration, we use its concentration (G) rather than using a mass fraction like for other variables. Moreover, due to the significant degradation ppGpp is subject to [19], we neglect dilution of ppGpp through growth and write the derivative of the ppGpp concentration as follows:

$$\frac{d}{dt}G = \Omega - \delta G, \tag{3.13}$$

with G the concentration of ppGpp, Ω the production of ppGpp, and δ a parameter describing the passive degradation of ppGpp. In the case of carbon limitation, production of ppGpp mostly happens through the ppGpp synthetase RelA sensing dwelling of ribosomes which are waiting for the right acyl-tRNA to continue translating [20]. Through this sensing, ppGpp is triggered by a slow down in translation [8]. For more details about this mechanism, we refer the reader to Chapter 1 section 1.2. We therefore write $\Omega = \Omega_0 R_{dwell}(R,T)$ with Ω_0 a newly introduced parameter and $R_{dwell}(R,T)$ the mass fraction of dwelling ribosomes. To explicit this mass fraction, we re-write the growth rate by including the translation elongation rate, namely the rate at which an average ribosome is assembling amino acids. We do so in the following way:

$$\mu = \mu_0 R \sigma(T) \qquad \text{With:} \quad \sigma(T) = \sigma_{max}\frac{T}{T+k_T}, \tag{3.14}$$

with μ_0 the growth rate per ribosome and per translation elongation unit and σ_{max} the maximal translation elongation rate. Apart from allowing us to compare our results to experimental translation elongation rate data, we can also now define dwelling ribosome

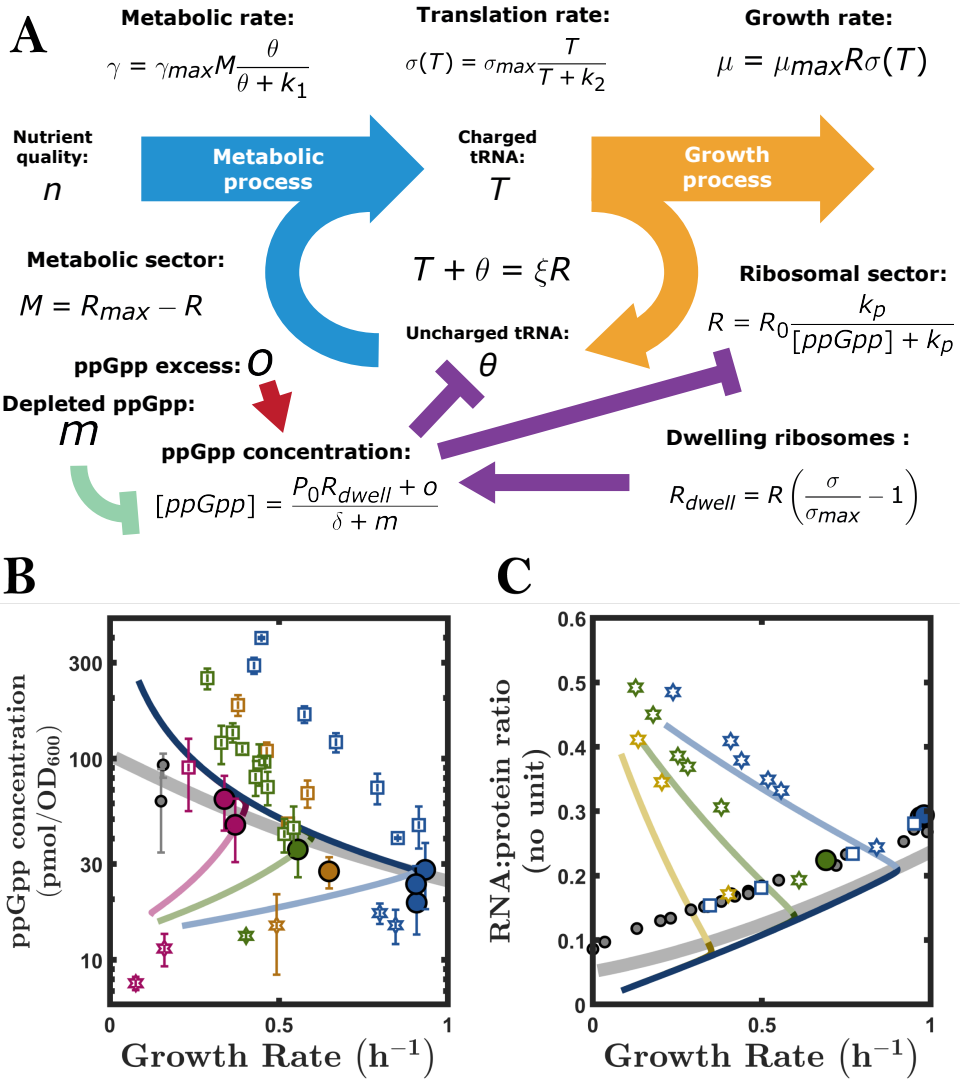


Figure 3.5: Implementing ppGpp regulation as a ribosomal allocation inhibitor activated by dwelling ribosomes is sufficient to find almost exact optimal allocation and predicts higher levels of ppGpp in excess ppGpp, with difference with carbon-limitation far smaller than experimental data. (A) Schematic view of the model with implementation of ppGpp regulation. (B-C) Model predictions and experimental data are displayed in the same way as earlier Figures 3.2-3.4 with the exception that dark and light colored lines now represent excess and depleted ppGpp respectively and induce lack and excess of ribosomes, and the grey line represents prediction for wild-type *E. coli*. The wild-type prediction is obtained scanning values of n with $o = 0$ and $m = 0$. Excess and depleted ppGpp predictions are obtained with $(o > 0; m = 0)$ and $(o = 0; m > 0)$ respectively at a fixed n value for each color corresponding to the different carbon conditions tested experimentally. (B) ppGpp concentrations predicted by the model in comparison with experimental data. (C) Comparison of predicted RNA:protein ratio predictions with experimental data.

mass fraction as follows:

$$R_{dwell}(R,T) = R\left(\frac{\sigma(T)}{\sigma_{max}} - 1\right) = R\frac{k_T}{T+k_T}. \tag{3.15}$$

The following equation describes ppGpp concentration dynamics:

$$\frac{d}{dt}G = \Omega_0 R_{dwell}(R,T) - \delta G. \tag{3.16}$$

ADDING EFFECT OF PPGPP ON RIBOSOME PRODUCTION

Next, we modeled the repressive effect of ppGpp on ribosome abundance. To do so, we define $\rho(G)$ as the fraction of newly made biomass that is ribosomes such that:

$$\frac{dR}{dt} = \mu\rho(G) - \mu R. \tag{3.17}$$

As ppGpp is known to bind to RNA polymerase attempting to transcript ribosomal operons and inhibit their activity, we assumed that this fraction $\rho(G)$ follows a Michaelis-Menten type inhibition relationship with ppGpp:

$$\rho(G) = R_0\frac{k_G}{G+k_G}, \tag{3.18}$$

where R_0 and k_G are newly introduced parameters. R_0 describes the maximal ribosomal mass fraction achieved at $G=0$ and k_G the concentration of ppGpp necessary to decrease ribosomal allocation to half of R_0. We also maintain the tradeoff relationship between the between metabolic and the ribosomal sector described by $M+R=R_{max}$. This leads to the following time derivatives for M and R:

$$\begin{cases} \frac{dR}{dt} = \mu R_0\frac{k_G}{G+k_G} - \mu R \\ \frac{dM}{dt} = \mu\left(R_{max} - R_0\frac{k_G}{G+k_G}\right) - \mu M \end{cases}. \tag{3.19}$$

IMPLEMENTING PPGPP PERTURBATION WITH TWO NEW INPUTS

Finally, to be able to simulate excess and depleted ppGpp, for which we produced experimental results described in Chapter 2, we add two inputs which we name o and m. The first one adds to the ppGpp production to describe expression of the catalytic domain of RelA used in experiments while the second one adds to the passive degradation of ppGpp to describe expression of MESH1. We can thus write the ppGpp concentration derivative including these inputs as follows:

$$\frac{d}{dt}G = \Omega_0 R_{dwell}(R,T) + o - (\delta + m)G. \tag{3.20}$$

SOLVING STRINGENT RESPONSE MODEL

To summarize, we have the following set of dynamic equations for this model:

$$\begin{cases} \frac{dT}{dt} = \gamma - \mu - \mu T \\ \frac{d\theta}{dt} = \mu - \gamma + \mu\xi R_0\frac{k_G}{G+k_G} - \mu\theta \\ \frac{dR}{dt} = \mu R_0\frac{k_G}{G+k_G} - \mu R \\ \frac{dM}{dt} = \mu\left(R_{max} - R_0\frac{k_G}{G+k_G}\right) - \mu M \\ \frac{d}{dt}G = \Omega_0 R_{dwell} + o - (\delta+m)G \end{cases} \quad \text{With:} \quad \begin{cases} \gamma = n\gamma_{max}(R_{max}-R)\frac{\theta}{\theta+k_\theta} \\ \mu = \mu_0 R\sigma \\ \sigma = \sigma_{max}\frac{T}{T+k_T} \\ R_{dwell} = R\left(\frac{\sigma(T)}{\sigma_{max}} - 1\right) \end{cases}. \tag{3.21}$$

These equations correspond to the steady-state values:

$$\begin{cases} T = \frac{\gamma}{\mu} - 1 \\ \theta = 1 - \frac{\gamma}{\mu} + \xi R \\ R = R_0 \frac{k_G}{G + k_G} \\ M = R_{max} - R \\ G = \frac{\Omega_0 R_{dwell} + o}{\delta + m} \end{cases} . \tag{3.22}$$

These equations, with the exception of the ones governing the ribosome mass fraction R and the addition of the ppGpp concentration G, are identical to the previous model analyzed in the previous section 3.3.2. The steady-state equations for this model lead, for $\mu(n, o, m)$, to a quintic equation which is to our knowledge not solvable. Even if one manages to solve this equation, the solution would likely be difficult to analyze. For these reasons, we solved this equation numerically by minimizing the time derivatives written above.

OPTIMIZING STRINGENT RESPONSE MODEL PARAMETERS

The number of parameters of this model makes their manual tuning very difficult. Exploring the parameter space in such a way to assess if the architecture of this model can reproduce the experimental result is a strenuous task. For these reasons, we optimized the parameters of this model by minimizing the square of the distance between all experimental data points shown in Figure 3.2 and the corresponding model predictions. For more details about the optimization process, please refer to section 3.6.

PPGPP SCALING WITH DWELLING RIBOSOMES APPROACHES OPTIMAL RIBOSOME CONTENT

We found that, with optimized parameters, the modeled ppGpp regulation finds a resource allocation that is not exactly the optimal one but well approximates it, as is visible in Figure 3.5B-C by the gray line intersecting colored lines close to their maximum in terms of growth rate. It is also able to reproduce reasonably well the RNA:protein ratio (Figure 3.5C). As the distance to the optimal allocation is minimal and likely close to experimental errors, we can say that modeling the ppGpp prodution proportionally to dwelling ribosomes only and inhibiting ribosomes with a Michealis-Menten type inhibition is sufficient to find optimal ribosomal allocation.

MODEL PREDICTS SLIGHTLY HIGHER LEVELS OF PPGPP NECESSARY TO HAMPER GROWTH THROUGH EXCESS PPGPP

Prediction for excess ppGpp yielded higher growth rate than limiting nutrient quality for the same concentration of ppGpp, see Figure 3.5B. This is because the predicted translation elongation rate is higher in excess ppGpp due to accumulation of acyl-tRNA, see Figure 3.7. The higher translation elongation rate is due, in the model, to a higher ribosome saturation because of more abundant acyl-tRNAs. However, the amplitude of the split between the predicted curve obtained when inducing excess ppGpp and the one obtained from decreasing nutrient quality is far from matching our data. Additionally, the predicted ppGpp concentrations with regards to growth rate in different carbon conditions quickly join and follows the same relationship. This is however not the case for our ppGpp concentration measurements, which seem to follow parallel lines with regards to growth rate.

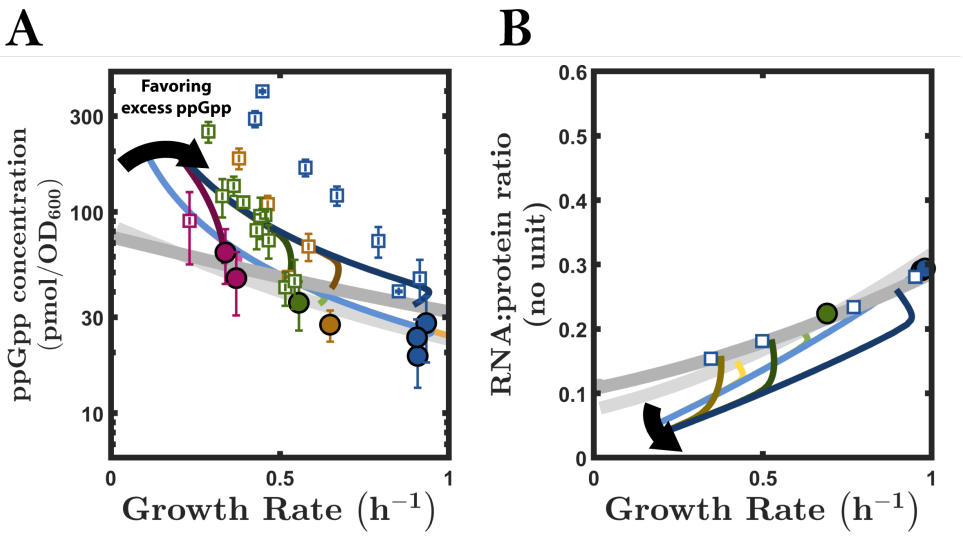


Figure 3.6: Larger difference in ppGpp level of carbon limitation and excess ppGpp through ribosomal saturation implies larger differences in RNA:protein ratio. Predictions of ppGpp concentration (A) and RNA:protein (B) of wild-type *E. coli* and under excess ppGpp, for two version of the model described in section 3.3.4. Predictions are obtained similarly as in Figure 3.5. experimental data and model predictions are displayed similarly as previously with the exception that predications for depleted ppGpp are omitted. All colored lines thus correspond to excess ppGpp. The first version, displayed with lighter colors corresponds to the one in Figure 3.5. The second one, displayed with darker colors, corresponds to the same model with different parameters obtained following an optimization favoring the match between predictions and experimental values of ppGpp concentration in when excess ppGpp is induced and ignoring distance between RNA:protein ratio predictions and experimental data in excess ppGpp.

3.3.4. RIBOSOME SATURATION CANNOT EXPLAIN BOTH PPGPP LEVELS AND RIBOSOME CONTENT IN EXCESS PPGPP

In Figure 3.5, we displayed predictions of our model concerning ppGpp concentration and RNA:protein ratio. We found that this model predicts that inducing excess ppGpp follows a relationship with growth rate with slightly higher ppGpp concentration than the natural level at a same growth rate. However, the difference between these two relationships linking ppGpp concentration and growth rate is much smaller than the one we observed experimentally. In an effort to understand why this model is unable to reproduce the concentrations of ppGpp seen when excess ppGpp is induced, we attempt in this section 3.3.4 to increase this split while conserving other predictions as close to experimental data as possible by using a parameter optimization favoring shorter distance to excess ppGpp concentrations data.

ATTEMPTING TO MATCH EXCESS PPGPP EXPERIMENTAL CONCENTRATIONS

We implemented the same optimization process as previous by giving a higher weight to distances between predicted and measured ppGpp concentrations in excess ppGpp. We also removed from the optimization the distance to experimental measurements

in excess ppGpp of RNA:protein ratio and translation elongation rate. This way, we favored parameters which reproduce the split in ppGpp versus growth rate between excess ppGpp and nutrient quality limitation, at the expense of RNA:protein ratio and translation elongation rate results.

MORE PPGPP REQUIRED TO INHIBIT GROWTH IMPLIES LARGER SPLIT BETWEEN OPTIMAL AND EXCESS PPGPP RIBOSOME CONTENT

With the optimization process favoring minimal distance to excess ppGpp concentrations described in the previous paragraph, we found that our model can approach excess ppGpp measurements slightly better in terms of ppGpp, but also that favoring this affects the RNA:protein ratio predictions, see Figure 3.6. Indeed, it seems that a larger split between excess ppGpp and nutrient quality limitation in the ppGpp versus growth rate plane implies a larger split in the $^{R}/_{P}$ versus μ plane, which conflicts with data from Zhu et al. [5]. Their data shows a clear overlap of the RNA:protein ratio when limiting growth through poorer carbon sources and excess ppGpp. This can also be interpreted mathematically from our equations. Indeed, our model achieves higher growth rates in excess ppGpp due to faster translation elongation, in other words $\sigma_{WT}(\mu) < \sigma_o(\mu)$ with WT corresponding to carbon limited wild-type *E. coli* and o referring to our engineered strain with excess ppGpp. As $\mu = \mu_0 R\sigma$, this necessarily leads to $R_{WT}(\mu) > R_o(\mu)$, meaning that for the same growth rate, the wild-type strain has more ribosomes, which leads to a higher RNA:protein ratio as ribosomes consitute a majority of bacterial RNA. Moreover, the bigger the difference $\sigma_o(\mu) - \sigma_{WT}(\mu)$ the larger the $R_{WT}(\mu) - R_o(\mu)$ difference.

EXPLAINING DIVERGENCE OF EXCESS PPGPP WITH THE CARBON LINE THROUGH RIBOSOME SATURATION IMPLIES FASTER TRANSLATION ELONGATION RATE IN EXCESS PPGPP

Our model can also make prediction for translation elongation rate σ. This rate has been shown experimentally to decrease in carbon conditions leading to lower growth rate. When excess ppGpp is induced from a carbon source supporting fast growth, translation elongation rate decreases with growth rate, showing a clear overlap with the relationship obtained by decreasing the quality of the carbon source [5]. Our model, on the other hand, explains faster growth in excess ppGpp through relatively faster translation, see Figure 3.7. The model does predict a decrease in translation elongation rate in excess ppGpp, which is due to a decrease in total tRNA content leading to lower amount of acyl-tRNA. However, due to higher acylation of tRNAs, excess ppGpp is predicted to keep, at similar growth rate, a higher translation elongation rate in comparison with nutrient limitation, in contradiction with experimental data. This shows that explaining divergence of excess ppGpp and the carbon line through ribosome saturation leads to discrepancy with the translation elongation rate data available. We can also note that $\mu = \mu_0 R\sigma$ leads to:

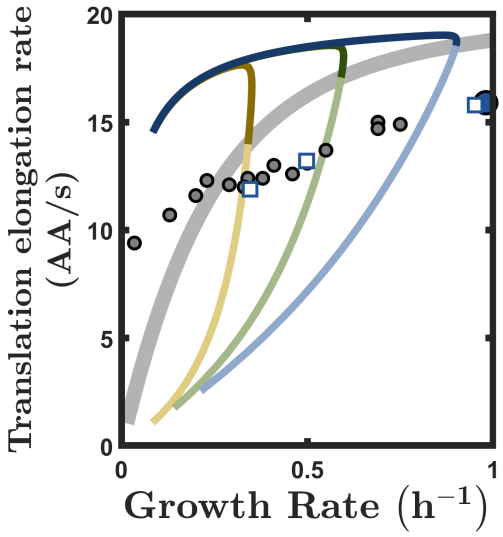


Figure 3.7: Ribosomal sub-saturation leads to higher translation elongation in excess ppGpp, contradicting experimental data. Predictions of translation elongation rate σ by the model described in Figure 3.5 in comparison with experimental results of carbon limitation and excess ppGpp. Experimental data and model predictions obtained and displayed similarly as in Figures 3.2 & 3.5

$$\sigma_{WT}(\mu) = \sigma_o(\mu) \iff R_{WT}(\mu) = R_o(\mu). \tag{3.23}$$

In other words, this model architecture can only obtain overlap of translation elongation rate in nutrient quality limitation and excess ppGpp if and only if ribosome content overlaps as well. As rRNA and tRNA keep a constant ratio, this is also true for the RNA:protein ratio. Therefore explaining the divergence in the ppGpp concentration measurements through ribosome saturation leads to predictions disregarding experimental evidence for both RNA:protein ratio and translation elongation rate.

OVERLAP OF EXCESS PPGPP RIBOSOME CONTENT WITH OPTIMUM NECESSITATES ADDITIONAL FACTORS INFLUENCING RIBOSOME CONTENT

Obtaining overlap of RNA:protein ratio versus growth rate in carbon-limited wild-type *E. coli* and when excess ppGpp is induced imposes mathematical constraints to the model, which we can derive. As tRNA and ribosomal RNA keep a constant ratio, this overlaps

requires that the $R(\mu)$ relationships also overlap in these conditions. In other words, $R_{WT}(\mu) = R_o(\mu)$. However, in our model, we have: $R(G) = R_0 \frac{k_G}{G+k_G}$. This equation shows that R is a single variable monotonous function of G. Therefore our model predicts the following:

$$R_{WT}(\mu) = R_o(\mu) \iff G_{WT}(\mu) = G_o(\mu). \tag{3.24}$$

This logical relationship shows that respecting the overlap of the RNA:protein experimental data under carbon-limitation and excess ppGpp requires the overlap of the ppGpp concentration in the same conditions. In other words, no matter the parameters, this model architecture cannot simultaneously reproduce RNA:protein ratio measurements and our ppGpp concentration measurements in excess ppGpp. The hypothesis that ribosomal saturation could be the cause of higher ppGpp levels required to tune down growth through excess ppGpp with such a model architecture is invalidated. Additionally, we identified that the assumption that the ribosomal allocation can be written as a monotonous function f of the single variable ppGpp concentration is wrong: $R \neq f(G)$. Indeed, at least in carbon limitation and excess ppGpp, ribosomal allocation does not follow the same relationship depending on ppGpp concentration. This can also be extracted from the data presented in our Chapter 2 data in combination with RNA:protein ratio data collected by Zhu et al. [5], as shown in Figure 3.8.

3.3.5. ADDING EFFECT OF NTPS ON RIBOSOME PRODUCTION IS INSUFFICIENT TO EXPLAIN EXCESS PPGPP BEHAVIOR

Apart from ppGpp, other compounds and processes have been reported to impact ribosomal operon transcription. Before direct effect of ppGpp on ribosome abundance was established and considered as the main regulation, many studies aiming to elucidate the mechanisms of ribosome abundance control and its correlation with growth rate focused on the effect of the nucleosides triphosphate ATP and GTP. The concentration of the NTP initiating this transcription, namely ATP or GTP depending on the ribosomal operon considered, have been reported to influence the transcription rate *in vitro* [21–24] and in some cases in vivo [25]. Following our observation that the two NTPs initiating ribosomal operons, ATP and GTP, slightly accumulate in excess ppGpp, we explore in this section 3.3.5 whether including activation by these NTPs, competitively inhibited by ppGpp, allows us to predict RNA:protein ratios measured in excess ppGpp.

NTP AS SUBSTRATE OF RIBOSOME PRODUCTION COMPETITIVELY INHIBITED BY PPGPP

To attempt to describe effect of NTPs and ppGpp on ribosome content, we write the variation of ribosome content as a competition between ribosome production Υ and the dilution through growth described by a term $-\mu R$, with μ the growth rate and R the mass fraction of ribosomes. This variation is thus defined as follows:

$$\frac{dR}{dt} = \Upsilon - \mu R. \tag{3.25}$$

We then explicit the ribosome production by including the elements known to play a role in ribosomal operon transcription. We therefore consider ribosome production as scaling with the transcription of the operons and neglect the role of translation, protein maturation and ribosome assembly in regulating ribosome content. This is likely a fair

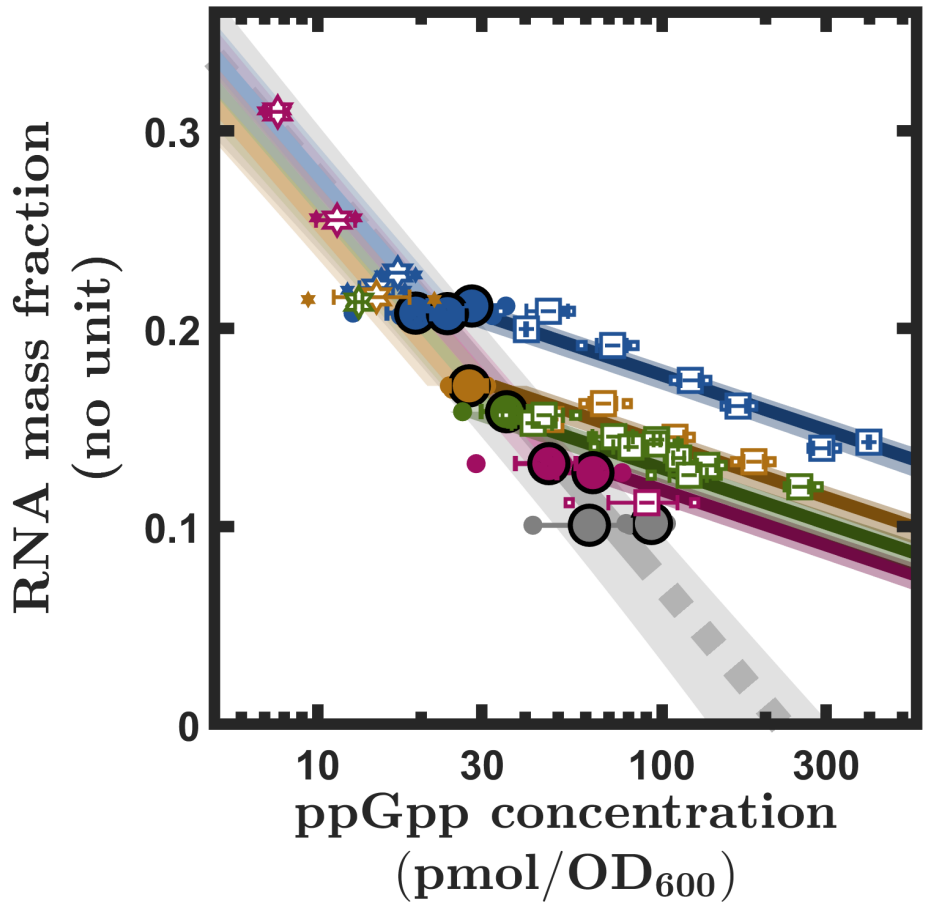


Figure 3.8: ppGpp levels do not alone determine RNA mass fraction. Results of RNA mass fraction adapted from Zhu et al. [5] in comparison with our results of ppGpp concentration. To relate our ppGpp concentration results to RNA mass fraction, we first converted RNA:protein ratio measurements obtained by Zhu et al. [5] into RNA mass fraction and then used single linear fit through this data for each strain and carbon source combination. Circles represent WT samples, stars represent depleted ppGpp and squares correspond to excess ppGpp. Each color corresponds to a different carbon source: blue for glucose, ochre for succinate, green for glycerol and magenta for acetate.

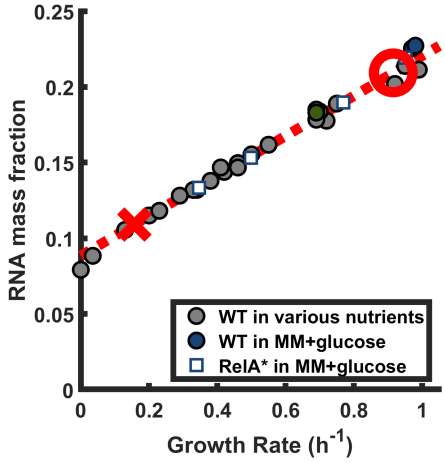


Figure 3.9: RNA mass fraction content attempted to be explained by ribosome production model. RNA mass fraction mesurements, compiled and adapted from Dai et al. [14] and Zhu et al. [5]. Graphical conventions are the same as in Figure 3.2, with the addition of the red cross and circle representing constraining points (μ_1, R_1) and (μ_2, R_2) respectively. These two points correspond to the expected RNA mass fraction in minimal media supplemented with glutamate and glucose respectively. They are obtained through the dashed line representing a fit of the RNA mass fraction data. This fit is used to extract these predicted values from the growth rate corresponding to our measurements of NTP and ppGpp concentration, displayed in Figure 3.9.

assumption, as the RNA mass fraction is composed mostly of ribosomal RNA and ribosomal proteins concentration thus vary in the same way. We consider RNA polymerase, described by its concentration P, as an enzyme operating ribosome production. This production thus scales with P. We then explicit the regulation on this production as a Michaelis-Menten type competitive inhibition with the concentration of initating NTP η as a substrate, and the ppGpp concentration G as a competitive inhibitor. We obtain the following ribosome production:

$$\Upsilon\left(G,\eta,P\right) = R_0 P \frac{\eta}{\eta + k_S\left(1+\frac{G}{k_G}\right)}. \tag{3.26}$$

This production leads to the variation of ribosome content that follows:

$$\frac{dR}{dt} = R_0 P \frac{\eta}{\eta + k_S\left(1+\frac{G}{k_G}\right)} - \mu R. \tag{3.27}$$

CONSTRAINING PARAMETERS TO RESPECT WILD-TYPE GROWTH LAW

The dynamic equation for $\frac{dR}{dt}$ leads to the following steady state for R:

$$R = \frac{R_0 P}{\mu} \times \frac{\eta}{\eta + k_S\left(1+\frac{G}{k_G}\right)}. \tag{3.28}$$

From this equation, we derive the expected concentration of NTP, corresponding to the concentration of ppGpp G, necessary to achieve the ribosome mass fraction R as follows:

$$\eta = k_S \frac{\mu R}{R_0 P - \mu R}\left(1+\frac{G}{k_G}\right). \tag{3.29}$$

This equation describes the substrate concentration if equation 3.28 is true. In other words, it is, for given $G(\mu)$ and $R(\mu)$, the necessary NTP concentration for ribosome content to be described as competitive inhibition of ppGpp on the substrate. We constrain our model through the points $(\eta_1; R_1; G_1)$ and $(\eta_2; R_2; G_2)$: the two extreme points of the wild-type curve, glutamate and glucose respectively, to make sure it respects the behavior of wild-type *E. coli* in various nutrient conditions. With this constraint, we have:

$$\begin{cases} \eta_1 = k_S \frac{\mu_1 R_1}{R_0 P_1 - \mu_1 R_1} \left(1 + \frac{G_1}{k_G}\right) \\ \eta_2 = k_S \frac{\mu_2 R_2}{R_0 P_2 - \mu_2 R_2} \left(1 + \frac{G_2}{k_G}\right) \end{cases}, \tag{3.30}$$

which allows us to write R_0 and k_S depending on $(k_G, \eta_1, R_1, G_1, \eta_2, R_2, G_2)$. These expressions are detailed in section 3.7.3.

IDENTIFICATION OF THE CONSTRAINED PARAMETERS REVEALS UNPHYSICALITY OF THE RIBOSOME PRODUCTION DEFINITION

Constraining parameters as described in the previous paragraph leaves k_G as the only free parameter. The two other parameters k_S and R_0 depend on k_G following expressions derived in section 3.7.3. These parameters also depend on the growth rate of the two constraining points, the corresponding concentrations of NTP and ppGpp as well as the corresponding ribosome mass fraction. For the growth rate and the concentration of these two molecules, we use our measurements in glutamate (point 1) and glucose (point 2) obtained in Chapter 2 and displayed in Figure 3.10 with the two constraining points represented with a red cross and circle respectively. We deduce the ribosome mass fraction using a single fit through the RNA mass fraction data adapted from Zhu et al. [5] and Dai et al. [14], displayed in Figure 3.9. When scanning positive values of k_G, we obtained only negative values for k_S and R_0, regardless of using ATP or GTP as the initiating NTP. In other words, the definition of ribosome used in equation 3.26 in combination with the two constraining points defined by the set of equations 3.30 is unable to reproduce, through the NTP and ppGpp regulation defined in this model, the ribosome fraction observed in wild-type *E. coli* and displayed in Figure 3.9.

PHYSICALITY OF THE RIBOSOME PRODUCTION DEFINITION REQUIRES STRONGER SCALING OF RIBOSOME PRODUCTION WITH GROWTH RATE

We investigated the conditions for physical solutions with $k_S > 0$ and $R_0 > 0$. We found that these conditions lead to the single condition $k_G < [k_G]_{max}$ (proof in section 3.7.3). For GTP, we get $[k_G]_{max} = -46.1\ pmol/mL/OD$. For ATP, we have $[k_G]_{max} = -81.1\ pmol/mL/OD$. Physicality could only be obtained with a higher value of $[k_G]_{max}$. We can see from the expression of $[k_G]_{max}$ (equation 3.61) that $[k_G]_{max}$ could be increased by increasing P_2 or η_2. In other words, the concentration of RNA polymerase and/or of the initiating NTP do not scale with growth rate strongly enough to reproduce the ribosome content scaling with growth rate with this definition of ribosome production.

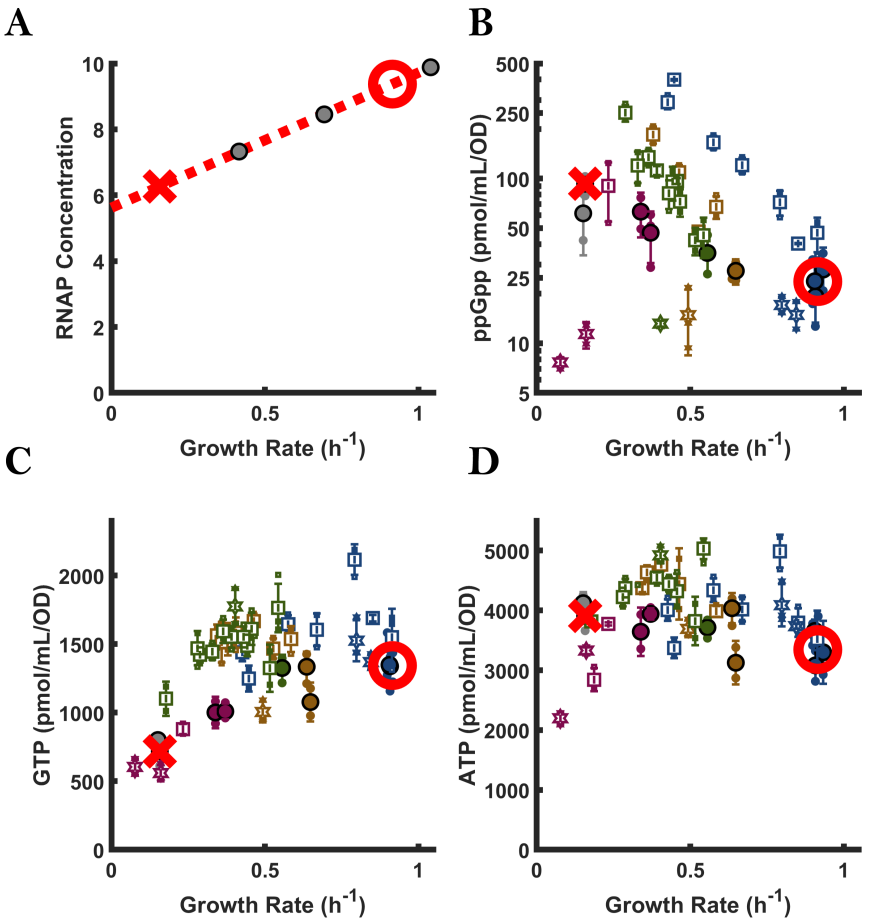


Figure 3.10: Quantities used for constraining ribosome production model. This figure display the concentrations used in the ribosome production model described in section 3.3.5. (A) RNA polymerase concentration adapted from Klumpp et al. [26]. Gray dots represent RNAP concentrations found in wild-type *E. coli* growing steadily in various nutrient conditions, as a function of growth rate. Red dashed line represents fit used to obtain constraining points. Red cross and circle represent expected concentration in minimal media supplemented with glutamate and glucose respectively, which respectively correspond to the points (μ_1, P_1) and (μ_2, P_2) used for constraining the ribosome production model. (B-D) Concentrations measured as described in our previous Chapter 2 as a function of growth rate. Graphical conventions are the same as in Figure 3.2, with the addition of the red cross and circle representing constraining points 1 and 2 respectively. (B) ppGpp concentration as a function of growth rate. (C) GTP concentration as a function of growth rate. (D) ATP concentration as a function of growth rate.

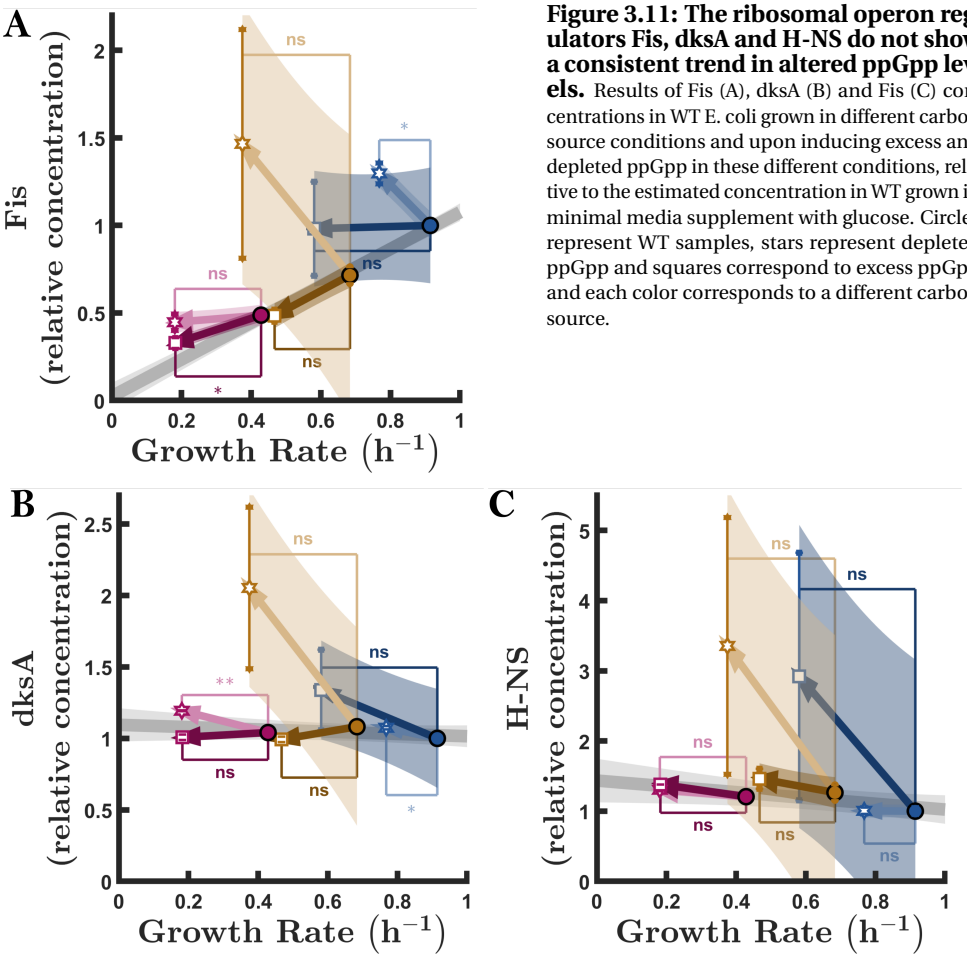


Figure 3.11: The ribosomal operon regulators Fis, dksA and H-NS do not show a consistent trend in altered ppGpp levels. Results of Fis (A), dksA (B) and Fis (C) concentrations in WT E. coli grown in different carbon source conditions and upon inducing excess and depleted ppGpp in these different conditions, relative to the estimated concentration in WT grown in minimal media supplement with glucose. Circles represent WT samples, stars represent depleted ppGpp and squares correspond to excess ppGpp and each color corresponds to a different carbon source.

3.3.6. KNOWN REGULATORS AND PROCESSES INFLUENCING RIBOSOME PRODUCTION DO NOT SHOW THE NECESSARY TREND TO COUNTERACT PPGPP

In this section 3.3.6, we investigate regulatory proteins and groups of proteins responsible for processes that might affect ribosomal production. We do so using results of untargeted proteomics in carbon limitation and ppGpp perturbations obtained in the previous Chapter 2. For details of the methodology on how these results were obtained, we refer the reader to Chapter 2. Briefly, the perturbations of ppGpp were performed in the same way as for our ppGpp concentration measurements and our measurements allow estimation of relative abundance of single proteins accross different conditions as well as the proteome mass fraction represented by a group of proteins.

THE REGULATOR FIS DOES NOT ACCUMULATE IN EXCESS PPGPP

The regulator Fis is known to upregulate ribosomal operons during nutrient upshifts by binding upstream from the all the seven ribosomal operons [27–29]. We display Fis concentration in the different conditions considered in Figure 3.11A. We found, as previously reported [30], that Fis correlates with growth rate in different nutrient conditions. When we induced excess ppGpp, we found that Fis concentration either stays constant or decreases. In none of the excess ppGpp conditions did we observe accumulation of Fis. Depleted ppGpp, on the other hand showed important accumulation of Fis in glucose and succinate but not in acetate.

DKSA AND H-NS DO NOT SHOW A CONSISTENT TREND IN EXCESS PPGPP

We also measured the concentration of the proteins dksA and H-NS, known to negatively impact transcription of ribosomal operons [31, 32]; dksA through enhancing the effect of ppGpp, H-NS through competing with Fis by binding DNA sequences overlapping with Fis binding site [31]. We display the relative concentration of these proteins in the same conditions in Figure 3.11B-C. Our WT strain, in various carbon conditions, keeps a rather constant concentration of these proteins. Perturbing ppGpp concentration did not show any significant trend for the abundance of these proteins. In particular, we did not observe depletion of these proteins in excess ppGpp.

CONCENTRATIONS OF PROTEINS RESPONSIBLE FOR DNA REPLICATION DO NOT INCREASE IN EXCESS PPGPP

As they might influence the ribosomal operon copy number, we extracted proteome mass fraction of proteins involved in the DNA replication process, as well as the ones initiating this process. We did so through the gene ontology (GO) terms "DNA replication" and "DNA replication initiation" respectively. In accordance with previous results, we find that WT *E. coli* has a slightly higher concentration of these two groups of proteins in carbon sources sustaining faster growth, as shown in Figure 3.12. We also observed that, in excess ppGpp, the proteins reponsible for DNA replication either kept a similar mass fraction as in the unperturbed WT strain or slightly decreased. Depleted ppGpp had little effect on the abundance of these proteins, except in acetate where they increase slightly.

PROTEINS INVOLVED IN DNA TOPOLOGY ARE POORLY AFFECTED BY PPGPP PERTURBATION

Because they might modulate the expression of some operons, including ribosomal ones, we were interested in quantifying the relative abundance of proteins involved in modifying DNA topology. To do so, we extracted proteins annotated with the GO term "DNA topological change". We display the concentration of these proteins in Figure 3.13. We found that WT *E. coli* has a rather unchanged abundance of these proteins in the tree carbon conditions investigated. ppGpp perturbation towards both excess or depleted ppGpp had little effect on the concentration of these proteins.

RNA POLYMERASE ABUNDANCE IS UNAFFECTED BY EXCESS PPGPP

Finally, we measured the proteome mass fraction of proteins forming RNA polymerases (RNAP) complexes through the GO terms "RNA polymerase complex" as RNAPs are the machineries synthesizing ribosomal RNA. Results are shown in Figure 3.14. We find that

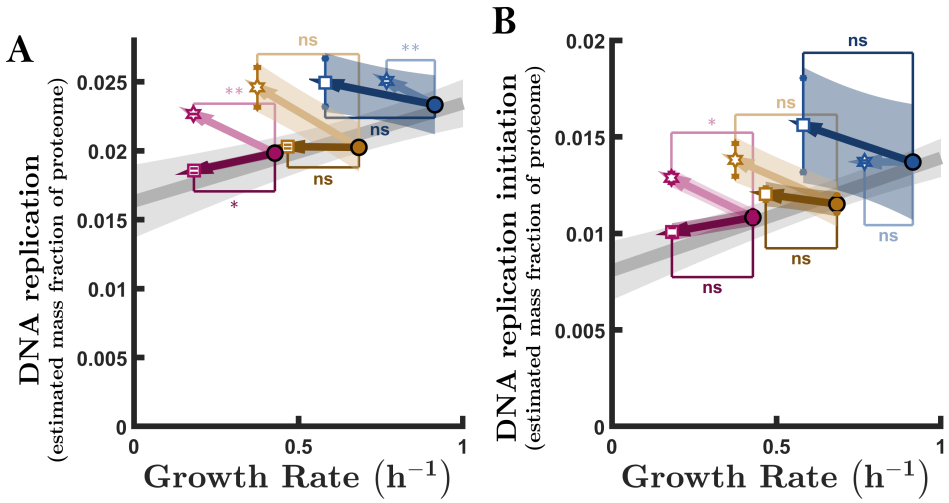
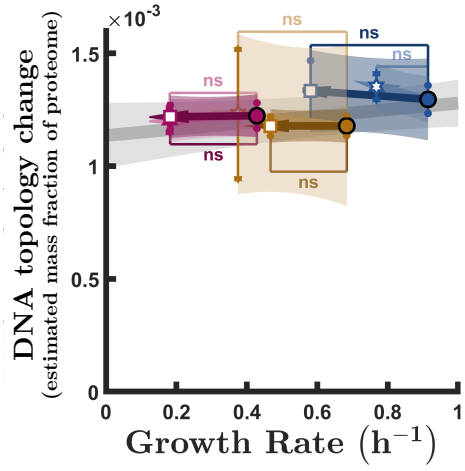


Figure 3.12: Proteins responsible for DNA replication and its initiation do not accumulate in excess ppGpp. Results of proteome mass fraction of proteins belonging to the GO terms "DNA replication" (A) and "DNA replication initiation" (B). Circles represent WT samples, stars represent depleted ppGpp and squares correspond to excess ppGpp and each color corresponds to a different carbon source.

Figure 3.13: Proteins responsible for modifying DNA topology poorly respond to ppGpp perturbations. Results of proteome mass fraction of proteins belonging to the GO terms "DNA topological change". Circles represent WT samples, stars represent depleted ppGpp and squares correspond to excess ppGpp and each color corresponds to a different carbon source.



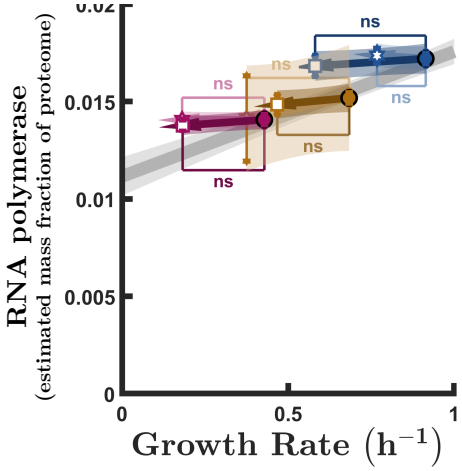


Figure 3.14: Proteins constituting RNA polymerases show a slight accumulation in excess ppGpp. Results of proteome mass fraction of proteins belonging to the GO terms "RNA polymerase complex". Circles represent WT samples, stars represent depleted ppGpp and squares correspond to excess ppGpp and each color corresponds to a different carbon source.

the abundance of RNAP proteins of our WT strain slightly increase with increasing growth rate. When both excess and depleted ppGpp were induced, the mass fraction of RNAP proteins stayed to a similar level as for the unperturbed natural case.

3.4. DISCUSSION

THE MAIN TRIGGER FOR PPGPP PRODUCTION IN STEADY GROWTH IS LIKELY DWELLING RIBOSOMES

In steady growth, the tRNA to total RNA ratio keeps a roughly constant value (Figure 3.2C) while the mass fraction of RNA increases (Figure 3.2B). This means that the abundance of tRNA increases with growth rate. Additionally, Dai et al [14] showed that several tRNAs show a constant aminoacylation ratio. This indicates that the abundance of deacyl-tRNA likely increases with faster growth. This would make deacyl-tRNA a poor candidate to be the main activation of the ppGpp response in steady growth. Our model considering only activation of ppGpp production by dwelling ribosomes was able to reproduce ppGpp concentrations and the optimality of WT concentrations. One could speculate that while dwelling ribosomes are the main trigger in steady growth, the identified additional activation by deacyl-tRNA [15] might play a role in sudden starvation, during which aminoacylation of tRNAs likely drop, allowing ppGpp to spike at much higher concentrations than in steady growth as has been reported [25].

EXCESS PPGPP DOES NOT INCREASE RIBOSOME SATURATION

In their study, the authors of Kohanim et al. [12] explain a first rapid increase of growth rate following a carbon source concentration upshift by the presence of sub-saturated ribosomes in the less abundant nutrient conditions. These ribosomes would be quickly put to use when nutrient availability suddenly increases and rapidly increase growth rate. Following this idea we investigated in the previous section 3.3 whether ribosome saturation can explain discrepancy between ppGpp versus growth rate in carbon limitation

and excess ppGpp. Using an adapted version of the model developed by Kohanim et al. [12] (section 3.3.1) and our model expliciting the repression by ppGpp on ribosome (section 3.3.3), we showed that explaining divergence of excess ppGpp and the carbon line in such a way leads to faster growth per ribosomes (Figures 3.3C and 3.6) in contrast with experimental evidence.

One year after the work of Kohanim et al., the study by Zhu et al. [5] showed that when inducing excess ppGpp, ribosome content decreases following the same relationship with regards to growth rate as lowering carbon source quality (Figure 3.2B). This contradicts the idea that higher ribosome saturation could explain divergence of excess ppGpp and the carbon line. Indeed, this idea implies that the same amount of ribosomes can yield a higher growth rate in excess ppGpp, which is in contradiction with observations from Zhu et al. [5]. They also measured translation rate per individual ribosome following excess ppGpp induction (Figure 3.2D), which followed the same curve as seen in carbon limitation in other measurements [8, 14]. Both in carbon limitation and excess ppGpp translation rate decreases towards about half its maximum at zero growth rate. This further disproves the hypothesis that ribosome saturation could explain our ppGpp vs. growth rate measurements. According to this hypothesis, the translation rate should increase in excess ppGpp. In other words, if ribosomes are sub-saturated in the poor nutrient conditions as Kohanim et al. [12] proposed in the case of carbon source concentration, then, inducing excess ppGpp does not increase this saturation and make ribosomes translate faster.

EXCESS PPGPP LIKELY DECREASES TRANSLATION RATE DIRECTLY OR THROUGH ITS EFFECT ON TRNA

The decrease in translation rate in excess ppGpp seems at first surprising as inducing excess ppGpp with a different system has been shown to trigger amino acid accumulation [33]. A possible explanation could be direct effect of ppGpp inhibiting translation [34]. Alternatively, downregulation of transfer RNA by ppGpp could also explain such result. tRNAs are necessary to deliver amino acids to ribosomes and allow them to serve translation. They are known to scale with ribosomes in carbon limitation [35] and some operons containing tRNA genes have been shown to be down-regulated by ppGpp [24, 36]. Their down-regulation by excess of ppGpp likely allows only a fraction of the accumulated amino acids to be bound to tRNA and participate to translation. This is consistent with measurements from Dai et al. [14] who showed that a major fraction of tRNAs around 80% are acylated in different carbon conditions. Decreasing total tRNA content with excess ppGpp likely leads to lack of deacyl-tRNA which results in a decrease in aminoacyl-tRNA content though the amino acid pool is increasing. This, in combination with the direct effect of ppGpp on translation, could explain why the accumulation of amino acids does not yield a faster translation rate.

EXCESS PPGPP MODELING CANNOT WRITE RIBOSOME CONTENT AS SOLE FUNCTION OF PPGPP LEVEL

In Figure 3.8, we confirmed that ribosome content anticorrelates with ppGpp concentration in various carbon conditions. On the other hand, when excess ppGpp is induced, the dependency of ribosome abundance to ppGpp level is weaker. Many attempts to model the ppGpp signaling response write ribosome abundance as a sole decreasing

function of the ppGpp level [8, 10]. In section 3.3.4, we have shown based on our data that such models cannot explain the effect of excess ppGpp on both ribosome abundance and growth rate. Further studies attempting to model this system and what happens when it is perturbed will need to identify the other regulators or processes influencing the ribosome content and implement this in a new model and, by doing so, investigate whether additional regulation on ribosome content in excess ppGpp is also relevant to the natural case.

NTPS ARE UNLIKELY TO COUNTERACT EXCESS PPGPP IMPACT ON RIBOSOMES

In section 3.3.5, we develop a model of ribosomal production proportional to the amount of RNA polymerases and regulated by ppGpp and NTPs, which highlighted the need to better understand other factors playing a role in ribosome abundance scaling with growth rate. This is key to understanding how natural ribosome content relationship with growth rate emerges before investigating how it is affected by excess ppGpp. While this does not explain whether NTP levels could explain the divergence of excess ppGpp, we would like to point out that the concentration of these NTPs is only increased in mild excess ppGpp and drop back to their natural value at stronger inductions, as shown in Chapter 2 Figure 2.12. For this reason, while the concentration of NTPs might play a role in increasing ribosome abundance at mild excess ppGpp induction, their accumulation is unlikely to fully explain what happens in excess ppGpp. Indeed, at strong inductions, the concentration of NTP is the same as for a poor carbon source while the ppGpp concentration is much larger (Chapter 2 Figure 2.1A) and the resulting ribosomal abundance is the same (Chapter 2 Figure 2.4 and Zhu et al. [5]).

KNOWN REGULATORS OF RIBOSOMAL OPERONS UNLIKELY TO COUNTERACT EXCESS PPGPP IMPACT ON RIBOSOMES

To counteract the effect of excess ppGpp and explain why higher ppGpp levels in this condition yield the same amount of ribosomes as in poor carbon, higher level of Fis in excess ppGpp would be required, which we did not observe (Figure 3.11). The two remaining known ribosomal operon regulators, apart from ppGpp, initiating NTPs and Fis, are dksA and H-NS. Both have a repressive effect on the transcription of ribosomal operons and counteraction of ppGpp through these proteins would thus require them to deplete in excess ppGpp, which we also did not observe. These known regulators of ribosomal operons are thus unlikely to counteract the effect of excess ppGpp on ribosome abundance. This phenomenon might be explained by another yet to be discovered regulator, or global effects modulating transcription.

CONCENTRATION OF RIBOSOMAL OPERON IS UNLIKELY TO INCREASE IN EXCESS PPGPP

Apart from regulators activating or repressing ribosomal operon, the ribosome abundance might also be influenced by the concentration of ribosomal operons. Our results of DNA replication proteins, displayed in Figure 3.12, show no accumulation of these proteins which seems to indicate that DNA replication is not increased in excess ppGpp. Moreover, excess ppGpp is known to downregulate the amount of DNA per cell [37], likely by stopping new rounds of replication. This same study also identified that the DNA origin of replication (ori), usually more abundant than the terminus (ter), is found in similar amount in excess ppGpp, likely as overlapping rounds of replication are stopped. Ribosomal operons, situated close to the origin of replication, are likely scaling in abundance like

this sequence. As both DNA content per cell and the ori/ter ratio decrease, the number of ribosomal operon per cell very likely decreases. The decreased cell size in excess ppGpp might slightly counteract this effect to increase the concentration of ribosomal operons but the results from Schreiber et al. [37] seem to indicate that the effect on DNA content and the ori/ter ratio is at least as drastic as the one on cell size. It seems thus quite unlikely that the concentration of ribosomal operon increases to counteract the effect of ppGpp on ribosome content when we induce excess ppGpp.

TOTAL ABUNDANCE OF RNAP IS UNLIKELY TO COUNTERACT EXCESS PPGPP

As RNA polymerases (RNAPs) are the machineries responsible for transcription, their abundance might influence the overall transcription rate, thus also the transcription of ribosomal operons. In Figure 3.14, we showed that the abundance of RNAP proteins is unchanged when excess ppGpp is induced. Therefore, the total abundance of RNAP counteracting the effect of excess ppGpp on ribosomes is very unlikely. On the other hand, at similar abundance, the availability of RNAPs can vary and also affect transcription of ribosomal operons, which we were not able to test with these measurements.

ADDITIONAL REGULATION ON RIBOSOME CONTENT MIGHT ARISE FROM PASSIVE CONSTRAINTS IMPOSED BY REGULATION OF OTHER PROTEOME SECTORS

After attempting to explain the counteraction of the effect of ppGpp on ribosome abundance in excess ppGpp through other ribosomal regulators or processes that might influence ribosomal transcription without success, we speculate another possible explanation. We believe that the downregulation of ribosome content in carbon limitation could arise from a combination of the effect of ppGpp and passive constraints arising from upregulation of other proteome sectors, such as the catabolic sector, by other regulators. This could for instance leave less RNAPs available to transcribe ribosomal operons. Excess ppGpp would decrease ribosomal content only through direct repression, without upregulating these sectors as is indeed observed in Chapter 2 for the proteome sector containing catabolic proteins. Additional modeling efforts, implementing such regulation, and experiments, to test RNAP availability for example, would be required to test this hypothesis.

3.5. CONCLUSIONS AND PERSPECTIVES

We have shown that simple laws are sufficient to explain reasonably well the behavior of wild-type *E. coli* growing in various carbon conditions in terms of ppGpp concentrations and RNA:protein ratio as well as the optimality of ppGpp concentrations (Figure 3.5). We thereby proposed a simple model architecture which could be used as a base for future modeling efforts: tradeoff between metabolic enzymes charging tRNAs with amino acids and ribosomes assembling amino acids into new biomass and the ppGpp response triggered by dwelling ribosomes and downregulating ribosomes.

Attempting to explain higher ppGpp concentrations measured when excess ppGpp is induced than in carbon limitation at similar growth rate through a higher aminoacyl-tRNA saturation of ribosomes, we found that this leads to higher growth rate per ribosome, in contrast with experimental evidences (section 3.3.4). We showed mathematically that the overlap of RNA:protein ratio in excess ppGpp and carbon limitation and the divergence

of ppGpp concentrations in excess ppGpp can only be achieved by models considering additional regulations on ribosome content than ppGpp only. This observation explained why our models were unable to capture both of these behaviors and gave directions for future models attempting to explain ppGpp perturbations: after identifying factors that influence ribosome content differently in excess ppGpp and carbon limitation, one could implement this regulation to the model we proposed and test whether this is sufficient to explain the various experimental results we compared our model to.

In that purpose, we built a more mechanistic model of ribosomal operon transcription including the reported activation by NTPs. This model was unable to explain observed ribosome content, highlighting the need to better understand the different factors that make ribosome content scale with growth rate. We investigated known ribosomal regulators and processes, other than ppGpp and NTPs, that could influence ribosome production: through our proteomics measurements and previously published studies, we found that most of them are unlikely to counteract the effect of excess ppGpp, with the notable exception of RNAP availability, which we were unable to investigate. Future studies could investigate RNAP availability in excess ppGpp or attempt to narrow down the search by placing a reporter gene under the control of a ribosomal promoter on a plasmid and measure its expression to understand if ribosomal promoters give rise to the same trends independently from the chromosome.

Following observations from our previous Chapter 2 that some proteome sectors such as the ones containing catabolic proteins are unaffected by excess ppGpp in contrast with carbon limitation, we speculated that ribosome content might, on top of being downregulated by ppGpp, be negatively impacted by the upregulation of other proteins in carbon limitation. The absence of this second passive regulation in excess ppGpp could explain why the same ppGpp concentration leads to higher ribosome content, for example through RNAP availability. Future modeling efforts could attempt to explain ribosome content in such a way to test this hypothesis, either by empirically using results of catabolic sector abundance, or by introducing its regulation into the model.

3.6. METHODS

STEADY STATE EQUATIONS SOLVING

To solve all presented models we expressed time derivatives of all entities of the model, depending on parameters and inputs. These entities may include $(R, C, x, \tau, T, [ppGpp])$. The first two of these entities are included in all models while the latter four are included only in some of the presented models. We minimized the time derivatives of the included entities using the least square non linear trust region reflective algorithm. We constrain mass fractions between 0 and 1 and $[ppGpp]$ between 0 and 10000 and use the criteria function tolerance, step tolerance and optimality tolerance equal to $2 \cdot 10^{-16}$. After checking that the algorithm exited due to reaching one of these criteria, we can extract the values of the variables which minimize the time derivatives. We use these values to compute all rates and quantities of interest. We repeat the process for all sets of parameters and inputs of interest. For example, to find the line describing the wild-type strain under carbon limitation, we scanned values of $n \in [0; 1]$ and ran the minimization algorithm for each of them.

PARAMETERS OPTIMIZATION

For the last and most complex model presented here, for which manually tuning parameters is a strenuous task, we numerically optimized parameters to best reproduce experimental data. To do so, we setup a function describing the distance between model predictions and experimental data points. This function first computes steady state values for a scan in nutrient quality n keeping other inputs $(o;m)$ at zero describing carbon-limited wild-type growth. Secondly, it also compute steady state values at fixed n values following scans of inputs o and m, keeping the one which is not scanned at zero. These scans describe ppGpp perturbations: $o > 0$ for excess ppGpp and $m > 0$ for depleted ppGpp. We define a value of n for each carbon source for which data for such perturbation exists, which we add to the list of parameters. Finally, for each experimental point i displayed in Figure 3.2 we computed the relative distance between the predicted line and the data point in the Y versus μ plane with Y the considered experimental quantity as follows:

$$d_i = min\left(\sqrt{\left(\frac{Y_{mod} - Y_{exp,i}}{\max(Y_{exp})}\right)^2 + \left(\frac{\mu_{mod} - \mu_{exp,i}}{\max(\mu_{exp})}\right)^2}\right), \tag{3.31}$$

with the following:

- μ_{mod} and Y_{mod} respectively the vectors describing the growth rate and the corresponding value for the quantity Y.
- Y_{exp} and μ_{exp} respectively the vectors describing the experimental values and the corresponding growth rate in the considered condition.
- $Y_{exp,i}$ and $\mu_{exp,i}$ the experimental value and growth rate of the considered point i.

We split conditions into wild-type in carbon-limitation, excess and depleted ppGpp in each carbon on source separately. For each excess and depleted ppGpp condition we also include the distance between the line corresponding to this perturbation and the wild-type point in order to force the perturbation line to anchor and intersect the carbon-limitation line at the right point. Finally we construct the vector we wish to minimize through parameter optimization by multiplying each distance d_i by a weight factor w_i as follows:

$$V = [w_1 d_1, \ldots, w_i d_i, \ldots, d_N]. \tag{3.32}$$

For the model results displayed in Figure 3.5 we used the following weights $w_i = 5$ for wild-type carbon limitation of RNA:protein ratio and $[ppGpp]$ and $w_i = 1$ for all other that points. For the model version favoring in Figure 3.6 we used $w_i = 10$ for RNA:protein ratio in carbon-limitation and $[ppGpp]$ in carbon-limitation or excess ppGpp, $w_i = 2$ for RNA:protein ratio in depleted ppGpp, $w_i = 0$ for RNA:protein ratio in excess ppGpp and all translation elongation points and $w_i = 1$ for the remaining.

To optimize parameters and avoid local minima we use rounds ten parallel runs of the Levenberg-Marquart algorithm. For each run we start with randomize each parameter as follows:

$$p_i = p_{i,0} \cdot 10^r \quad \text{With: ,} \tag{3.33}$$

with the following:

- $p_{i,0}$ the initial value of the considered parameter.
- p_i the new value of the parameter after randomization.
- j the number of the round of parallel optimizations.
- $r \equiv \mathcal{N}\left(0, \frac{1}{j}\right)$ a random variable.
- $\mathcal{N}(a, s)$ the gaussian distribution of mean a and standard deviation s.

For each round we select the parameters obtained from the optimization that gave the lowest sum of least squares. After ten rounds of parallel optimization the value of parameters found is usually the same, which likely corresponds to the global minimum. In other words, the value of the parameters that give the lowest sum of weighted square distances to each experimental data points.

3.7. SUPPLEMENTARY MATERIALS

3.7.1. MODEL DEFINITION THROUGH MASS FRACTION

Let's consider the simplified example with the mass of nutrient m_n, of precursor m_x, and of biomass m_B and respectively the metabolic and growth fluxes Γ and G in mass/time (or g/h) as such:

$$m_n \xrightarrow{\Gamma} m_x \xrightarrow{G} m_B. \tag{3.34}$$

Then we have $\frac{dm_x}{dt} = \Gamma - G$ and $\frac{dm_B}{dt} = G$. Thus:

$$\begin{cases} \frac{d}{dt}\left(\frac{m_x}{m_B}\right) = \frac{\frac{dm_x}{dt} m_B - m_x \frac{dm_B}{dt}}{m_B^2} \\ \frac{d}{dt}\left(\frac{m_x}{m_B}\right) = \frac{\Gamma - G}{m_B} - \frac{m_x}{m_B}\frac{G}{m_B} \end{cases}. \tag{3.35}$$

We define mass fraction of precursor $x = \frac{m_x}{m_B}$, catabolic rate $\gamma = \frac{\Gamma}{m_B}$ and growth rate $\mu = \frac{G}{m_B}$ and thus have:

$$\frac{dx}{dt} = \gamma - \mu - \mu x. \tag{3.36}$$

This reasoning can be generalized to any mass fraction $X = \frac{m_X}{m_B}$.

If the mass m_X is produced with flux P_X and consumed with flux C_X we then have $\frac{dm_X}{dt} = P_X - C_X$ and:

$$\begin{cases} \frac{dX}{dt} = \frac{d}{dt}\left(\frac{m_X}{m_B}\right) = \frac{\frac{dm_X}{dt} m_B - m_X \frac{dm_B}{dt}}{m_B^2} \\ \frac{dX}{dt} = \frac{P_X - C_X}{m_B} - \frac{m_X}{m_B}\frac{G}{m_B} \end{cases}. \tag{3.37}$$

We can then define production $p_X = \frac{P_X}{m_B}$ and consumption rate $c_X = \frac{C_X}{m_B}$ and we have:

$$\frac{dX}{dt} = p_X - c_X - \mu X. \tag{3.38}$$

We can thus apply this mass fraction definition to any quantities in our system, for example the deacyl-tRNA mass fraction θ or the acyl-tRNA mass fraction T.

3.7.2. SOLUTION FOR STEADY STATE GROWTH OF MODEL IN FIGURE 3.4

In this annex we present the analytical solving for the steady-state solution $\mu = \mu(R, n)$ for the model described in Figure 3.4.

Steady state equations:

$$\gamma = \gamma_{max} n (R_{max} - R) \frac{\theta}{\theta + k_\theta}; \tag{3.39}$$

$$\mu = \mu_{max} R \frac{T}{T + k_T}; \tag{3.40}$$

$$\frac{dT}{dt} = \gamma - \mu - \mu T = 0; \tag{3.41}$$

$$T + \theta = \xi R. \tag{3.42}$$

Solving:

$$(3.40) \rightarrow T = \frac{k_T \mu}{\mu_{max} R - \mu}; \tag{3.43}$$

$$(3.42)\&(3.43) \rightarrow \theta = \frac{\xi \mu_{max} R^2 - \xi R \mu - k_T}{\mu_{max} R - \mu}; \tag{3.44}$$

$$(3.41)\&(3.39) \rightarrow \gamma_{max} n (R_{max} - R) \frac{\theta}{\theta + k_\theta} = \mu (1 + T); \tag{3.45}$$

$$\begin{aligned} (3.44)\&(3.45) \rightarrow & \gamma_{max} n (R_{max} - R)(\xi \mu_{max} R^2 - \xi R \mu - k_T \mu) \\ & = \mu (\xi \mu_{max} R^2 - \xi R \mu - k_T \mu + k_\theta \mu_{max} R - k_\theta \mu) \frac{\mu_{max} R - \mu + k_T \mu}{\mu_{max} R - \mu}; \end{aligned} \tag{3.46}$$

$$\begin{aligned} & \gamma_{max} n (R_{max} - R)(\xi \mu_{max} R^2 - \xi R \mu - k_T \mu)(\mu_{max} R - \mu) \\ & = \mu (\xi \mu_{max} R^2 - \xi R \mu - k_T \mu + k_\theta \mu_{max} R - k_\theta \mu)(\mu_{max} R - \mu + k_T \mu); \end{aligned} \tag{3.47}$$

$$\begin{aligned} & \gamma_{max} n \xi \mu_{max}^2 (R_{max} - R) R^3 - \gamma_{max} n \xi \mu_{max} (R_{max} - R) R^2 \mu \\ & - \gamma_{max} \gamma_{max} n \mu_{max} (R_{max} - R)(\xi R + k_T) R \mu + \gamma_{max} n (R_{max} - R)(\xi R + k_T) \mu^2 \\ = & (\xi R + k_T + k_\theta)(1 - k_T) \mu^3 \\ & - \left[(\xi R + k_\theta)(1 - k_T) \mu_{max} R + (\xi R + k_T + k_\theta) \mu_{max} R \right] \mu^2 \\ & (\xi R + k_\theta) \mu_{max}^2 R^2; \end{aligned} \tag{3.48}$$

$$
\begin{aligned}
&(\xi R + k_T + k_\theta)(1-k_T)\mu^3 \\
&-\left[(\xi R + k_\theta)(1-k_T)\mu_{\max} R + (\xi R + k_T + k_\theta)\mu_{\max} R \right. \\
&\quad \left. +\gamma_{\max} n (R_{\max} - R)(\xi R + k_T)\right]\mu^2 \\
&+\left[(\xi R + k_\theta)\mu_{\max}^2 R^2 + \gamma_{\max} n \xi \mu_{\max}(R_{\max} - R)R^2 \right. \\
&\quad \left. +\gamma_{\max} n \mu_{\max}(R_{\max} - R)(\xi R + k_T)R\right]\mu \\
&-\gamma_{\max} n \xi \mu_{\max}^2 (R_{\max} - R)R^3 = 0
\end{aligned}
\tag{3.49}
$$

$$
\begin{aligned}
&\mu^3 - \frac{(\xi R + k_\theta)(1-k_T)\mu_{max} R + (\xi R + k_T + k_\theta)\mu_{max} R + \gamma_{max} n (R_{max} - R)(\xi R + k_T)}{(\xi R + k_T + k_\theta)(1-k_T)}\mu^2 \\
&+ \frac{(\xi R + k_\theta)\mu_{max}^2 R^2 + \gamma_{max} n \xi \mu_{max}(R_{max} - R)R^2 + \gamma_{max} n \mu_{max}(R_{max} - R)(\xi R + k_T)R}{(\xi R + k_T + k_\theta)(1-k_T)}\mu \\
&- \frac{\gamma_{max} n \xi \mu_{max}^2 (R_{max} - R)R^3}{(\xi R + k_T + k_\theta)(1-k_T)} = 0;
\end{aligned}
\tag{3.50}
$$

$$
\mu^3 + \mathscr{B}\mu^2 + \mathscr{C}\mu + \mathscr{D} = 0, \tag{3.51}
$$

with:

$$\mathscr{B} = -\frac{(\xi R + k_\theta)(1-k_T)\mu_{max} R + (\xi R + k_T + k_\theta)\mu_{max} R + \gamma_{max} n (R_{max} - R)(\xi R + k_T)}{(\xi R + k_T + k_\theta)(1-k_T)}$$

$$\mathscr{C} = \frac{(\xi R + k_\theta)\mu_{max}^2 R^2 + \gamma_{max} n \xi \mu_{max}(R_{max} - R)R^2 + \gamma_{max} n \mu_{max}(R_{max} - R)(\xi R + k_T)R}{(\xi R + k_T + k_\theta)(1-k_T)}$$

$$\mathscr{D} = -\frac{\gamma_{max} n \xi \mu_{max}^2 (R_{max} - R)R^3}{(\xi R + k_T + k_\theta)(1-k_T)}.$$

$$
\mathscr{X} = \sqrt[3]{\frac{2\mathscr{B}^3 - 9\mathscr{B}\mathscr{C} + 27\mathscr{D} \pm \sqrt{(2\mathscr{B}^3 - 9\mathscr{B}\mathscr{C} + 27\mathscr{D})^2 - 4(\mathscr{B}^2 - 3\mathscr{C})^3}}{2}}, \tag{3.52}
$$

choice of $\pm$ does not matter for this term, same three solutions are found afterwards.

Different solutions are:

$$
\mu_k = -\frac{1}{3}\left(\mathscr{B} + z^k \mathscr{X} + \frac{\mathscr{B}^2 - 3\mathscr{C}}{z^k \mathscr{X}}\right), \tag{3.53}
$$

with $z = \frac{-1+\sqrt{-3}}{2}$.

The physical solution is found for $z^k = 1$ as other two solutions lead to complex value of μ. The physical solution can thus be written as follows:

$$
\mu = \mu_0 = -\frac{1}{3}\left(\mathscr{B} + \mathscr{X} + \frac{\mathscr{B}^2 - 3\mathscr{C}}{\mathscr{X}}\right), \tag{3.54}
$$

with: $\mathscr{X} = \sqrt[3]{\frac{2\mathscr{B}^3 - 9\mathscr{B}\mathscr{C} + 27\mathscr{D} + \sqrt{(2\mathscr{B}^3 - 9\mathscr{B}\mathscr{C} + 27\mathscr{D})^2 - 4(\mathscr{B}^2 - 3\mathscr{C})^3}}{2}}$.

3.7.3. DERIVATION OF THE CONDITION FOR PHYSICALITY OF THE NTP-REGULATED RIBOSOME PRODUCTION MODEL

We defined parameter constraints as follows:

$$\begin{cases} \eta_1 = k_S \frac{\mu_1 R_1}{R_0 P_1 - \mu_1 R_1} \left(1 + \frac{G_1}{k_G}\right) \\ \eta_2 = k_S \frac{\mu_2 R_2}{R_0 P_2 - \mu_2 R_2} \left(1 + \frac{G_2}{k_G}\right) \end{cases}. \tag{3.55}$$

These constraints lead to the following expressions for k_S and R_0:

$$\begin{cases} k_S = \frac{\eta_1 (R_0 P_1 - \mu_1 R_1) k_G}{\mu_1 R_1 (G_1 + k_G)} \\ R_0 = \mu_2 R_2 \mu_1 R_1 \frac{\eta_1 (k_G + G_2) - \eta_2 (k_G + G_1)}{P_1 \eta_1 \mu_2 R_2 (k_G + G_2) - P_2 \eta_2 \mu_1 R_1 (k_G + G_1)} \end{cases}. \tag{3.56}$$

The conditions for physicality are $k_S > 0$ and $R_0 > 0$. The first one, $k_S > 0$, leads to:

$$R_0 > P_1 \mu_1 R_1. \tag{3.57}$$

This condition implies $R_0 > 0$ and is thus the only condition to have a physical solution for the two parameters k_S and R_0. It leads to:

$$\frac{\mu_2 R_2 \left[\eta_1 (k_G + G_2) - \eta_2 (k_G + G_1)\right]}{P_1 \left[P_1 \eta_1 \mu_2 R_2 (k_G + G_2) - P_2 \eta_2 \mu_1 R_1 (k_G + G_1)\right]} > 1. \tag{3.58}$$

For the denominator of the fraction on the left side of this inequation to be positive, is required:

$$k_G > \frac{P_2 \eta_2 \mu_1 R_1 G_1 - P_1 \eta_1 \mu_2 R_2 G_2}{P_1 \eta_1 \mu_2 R_2 - P_2 \eta_2 \mu_1 R_1}. \tag{3.59}$$

For GTP, this means $k_G > -0.3 pmol/mL/OD$ and for ATP: $k_G > -14.7 pmol/mL/OD$. In other words, the denominator in inequation 3.58 is always positive for any physical value of k_G. Using this knowledge, we can derive for 3.58 the condition on k_G for physicality, which is:

$$\frac{\mu_2 R_2}{P_1} \left[\eta_1 (k_G + G_2) - \eta_2 (k_G + G_1)\right] > P_1 \eta_1 \mu_2 R_2 (k_G + G_2) - P_2 \eta_2 \mu_1 R_1 (k_G + G_1). \tag{3.60}$$

This conditions leads to:

$$k_G < [k_G]_{max} = \frac{P_2 \eta_2 \mu_1 R_1 G_1 - {P_1}^2 \eta_1 \mu_2 R_2 G_2 - \mu_2 R_2 G_2 - \mu_2 R_2 (\eta_2 - \eta_1)}{{P_1}^2 \eta_1 \mu_2 R_2 + P_1 P_2 \eta_2 \mu_1 R_1 + \mu_2 R_2 (\eta_2 - \eta_1)}. \tag{3.61}$$

REFERENCES

[1] F. Dyson et al., *A meeting with enrico fermi,* Nature **427**, 297 (2004).

[2] W. Ross, C. E. Vrentas, P. Sanchez-Vazquez, T. Gaal, and R. L. Gourse, *The magic spot: a ppgpp binding site on e. coli rna polymerase responsible for regulation of transcription initiation,* Molecular cell **50**, 420 (2013).

[3] J. J. Lemke, P. Sanchez-Vazquez, H. L. Burgos, G. Hedberg, W. Ross, and R. L. Gourse, *Direct regulation of escherichia coli ribosomal protein promoters by the transcription factors ppgpp and dksa,* Proceedings of the National Academy of Sciences **108**, 5712 (2011).

[4] F. Büke, J. Grilli, M. C. Lagomarsino, G. Bokinsky, and S. J. Tans, *ppgpp is a bacterial cell size regulator,* Current Biology **32**, 870 (2022).

[5] M. Zhu and X. Dai, *Growth suppression by altered (p) ppgpp levels results from non-optimal resource allocation in escherichia coli,* Nucleic acids research **47**, 4684 (2019).

[6] M. Zhu and X. Dai, *Stringent response ensures the timely adaptation of bacterial growth to nutrient downshift,* Nature Communications **14**, 467 (2023).

[7] M. J. Noga, F. Büke, N. J. van den Broek, N. C. Imholz, N. Scherer, F. Yang, and G. Bokinsky, *Posttranslational control of plsb is sufficient to coordinate membrane synthesis with growth in escherichia coli,* MBio **11**, e02703 (2020).

[8] C. Wu, R. Balakrishnan, N. Braniff, M. Mori, G. Manzanarez, Z. Zhang, and T. Hwa, *Cellular perception of growth rate and the mechanistic origin of bacterial growth law,* Proceedings of the National Academy of Sciences **119**, e2201585119 (2022).

[9] D. W. Erickson, S. J. Schink, V. Patsalo, J. R. Williamson, U. Gerland, and T. Hwa, *A global resource allocation strategy governs growth transition kinetics of escherichia coli,* Nature **551**, 119 (2017).

[10] G. Chure and J. Cremer, *An optimal regulation of fluxes dictates microbial growth in and out of steady-state,* Elife **12**, e84878 (2023).

[11] E. Bosdriesz, D. Molenaar, B. Teusink, and F. J. Bruggeman, *How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization,* The FEBS journal **282**, 2029 (2015).

[12] Y. K. Kohanim, D. Levi, G. Jona, B. D. Towbin, A. Bren, and U. Alon, *A bacterial growth law out of steady state,* Cell reports **23**, 2891 (2018).

[13] J. Ryals, R. Little, and H. Bremer, *Control of rrna and trna syntheses in escherichia coli by guanosine tetraphosphate.* Journal of bacteriology **151**, 1261 (1982).

[14] X. Dai, M. Zhu, M. Warren, R. Balakrishnan, V. Patsalo, H. Okano, J. R. Williamson, K. Fredrick, Y.-P. Wang, and T. Hwa, *Reduction of translating ribosomes enables escherichia coli to maintain elongation rates during slow growth,* Nature microbiology **2**, 1 (2016).

[15] W. A. Haseltine and 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,* Proceedings of the National Academy of Sciences **70**, 1564 (1973).

[16] R. Rosset, J. Julien, and R. Monier, *Ribonucleic acid composition of bacteria as a function of growth rate,* Journal of molecular biology **18**, 308 (1966).

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

[18] H. Bremer and P. P. Dennis, *Modulation of chemical composition and other parameters of the cell at different exponential growth rates,* EcoSal Plus **3** (2008).

[19] D. K. Murray and H. Bremer, *Control ofspot-dependent ppgpp synthesis and degradation inescherichia coli,* Journal of molecular biology **259**, 41 (1996).

[20] K. Potrykus, H. Murphy, N. Philippe, and M. Cashel, *ppgpp is the major source of growth rate control in e. coli,* Environmental microbiology **13**, 563 (2011).

[21] T. Gaal, M. S. Bartlett, W. Ross, C. L. Turnbough Jr, and R. L. Gourse, *Transcription regulation by initiating ntp concentration: rrna synthesis in bacteria,* Science **278**, 2092 (1997).

[22] M. S. Bartlett, T. Gaal, W. Ross, and R. L. Gourse, *Regulation of rrna transcription is remarkably robust: Fis compensates for altered nucleoside triphosphate sensing by mutant rna polymerases at escherichia coli rrn p1 promoters,* Journal of Bacteriology **182**, 1969 (2000).

[23] D. A. Schneider, T. Gaal, and R. L. Gourse, *Ntp-sensing by rrna promoters in escherichia coli is direct,* Proceedings of the National Academy of Sciences **99**, 8602 (2002).

[24] B. J. Paul, M. M. Barker, W. Ross, D. A. Schneider, C. Webb, J. W. Foster, and R. L. Gourse, *Dksa: a critical component of the transcription initiation machinery that potentiates the regulation of rrna promoters by ppgpp and the initiating ntp,* Cell **118**, 311 (2004).

[25] H. D. Murray, D. A. Schneider, and R. L. Gourse, *Control of rrna expression by small molecules is dynamic and nonredundant,* Molecular cell **12**, 125 (2003).

[26] S. Klumpp and T. Hwa, *Growth-rate-dependent partitioning of rna polymerases in bacteria,* Proceedings of the National Academy of Sciences **105**, 20245 (2008).

[27] C. A. Hirvonen, W. Ross, C. E. Wozniak, E. Marasco, J. R. Anthony, S. E. Aiyar, V. H. Newburn, and R. L. Gourse, *Contributions of up elements and the transcription factor fis to expression from the seven rrn p1 promoters in escherichia coli,* Journal of bacteriology **183**, 6305 (2001).

[28] L. Nilsson, H. Verbeek, E. Vijgenboom, C. van Drunen, A. Vanet, and L. Bosch, *Fis-dependent trans activation of stable rna operons of escherichia coli under various growth conditions,* Journal of bacteriology **174**, 921 (1992).

[29] C. A. Ball, R. Osuna, K. Ferguson, and R. C. Johnson, *Dramatic changes in fis levels upon nutrient upshift in escherichia coli,* Journal of bacteriology **174**, 8043 (1992).

[30] P. Mallik, B. J. Paul, S. T. Rutherford, R. L. Gourse, and R. Osuna, *Dksa is required for growth phase-dependent regulation, growth rate-dependent control, and stringent control of fis expression in escherichia coli,* Journal of bacteriology **188**, 5775 (2006).

[31] J. D. Gralla, *Escherichia coli ribosomal rna transcription: regulatory roles for ppgpp, ntps, architectural proteins and a polymerase-binding protein,* Molecular microbiology **55**, 973 (2005).

[32] H. Afflerbach, O. Schröder, and R. Wagner, *Effects of the escherichia coli dna-binding protein h-ns on rrna synthesis in vivo,* Molecular microbiology **28**, 641 (1998).

[33] G. Bokinsky, E. E. Baidoo, S. Akella, H. Burd, D. Weaver, J. Alonso-Gutierrez, H. García-Martín, T. S. Lee, and J. D. Keasling, *Hipa-triggered growth arrest and β-lactam tolerance in escherichia coli are mediated by rela-dependent ppgpp synthesis,* Journal of bacteriology **195**, 3173 (2013).

[34] S. Diez, J. Ryu, K. Caban, R. L. Gonzalez Jr, and J. Dworkin, *The alarmones (p) ppgpp directly regulate translation initiation during entry into quiescence,* Proceedings of the National Academy of Sciences **117**, 15565 (2020).

[35] H. Bremer, P. P. Dennis, et al., *Modulation of chemical composition and other parameters of the cell by growth rate,* Escherichia coli and Salmonella: cellular and molecular biology **2**, 1553 (1996).

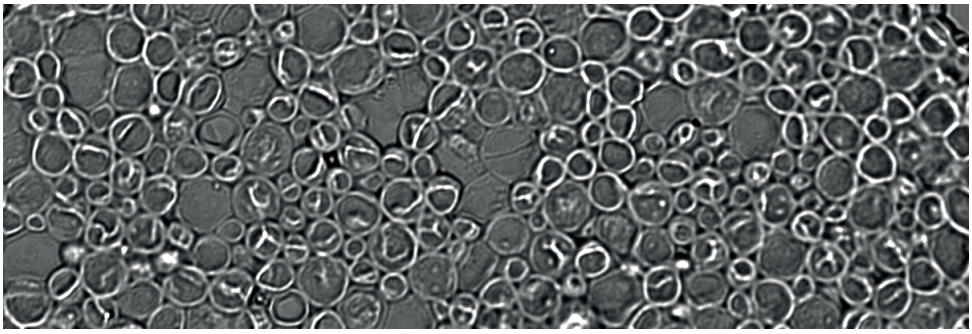
[36] J. Mizushima-Sugano and Y. Kaziro, *Regulation of the expression of the tufb operon: Dna sequences directly involved in the stringent control.* The EMBO journal **4**, 1053 (1985).

[37] G. Schreiber, E. Z. Ron, and G. Glaser, *ppgpp-mediated regulation of dna replication and cell division in escherichia coli,* Current microbiology **30**, 27 (1995).

4

A dormancy story: Conditions and mechanisms for long-term growth arrest and recovery of dry yeast

Milan Lacassin, Pim van Beeck, Yaroslav Blanter, Hyun Youk



Il dort dans le soleil, la main sur sa poitrine
Tranquille. Il a deux trous rouges au côté droit.

In the sun he sleeps, hand on his chest
Peaceful. He has two red holes on his left side.

Arthur Rimbaud, in *Le Dormeur du Val* [1], October 1870, depicting a, first thought asleep dead soldier, highlights our difficulty in distinguishing death from dormancy.

Several organisms are able to have their lives halted in a dried state and resume normal dividing activity after being rehydrated. While many of these organisms are unicellular such as yeast and some species of bacteria, some cells from more complex organisms or even multicellular organisms can also survive dehydration. However, these organisms are rare exceptions as most cells and organisms cannot be revived after they are desiccated. What principles enable some organisms to restart their lives from a dormant dry state after they are rehydrated remains a poorly understood question. By addressing this question, we may be able to desiccate, preserve and revive organisms or cells for which it is not yet possible. After freeze-drying and rehydrating budding yeast while inducing a reporter gene, we show, using epifluorescence microscopy, RNA FISH and rhodamine 123 staining, that, among the small proportion that conserve an intact outer membrane, only some yeast cells survive and that they exhibit a lag-phase of variable duration during which they are unable to divide and transcribe and their mitochondrial activity is low. We then hypothesize explanations for this behavior.

4.1. INTRODUCTION

First noticed by Anthony van Leeuwenhoek in 1702 in nematodes and rotifers [2], anhydrobiosis, the ability of some organisms to survive in a desicated state, was further explored in the second half of the 20th century. The preservation of brewer's dry yeast after freeze-drying was reported by [3–5] and successful protocols were then described and used for industrial purposes like long preservation yeast culture starters for beer and bread. The mechanisms allowing yeast to be preserved in the form of a dry powder, while most organisms cannot survive desiccation, were only later described. Yeast accumulates large amounts of trehalose. This disaccharide replaces proteins hydrogen bonding with water preserving their structure, avoids phospholipids melting transition during rehydration preserving membranes, and maintains vitrification of the preserved cells preventing crystal formation inside the cells [6, 7]. While most multicellular organisms are unable to survive desiccation, there are a few exceptions [8]. The study of these more complex anhydrobiotic organisms revealed an additional strategy to survive in a dry state. Intrinsically disordered proteins help the brine shrimp *Artemia fransiscana* [9] and tardigrades [10] to do so. Using those discoveries, attempts were made to extend desiccation tolerance to mammalian cells. It was shown that trehalose addition allows to dry and stabilize human blood platelets. Attempts to stabilize nucleated cells remained unsuccessful but hinted that a stress protein from artemia allows maintaining stability further in the drying process [11]. Interestingly, a more recent study [12] shows that mammalian cells might face another challenge in surviving desiccation: the preservation of DNA integrity. In the last decade, while decreasing interest for desiccation has been shown, many new molecular biology methods have flourished. In this work, using synthetic biology and single-cell observations to go back to the firstly studied anhydrobiotic organism budding yeast as a model, we attempt to unravel the mechanisms that allows it or not to resume a living activity after rehydration, hoping to find principles that are also true for other organisms.

4.2. RESULTS

ADDITION OF PROTECTANT GREATLY INCREASES THE LOW PROPORTION OF CELLS SURVIVING FREEZE-DRYING

First observation we made when we dried and rehydrated yeast cells, is that only some fraction of them survive. To be able to imagine better applications, and easily perform experiments, maximizing this survival is an important first step. It is also relevant for multicellular anhydrobiotic organisms as they are likely to need the vast majority of their cells to stay intact to survive. As mentioned above, trehalose is known to have several properties facilitating survival to freezing and drying. When starved or stressed, yeast cells accumulate trehalose which allows them to survive drying conditions. Its ability and natural accumulation makes trehalose a good candidate as a protectant to increase the proportion of cells that survive freeze-drying. We image yeast cells from a GFP-inducible strain right after rehydration in synthetic complete media with dextrose (SD) to assess the ability of these cells to resume two of their main functions: expressing a gene and dividing. We observe that adding a solution of trehalose in phosphate buffered saline (PBS) before freeze-drying greatly increases the proportion of cells that survive compared to directly freeze-drying those cells after removing media. (see figure 4.1).

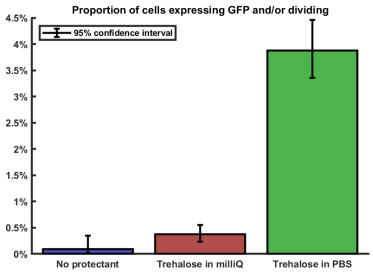
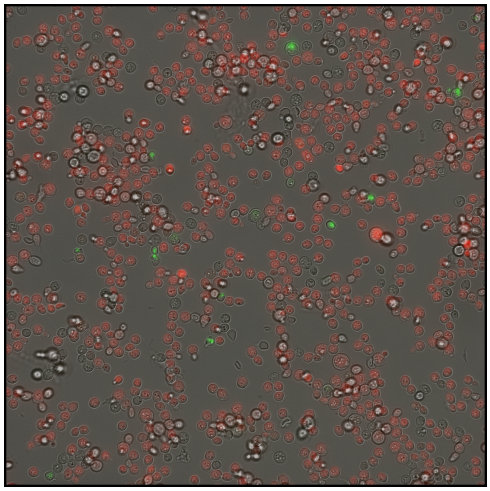


Figure 4.1: Survival after freeze-drying with different protecting media. Survival is quantified by the proportion of cells able to divide and/or expressing a constitutive reporter gene (GFP).

OUTER MEMBRANE DISRUPTION, MAIN BUT NOT ONLY CAUSE OF DEATH

Even with the addition of the protectant trehalose to increase the proportion of surviving cells, around 96% of the cells do not express GFP nor divide. In an effort to understand the limitations of anhydrobiosis, we investigate the cause of death of this majority of non-survivor cells. Membrane integrity is key to survival and often used as a proxy for viability [13]. To assess the role of membrane disruption we stain freeze-dried cells right after being rehydrated with propidium iodide (PI) known to stain cells with a leaky damaged outer membrane. We found that outer membrane disruption only explains a part of the deaths. In other words, some cells with an intact membrane are unable to express GFP or divide, which suggests another cause of death (Figure 4.2).

We observed a few cells taking up PI within the first hour after the addition of media to the dry cells suggesting either a slow rehydration process during which the membrane is damaged or another process leading to cell death. We also identified a proportion of apoptotic cells, stained by annexin V, among PI-unstained cells. As annexin V compromises viability, we cannot grow cells following this treatment. However, we hypothesize that these apoptotic cells might correspond to the ones taking up PI later on in the first hour after rehydration instead of right after rehydration. Note that as apoptosis eventually leads to membrane disruption and thus uptake of PI, these apoptotic cells are



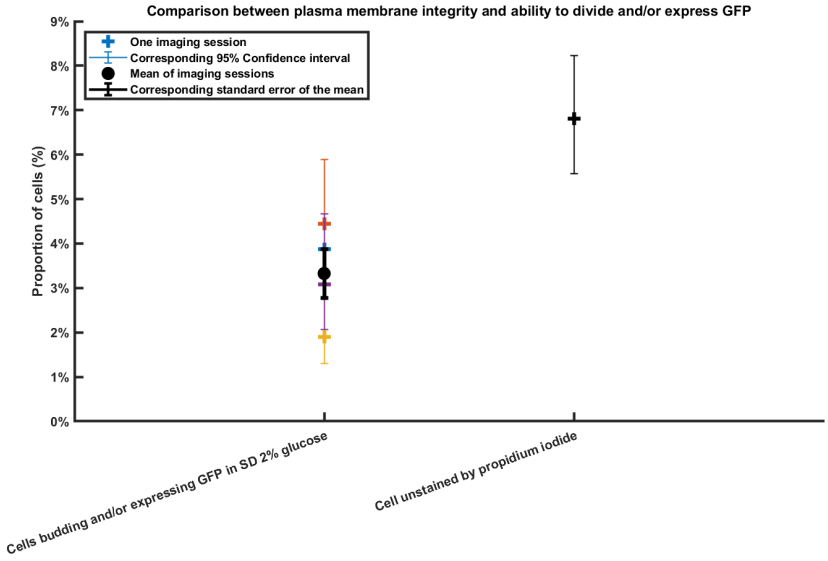


Figure 4.2: **(top)** Propidium iodide and annexin V staining. Accidentally dead cells are stained only by propidium iodide (red). Late apoptotic are stained by both annexin V and propidium iodide (yellow). Early apoptotic cells are stained only by annexin V (green). **(bottom)** Comparison between the number of cells with an intact outer membrane and the number of cells expressing GFP and/or dividing.

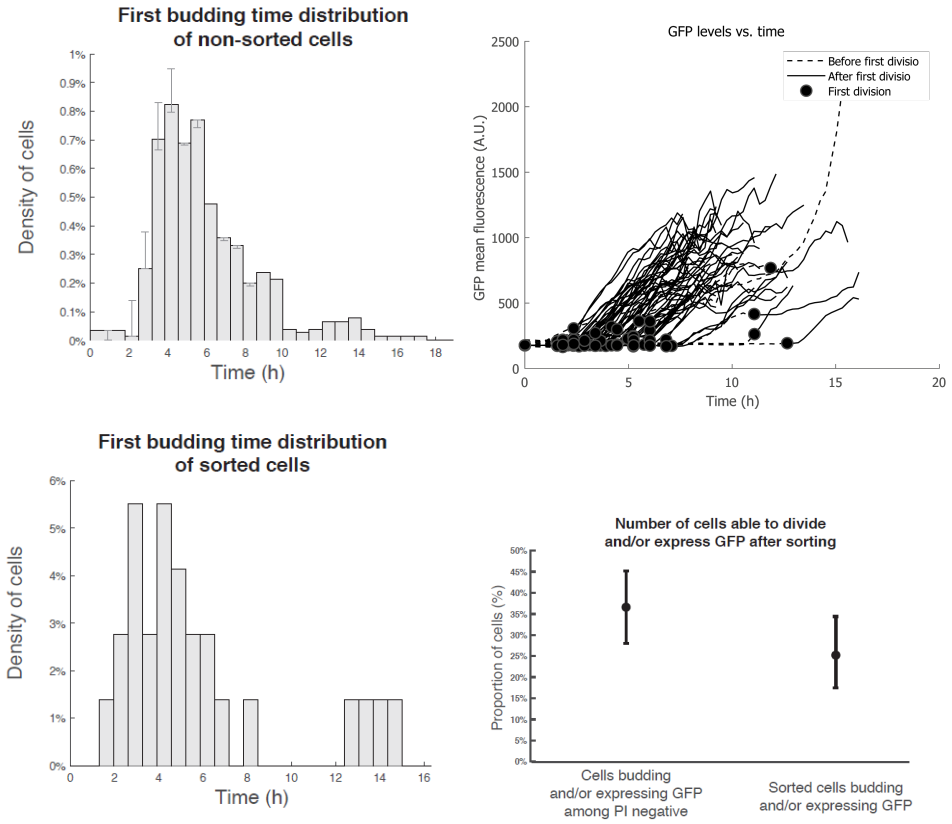


Figure 4.3: **(top left)** Distribution of lag time of dividing cells in 2% glucose. **(top right)** GFP level of freeze-dried rehydrated cells in 2% glucose. **(bottom left)** Lag time distribution of sorted freeze-dried cells. **(bottom right)** Comparison of survival in sorted and unsorted rehydrated freeze-dried cells, assessed by the ability to divide and/or express GFP.

unlikely to account for the intact cells unable to divide which have been imaged for several hours without expressing GFP, dividing or taking up PI. What caused their death did not compromise their outer membrane or lead to apoptosis.

SURVIVOR CELLS WAKE UP FROM DRYING AFTER A HIGHLY VARIABLE LAG TIME DURING WHICH GENE EXPRESSION IS HALTED

Studying how the few survivor cells recover might reveal the requirements to reestablishing the main cellular functions and help us understand the so far unexplained deaths. To do so we image yeast cells after rehydration in growth media while inducing GFP expression. We extract from these images the time of the first division the time of the apparition of the bud leading to the first division and the level of GFP over time. We show that rehydrated freeze-dried yeast cells exhibit a highly variable lag time after which they initiate their first division, with most cells budding between three and eight hours after rehydration up to the slowest recorded recovery around 17 hours (see figure 4.3).

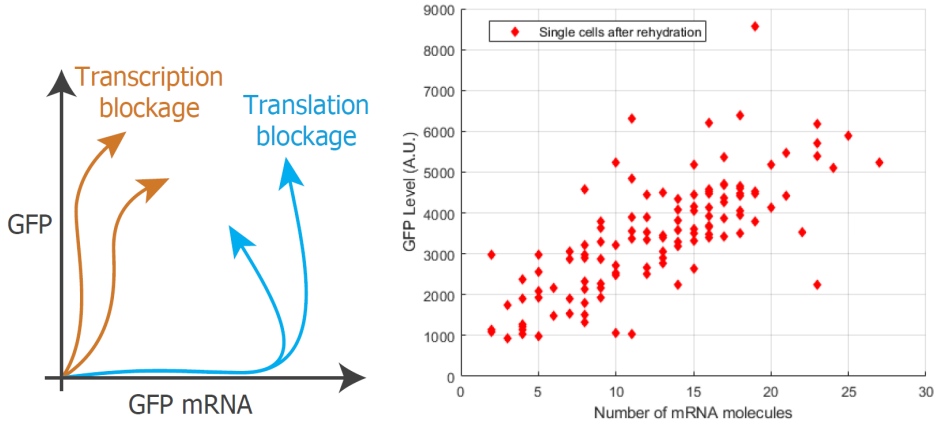


Figure 4.4: (left) Two possible scenarios, transcriptional and translational blockage, and corresponding GFP mRNA / GFP trajectories. (right) GFP mRNA and GFP levels over time, each point corresponding to a cell. GFP mRNA number of molecules is measured by treating fluorescence images with FISH_quant software for FISH experiment dot detection. GFP level is measured as mean GFP fluorescence over the cell area.

The typical doubling time of *S. Cerevisiae* in SD being around or slightly lower than two hours, these results imply the existence of a stochastic process preventing these cells from dividing earlier. Interestingly, we observe that, regardless of the lag time, most cells start expressing GFP and dividing simultaneously. This suggests that the process preventing survivor cells from dividing earlier is linked to gene expression, or at least, also blocks this fundamental ability.

TRANSCRIPTION IS BLOCKED DURING LAG PHASE BEFORE CELL DIVISION RESUMPTION

To further investigate the blockage that prevents survivor cells from expressing a reporter gene, we try to distinguish at which level, transcription or translation, this blockage acts. To do so, we perform RNA fluorescence in situ hybridization (RNA FISH) on freeze-dried yeast cells at different times after rehydration in growth media and before more of them divide (<4h). If survivor cells are able to transcribe throughout this lag phase during which they can't express GFP, we should observe cells with GFP-coding mRNAs and basal level GFP. In figure 4.4 we show that we do not observe such cells and that the GFP and GFP-coding mRNA follow a behavior that seems to indicate that transcription is blocked during the lag phase and resumes shortly before apparition of the first bud.

SORTING CELLS WITH INTACT OUTER MEMBRANE ALLOWS TO INCREASE THE PROPORTION OF SURVIVORS

Many of the previously presented results are obtained on a small proportion of survivor cells. Most cells have a damaged outer membrane, thus no chance of waking up. By ruling out these cells, we can observe many more cells with an intact membrane. This could allow us, among other things, to space cells more in growth after rehydration microscopy experiments, allowing to image them for longer before they are hidden by a neighboring colony originating by a cell with a shorter lag time, or to image more cells of interest using

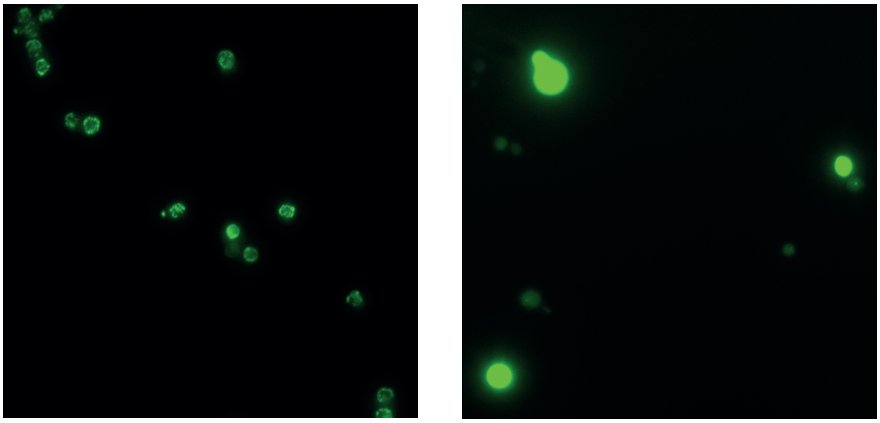


Figure 4.5: Comparison of rhodamine 123 stained starved cells (left) and freeze dried cells right after rehydration (right). Highly fluorescent cells correspond to necrotic cells.

RNA FISH, for which the protocol requires a significant amount of cells and the obtained slides can easily bleach. For example, in the previously presented FISH experiment without sorting, we only observe very few cells with GFP or GFP-coding mRNA per image, making the analysis of the GFP versus number or GFP coding mRNA tedious.

POST-REHYDRATION MITOCHONDRIAL ACTIVITY IS LOWER THAN DURING STARVATION PHASE

Lastly, we test another key component of yeast metabolism: mitochondria. Rhodamine 123 staining probes mitochondrial membrane potential[14], which is the driving force for ATP synthesis[15]. We stain starved yeast cells and freshly rehydrated freeze-dried yeast cells with rhodamine 123. In starved cells, we observe dot-shaped and tubular structures, corresponding to active mitochondria. In rehydrated freeze-dried, we observe very few low intensity dot shapes (see Figure 4.5), indicating a much lower membrane potential than during starvation phase. We can thus hypothesize that mitochondrial function is set to a dormant state when cells are dried and takes time to re-activate, or that mitochondria face some damages that need to be repaired before recovering their activity.

4.3. DISCUSSION

Based on the observation that only a small proportion of yeast cells survive the process of freeze-drying, we presented preliminary results that hint towards the different causes of deaths that can occur during drying and rehydration and the challenges that cells face in restarting their gene expression and dividing activity. While membrane disruption seems to be the most common cause of death, apoptosis also appears to play a part. Interestingly, CDC48 mutants were linked to apoptosis and abnormal shape mutants [16]. We also observe numerous colonies with similar abnormal shapes, which in two

rare occasions did not express the GFP reporter gene at all. We hypothesize, inspired by a study that shows that dry mammalian cells face DNA damages[12], explanations that may motivate further studies: the appearance of those unusually shaped colonies might be due to DNA damages in CDC48 gene or other cell cycle related genes. Similar damages might as well disrupt the inserted genetic construct explaining the few cells that do not express the reporter gene. Cells suffering too severe DNA damages or in key genes, while they may have an intact outer membrane, are unable to express GFP or divide. Finally, these DNA damages might be the origin of the observed apoptotic events. Cells that avoided the two identified possible deaths, accidental membrane disruption and apoptosis, seem to be stuck in a lag phase of variable duration during which their gene expression machinery is halted at the transcription level. If this blockage is in any way linked to DNA damages that cells might face or not, remains an open question. While our results may not be sufficient to support such hypothesis, we suggest further studies to focus on the mechanisms underlying this gene expression blockage and potential DNA damages, as it might reveal principles that are true for other more complex organisms[12].

4.4. METHODS

YEAST STRAIN

All yeast strains in this work are using the haploid GFP-inducible TT7 strain described in [17], based on the wild-type strain W303-K6001 [18] it has the following genome: $MATa; his3{-}11\_15; leu2{-}3\_112; ura3{-}1; ADE2; can1{-}100; HygB; trp\Delta 2; ura3{-}1; pADH1{-}rtTA; pTET07-GFP$. ADE2 and HygB genes were added for selection purposes.

CELL CULTURE, FREEZE-DRYING AND REHYDRATION

Yeast cells were grown overnight in aerobic conditions in YNB media and conserved frozen at $-80°C$ after addition of 50% glycerol. From this frozen glycerol stock, colonies were obtained on a minimum synthetic medium with 2% dextrose (SD) agar plate. Inoculated from a single colony from this plate, 200mL culture was grown in SD in a flask under constant shaking at $30°C$ over 48 hours due to large volume compared to an usual overnight culture. The 200mL of saturated culture were separated in four 50mL flacon tubes and centrifuged at 4000rpm for 5 minutes. Each pellet was resuspended in 10mL PBS and gathered in a single tube. Cells were centrifuged and washed in PBS three times to remove remaining media. After a last centrifugation, washing PBS solution was removed as much as possible. In all presented results except "no protectant" and "trehalose in milliQ" 4.1, cell pellet was resuspended in 1mL of 10% trehalose solution in PBS was added. For the two previously mentioned results, respectively no protectant and 1mL of 10% trehalose in milliQ water was added. The tube containing the resuspended pellet was frozen at $-80°C$ in a diagonal position to maximize surface. The lid was replaced by aluminum foil and the cells were freeze-dried in the same tube and positioned in a Christ Alpha 1-2 LD plus freeze-dryer. Obtained powder of FD yeast cells was stored at $4°C$ for up to a month. For each experiment, a small spatula of the obtained freeze-dried yeast powder was rehydrated by adding the desired medium, pre-warmed at $30°C$ up to optical density 0.3 at 600nm and then diluted to the desired concentration.

MICROSCOPY DIVIDING ABILITY, GFP EXPRESSION AND LAG TIME MEASUREMENTS

After rehydration yeast cells were attached to the concanavalin A coated glass bottom of a 200μL well of a pre-warmed Zell-Kontakt 96-well Fluorcarbon Film Imaging Plate by centrifugating 200μL of rehydrated yeast cells at the desired dilution at 4000rpm for 5 minutes. The plate was maintained at 30°C using a Okolab frame heating system and different non-overlapping fields of view were recorded in brightfield, GFP channel, and mCherry channel for PI staining using a Nikon eclipse Ti inverted microscope. The ability to divide and or express GFP was manually recorded. All cells that were able to divide and/or express GFP were tracked using interpolation between manually drawn ellipse masks. The mean GFP over the mask was extracted over time.

PROPIDIUM IODIDE STAINING AND ANNEXIN V STAINING

For PI and annexin V co-staining we used Tali apoptosis kit and followed the protocol described with this kit, adapting the first steps to use rehydrated freeze-dried cells and the last steps to observe the cells in a 96-well imaging plate. We also stained cells for growth experiment with PI only, as annexin V compromises viability. We used the same PI concentration of 1$\mu g/mL$ as in the kit protocol.

FLUORESCENCE-ACTIVATED CELL SORTING

After staining cells with PI as described above, we sorted them according to mCherry fluorescence using a BD FACSMelody Cell Sorter. Cells were directly collected into pre-warmed desired media, ready to be used with the microscopy method described previously.

RNA FISH

For GFP mRNA FISH in yeast we used the protocol described by stellaris in Stellaris RNA FISH Protocol for S.Cerevisae and in [19], applied in [17]. FISH probes corresponding to GFP mRNA were ordered from Stellaris.

RHODAMINE 123 STAINING AND MITOCHONDRIAL MEMBRANE POTENTIAL

Mitochondrial membrane potential was assessed using the rhodamine 123 staining method described in [20].

REFERENCES

[1] A. Rimbaud, Le dormeur du val (Prometheus Press, 1983).

[2] A. Tunnacliffe and J. Lapinski, *Resurrecting van leeuwenhoek's rotifers: a reappraisal of the role of disaccharides in anhydrobiosis,* Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **358**, 1755 (2003).

[3] L. J. Wickerham and A. Andreasen, *The lyophil process: Its use in the preservation of yeasts,* Wallerstein Laboratories Commununications **5**, 165 (1942).

[4] L. Atkin, W. Moses, and P. P. Gray, *The preservation of yeast cultures by lyophilization,* Journal of Bacteriology **57**, 575 (1949).

[5] J. Wynants, *Preservation of yeast cultures by lyophilization,* Journal of the Institute of Brewing **68**, 350 (1962).

[6] M. A. Mensink, H. W. Frijlink, K. van Der Voort Maarschalk, and W. L. Hinrichs, *How sugars protect proteins in the solid state and during drying (review): Mechanisms of stabilization in relation to stress conditions,* European Journal of Pharmaceutics and Biopharmaceutics **114**, 288 (2017).

[7] J. H. Crowe, L. M. Crowe, and S. A. Jackson, *Preservation of structural and functional activity in lyophilized sarcoplasmic reticulum,* Archives of Biochemistry and Biophysics **220**, 477 (1983).

[8] P. Alpert, *Constraints of tolerance: why are desiccation-tolerant organisms so small or rare?* Journal of Experimental Biology **209**, 1575 (2006).

[9] B. Janis, V. N. Uversky, and M. A. Menze, *Potential functions of lea proteins from the brine shrimp artemia franciscana–anhydrobiosis meets bioinformatics,* Journal of Biomolecular Structure and Dynamics **36**, 3291 (2018).

[10] T. C. Boothby and G. J. Pielak, *Intrinsically disordered proteins and desiccation tolerance: elucidating functional and mechanistic underpinnings of anhydrobiosis,* BioEssays **39**, 1700119 (2017).

[11] J. H. Crowe, L. M. Crowe, W. F. Wolkers, A. E. Oliver, X. Ma, J.-H. Auh, M. Tang, S. Zhu, J. Norris, and F. Tablin, *Stabilization of dry mammalian cells: lessons from nature,* Integrative and Comparative Biology **45**, 810 (2005).

[12] M. Zhang, H. Oldenhof, B. Sydykov, J. Bigalk, H. Sieme, and W. F. Wolkers, *Freeze-drying of mammalian cells using trehalose: preservation of dna integrity,* Scientific reports **7**, 1 (2017).

[13] D. Carmona-Gutierrez, M. A. Bauer, A. Zimmermann, A. Aguilera, N. Austriaco, K. Ayscough, R. Balzan, S. Bar-Nun, A. Barrientos, P. Belenky, et al., *Guidelines and recommendations on yeast cell death nomenclature,* Microbial Cell **5**, 4 (2018).

[14] A. Baracca, G. Sgarbi, G. Solaini, and G. Lenaz, *Rhodamine 123 as a probe of mitochondrial membrane potential: evaluation of proton flux through f0 during atp synthesis,* Biochimica et biophysica acta (BBA)-bioenergetics **1606**, 137 (2003).

[15] L. D. Zorova, V. A. Popkov, E. Y. Plotnikov, D. N. Silachev, I. B. Pevzner, S. S. Jankauskas, V. A. Babenko, S. D. Zorov, A. V. Balakireva, M. Juhaszova, et al., *Mitochondrial membrane potential,* Analytical biochemistry **552**, 50 (2018).

[16] F. Madeo, E. Fröhlich, and K.-U. Fröhlich, *A yeast mutant showing diagnostic markers of early and late apoptosis,* The Journal of cell biology **139**, 729 (1997).

[17] T. Maire, T. Allertz, M. A. Betjes, and H. Youk, *Dormancy-to-death transition in yeast spores occurs due to gradual loss of gene-expressing ability,* Molecular Systems Biology **16**, e9245 (2020).

[18] M. Ralser, H. Kuhl, M. Ralser, M. Werber, H. Lehrach, M. Breitenbach, and B. Timmermann, *The saccharomyces cerevisiae w303-k6001 cross-platform genome sequence: insights into ancestry and physiology of a laboratory mutt,* Open biology **2**, 120093 (2012).

[19] A. Raj, P. Van Den Bogaard, S. A. Rifkin, A. Van Oudenaarden, and S. Tyagi, *Imaging individual mrna molecules using multiple singly labeled probes,* Nature methods **5**, 877 (2008).

[20] P. Ludovico, F. Sansonetty, and M. Côrte-Real, *Assessment of mitochondrial membrane potential in yeast cell populations by flow cytometry,* Microbiology **147**, 3335 (2001).

5

CONCLUSION

Throughout this thesis, we have studied systems and phenomena that allow micro-organisms to maximize their growth in challenging nutritional conditions, but also to completely halt growth and withstand the removal water, the most essential compound for life. We reviewed, in Chapter 1, knowledge on the bacterial guanosine tetraphosphate (ppGpp) signaling system, responsible for bacterial response to stress and starvation, with a focus on its regulation on compounds or processes playing a role in bacterial steady-state growth. Then, we presented in Chapter 2 novel results of impact of ppGpp level perturbation on concentrations of proteins and metabolites which are of significant importance to bacterial growth. With this result, we determined the scope of ppGpp regulation in steady-state growth independently from nutritional effects. In Chapter 3 we developed a coarse-grained model of bacterial growth, the effect of ppGpp on processes relevant to growth, and the perturbations of ppGpp concentrations. By doing so, we refuted some of our intuitions and identified conditions that a model should verify to be able to explain the effect of ppGpp perturbation on bacterial growth rate. Finally, in Chapter 4 we freeze-dried *Saccharomyces cerevisiae* and rehydrated them while inducing expression of a reporter gene under the microscope to identify what allows some of these yeast cells to resume their lives.

PPGPP IS REPONSIBLE FOR ONLY PART OF PROTEOME CHANGES IN LIMITING NUTRIENTS

In Chapter 2, we confirmed that artificially increasing ppGpp leads to less abundant ribosomes. While this is a significant change in the way resources are allocated in the proteome, it is only part of the changes observed when growth is limited by the quality of the available carbon source. Changes in abundance of other groups of proteins happening in different nutrient conditions are in large parts happening independently of ppGpp concentration.

MASS OF CATABOLIC PROTEINS IS REGULATED INDEPENDENTLY OF PPGPP

One of the most surprising result we obtained when inducing an excess of ppGpp, is the unchanged mass fraction of the proteome responsible for carbon input. This conflicts with the previously proposed idea that ppGpp balances growth and metabolism and

reveals the limits of the ppGpp regulation: the proteome mass fraction of catabolic proteins is regulated independently of ppGpp.

PPGPP DOES NOT CONTROL AMINO ACID BIOSYNTHESIS IN STEADY-STATE GROWTH

While ppGpp is known to be necessary to synthesize some of amino acid biosynthesis proteins, artificially titrating its concentration has no effect on their total abundance. This surprising result seems to indicate that only a small concentration of ppGpp is enough for activation of these genes by ppGpp. The increase in abundance of these proteins in faster growth is thus likely driven by other processes such as the known feedback inhibition of amino acids on the pathways operating their own synthesis.

INSUFFICIENT PPGPP CONCETRATION MIGHT LEAD TO SINGLE AMINO ACID STARVATION

While the reason for slower growth when ppGpp concentration is artificially increased is easy to grasp (translational flux is limited by insufficient ribosomes), the ones leading to a slow down in depleted ppGpp remained more mysterious. Our results show that this perturbation leads to a strong upregulation of proteins responsible for a single nutritional condition dependent amino acid pathway. This suggests that depleting ppGpp might lead to starvation of a single amino acid depending on the condition.

5

ABSOLUTE PPGPP CONCENTRATION DOES NOT FULLY DETERMINE RIBOSOME ABUNDANCE

In Chapter 3, we seeked an explanation to the observation that much larger concentrations of ppGpp are required to slow down growth trough excess ppGpp than the ones found in the natural case. We found that this phenomenon happens because the ppGpp concentration does not fully determine ribosome abundance, which is likely the main determinant for growth rate in this condition. This finding challenges models of the ppGpp response assuming that ribosome content depends on the concentration of ppGpp only.

RIBOSOME ABUNDANCE MIGHT BE REGULATED PASSIVELY BY UPREGULATION OF CATABOLIC GENES

We also investigated, through modeling and experimental results, compounds and processes that might play a role in the regulation of ribosome content. We found that most known regulators and processes known or suspected to impact ribosome abundance do not seem likely to give rise to the observed ribosome content in excess ppGpp. Following this observations, we suggested that the upregulation of catabolic proteins in poor nutrients passively downregulates ribosome abundance, for example through lower availability of RNA polymerases. This regulation would not be present in excess ppGpp and explain why higher ribosome content and growth rates are observed than in the natural case for a similar ppGpp concentration.

HAS THE SPOT LOST ITS MAGIC?

Because of its first unknown nature and its large amount of targets, ppGpp gained the nickname of "magic spot". With our results nuancing the scope and strength of ppGpp regulation comes the question: has it lost its magic? It seems that ppGpp's main role is to regulate translation-related proteins to satisfy a offer and demand tradeoff: having the right amount of these proteins to sustain the translation flux to grow as fast as possible in the current conditions. Not so magic after all? It seems that the magic now lies mostly around the famous regulator. What are the quantitative laws describing its coordination with regulators of catabolic proteins and amino acid synthesis? What else is regulating

ribosome content when ppGpp concentration is artificially increased? Does this regulator or process also participate in scaling ribosome content with growth rate?

Yeasts need to preserve their cellular membrane and their ability to transcribe to survive desiccation

Finally, we presented another story investigating the conditions and mechanisms for survival of yeast cells in drying conditions. By freeze-drying yeast cells, and imaging them right after rehydration while expressing a reporter gene, we identified two challenges yeast cells face when attempting to survive in a dry state during which their replicative life is halted. Firstly, a vast majority of these cells seem to loose integrity of their outer membrane, resulting in their death. However, not all cells preserving a intact membrane are able to divide. We show that some cells with an intact membrane are unable to divide and to express the reporter gene due to a blockage at the transcription level. In surviving yeasts, this blockage seems to be lifted after a highly variable lag time.

Final thoughts

Through these three studies, we brought insights on the existing strategies some microorganisms adopt to maximize growth and completely halt it to survive one of the harshest stress. We also suggested some directions to investigate these systems and the questions our results raised. Hopefully, we also gave the reader a glimpse of how much is still to unveil about growth and survival of microorganisms in challenging conditions.

LIST OF PUBLICATIONS

1. Sallé, J., Xie, J., Ershov, D., **Lacassin, M.**, Dmitrieff, S., & Minc, N. (2019). Asymmetric division through a reduction of microtubule centering forces. Journal of Cell Biology, 218(3), 771–782.

2. Ripamonti, A.[1], **Lacassin, M.**[1], Droghetti, R., Bokinsky, G.[2], & Lagomarsino, M. C.[2] (2025). Transcriptional competition biases the effects of second messengers in *E. coli*. bioRxiv, 2025-04 (under revision).

[1] Equal contribution.
[2] Equal contribution.

ACKNOWLEDGEMENTS

Setting in words results, methods and conclusions in a book makes them appear somewhat linear. However, those who have followed me through it and have worked on similarly large and personal projects know the wanderings, the slips, and the tumbles. What they often realize less, is how much they have contributed to this work through mentoring, support of many kinds and various fulfilling moments that allow resilience. No word could cover everything people have done for me during this project, but my hope is that you find in them a moment of happiness, pride, and motivation to keep being the great people that you are. If you are reading this and you do not find your name, you are still in my thoughts.

First of all, I would like to adress my thanks and compliments to my daily supervisor **Greg**. You jumped in to literally save my PhD and I truly believe you did it out of kindness and responsibility towards fellow scientists. Your humility, integrity and openness were an absolute gift. Your huge dedication will stay an inspiration for me and many others. I wish you all the fun and success with your personal and scientific projects! Another massive source of support and inspiration came from you, **Yaroslav**. Your down-to-earthness, constancy and massive desire to help were such a big relief. I would never think of orienting a PhD candidate in need of well-chosen advice to someone else than you. Your commitment to speaking French with me and learning about biological concepts is admirable. Mentoring is hard and one does not always realize how much impact they have when helping someone going through a difficult time. **Arjen**, my deepest thanks for your advice and well-chosen words.

The kindest and most soothing of all support came from you, **Konstantina**. No words can thank you enough for the endless generosity and tolerance you have blessed me with. Your rock-hard resilience and expressive communication not only allowed me to get through the downs, they also brought tremendous amounts of joys and laughs. The cover of this thesis, like everything with your touch, is magnificent. My biggest wish is that you find the successes and rewards you deserve. Thank you so very much!

Behind every scientist, there is an education that fosters curiosity from the very beginning. My parents, **Robin** & **Dominique**, among a billion other things, gave me a big dose of that. They also gave me loads of trust, support, and freedom which have helped me move to a different country while learning so much on many levels.

This work would have been impossible without the support of my labmates. I would like to address my thanks to all my colleagues and in particular to **Adja** for your precious lab and emotional support that came with a whole bunch of laughs; to **Jaïrus** for all the fun and interesting talks; to **Ferhat** for teaching me how to collect ppGpp samples instead of learning it from the method section of an article, which, on top of being pretty much impossible, would certainly not be as much fun as learning from you. Many thanks to **Flora** for teaching me how to use the LC-MS with great patience and uplifting enthousiasm. **Marek**, your massive knowledge, attention to detail and time helping your

former colleagues were so helpful. **Théo**, for making me feel welcome and supported from my very first day in the Netherlands, your inspiring curiosity, and all the fun around our shared passion for plants; **Hirad**, for the tears that came with your overflowing humour; **Diego**, for all the fun during our outings; **Diederik**, for the lovely chats we had around delicious beers; **Yiteng**, for sharing nice chats and skiing adventures with me; **Lars**, for the inspiration that came from your never-ending energy and trueness.

Science is nothing without good collaborations. Thankfully, I was lucky enough to have great collaborators. **Rossana** & **Marco**, thank you for the very interesting and helpful talks trying to make sense of the data I collected, as well as the meticulous work of testing our intuitions and proving the ones that turned out to be true. **Andrea**, I am very grateful for your hard work writing down our article. Best measurements come from experts like **Maxime** and **Martin**. Thank you for the precious help measuring the proteins in my samples. Collaborating with you has been both very helpful and fun.

I would also like to thank all my new ALTEN colleagues and in particular **Jelmer**, **Dexter**, **Bernhard**, **Shelly**, and **Jorn**. Your trust and understanding ear provide a great environment for me to keep trying new things and grow from it. I am very grateful for it.

Some people have a friend that is family because they have been there for you for so long and you know you can always count on each other. I am lucky to be one of these people. **Frédéric**, your reliability is amazing. During a Phd, active moments of evasion are precious. **Koen** & **Bénédicte**, thank you so much for the lovely adventures and invitations to visit you, always in the most lovely places, from Cascais to Utrecht.

All doctors know that an important part of a PhD is playing hard when possible. For that, I believe, there are no better partners than my chosen family. With you, **Nicole**, hilarious jokes only end when the next one starts and you always have an understanding ear for your friends. **Quentin**, you are a never ending source of positive energy and sweet support. **Sonakshi**, we have lived so many fun adventures together and your determination is an inspiration. **Luca**, your smart and approachable mind is always lovely to be around of. **Stephan**, your curiosity and imagination is bringing me so much. **Thomas**, thank you so much for all the fun and positivity you carry around and throw on all of us. **Pauline**, thank you for being your constant hilarious self, amazing to have you around. **Deniz**, your spontaneity and discipline makes us all feel welcome and supported. **Eleni**, you bring some extra fun into every moment. **Rajarshi**, you are just a ball of sweetness. **Will**, always the goofiest funniest word comes from you. **Gert-Jan**, always a big dose of fun hanging out with you. **Deepika**, you have such a sweet and kind soul. **Su**, you always make time fly in a light happy mood. **Nina**, you always have a sweet word and some energy for all of us.

A fun, cosy and accepting life at home is key to a successful PhD. Because of that, my Delft housemates have been an important brick in my PhD track. **Stefan**, thank you for all the chats, walks, dinners and parties. Your presence is both pure fun and enriching. **Michiel**, you are so kind and full of calmness. Spending time with you at home or bouldering together is always a lot of fun. **Daniel**, thank you so much for the hilarious jokes and lovely chats we had.

Finally, if we are to talk about life at home and though you will never read this – dear reader don't worry no one died... **Woody** is just my cat! I would like to thank you because your insouciance is heartwarming and you are such sweet company.

Я
Я
Я
Я
Я